

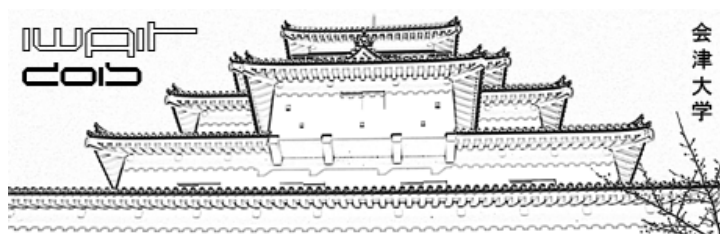


***Proceedings of the
International
Workshop on
Applications in
Information
Technology
(IWAIT-2015)***

October 8 – 10, 2015
University of Aizu
Aizu-Wakamatsu
Japan

The University of Aizu Press
2015

Proceedings of the
**International Workshop on
Applications in
Information Technology**



Aizu-Wakamatsu, Japan October 8 – 10, 2015

Edited by
Evgeny Pyshkin
Vitaly Klyuev

Hosted by
The University of Aizu, Japan

In Cooperation with
St. Petersburg State University, Russia
Peter the Great St. Petersburg Polytechnic University, Russia



The University of Aizu Press
2015



会津大学

Copyright © The University of Aizu Press, 2015
All rights reserved

ISBN 978-4-900721-03-6

Additional copies may be ordered from:
The University of Aizu Press
University of Aizu
Aizu-Wakamatsu, Fukushima-ken 965-8580
Japan

International Workshop on Applications in Information Technology

October 8 – 10, 2015



Table of Contents

Organizing and Program Committees	vi
Greeting from the President of the University of Aizu	x
Welcome from the Workshop and Program Committee Chairs	xi
Workshop Information.....	xii
Audio Files Compression with the Variational Method of Identification of Modeling Difference Equations	
<i>Eugenia M. Khassina, Andrei A. Lomov</i>	1
Visualization of Execution Paths for Concurrent Programs	
<i>Andrei Eleshevich, Marat Akhin</i>	5
Intermittent Control Properties of Car Following: I. Driving Simulator Experiments	
<i>Hiromasa Ando, Ryoji Yamauchi, Ihor Lubashevsky, Arkady Zgonnikov</i>	8
Intermittent Control Properties of Car Following: II. Dynamical Trap Model	
<i>Ryoji Yamauchi, Hiromasa Ando, Ihor Lubashevsky, Arkady Zgonnikov</i>	10
Admissible Trajectory Planning for Wheeled Mobile Robot	
<i>Nelli Korotkova , Evgeny Veremey</i>	13
Constraining Operation Delay for Dynamic Power Optimization of Asynchronous Circuits	
<i>Shunya Hosaka, Hiroshi Saito</i>	16
Hybrid Tsunami Modeling Infrastructure: Tsunami Source Data and Bathymetry Editor	
<i>Shunsuke Takano, Kensaku Hayashi, Alexander Vazhenin, Andrey Marchuk</i>	21
Traveler Guide Assistant: Introducing an Application for an OpenStreetMap Based Travel Itinerary Construction	
<i>Alexander Baratynskiy, Evgeny Pyshkin</i>	25

Automated Leisure Walk Route Generation for an Interactive Travel Planner <i>Boris Skripal, Evgeny Pyshkin</i>	29
Human Response Delay in Balancing Virtual Pendulums with Overdamped Dynamics <i>Takashi Suzuki, Shigeru Kanemoto, Ihor Lubashevsky</i>	33
Effect of Extended Visual Feedback on Human Information Processing in Virtual Stick Balancing <i>Takashi Suzuki, Arkady Zgonnikov, Shigeru Kanemoto, Ihor Lubashevsky</i>	35
Is the Modern Theory of Stochastic Processes Complete? Example of Markovian Random Walks with Constant Non-Symmetric Diffusion Coefficients <i>Kosuke Hijikata, Ihor Lubashevsky, Alexander Vazhenin</i>	39
Asymptotics of Psychometric Function of Gray Color Categorization <i>Marie Watanabe, Ren Namae, Ihor Lubashevsky</i>	41
3D Automated Anatomically Correct Face Reconstruction: Facial Feature Points Positioning and Motion Analysis <i>Iurii Dorofeev, Mikhail Glukhikh</i>	43
Design of Automatic Speech Emotion Recognition System <i>Elena Dmitrieva, Kirill Nikitin</i>	47
Using Dynamic Predicate Logic for Pronominal Anaphora Resolution in Russian Texts <i>Nikita Gerasimov, Evgeny Pyshkin</i>	51
A Program Complex for Learning of Optimal Control Problems <i>Andrey Klyuenkov</i>	55
Study of Patterns in the Hyperlink Structure of Large Sites <i>Evgenii Klemeshov, Ivan Blekanov, Sergei Sergeev</i>	58
Hierarchical Clustering of Large Text Datasets Using Locality-Sensitive Hashing <i>Vasilii Korelin, Ivan Blekanov</i>	61
Developing a Mobile Application for Wine Amateurs <i>Anton Kiselev, Andrey Kuznetsov</i>	65
Multicriteria Regulation of Investments in the Economy of the Russian Federation <i>Alexander Popkov</i>	68
Modeling Technical Systems with smartIfflow for Safety Related Tasks <i>Philipp Hönig, Rüdiger Lunde, Florian Holzapfel</i>	71

Fragments Video Protection Mechanism for Video on Demand Server <i>Aleksandr N. Matlash, Sergey Y. Sevryukov</i>	76
Towards Self-Learning AI for the Videogame of Tennis <i>Akane Yamada, Maxim Mozgovoy</i>	79
Realistic Ball Motion Model for a Tennis Videogame <i>Andrey Lopukhov, Alexander Serikov</i>	81
An Evaluation of Classification Accuracy in a Multilayer Perceptron <i>Hiroaki Yui, Subhash Bhalla</i>	84
Compact Algorithm of Extrema Count in the Interference Measurement of Transparent Film's Thickness <i>Alexander V. Boldyrev, Igor V. Gonchar, Alexey S. Ivanov, Alexander B. Fedortsov</i>	87
OD-matrix estimation for urban traffic area control <i>Anastasiya Raevskaya, Alexander Krylatov</i>	91
Developing Google Chrome Extensions: Case Study <i>Kentaro Murai, Vitaly Klyuev</i>	94
3D Interface for Easy Operation of Disaster Management Robot Teleportation <i>Venushka Dharmasiri, Thilak Chamindra</i>	97
Entity Resolution using Co-occurrence Graph and Continuous Learning <i>Anoop Kumar Pandey, Srinath Srinivasa</i>	101
Data Collection for Investigation of Reliable Reviews <i>Jun Kikuchi, Vitaly Klyuev</i>	104
Personal Data Leakage: Android Case Study <i>Yoichi Saito, Vitaly Klyuev</i>	107
A Consideration of Query Improvement to the SMOKA Astronomical Archive System <i>Yilang Wu, Wanming Chu, Subhash Bhalla</i>	109
The Time Synchronization Model of Sensors and Reference-Transponders for the Aircraft Navigation System <i>Dmitriy Novopashin</i>	111
Author Index.....	114

Organizing and Program Committees



Honorary Chairs

Ryuichi Oka

*President,
University of Aizu, Japan*

Leon Petrosjan

*Dean,
Faculty of Applied mathematics –
Control Processes,
St. Petersburg State University,
Russia*

Mikhail Okrepilov

*Department Chair,
Measuring Information
Technologies Department,
Peter the Great St. Petersburg
Polytechnic University, Russia*

Workshop Chair

Vitaly Klyuev

University of Aizu, Japan

Program Committee Chairs

Evgeny Pyshkin

*Peter the Great St. Petersburg
Polytechnic University, Russia*

Maxim Mozgovoy

University of Aizu, Japan

Nikolay Smirnov

*St. Petersburg State University,
Russia*

Program Committee

Marat Akhin

*Peter the Great St. Petersburg
Polytechnic University, Russia*

Subhash Bhalla

University of Aizu, Japan

Ivan Blekanov

*St. Petersburg State University,
Russia*

John Brine

University of Aizu, Japan

**Program Committee
(cont.)**

Alfredo Capozucca	<i>University of Luxembourg, Luxembourg</i>
Vlatko Davidovski	<i>Cognizant Business Consulting, Switzerland</i>
Vladimir Dobrynin	<i>St. Petersburg State University, Russia</i>
Vladimir Filipovskii	<i>Peter the Great St. Petersburg Polytechnic University, Russia</i>
Maria Ganzha	<i>System Research Institute, Polish Academy of Sciences, Poland</i>
Mikhail Glukhikh	<i>JetBrains, Peter the Great St. Petersburg Polytechnic University, Russia</i>
Nicolas Guelfi	<i>University of Luxembourg, Luxembourg</i>
Mohamed Hamada	<i>University of Aizu, Japan</i>
Yannis Haralambous	<i>Institut Mines-Télécom, Télécom Bretagne, France</i>
Vladimir Itsykson	<i>Peter the Great St. Petersburg Polytechnic University, Russia</i>
Qun Jin	<i>Waseda University, Japan</i>
Vitaly Klyuev	<i>University of Aizu, Japan</i>
Andrey Kuznetsov	<i>Motorola Solutions Inc., Peter the Great St. Petersburg Polytechnic University, Russia</i>
Rüdiger Lunde	<i>Ulm University of Applied Sciences, Germany</i>

**Program Committee
(cont.)**

Viacheslav Marakhovsky	<i>Peter the Great St. Petersburg Polytechnic University, Russia</i>
Maxim Mozgovoy	<i>University of Aizu, Japan</i>
Kirill Nikitin	<i>Peter the Great St. Petersburg Polytechnic University, Russia</i>
Kendall Nygard	<i>North Dakota State University, USA</i>
Kohei Ohno	<i>Meiji University, Japan</i>
Mikhail Okrepilov	<i>Peter the Great St. Petersburg Polytechnic University, Russia</i>
Marcin Paprzycki	<i>System Research Institute, Polish Academy of Sciences, Poland</i>
Evgeny Pyshkin	<i>Peter the Great St. Petersburg Polytechnic University, Russia</i>
Benoît Ries	<i>University of Luxembourg, Luxembourg</i>
Hiroshi Saito	<i>University of Aizu, Japan</i>
Sergei Sergeev	<i>St. Petersburg State University, Russia</i>
Sergei Sevryukov	<i>St. Petersburg State University, Russia</i>
Nikolay Smirnov	<i>St. Petersburg State University, Russia</i>
Kari Smolander	<i>Lappeenranta University of Technology, Finland</i>
Margarita Sotnikova	<i>St. Petersburg State University, Russia</i>
Alexander Vazhenin	<i>University of Aizu, Japan</i>

Program Committee (cont.)	Evgeny Veremey	<i>St. Petersburg State University, Russia</i>
	Vladimir Zaborovsky	<i>Peter the Great St. Petersburg Polytechnic University, Russia</i>
Proceedings Editors	Evgeny Pyshkin	<i>Peter the Great St. Petersburg Polytechnic University, Russia</i>
	Vitaly Klyuev	<i>University of Aizu, Japan</i>
Editor-in-Chief	Yasuhiro Abe	<i>The University of Aizu Press, University of Aizu, Japan</i>
Proceedings Cover Designer	Evgeny Pyshkin	<i>Peter the Great St. Petersburg Polytechnic University, Russia</i>
Local Arrangements Team	Yuriko Nagashima	<i>University of Aizu, Japan</i>
	Aya Anazawa	<i>University of Aizu, Japan</i>
St. Petersburg Session Manager	Nina Tamm	<i>Peter the Great St. Petersburg Polytechnic University, Russia</i>



Greeting from the President of the University of Aizu

It is my pleasure to welcome you to the International Workshop on Applications in Information Technology.

In 2014, the University of Aizu was selected for the MEXT's Top Global University Program. This workshop is organized within the framework of a new university educational program.

Initially, we conceived this workshop as a place for the first scientific presentations of the best students of the University of Aizu and partner universities from Russia. This workshop creates the necessary conditions to keep international scientific contacts at the student level. This event is a training ground for both Japanese and foreign students on their way to becoming scientists. Common discussions between professors and students at the workshop allow our students to learn how to conduct research, how to prepare results, how to present them to the public as well how to prove the results of their achievements.

We are glad to know that the rationale for this workshop motivated students from Germany, India, and Sri-Lanka to participate in it.

We want to offer our special thanks to the professors, researchers and students of Saint-Petersburg State University (Russia), Peter the Great Saint-Petersburg Polytechnic University (Russia), Novosibirsk National Research State University (Russia), Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences, Ulm University of Applied Sciences (Germany), Motorola Solutions Inc., the Informatics Institute of Technology (Sri Lanka), Saint-Petersburg State Mining Institute (Russia), International Institute of Information Technology Bangalore (India), and National University of Science and Technology (Russia) for their participation in this event. Their contributions have been very helpful in making this workshop successful.

We also extend special thanks to our international team of reviewers representing many respected universities from all over the world. Their comments and suggestions were very helpful in validating the quality of the submissions.

Finally, we would like to thank all of the participants for their contributions and efforts to maintain the high scientific level of the International Workshop on Applications in Information Technology.

A handwritten signature in black ink, reading 'Ryuichi Oka' in a cursive style.

Ryuichi Oka, Ph.D.

President of the University of Aizu



Welcome from the Workshop and Program Committee Chairs

The International Workshop on Applications in Information Technology 2015 (IWAIT-2015) is an excellent example of the ongoing and productive collaboration among the international participants – the University of Aizu, Saint-Petersburg State University, and Peter the Great Saint-Petersburg Polytechnic University. Moreover, a desirable component of education at any university is the involvement of students in international scientific research. We are pleased that a significant outcome of our inter-university cooperation is the international participation among our students. This development is an important step forward and anticipates the future growth of the shared research activities among our partners.



This workshop brings together international students who work on the design, development, use, and evaluation of information technology applications.

IWAIT-2015 received a number of student submissions from different countries. All submissions were peer-reviewed in a single-blind review process. A total of 35 papers were selected for presentation at the workshop. Most reviewers provided detailed comments to the authors, making it a valuable experience for the workshop participants. Students from Japan, Russia, Germany, India, and Sri-Lanka have a great opportunity to discuss the results obtained, to exchange their opinions and views, and as the most important outcome, to establish new international scientific contacts. We believe that some students will maintain their cooperative relationship into the future and perhaps throughout their careers.

We would like to thank our outstanding team of volunteer educators and support staff who have worked very hard to put together IWAIT-2015. We acknowledge the invaluable assistance of the international referees. The assistance and support of the aforementioned teams have been very helpful in making IWAIT-2015 successful.

We hope you enjoy IWAIT-2015!

Vitaly Klyuev, IWAIT-2015 Workshop Chair

Ph.D., Professor, University of Aizu

Evgeny Pyshkin, IWAIT-2015 Program Committee Chair

Ph.D., Senior Associate Professor, Peter the Great St. Petersburg Polytechnic University

Workshop Information



MISSION STATEMENT

The main objective of this workshop is to foster rich creativity in students' research works and to encourage students to participate actively in open discussions with their colleagues. This event is a place for the first scientific presentations of the best students of the universities participating in this project. This workshop creates the necessary conditions to keep international scientific contacts at the student level. We do hope that some of students participating in this workshop will tightly collaborate in their future projects.

2015 CALL FOR PAPERS SUMMARY

The 2015 International Workshop on Applications in Information Technology invited student papers presenting new advances and research results in the fields of information technology applications. The workshops is aimed to bring together under-graduate, graduate, master and Ph.D. students, academic scientists, researchers and scholars in the domains which include, but are not limited to:

- Software Engineering
- Computer Systems and Architectures
- Web Technology
- Control Systems
- Cloud Computing
- Human-Centric Computing
- Information Retrieval and Data Mining
- Information Technology
- Intelligent Systems
- Mobile Applications
- Knowledge Engineering
- Machine Learning
- Machine Vision and Pattern Recognition
- Modeling and Simulation
- Game Development
- Visualization
- Computer Assisted Learning
- Learning Environments and Applications



WORKSHOP INFORMATION SPONSOR

St. Petersburg State Polytechnic University Journal.

Computer Science, Telecommunication and Control Systems

<http://ntv.spbstu.ru/telecom/>

Audio Files Compression with the Variational Method of Identification of Modeling Difference Equations*

Eugenia M. Khassina
Novosibirsk National Research State University
Department of Information Technologies
Novosibirsk, Russia
jenya-100@yandex.ru

Andrei A. Lomov[†]
Sobolev Institute of Mathematics of the Siberian
Branch of the Russian Academy of Sciences
Novosibirsk National Research State University
Novosibirsk, Russia
lomov@math.nsc.ru

ABSTRACT

In this paper we consider the variational audio compression algorithm based on signal modeling with solutions of linear difference equations from a certain parametric family in the time domain using the STLS cost function. The identification of a modeling difference equation parameters for each frame of an audio file signal allows one to perform compression of the file representing the signal frames in the adaptive Laplace basis of exponentially damped sinusoids. Such an approach better reflects the physics of audio signals generated by real musical instruments than the traditional Fourier representation of the signals with harmonics that is used in some popular audio codecs such as OGG Vorbis and MP3.

Categories and Subject Descriptors

G.1.2 [Numerical Analysis]: Approximation—*structured total least squares approximation*; G.1.6 [Numerical Analysis]: Optimization—*least squares methods*; H.5.5 [Information interfaces and presentation]: Sound and Music Computing—*modeling*

Keywords

audio signals modeling, audio codec, exponential sinusoidal model, parametric identification, difference equations, variational identification method, structured total least squares

1. INTRODUCTION

Lossy audio codecs such as MPEG-1 codecs (MP1, MP2 and MP3) and OGG Vorbis tend to decompose an audio signal into harmonics. However, the method often does not correspond to the physical nature of sounds produced by

*The work has been supported by the Russian Foundation for Basic Research (project no. 13-01-00329).

[†]Research supervisor of the work.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

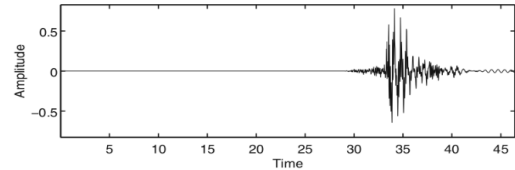


Figure 1: Transient [1].

conventional musical instruments, for which the presence of a considerable quantity of transients (high-amplitude, short-duration sound segments followed by an exponential decay) is common. If a piece of a signal contains transients it can not be considered as a quasi-stationary episode. That is why MP3 audio codec, for instance, has to use the Modified Discrete Cosine Transform (MDCT) with windows of varied length while processing a signal. Using a short-window mode of the encoding scheme allows to avoid what is commonly referred to as a pre-echo artifact [1].

In Figure 1 a transient can be seen. Though it is more natural to decompose such signals into a sum of exponentially damped sinusoids, rather than to represent them as a Fourier series, such an approach requires a considerable amount of CPU resources. This is the reason why the creation and the usage of codecs based on the principle have only recently become justified.

In formula (1) below, audio signal $s(n)$ is represented as a superposition of slowly time-varying exponentially weighted sinusoids and quasi-stationary noise $\eta(n)$. The signal model is called the Exponential Sinusoidal Model (ESM). The frequencies ω_i , phases ϕ_i , amplitudes a_i and damping parameters γ_i can be obtained without switching to the frequency domain.

$$s(n) \approx \sum_{i=1}^K a_i(n) e^{-\gamma_i(n)n} \sin(\omega_i(n)n + \phi_i(n)) + \eta(n). \quad (1)$$

In [2] a Total Least Squares (TLS) problem of order $2K$ is solved to estimate ω_i and γ_i , where K is a number of available exponentially damped sinusoids predefined by a user. On the basis of the TLS-ESM scheme an experimental audio codec was created and tested [2]. An essential disadvantage of the method, however, is that TLS tries to make the original and the modeled signals as close to each other as possible in the time domain not on the whole frame of the signal but on separate sets of $2K$ samples along the frame, considering

the sets independent.

The goal of the work is to research an audio compression algorithm which decomposes a signal into exponentially damped sinusoids in the time domain without switching to the frequency domain. We use an approximation of an audio signal on the whole audio frame using the variational identification method [5, 6, 7], close to the Structured Total Least Squares (STLS) [8] and the Global Total Least Squares (GTLS) [9] methods.

In [3] a vocoder based on ESM-STLS scheme was proposed. The testing of the vocoder was performed in comparison to Code-excited linear prediction (CELP), a standard speech coding algorithm. The results of the testing showed that, providing a similar compression ratio, the new vocoder has a substantially higher signal-to-noise ratio (SNR). However, speech spectrum is rather simple that makes speech signals easily compressible. Our experiments showed that Newton's iterative algorithms, to which STLS1 and STLS2 used in [3] belong, have bad convergence when solving the STLS problem for music audio files with wider spectra.

2. THEORETICAL ASPECTS

2.1 Coding of an audio frame

Our algorithm divides the whole signal of an audio file into frames of N samples each, processing the frames one by one. In section 3 we will consider how N value and other parameters are chosen. Conventionally N equals 100. By $s[k]$, $k = 1, N$ denote a frame of samples.

We will treat the vector $s \doteq (s[1]; \dots; s[N])$ as a perturbed observation of a solution process $z \doteq (z[1]; \dots; z[N])$ of a certain homogeneous linear difference equation with real coefficients. We will solve the inverse problem of identifying the unknown coefficients of the equation. Let us take as an example a difference equation of order $p = 3$:

$$\begin{aligned} z[k+3] + \alpha_2 z[k+2] + \alpha_1 z[k+1] + \alpha_0 z[k] &= 0, \\ k &= \overline{1, N-3}. \end{aligned} \quad (2)$$

Denote the characteristic roots of the system (2) by ξ_i . The characteristic polynomial of the system is real, hence all the roots that are complex should occur in complex-conjugate pairs. For the present example, suppose that the single real root of the characteristic polynomial is ξ_2 . We are interested only in real solutions of the difference equation, as audio samples of observation s are real. Therefore, we can transform the general solution of (2) to the real form as follows:

$$\begin{aligned} z[k] &= C_1 \xi_1^k + \bar{C}_1 \bar{\xi}_1^k + C_2 \xi_2^k = \\ &= A_1 \rho_1^k \cos(k\omega_1) + A_2 \rho_1^k \sin(k\omega_1) + A_3 \rho_2^k, \quad \forall i \ A_i \in \mathbb{R}. \end{aligned} \quad (3)$$

We introduce the following notation for the vector of the coefficients of the difference equation:

$$\gamma \doteq (\alpha_0 \ \alpha_1 \ \alpha_2 \ 1)^\top.$$

Let us define the objective function for the identification of vector γ :

$$J(\gamma) = \|s - z(\gamma)\|^2, \quad z(\gamma) \doteq \arg \min_{z: (2)} \|s - z\|^2. \quad (4)$$

This variational problem was first formulated and solved by A. O. Egorshin [5, 6]. For its numerical solution we apply the iterative algorithm with an updating inverse matrix proposed by A. O. Egorshin and, independently, by M. R. Osborne

[10]. As the initial γ for the iterative algorithm we use the least-squares estimate γ_{LS} [11].

To describe the minimization iterative algorithm, first, let us notice that the difference equation (2) can be transformed to the matrix form:

$$\underbrace{\begin{pmatrix} \alpha_0 & \alpha_1 & \alpha_2 & 1 & 0 \\ 0 & \alpha_0 & \alpha_1 & \alpha_2 & 1 \\ \vdots & & \ddots & \ddots & \ddots \\ 0 & & & \alpha_0 & \alpha_1 & \alpha_2 & 1 \end{pmatrix}}_G \underbrace{\begin{pmatrix} z[1] \\ z[2] \\ \vdots \\ z[N] \end{pmatrix}}_z = 0. \quad (5)$$

Then we use the identity $G_\gamma s \equiv V(s)\gamma$, where V is a Hankel matrix:

$$\begin{aligned} &\underbrace{\begin{pmatrix} \alpha_0 & \alpha_1 & \alpha_2 & 1 & 0 \\ 0 & \alpha_0 & \alpha_1 & \alpha_2 & 1 \\ \vdots & & \ddots & \ddots & \ddots \\ 0 & & & \alpha_0 & \alpha_1 & \alpha_2 & 1 \end{pmatrix}}_{G_\gamma} \underbrace{\begin{pmatrix} s[1] \\ s[2] \\ \vdots \\ s[N] \end{pmatrix}}_s \equiv \\ &\equiv \underbrace{\begin{pmatrix} s[1] & s[2] & s[3] & s[4] \\ s[2] & s[3] & s[4] & s[5] \\ \vdots & \vdots & \vdots & \vdots \\ s[N-3] & s[N-2] & s[N-1] & s[N] \end{pmatrix}}_{V_1} \underbrace{\begin{pmatrix} s[4] \\ s[5] \\ \vdots \\ s[N] \end{pmatrix}}_{V_2} \underbrace{\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ 1 \end{pmatrix}}_\gamma. \end{aligned} \quad (6)$$

Now, the iterations with an updating inverse matrix which solve the variational identification task (4) are:

1. The initial value: $\gamma = \gamma(0) = \gamma_{LS}$.
2. For $k \geq 0$

$$\begin{cases} \tau = \left(V(s)^\top (G_{\gamma(k)} G_{\gamma(k)}^\top)^{-1} V(s) \right)^{-1} \cdot \gamma(k), \\ \gamma(k+1) = \frac{1}{(0 \dots 01)\tau} \tau. \end{cases} \quad (7)$$

The last row means the division of the whole auxiliary vector τ by its last element in order to make the last element of vector $\gamma(k+1)$ equal to unity. The main difference of the Egorshin—Osborne iterations from the computational TLS algorithm consists in the presence of the inverse matrix $(G_{\gamma(k)} G_{\gamma(k)}^\top)^{-1}$ which is updated on each iteration.

Using the calculated estimate for γ , we find the modeling process $z(\gamma)$ nearest to the observation s as the linear projection [5, 6]:

$$z(\gamma) = \left(I - G_\gamma^\top (G_\gamma G_\gamma^\top)^{-1} G_\gamma \right) \cdot s, \quad (8)$$

where I is an identity matrix.

The obtained coefficient vector γ and the corresponding process $z(\gamma)$ that fit the observation s are used in the further course as we will show in subsection 2.2.

Note also that in the section the order p of the difference equation is considered known. A way of choosing p and problems encountered when using the iterations (7) at the implementation stage will be conveyed in the next section.

2.2 Decoding of an audio frame

From (3) we can see that, in order to restore process z , for each complex-conjugate pair of the roots of the characteristic polynomial of the difference equation we need to know the real argument and the real modulus of the polar form of

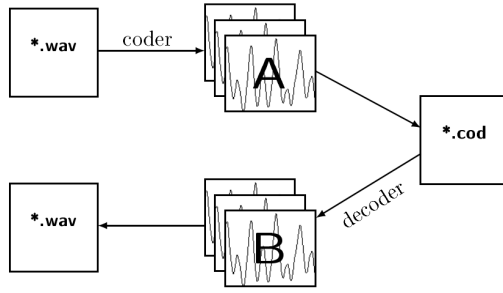


Figure 2: General scheme of work of the codec.
A. Original frames. B. Decompressed frames.

that root of the pair that lies above the real axis and we also need to know the set of real coefficients A_i . Thus, we should keep $2p$ float numbers to restore one audio frame. Besides, we should also keep one byte (or two/three for the last frame of an audio file) of service information for each audio frame such as an order of a difference equation and the frame size type. At the paper we will not describe in detail the structure of the frame service information byte.

The only thing left for us to understand is how to get coefficients A_i , $i = \overline{1, p}$. Let us transform the expression (3) to the matrix form:

$$\begin{pmatrix} z[1] \\ \vdots \\ z[N] \end{pmatrix} = \underbrace{\begin{pmatrix} \rho_1 \cos(\omega_1) & \rho_1 \sin(\omega_1) & \rho_2 \\ \rho_1^2 \cos(2\omega_1) & \rho_1^2 \sin(2\omega_1) & \rho_2^2 \\ \vdots & \vdots & \vdots \\ \rho_1^N \cos(N\omega_1) & \rho_1^N \sin(N\omega_1) & \rho_2^N \end{pmatrix}}_H \cdot \underbrace{\begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}}_d = Hd. \quad (9)$$

Using the iterations (7) we have found the model G and the process z , corresponding to the original observation s , such that $Gz = 0$. Knowing G (the coefficients of the difference equation), we can find the characteristic roots of the equation and, thus, the matrix H . Note that the next expression is true:

$$Gz = GHd = 0, d \neq 0 \Rightarrow G \perp H.$$

The needed vector d containing coefficients A_i can be found with the least squares method:

$$d = (\bar{H}^\top \bar{H})^{-1} \bar{H}^\top z, \quad (10)$$

where \bar{H} is a submatrix composed of $\geq p$ rows of matrix H .

3. CODEC IMPLEMENTATION

We have realized an audio codec (consisted of two modules: a coder and a decoder) based on the theory described above in Scilab environment (<http://www.scilab.org>), similar to MATLAB. The codec realization and testing were performed on the Debian GNU/Linux operating system. As input the codec accepts a mono WAV audio file with sample rate 44.1 KHz and bit depth 16 bits. As output the codec produces a compressed file, whose structure was defined by us, with a new extension .cod. In Figure 2 you can see the general scheme of work of the codec.

An original audio signal to be compressed is divided into frames of N samples with a shift of $(N - M)$ samples. That is, each pair of consequent frames overlap by M samples to be glued smoothly after their decompression. After the

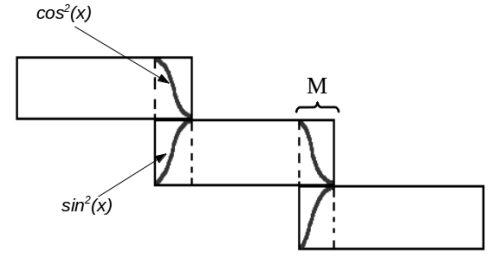


Figure 3: Gluing of frames after decompression.

decoding procedure two neighboring frames are summed on the gluing area preliminarily weighted. The weighting coefficients vary from 0 to 1. The weighting functions we use are $\sin^2(ck)$ and $1 - \sin^2(ck)$ where time index k runs from 0 to $M - 1$ and $c(M - 1) = \pi/2$. In Figure 3 you can see the gluing scheme.

The values of N and M can be predefined by a user. It was discovered that for frames longer than 150 samples the coding procedure often fails because the iterations (7) do not converge, and if we take M value < 5 an audible noise appears on joints of neighboring frames. We chose the average values: $N = 100$ and $M = 10$.

When coding a frame we increment a model order p in a cycle from $p_{min} = 2$ to $p_{max} = 13$. For each p the identification of coefficients of the equation (2) is performed and the model process (8) is found. After this, the relative modeling error is counted as follows:

$$\frac{\|z - s\|}{\|s\|} \leq 5\%, \quad (11)$$

where the value 5% is the relative error threshold. If the relative error counted does not exceed the threshold then we break the p cycle and write the found $2p$ float numbers and a service information byte to the output .cod file. Obviously, the least suitable model order is preferable to make the compressed file as small as possible. If the coder fails for any order p with the chosen relative precision 5% then it divides the frame in half and tries to model each of the two smaller frames again. If the modeling process for a smaller frame is not successful anyway, then the frame is written to .cod file directly without compression. The relative error threshold is predefined by a user and, in general, can be set to an arbitrarily small number but it would lead to a low compression ratio.

4. TESTING

The testing was performed over 20 piano audio files and 20 electric guitar files of 44100 samples each (one second duration) for our audio codec and also for LAME MP3 [version 3.99.5] [x86] codec with constant bit rates (CBR) 128 Kbps and 256 Kbps in order for us to be able to assess the effectiveness of our codec comparing to it. You can see the results of the testing in Table 1.

We compressed each original WAV audio file with a coder to .cod or MP3 file and then decoded the compressed file to a WAV file again. After that, we counted the relative error between the audio signal of the original WAV file and the signal of the decompressed WAV file as shown in (11). The relative error values presented in the Table 1 are average for

files type	files number	Our codec		LAME MP3 128 kbps		LAME MP3 256 kbps	
		compression ratio	relative error	compression ratio	relative error	compression ratio	relative error
piano	20	3.334	0.028	5.15	0.052	2.575	0.001
electric guitar	20	1.451	0.028	5.15	0.056	2.575	0.003

Table 1: Testing results.

the both sets of 20 audio files. The relative error threshold in our codec is 5 %, thus, the average relative error values relating to our codec are less than 0.05.

Using a bit rate of 128 Kbps usually results in a sound quality equivalent to what we would hear on the radio. As you can see the relative error for our codec is less than the one for LAME MP3 codec with CBR 128 Kbps. However, the relative error is only an objective sound quality measurement. When we were assessing the subjective perceptual quality of the sound produced by our codec by listening to it in headphones, the sound happened to be distinctly worse than the sound of the audio files produced by LAME MP3 codec with CBR 128 Kbps.

The compression ratio values relating to our codec are average for the both sets of 20 audio files in the Table 1. The compression ratio reached by LAME MP3 was identical for all the audio files (5.15 times for 128 Kbps and 2.575 times for 256 Kbps) as we used it in the constant bit rate mode.

Considering the work of our codec, one can also notice that the average compression ratio reached by the codec for "simple" piano files is two times bigger than the ratio for "complicated" electric guitar files. The reason of it is that the iterations (7) converge worse for the latter. Therefore, more frames of an electric guitar file are written fully to an output compressed file .cod increasing its size.

5. CONCLUSIONS

We consider the codec as an interesting application of parametric identification methods in the time domain. The key point of its work is the variational (STLS) objective function (4) that is minimized in our modeling algorithm. The iterations (7) minimize the function over a difference equation coefficients effectively for simple piano music files and the algorithms STLS1 and STLS2 used in [3] also solve the STLS problem well for speech signals. However, the STLS approach does not work properly for more complicated music files. We are going to handle the problem by dividing an audio signal into frequency subbands and coding each of them independently. Besides, we search for more efficient ways of minimization of the variational objective function. For instance, we try to do it over the roots of the characteristic polynomial of the difference equation.

References

- [1] Yuli You. *Audio Coding Theory and Applications*. New York: Springer, 2010.
- [2] Kris Hermus et al. "Perceptual Audio Modeling with Exponentially Damped Sinusoids". In: *Signal Processing* 85.1 (2005), pp. 163–176.
- [3] Philippe Lemmerling, Nicola Mastronardi, and Sabine Van Huffel. "Efficient implementation of a structured total least squares based speech compression method". In: *Linear Algebra and its Applications* 366 (2003), pp. 295–315.
- [4] Pieter P. N. de Groen. "An Introduction to Total Least Squares". In: *Nieuw Archief voor Wiskunde* 14.4 (1996), pp. 237–253.
- [5] Andrei A. Lomov. "Variational identification methods for linear dynamic systems and the local extrema problem". In: *Upravlenie Bol'shimi Sistemami* 39 (2012). <http://ubs.mtas.ru/upload/library/UBS3903.pdf> [Last accessed 26 June 2015], pp. 53–94.
- [6] Alexey O. Egorshin. "Computational closed algorithms of identification of linear objects". In: *Optimal and Self-Adjusting Systems*. Novosibirsk: ed. of Institute of Automation and Electrometry of the USSR Academy of Sciences, 1971, pp. 40–53.
- [7] Alexey O. Egorshin. "Least square method and fast algorithms in variational problems of identification and filtration (VI method)". In: *Avtometriya* 1 (1988), pp. 30–42.
- [8] Bart De Moor. "Structured total least squares and L_2 approximation problems". In: *Linear Algebra and its Applications* 188-189 (1993), pp. 163–207.
- [9] Berend Roorda and Christiaan Heij. "Global total least squares modelling of multivariable time series". In: *the IEEE Transactions on Automatic Control* AC-40 (1995), pp. 50–63.
- [10] Michael Robert Osborne and Robert Scott Anderssen. "A class of nonlinear regression problems". In: *Data Representation*. Saint Lucia: University of Queensland Press, 1970, pp. 94–101.
- [11] Eugenia M. Khassina. "File compression by use of the method of variation identification of modeling difference equations". In: *Proceedings of the X International Conference "System Identification and Control Problems" SICPRO '15*. V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences. Moscow, Russia, Jan. 2015, pp. 648–658.
- [12] Eugenia M. Khassina. "Modifications of Vorbis algorithms for audio file compression". In: *Materials of LII International Scientific Student's Conference "The Student and scientific and Technical Progress"*. Section "Mathematical Modeling". Novosibirsk State University. Novosibirsk, Russia, 2014, p. 151.

Visualization of Execution Paths for Concurrent Programs

Andrei Eleshevich
 Peter the Great St. Petersburg Polytechnic
 University
 29 Polytechnicheskaya st.
 195251 St. Petersburg Russia
 ordronus@gmail.com

Marat Akhin
 Peter the Great St. Petersburg Polytechnic
 University
 29 Polytechnicheskaya st.
 195251 St. Petersburg Russia
 akhin@kspt.icc.spbstu.ru

ABSTRACT

Understanding concurrent program behaviour is very hard even for experienced developers, let alone students, because of different possible thread interleavings, which are often not so obvious *prima facie*. In this paper we present a visualization system intended to help students in this difficult task. It collects all possible execution traces with the help of Java PathFinder and visualizes them as UML sequence diagrams, thus allowing one to discern possible execution schedules for a given concurrent program. We believe this kind of visualization could be of great use in teaching concurrent programming.

Categories and Subject Descriptors

K.3.2 [Computers and Education]: Computer and Information Science Education—*computer science education*; D.2.5 [Software Engineering]: Testing and Debugging—*monitors, testing tools, tracing*; D.1.3 [Programming Techniques]: Concurrent Programming—*parallel programming*

Keywords

Education, Visualization, Concurrency

1. INTRODUCTION

It is quite difficult to teach programming to students [4], and it is even more so with concurrent programming. At the same time, concurrency is one of the most required skills for a developer nowadays, as more and more problems stop fitting on a single core every year.

The main problem with teaching concurrent programming is that in a single-threaded program execution order is easily defined and clear to the naked eye, whereas in a multi-threaded program, because of context switches which can happen arbitrarily, there are many possible thread interleavings, some of which might lead to bugs such as race conditions and deadlocks.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.
 Copyright 2015 University of Aizu Press.

In many cases concurrency errors stay hidden most of the time and surface only in several specific thread interleavings (so called Heisenbugs). When teaching, this can lead to students becoming overcautious and starting to abuse synchronization, which in turn nullifies the advantages of concurrent programming.

We propose to enhance concurrent programming teaching by providing a visualization tool which captures all possible execution orders and shows them to the student, so that she can explore them and reason about side effects of different thread interleavings on the program results. The prototype implementation is based on Java PathFinder (JPF) [5] and uses it to collect interesting execution traces which are later rendered as UML sequence diagrams; if JPF finds a possible concurrency error, it is shown to the student together with the schedule that caused it.

The rest of the paper is organized as follows. Section 2 talks about related work in the area of visualization of concurrent program. We present our approach, the architecture of our prototype tool and show the diagram of a program containing a race condition in sections 3 and 4. Possible future work and conclusions are discussed in section 5.

2. RELATED WORK

There have been a lot of research in the area of concurrent program visualization for program understanding. JThreadSpy [3] uses instrumentation to collect execution traces and renders them as UML sequence diagrams, in which it is very similar to our approach. The differences are that it only instruments class methods and uses runtime information (therefore cannot explore all reachable thread interleavings). Our approach captures all information about program execution and explores every possible execution trace due to the power of JPF.

Atropos [1] also uses instrumentation for trace collection, but it records all operations and data being manipulated by them. It then creates a data dependence graph (DDG) for visualization; the user can analyze DDG to find which data dependencies caused an error in the program. Atropos does not allow the user to analyze different possibilities, as it uses runtime information to build the DDG, i.e. captures only one concrete execution.

Thread Interleaving Explorer (TIE) [2] is a tool most similar to ours and actually was the main inspiration behind this work. It is a debugger for erroneous execution JPF traces that allows the user to select a threading schedule and study its influence on the program execution. Unlike our approach, it focuses on exploring one (erroneous) trace at a time and

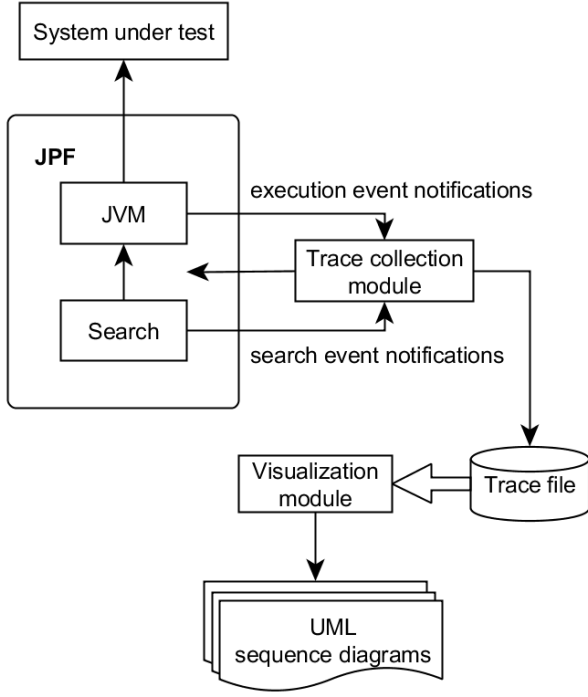


Figure 1: Prototype architecture

cannot visualize interactions between context switches (i.e. show several executions at the same time in detail).

3. APPROACH

Our main idea is to exhaustively exercise all possible thread interleavings using JPF¹. By employing various model checking techniques, JPF can efficiently backtrack to already visited program states (should they appear during exploration) without the need to re-execute the program. We collect the information about these unique program states and combine them later to create different thread schedulings. These schedulings are presented to the user, so that she can study them to gain insights to how her concurrent program actually works.

We utilize well-known UML sequence diagrams (USD) to visualize possible program executions and follow their standard notation. USD objects correspond to program objects, USD lifelines — to different object activities. Static methods and fields are represented via special dummy objects (one per class).

Method calls are mapped to USD messages from the caller to the callee objects; synchronized calls are represented with solid arrows, open arrows are used otherwise. Object allocations and thread starts are shown as special method calls (which maps nicely to their actual semantics).

Fields are viewed as properties with dummy `get/set` methods that are also visualized as special method calls (for which we track the current field values). At the moment, our prototype supports only fields of primitive types and does not handle `volatile/final` fields w.r.t. concurrency quirks; this

¹Here JPF is only a tool, our approach can be based on any other system that provides information about possible concurrent executions.

Listing 1: Racer program

```
public class Racer extends Thread {
    int d = 42;

    public void run() {
        doSomething(1001);
        d = 0;
    }

    public static void main(String[] args) {
        Racer racer = new Racer();
        racer.start();

        doSomething(1000);
        int c = 420 / racer.d;
        System.out.println(c);
    }

    static void doSomething(int n) {
        try {
            Thread.sleep(n);
        } catch (InterruptedException ix) {}
    }
}
```

is one of the possible areas for future work.

To represent thread information, we extended USD by adding color annotations to the activation boxes — every thread is assigned a unique color that is used to mark methods run by the corresponding thread. If a thread execution is paused, it is shown by a darker shade of the thread's color.

Another USD extension is to support execution branches. If a method execution creates several interesting schedulings, its activation box is labeled with a number of branch options JPF found during program exploration. From this label the user can open another possible scheduling in a separate window.

4. PROTOTYPE ARCHITECTURE

Our prototype implementation consists of two modules: trace collection module and visualization module 1. Let us discuss them in more detail.

4.1 Trace Collection

We collect possible execution traces from JPF using a custom JPF listener and capture such information as method invocations, field accesses and object allocations. We also record different execution branches and their unique identifiers (UID) and context switches possible in the program. An execution branch in our approach roughly corresponds to JPF choice during its model checking phase; if several executions lead to the same JPF state, we consider them to be the same branch (though there would be several different thread interleavings leading to it).

4.2 Visualization

After collecting all execution traces, visualization is done using a stand-alone GUI program. It carefully aggregates possible execution branches (by unpacking branch combinations possible in the program to execution paths) and builds UML sequence diagrams from them. The first diagrams represent correct (w.r.t. JPF) executions, the last one shows an execution with possible errors. The user can open additional diagrams by selecting a branching point and picking one of the possible branching options there.

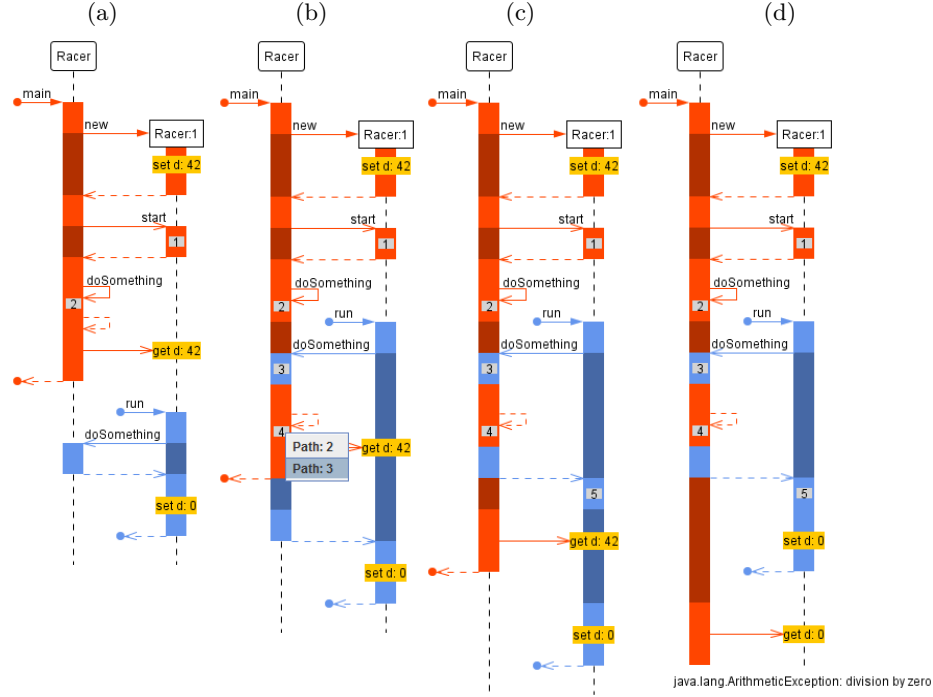


Figure 2: Execution paths for Racer program

4.3 Example

Let us consider a classic Racer example — a program with race condition shown in listing 1.

The results of thread interleaving exploration for Racer program using our prototype tool are shown in figure 2. You can see that several execution paths were explored (figures 2a–2c) before finding an error (figure 2d).

The bug is a race condition between read access to `racer.d` in `main` method and write access to it in `run` method. As they are not synchronized, `d = 0` could happen before division which would cause a division-by-zero error. As seen from figure 2, JPF exploration is depth-first as it tries to complete the execution before backtracking to a branching point. If a branching point creates non-distinct executions (up to the next branching point), we consider it to be non-interesting and does not show it in the interface.

5. CONCLUSIONS AND FUTURE WORK

We proposed an approach to the visualization of concurrent executions based on JPF for thread interleaving exploration. It creates UML sequence diagrams which allows the students to easily reason about different concurrent executions and witness the possible bugs first-hand.

Of course, our prototype leaves much to be desired. For example, visualization lacks object or method filters which would help with reducing the sequence diagram's size. The sequence diagram itself could be improved by adding explicit notation for advanced synchronization primitives, e.g. `synchronized` blocks, `wait/notify` calls or `volatile` accesses.

Another idea would be to employ slicing to compact the diagram leaving only variables and interactions interesting to the user.

6. REFERENCES

- [1] J. Lönnberg, M. Ben-Ari, and L. Malmi. Java replay for dependence-based debugging. In *Proceedings of the Workshop on Parallel and Distributed Systems: Testing, Analysis, and Debugging*, PADTAD '11, pages 15–25, New York, NY, USA, 2011. ACM.
- [2] G. Maheswara, J. S. Bradbury, and C. Collins. TIE: An interactive visualization of thread interleavings. In *Proceedings of the 5th International Symposium on Software Visualization*, SOFTVIS '10, pages 215–216, New York, NY, USA, 2010. ACM.
- [3] G. Malnati, C. M. Cuva, and C. Barberis. JThreadSpy: Teaching multithreading programming by analyzing execution traces. In *Proceedings of the 2007 ACM Workshop on Parallel and Distributed Systems: Testing and Debugging*, PADTAD '07, pages 3–13, New York, NY, USA, 2007. ACM.
- [4] E. Pyshkin. Teaching programming: What we miss in academia. In *Software Engineering Conference in Russia (CEE-SECR), 2011 7th Central and Eastern European*, pages 1–6. IEEE, 2011.
- [5] W. Visser, K. Havelund, G. Brat, S. Park, and F. Lerda. Model checking programs. *Automated Software Engg.*, 10(2):203–232, Apr. 2003.

Intermittent Control Properties of Car Following: I. Driving Simulator Experiments

Hiromasa Ando,¹ Ryoji Yamauchi,² Ihor Lubashevsky,³ Arkady Zgonnikov⁴
University of Aizu

Ikki-machi, Aizu-Wakamatsu, Fukushima 965-8560, Japan

¹)m5181109@u-aizu.ac.jp, ²)s1200149@u-aizu.ac.jp, ³)i-lubash@u-aizu.ac.jp, ⁴)arkady@u-aizu.ac.jp

ABSTRACT

A rather simple car driving simulator was created based on the available open source engine TORCS and used to analyze the basic features of human behavior in car driving within the car-following and free-driving setups. Four drivers with different skill in driving real cars participated in these experiments. They were instructed to driver a virtual car without overtaking the lead car driven by computer at a fixed speed and not to lose sight of it as well as to drive a virtual car on empty road in a style convenient individually. Based on the collected data the distribution of the headway, velocity, acceleration, and jerk are constructed and compared with available experimental data collected previously by the analysis of the real traffic flow. As the main results we draw a conclusion that the human behavior in car driving should be categorized as a generalized intermittent control with noise-driven activation of the active phase. Besides, we hypothesize that the car jerk is an individual phase variable required for describing car dynamics.

Categories and Subject Descriptors

J.4 [Social and Behavioral Sciences]: psychology; H.1.2 [User/Machine Systems]: human factors

General Terms

Experiment

Keywords

Human behavior, status quo bias, intermittent control, car-following dynamics, car-driving simulator

1. INTRODUCTION

In the last decades a new concept of how human operators stabilizing mechanical systems, called human intermittent control, was developed (e.g., [1]). It considers human operators not to be capable of controlling system dynamics

continuously and, as a result, their actions must be a sequence of alternate phases of active and passive behavior, with the switching between these phases being event-driven. Recently, we developed a concept of noise-driven control activation as a more advanced alternative to the conventional threshold-driven activation [2]. In this concept the transition from passive to active phases is probabilistic and reflects human perception and fuzzy evaluation of the current system state before making decision concerning the necessity of correcting the system dynamics. During the passive phase the control is halted and the system moves on its own, broadly speaking, during the passive phase the operator accumulates the information about the system state. The periods of active phase can be regarded as fragments of open-loop control, which is due to the delay in human reaction (e.g., [1]).

Driving a car in following a lead car is a characteristic example of human control, which allows us to suppose that the intermittency of human control should be pronounced in the driver behavior and affect the motion dynamics essentially. We have expected that the characteristics of human intermittent control found in stick balancing [2] should be pronounced also in car driving.

2. EXPERIMENTAL SETUP



Figure 1: Car-following setup.

To verify this statement we created a rather simple driving simulator using the open source engine 'TORCS' [3] and conducted virtual experiments on car following and driving on empty road. A screenshot in Fig. 1 illustrates the typical situation in the car-following experiments and the track geometry. Four drivers with different skill of driving real cars were involved. The results of these experiments were

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

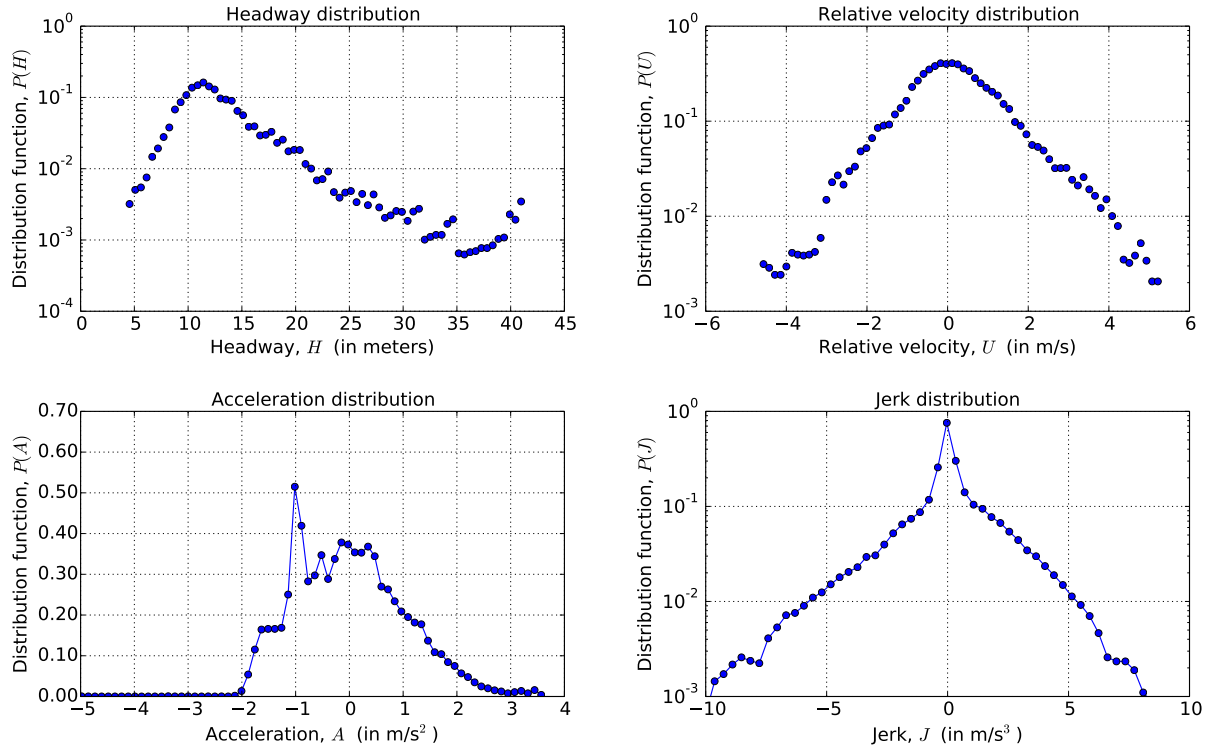


Figure 2: Some characteristics of car-following found in the driving simulator experiments. The shown forms are typical for experienced drivers.

compared with our previous results obtained in balancing a virtual over-damped pendulum to single out the characteristic features caused by the control intermittency.

3. RESULTS AND CONCLUSION

Figure 2 depicts some results obtained by experienced drivers in following a lead car moving with an effective speed about 150 km/h in the virtual environment. The obtained distributions of the headway distance and the car velocity are rather close to the results found for single car data as well as GPS data [4]. In particular, the shown forms demonstrate that the headway distance and the velocity difference are distributed according to the asymmetric Laplace law; the experience of drivers is reflected mainly in the scales. It allows us to state that the anomalous shape of these characteristics is mainly due to the basic properties of human perception, whereas mechanical details of real cars are not too significant.

The distribution functions of the acceleration and jerk (two lower rows in Fig. 2) depend substantially on the driver's individuality and reflect their personal styles of driving which seem to be rather similar for the three drivers with some driving experience. Nevertheless, as a general feature, we note the bimodal form of the acceleration distribution and a certain rather sharp peak all the jerk distributions possess at the origin, $j = 0$. The existence of this peak is a characteristic feature of human control intermittency [2]. The jerk distribution peak signals that the jerk is the main control parameter changing which drivers govern the car motion. For the real traffic data the acceleration distribution is also

of bimodal form [4].

Based on the obtained results we draw a conclusion that the car dynamics, at least, in the car-following setup has to be described using *four* phase variables, namely, the car position, velocity, acceleration, and jerk. It also meets our previous hypothesis based on the theory of rational drive behavior [5]. The distribution of car velocity found for the free-motion setup possesses a number of anomalous properties, in particular, is characterized by an extremely large thickness.

4. REFERENCES

- [1] Ian D. Loram, Henrik Gollee, Martin Lakie, and Peter J. Gawthrop. Human control of an inverted pendulum: is continuous control necessary? Is intermittent control effective? Is intermittent control physiological? *The Journal of Physiology*, 589(2):307–324, 2011.
- [2] Arkady Zgonnikov, Ihor Lubashevsky, Shigeru Kanemoto, Toru Miyazawa, and Takashi Suzuki. To react or not to react? Intrinsic stochasticity of human control in virtual stick balancing. *Journal of The Royal Society Interface*, 11:20140636, 2014.
- [3] The official site of TORCS. <http://torcs.sourceforge.net/index.php>.
- [4] Peter Wagner and Ihor Lubashevsky. Empirical basis for car-following theory development. *arXiv preprint cond-mat/0311192*, 2003.
- [5] Ihor Lubashevsky, Peter Wagner, and Reinhard Mahnke. Rational-driver approximation in car-following theory. *Physical Review E*, 68(5):056109, 2003.

Intermittent Control Properties of Car Following: II. Dynamical Trap Model

Ryoji Yamauchi,¹ Hiromasa Ando,² Ihor Lubashevsky,³ Arkady Zgonnikov⁴
University of Aizu

Ikki-machi, Aizu-Wakamatsu, Fukushima 965-8560, Japan

¹⁾s1200149@u-aizu.ac.jp, ²⁾m5181109@u-aizu.ac.jp, ³⁾i-lubash@u-aizu.ac.jp, ⁴⁾arkady@u-aizu.ac.jp

ABSTRACT

A new model for car-following is proposed to capture the found properties in our previous experiments. It is based on the experimental results showing that (i) human behavior in car driving should be categorized as a generalized intermittent control with noise-driven activation of the active phase and (ii) the extended phase space required for modeling human actions in car driving has to comprise four phase variables, namely, the headway distance, the velocity of car, its acceleration, and the car jerk, i.e., the time derivative of the car acceleration.

Categories and Subject Descriptors

J.4 [Social and Behavioral Sciences]: psychology; H.1.2 [User/Machine Systems]: human factors

General Terms

Theory

Keywords

Human behavior, status quo bias, intermittent control, car-following dynamics, dynamical traps

1. INTRODUCTION

In the last decades a new concept of how human operators stabilizing mechanical systems, called human intermittent control, was developed (e.g., [1]). It considers human operators not to be capable of controlling system dynamics continuously and, as a result, their actions must be a sequence of alternate phases of active and passive behavior, with the switching between these phases being event-driven. Recently, we developed a concept of noise-driven control activation as a more advanced alternative to the conventional threshold-driven activation [2]. In this concept the transition from passive to active phases is probabilistic and reflects human perception and fuzzy evaluation of the current

system state before making decision concerning the necessity of correcting the system dynamics. During the passive phase the control is halted and the system moves on its own, broadly speaking, during the passive phase the operator accumulates the information about the system state. The periods of active phase can be regarded as fragments of open-loop control, which is due to the delay in human reaction (e.g., [1]).

Driving a car in following a lead car is a characteristic example of human control, which allows us to suppose that the intermittency of human control should be pronounced in the driver behavior and affect the motion dynamics essentially. Previously [3] we reported the results obtained in our experiments on car-following based on a car driving simulator created using the open source engine TORCS [4]. As the main results we have drawn a conclusion that the human behavior in car driving should be categorized as a generalized intermittent control with noise-driven activation of the active phase. Besides, we have argued for the hypothesis that the car jerk is an individual phase variable required for describing car dynamics.

2. FOUR-VARIABLE MODEL OF CAR-FOLLOWING

In this paper we discuss a mathematical model for car-following that employs the results of the experiments noted above. It is based on the assumption that to describe the driver behavior the extended phase comprising four independent variables is required; this idea was partly elaborated in [5]. A driver is not able to change the car position and its velocity directly; he can only vary the car acceleration by pressing the gas or break pedal. In real driving the car acceleration on its own is an important characteristic of car motion. Therefore, in describing the car dynamics we have to include the car acceleration in the list of the phase variables [6, 7, 8]. However, according to the found characteristics of the jerk distribution [3, 5], *the jerk on its own is also an independent phase variable* or another additional variable combining the headway distance h , the car velocity v , acceleration a , and jerk $j = da/dt$ within a certain relationship should be introduced. In the proposed model using a simplified description of car motion control, this fourth variable is the position θ of an effective pedal combining the gas and break pedals into one control unit.

Namely, the model is specified as follows. The car ahead is assumed to move at a fixed velocity V and the dynamics

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

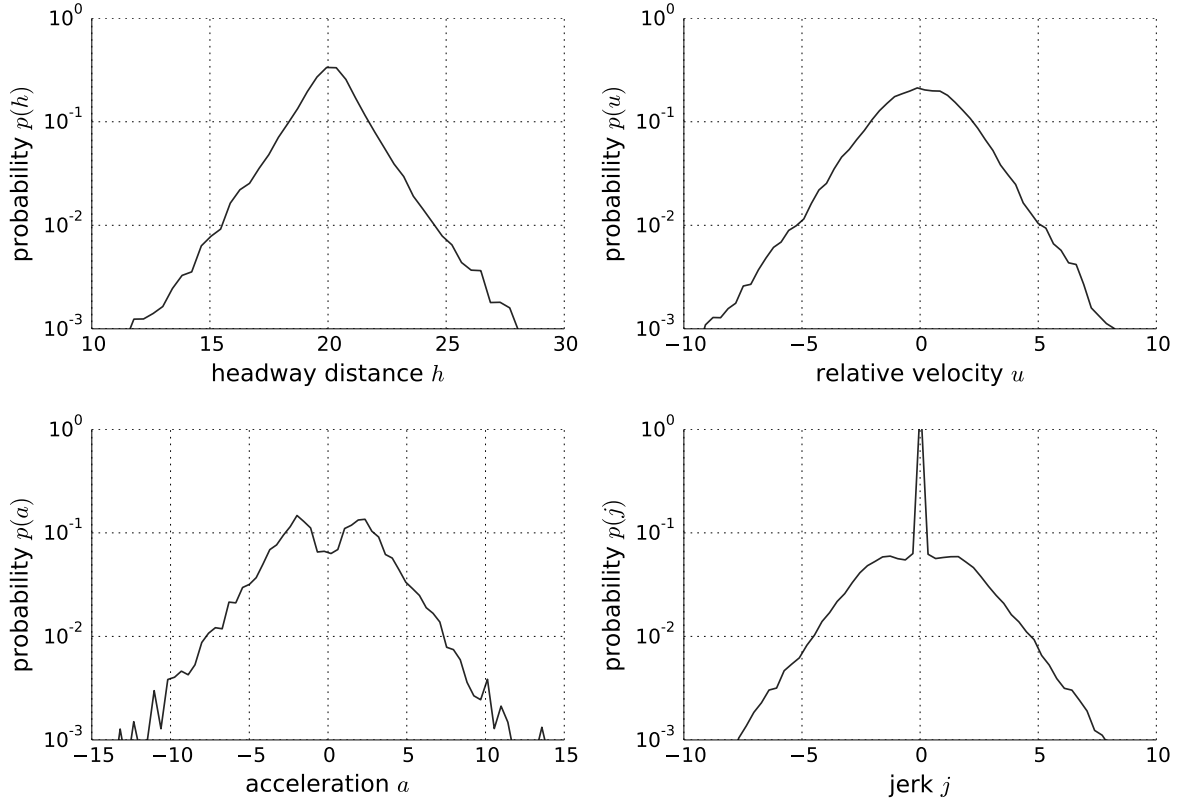


Figure 1: The distributions of the headway distance h , relative velocity $u = v - V$, acceleration a , and the jerk $j \propto (\theta - v)$ obtained by numerical solution of model (1)–(4). In simulation the following values were used, $v_{\max} = 30$ m/s (about 100 km/h), $V = 15$ m/s, $D = 20$ m, $a_{th} = 0.1$ m/s², $\tau_h = \tau_\theta = 0.2$ s, $\tau_v = 1$ s, and $\epsilon = 0.005$ m/s^{1.5}. The numerical labels at axes are given in the corresponding units composed of meters and seconds.

of the following car is given by the equations

$$\frac{dh}{dt} = V - v, \quad (1)$$

$$\frac{dv}{dt} = a, \quad (2)$$

$$\tau_\theta \frac{d\theta}{dt} = \theta - a, \quad (3)$$

$$\tau_h \frac{d\theta}{dt} = \Omega(a - \theta) \cdot [a_{\text{opt}}(h, v) - a] + \epsilon \xi(t). \quad (4)$$

Equations (1) and (2) are just simple kinematic relations between the variables h , v , and a , equation (3) describes the mechanical properties of the car engine and its response with some delay τ_θ to the position θ of the control unit measured here in units of acceleration. The last equation (4) describes the driver behavior. It combines the basic ideas of noise-driven activation in human intermittent control [2] and the concept of action dynamical traps for systems with inertia [8]. The driver is able to control directly only the position θ of the control unit and the difference $(\theta - a)$ between the desired acceleration θ and the current car acceleration a is the parameter quantifying the difference between his active and passive behavior. The bounded capacity of driver cognition is described in terms of action dynamical traps via

the introduction of cofactor

$$\Omega(a - \theta) = \frac{(a - \theta)^2}{(a - \theta)^2 + a_{th}^2} \quad (5)$$

similar to fuzzy reaction coefficients. Here a_{th} is the driver perception threshold of car acceleration. The ansatz

$$a_{\text{opt}}(h, v) = \frac{1}{\tau_v} \left[v_{\max} \frac{h^2}{h^2 + D^2} - v \right] \quad (6)$$

determines the optimal acceleration with which the strictly rational driver with perfect perception would drive the car. This expression inherits the optimal velocity mode widely used in modeling traffic flow (see, e.g., [9]). Here τ_v is the human response delay time, v_{\max} is the maximal velocity acceptable for safety reasons on a given road without neighboring cars, and D is the characteristic headway distance when drivers consider it necessary to slow their cars down as the headway distance decreases. The last term in equation (4) is the random Langevin force, where $\xi(t)$ is white noise of unit amplitude and ϵ is the Langevin force intensity. The interplay between the fuzzy perception function $\Omega(\theta - a)$ and this Langevin force are two main components of the noise-induced activation model elaborated in [2] for describing human balance of overdamped pendulum. Fi-

nally, the difference $[a_{\text{opt}}(h, v) - a]$ quantifies the stimulus for the driver to correct the current state of car motion.

It should be noted that this approach to describing the effects caused by the bounded capacity of human cognition inherits the general formalism developed previously and called the dynamical traps [10, 11, 12, 13]. It assumes that individuals (operators) governing the dynamics of a certain system try to follow an optimal strategy in controlling its motion but fail to do this perfectly because similar strategies are indistinguishable for them. In systems, where the optimal dynamics implies the stability of a certain equilibrium point in the corresponding phase space, the human fuzzy rationality gives rise to some neighborhood of the equilibrium point, the region of dynamical traps, wherein each point is regarded as an equilibrium one by the operator. So, when the system enters this region and while it is located in it, maybe for a long time, the operator control is suspended. In this case the system can leave the dynamical trap region only because of the mismatch between actions which may be treated as some random factor.

In the given paper we actually present a preliminary investigation of this model and its goal is to demonstrate a potential capability of such an approach to describing complex properties of real traffic flow. Figure 1 depicts the results of numerical solution of model (1)–(4) using the characteristic values of the systems parameters employed by other models, at least, being of the same order (cf., e.g., [9]). It should be noted that the distributions obtained by numerical simulation of the developed model and constructed based on the experimental data collected by subjects with experience of driving real cars [3, 5] look rather similar.

3. CONCLUSION

The experiments conducted previously [3, 5] have demonstrated that the behavior of subjects involved into driving virtual cars should be categorized as the generalized intermittent control over mechanical systems. It consists of a sequence of alternate fragments of active and passive phases of driver behavior. The passive phase is characterized by the fact that during the corresponding time interval a driver does not change the position of the gas or break pedal. In this case the jerk plays the role of the parameter controlled directly by the driver and, so, has to be regarded as an independent phase variable determining the car dynamics. It enabled us, keeping in mind also driving real cars, to pose the hypothesis that a sophisticated description of car motion controlled by human actions requires the introduction of four dimensional phase space, where the car position, velocity, acceleration, jerk are the independent variables.

In this paper a new model for the car-following that allows for these features has been proposed. Its numerical simulation has demonstrated that the combination of the concepts of the noise-driven activation in human intermittent control and the action dynamical traps caused by the bounded capacity of human cognition can reproduce, at least, qualitatively the results of the conducted experiments.

4. REFERENCES

- [1] Ian D. Loram, Henrik Gollee, Martin Lakie, and Peter J. Gawthrop. Human control of an inverted pendulum: is continuous control necessary? Is intermittent control effective? Is intermittent control physiological? *The Journal of Physiology*, 589(2):307–324, 2011.
- [2] Arkady Zgonnikov, Ihor Lubashevsky, Shigeru Kanemoto, Toru Miyazawa, and Takashi Suzuki. To react or not to react? Intrinsic stochasticity of human control in virtual stick balancing. *Journal of The Royal Society Interface*, 11:20140636, 2014.
- [3] Hiromasa Ando, Ryoji Yamauchi, Ihor Lubashevsky, and Arkady Zgonnikov. Anomalous Properties of Car Following: I. Driving Simulator Experiments. *this issue*, 2015.
- [4] The official site of TORCS. <http://torcs.sourceforge.net/index.php>.
- [5] Hiromasa Ando, Ihor Lubashevsky, Arkady Zgonnikov, and Yoshiaki Saito. Statistical Properties of Car Following: Theory and Driving Simulator Experiments. In *Proceedings of the 46th ISCIE International Symposium on Stochastic Systems Theory and Its Applications Kyoto, Nov. 1-2, 2014*, pages 149–155, Kyoto, 2015. Institute of Systems, Control and Information Engineers (ISCIE).
- [6] Ihor Lubashevsky, Peter Wagner, and Reinhard Mahnke. Rational-driver approximation in car-following theory. *Physical Review E*, 68(5):056109, 2003.
- [7] I. Lubashevsky, P. Wagner, and R. Mahnke. Bounded rational driver models. *The European Physical Journal B-Condensed Matter and Complex Systems*, 32(2):243–247, 2003.
- [8] Arkady Zgonnikov and Ihor Lubashevsky. Extended phase space description of human-controlled systems dynamics. *Progress of Theoretical and Experimental Physics*, 2014(3):033J02, 2014.
- [9] Martin Treiber and Arne Kesting. *Traffic Flow Dynamics: Data, Models and Simulation*. Springer, Heidelberg, 2012.
- [10] I. A. Lubashevsky, V. V. Gafiychuk, and A. V. Demchuk. Anomalous relaxation oscillations due to dynamical traps. *Physica A: Statistical Mechanics and its Applications*, 255(3):406–414, 1998.
- [11] I. Lubashevsky, M. Hajimahmoodzadeh, A. Katsnelson, and P. Wagner. Noise-induced phase transition in an oscillatory system with dynamical traps. *The European Physical Journal B-Condensed Matter and Complex Systems*, 36(1):115–118, 2003.
- [12] I. Lubashevsky, R. Mahnke, M. Hajimahmoodzadeh, and A. Katsnelson. Long-lived states of oscillator chains with dynamical traps. *The European Physical Journal B-Condensed Matter and Complex Systems*, 44(1):63–70, 2005.
- [13] I. Lubashevsky. Dynamical Traps Caused by Fuzzy Rationality as a New Emergence Mechanism. *Advances in Complex Systems*, 15(08):1250045, 2012.

Admissible Trajectory Planning for Wheeled Mobile Robot

Nelli Korotkova
Saint-Petersburg State University
7-9, Universitetskaya nab.,
St. Petersburg, 199034, Russia
nelly_kor@mail.ru

Evgeny Veremey
Saint-Petersburg State University
7-9, Universitetskaya nab.,
St. Petersburg, 199034, Russia
e_veremey@mail.ru

ABSTRACT

In this paper the problem of planning a trajectory for driving a mobile robot from some initial position to an arbitrary goal that satisfies limitations of control signals is considered. For this purpose continuous-curvature turns are used within the Dubins's scheme. It is also offered to use presented trajectory planning algorithm along with a stabilizing feedback.

Categories and Subject Descriptors

I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, and Search—*control theory*; I.2.9 [Artificial Intelligence]: Robotics

General Terms

Theory

Keywords

Mobile robot, trajectory planning

1. INTRODUCTION

Wheeled mobile robots are subjects to nonholonomic constraints. It means that they need to perform some specific maneuver in order to reach an arbitrary configuration. Planning an admissible trajectory is a common approach to solving the task of finding such maneuver. Some physical constraints are also present and apply further limitations on movement of the system.

The first goal of this paper is to design a method of planning a path considering all constraints including acceleration bounds. Particularly the problem of meeting linear acceleration bounds is solved by adding straight segments to a path. This problem is not addressed in previous works. The second goal is to describe a way of building corresponding admissible control signals. Thus we implicitly associate path as a geometric curve with a timing law and make it a trajectory. Also the application of the described method along with a stabilizing feedback is presented.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

2. PROBLEM STATEMENT

A unicycle-type mobile robot can be represented with a following system:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} \cos \theta & 0 \\ \sin \theta & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v \\ \omega \end{pmatrix} \quad (1)$$

Its position is described with three values where (x, y) are plane coordinates and θ is orientation. v and ω are inputs, linear and angular velocities respectively. Due to mentioned above physical constraints their values are limited as well as values of their derivatives, linear and angular accelerations:

$$\begin{aligned} |v| &\leq v_{\max}, \quad |\omega| \leq \omega_{\max}, \\ |\dot{v}| = |\dot{a}| &\leq a_{\max}, \quad |\dot{\omega}| = |\dot{\varepsilon}| \leq \varepsilon_{\max} \end{aligned} \quad (2)$$

The aim is to build control signals able to drive a robot from an initial position $q_s = (x_s, y_s, \theta_s)^T$ to an arbitrary goal q_g .

3. TRAJECTORY PLANNING

Let us solve the trajectory planning task by planning a path as a plane geometric curve and simultaneously building admissible control laws that realize motion along this path.

As it is commonly done, let us simplify the path planning task by assuming that the robot moves with a constant maximum speed the whole way. In order to satisfy linear acceleration limit, segments that allow robot to gain speed and to slow down must be added to the beginning and ending of the path. These can be line segments of length $S = v_{\max}^2 / 2a_{\max}$. Then the rest of the path must be built to connect shifted configurations

$$\begin{aligned} q'_s &= q_s + (S \cos \theta_s, S \sin \theta_s, 0)^T, \\ q'_g &= q_g - (S \cos \theta_g, S \sin \theta_g, 0)^T. \end{aligned}$$

The control signal for a linear velocity is then a trapezoid function

$$v(t) = \begin{cases} a_{\max} t, & 0 \leq t < \tau, \\ v_{\max}, & t \leq T - \tau, \\ v_{\max} - a_{\max} t, & T - \tau < t \leq T, \end{cases}$$

where T is total time of motion and $\tau = v_{\max} / a_{\max}$.

3.1 Dubins's Paths

A method of planning a path that connects two specified configurations using straight lines and circle arcs of minimum radius assuming that robot moves with a constant speed was proposed in [1]. It has been proved that the shortest path belongs to one of two families: CSC which stands

for paths that consist of an arc, a tangent straight line and another arc; CCC which denotes paths combined of three arcs; or is a subpath of any of these. Considering that a curve segment can correspond to different rotations, left or right turn, these define a set of six paths.

The problem of these paths is that in the intermediate points between consequent segments of such path an angular acceleration is required to be infinite. Thus these paths fail to satisfy the angular acceleration limit.

3.2 Continuous-Curvature Turns

This problem was solved in [2] by offering continuous-curvature turns that provide a transitional phase between a straight line and a circle arc or two arcs with different direction of rotation. The main idea is to use clothoid arcs with a sharpness less or equal to the maximum sharpness σ_{\max} for this purpose. σ_{\max} and radius of circular arcs r_{\min} depend on limitations (2):

$$r_{\min} = \frac{v_{\max}}{\omega_{\max}}, \quad \sigma_{\max} = \frac{\varepsilon_{\max}}{v_{\max}^2}. \quad (3)$$

The set of positions that are reachable from some initial position is a circle of radius r . At each point of such circle an angle between orientation of the robot and tangent to the circle is equal to a constant value μ . r and μ can be found using (3) and equations from [2]. These circles are called CC circles. For any initial position q four CC circles can be defined: $C_l^+(q)$, $C_r^+(q)$, $C_l^-(q)$, $C_r^-(q)$. Here sign "+" or "-" denotes direction of motion, forward or backward, r or l denotes direction of rotation. CC circles with "-" can be also interpreted as sets of configurations from which q is reachable by forward CC turns. Centers (x_Ω, y_Ω) of CC circles can be found as following

$$\begin{aligned} C_l^\pm(q) : \quad & x_\Omega = x \pm r \sin(\mu \mp \theta), \\ & y_\Omega = y + r \cos(\mu \mp \theta); \\ C_r^\pm(q) : \quad & x_\Omega = x \pm r \sin(\mu \pm \theta), \\ & y_\Omega = y - r \cos(\mu \pm \theta). \end{aligned} \quad (4)$$

CC turns can be used within Durbins's scheme by substituting regular circles representing turns with CC circles. An algorithm is to use (4) to find centers of $C_l^+(q_s)$, $C_r^+(q_s)$ for the initial position and $C_l^-(q_g)$, $C_r^-(q_g)$ for the goal and then connect them with tangent lines or circles considering directions of rotations.

The section of control signal for an angular velocity corresponding to a CC turn of angle δ is a trapezoid or triangular function. Let δ_{\min} be the angle turning by which requires a CC turn combined only of two symmetrical clothoid arcs of maximum sharpness.

Then if $\delta > \delta_{\min}$

$$\omega(t) = \begin{cases} \varepsilon_{\max} t, & t_0 \leq t < t_0 + \tau \\ \omega_{\max}, & t_0 + \tau \leq t \leq t_1 - \tau \\ \omega_{\max} - \varepsilon_{\max} t, & t_1 - \tau < t \leq t_1 \end{cases}$$

where t_0 and t_1 are moments of time of beginning and ending a CC turn respectively and $\tau = \omega_{\max}/\varepsilon_{\max}$.

If $\delta \leq \delta_{\min}$

$$\omega(t) = \begin{cases} \varepsilon t, & t_0 \leq t < \frac{t_0+t_1}{2} \\ -\varepsilon t, & \frac{t_0+t_1}{2} \leq t \leq t_1 \end{cases}$$

where $\varepsilon = \sigma v_{\max}$, $\sigma \leq \sigma_{\max}$.

Let us note that for building control signals we actually need to know only turning angles and lengths of line segments. The length of CC turn curve must be also known to find a total length of a path in order to determine the shortest one. This value can be computed knowing turning angle δ .

3.3 CSC-type Paths

Here and further let us assume that two centers $\Omega_1 = (x_1, y_1)$ and $\Omega_2 = (x_2, y_2)$ of appropriate CC circles are known.

In the case of CC circles tangency between such circle and a line is quite different from a classical tangency. μ -tangency, conditions of existence of those lines and ways to find lengths of line segments are described in [2]. So let us mention here only a way to determine turning angles for a first and a second turns.

For an external μ -tangent which is associated with left-straight-left (LSL) and right-straight-right (RSR) paths orientation of the robot while moving along a line segment is

$$\theta = \alpha = \arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right).$$

For an internal μ -tangent associated with a LSR path

$$\theta = \alpha - \beta + \frac{\pi}{2},$$

for a RSL path

$$\theta = \alpha - \beta,$$

where $\beta = \arcsin(2r \cos \mu / \rho)$, ρ is distance between centers.

Using θ_s , θ and θ_g and considering direction of rotation, angles of first and second turns can be obtained.

3.4 CCC-type Paths

For paths consisting of three curve segments, RLR or LRL, center of a CC circle tangent to both known circles must be found. In this case tangency between circles is classical though orientation of a robot is not aligned with a tangent line in the point of tangency.

For a pair of circles two options of a tangent circle are available. Their centers $\Omega = (x, y)$ can be found as following

$$x = x_1 + \cos \gamma, \quad y = y_1 + \sin \gamma, \quad \gamma = \alpha \pm \arccos(\rho / (4r)).$$

θ at the intermediate point between two circles with centers $\Omega'_1 = (x'_1, y'_1)$ and $\Omega'_2 = (x'_2, y'_2)$ is

$$\theta = \alpha' - \mu + \frac{\pi}{2}, \quad \alpha' = \arctan\left(\frac{y'_2 - y'_1}{x'_2 - x'_1}\right)$$

for a LR transition and

$$\theta = \alpha' + \mu - \frac{\pi}{2}$$

for a RL transition.

Again we can obtain angles of each of three turns using θ_s , θ_g and robot's orientations at the intermediate points.

3.5 Backward Motion

Also the possibility of backward motion is considered. In order to build paths for moving backwards it is convenient to virtually change orientations of q_s and q_g to opposite by adding π to corresponding values. Then the same computations must be performed considering that rotation directions are changed to opposite. Finally, signs of both control laws must be changed to opposite.

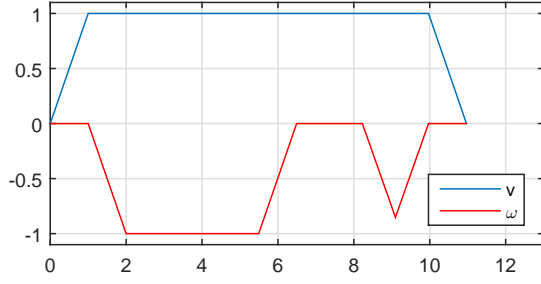


Figure 1: Control signals.

4. FEEDBACK

In order to provide stability of the motion path planning is used along with the tracking control offered in [3]. This approach supposes using a reference object described by model (1). Previously built control laws are acting as inputs to the reference object and its output is a trajectory function $q_r(t) = (x_r(t), y_r(t), \theta_r(t))$. Then control laws that go to robot's input are following:

$$\begin{aligned} v(t) &= v_r(t) + C_2 \{ [x_r(t - x(t))] \cos \theta + [y_r(t) - y(t)] \sin \theta \}, \\ \omega(t) &= \omega_r(t) + C_2 [\theta_r(t) - \theta(t)]. \end{aligned}$$

In [3] it is proved that if $C_1 > 0$, $C_2 > 0$ then a closed-loop system is exponentially stable.

5. EXPERIMENTAL RESULTS

To demonstrate the performance of proposed algorithm the Simulink model of the closed-loop system has been made. The robot is modeled with presence of disturbance and measurement noise.

The model is also utilized for solving the problem of choosing C_1 and C_2 that provide the best performance. The maximum of mean squared error for each degree of freedom in N moments of time is chosen as motion quality measure:

$$e(C_1, C_2) = \max_{\xi=x,y,\theta} \left[\frac{1}{N} \sum_{i=0}^{N-1} \left(\xi_r \left(\frac{T}{N} i \right) - \xi \left(\frac{T}{N} i \right) \right)^2 \right]$$

where $q_r(t)$ is reference trajectory and $q(t)$ is simulated trajectory of the robot.

Then the quality matrix Q is built for some grid of C_1 and C_2 values. The pair of coefficients that gives the least $e(C_1, C_2)$ is taken as an answer.

Quality matrices can be built for some set of trajectories, for example, a set that includes at least one of each type, forward and backward. Then coefficients are chosen using mean of all matrices Q for this set. The other way is to consider 12 sets, each containing trajectories of the same type, and get 12 pairs (C_1, C_2) . Which values to actually use must be decided after the path planning stage when it is known path of which type has turned out to be optimal.

Now let us demonstrate performance of the system for specified initial and goal positions

$$q_s = (0, 0, 0)^T, q_g = (0, 3, \pi/3)^T$$

and limitations

$$v_{\max} = 1, \omega_{\max} = 1, a_{\max} = 1, \varepsilon_{\max} = 1.$$

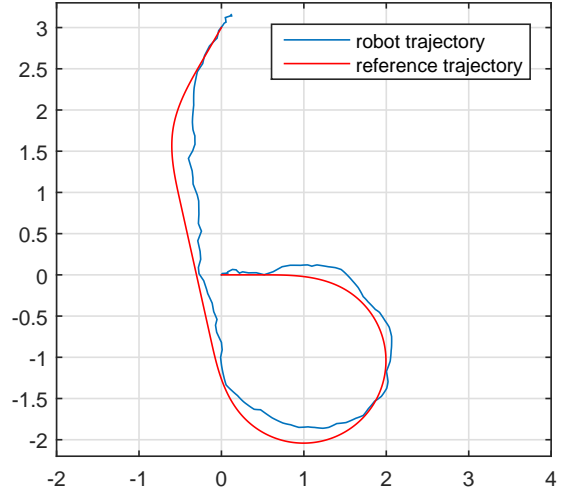


Figure 2: Robot trajectory and reference trajectory.

Values of feedback coefficients are taken

$$C_1 = 1.5, C_2 = 0.2.$$

Results are shown on figures 1 and 2. Robot stopped at position $q(t) = (0.04, 3.23, 1.14)^T$ and whole motion took $T = 10.97$ seconds.

6. CONCLUSIONS

Presented approach allows planning an admissible trajectory that is guaranteed to be realized by control laws that satisfy all considered constraints and building feedback control laws which drive robot to specified goal precisely enough in spite of presence of disturbance and measurement noise.

A subject to further research can be the task of creating more universal approach to choosing feedback law coefficients. Another direction is designing an algorithm to build a full set of continuous-curvature Reeds and Shepp's [4] paths which unlike Dubins's paths include cusps, points where direction of motion is changed. The main problem here is to satisfy linear acceleration limit at cusp points.

7. REFERENCES

- [1] L. E. Dubins. On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents. *American Journal of mathematics*, pages 497–516, 1957.
- [2] T. Fraichard and A. Scheuer. From reeds and shepp's to continuous-curvature paths. *Robotics, IEEE Transactions on*, 20(6):1025–1035, 2004.
- [3] E. Panteley, E. Lefeber, A. Loria, and H. Nijmeijer. Exponential tracking control of a mobile car using a cascaded approach. In *Proceedings of the IFAC workshop on motion control*, pages 221–226. Pergamon Grenoble, France, 1998.
- [4] J. Reeds and L. Shepp. Optimal paths for a car that goes both forwards and backwards. *Pacific journal of mathematics*, 145(2):367–393, 1990.

Constraining Operation Delay for Dynamic Power Optimization of Asynchronous Circuits

Shunya Hosaka
University of Aizu
Aizuwakamatsu 965-8580, Japan
m5191127@u-aizu.ac.jp

Hiroshi Saito
University of Aizu
Aizuwakamatsu 965-8580, Japan
hiroshis@u-aizu.ac.jp

ABSTRACT

In this paper, we propose a dynamic power optimization method for asynchronous circuits with bundled-data implementation constraining operation delay. In asynchronous circuits, we can change the execution time of each operation freely under a given latency constraint. Therefore, the proposed method relaxes the execution time of operations which consume more power while it tightens the execution time of operations which consume less power. In the experiments, we evaluate the effects of the proposed method by comparing area, execution time, power consumption, and energy consumption among synchronous circuits, asynchronous circuits without the proposed method, and asynchronous circuits with the proposed method. As a result, we can reduce dynamic power and energy consumption about 11% and 14% on average.

Categories and Subject Descriptors

VLSI [RTL Design]: Asynchronous Circuits

General Terms

Theory

Keywords

Dynamic Power Optimization, Mobility

1. INTRODUCTION

Current VLSIs are mostly based on synchronous circuits where circuit components are controlled by global clock signals. However, synchronization failures caused by clock skews and power consumption on the clock network are getting to be the critical problems when the integration technology of VLSIs is more and more advanced. There are no problems related to clock signals in asynchronous circuits since they do not use global clock signals. Instead, circuit components in asynchronous circuits are controlled

by local handshake signals. As circuit components are executed when required, asynchronous circuits are potentially low power consumption and low electro-magnetic interference.

However, it is well known that designing asynchronous circuits is more difficult than designing synchronous circuits. We have to consider delay model and data encoding when we design asynchronous circuits. If the delay model and data encoding are not appropriate, power reduction by asynchronous circuits may be small. To reduce power consumption, it is necessary for asynchronous circuits to optimize power.

In this paper, we propose a dynamic power optimization method for asynchronous circuits with bundled-data implementation constraining operation delay. In asynchronous circuits, we can change the execution time of operations freely under a given latency constraint. Therefore, the proposed method relaxes the execution time of operations which consume more power while it tightens the execution time of operations which consume less power.

Several low power design methods for asynchronous circuits have been proposed. Huang et al. [2] proposed a power optimization method which isolates the operands of functional units if operations are not required. Hansen et al. [1] proposed an optimization method to optimize latency, area, and energy considering many-to-many functional units mappings using the branch-and-bound algorithm. Jeong et al. [3] proposed an optimization method to reduce the overhead of dual-rail asynchronous circuits based on eager evaluation using multiple outputs block level relaxation. Plana et al. [4] proposed a throughput optimization method for non-pipelined asynchronous circuits avoiding data hazards and glitches using a concurrent sequencer. The improvement of the throughput results in energy optimization. Compared to these methods, our proposed method reduces dynamic power constraining operation delays.

The rest of this paper is as follows. In section II, we describe asynchronous circuits with bundled-data implementation, data flow graph, and the calculation of mobility. In section III, we describe the proposed dynamic power optimization method. In section IV, we evaluate the proposed method in terms of area, performance, dynamic power, and energy, comparing with synchronous circuits and asynchronous circuits without the proposed method. Finally, in section V, we describe conclusions and future work.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

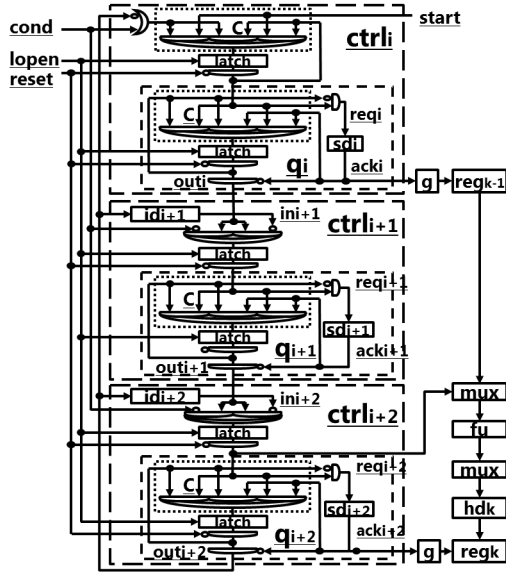


Figure 1: Bundled-data implementation

2. BACKGROUND

2.1 Bundled-data Implementation

Bundled-data implementation represents N-bit data using N+2 signals called "bundle". Additional two signals represent local handshake signals, request and acknowledge. In bundled-data implementation, delay elements whose delays are longer than data-path delays are put on request signals to guarantee the completion of operations and the timing of register writing. Therefore, the performance of bundled-data implementation depends on the delay of the control circuit including delay elements. The detail of bundled-data implementation is described in [5].

2.2 Data Flow Graph

Data Flow Graph (DFG) represents data flow of an application. DFG is defined by the following expression.

$$DFG = \langle N, E \rangle \quad (1)$$

N is the set of nodes. A node n_m shows an operation in the application. Node n_m has a label of allocated resource with the operation delay. E is the set of edges. An edge represents data dependence, resource sharing, or control dependence. A data dependence represents dependence of data between operations. A resource dependence represents resource sharing of operations. A control dependence represents which state an operation is executed. Fig.2(a) shows an example of DFG. In Fig.2(a), source and sink shows start and end nodes. There are no operations.

2.3 Mobility

Mobility of operations used for dynamic power optimization in the proposed method is defined using As Soon As Possible (ASAP) scheduling and As Late As Possible (ALAP) scheduling. ASAP scheduling schedules each operation as soon as possible when they can be scheduled. Fig.2(b) shows ASAP scheduling for the DFG in Fig.2(a). ALAP scheduling schedules each operation as late as possible under a given

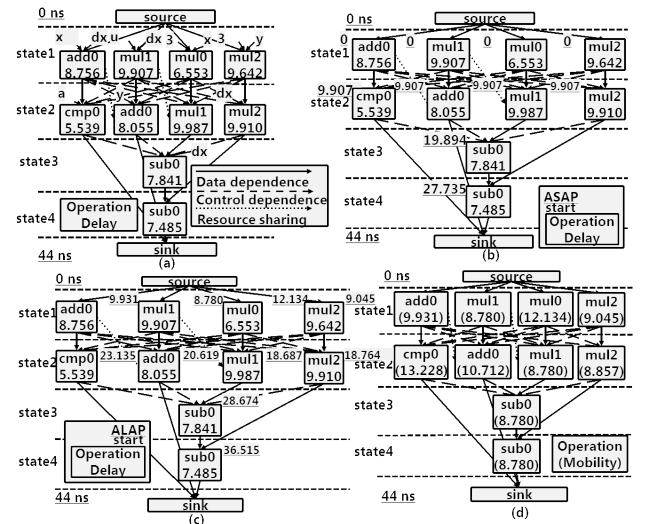


Figure 2: (a) DFG, (b) ASAP scheduling, (c) ALAP scheduling, (d) Mobility

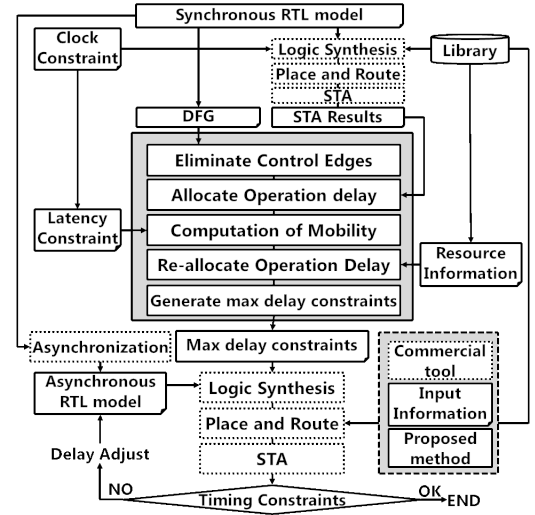


Figure 3: Design flow

latency constraint. Fig.2(c) shows ALAP scheduling for the DFG in Fig.2(a). The mobility of node n_m is defined by the following expression.

$$mobility_{n_m} = t_{alap_{n_m}} - t_{asap_{n_m}} \quad (2)$$

$t_{alap_{n_m}}$ shows ALAP start time of node n_m . $t_{asap_{n_m}}$ shows ASAP start time of node n_m . Fig.2(d) shows calculated mobility for nodes in the DFG of Fig.2(a).

3. PROPOSED METHOD

3.1 Design Flow

The proposed method generates max delay constraints to relax the execution time of operations which consume more power while tightening the execution time of opera-

tions which consume less power. Fig.3 shows a design flow with the proposed method. The proposed method is used in the process of generating an asynchronous RTL model from a synchronous RTL model. As inputs of the proposed method, we need to prepare a DFG corresponding to the synchronous RTL model's data-path circuit, data-path delays from Static Timing Analysis (STA), resource information file, and latency constraint. First, we allocate data-path delays from STA to each node in DFG. Next, we compute the mobility of operations from ASAP scheduling and ALAP scheduling. Then, we re-allocate operation delays to utilize mobility well. Finally, we generate max delay constraints for each data-path.

We consider two cases for re-allocation of operation delays.

- Case 1. Keep the state that operations are executed.
- Case 2. To maximize the use of mobility, we change the state of operations in non-critical paths to be executed.

Case1 reduces dynamic power by changing each state's execution time. In addition to Case1, Case2 reduces dynamic power by changing the state of operations in non-critical paths to be executed.

3.2 Inputs

We need to prepare a DFG corresponding to the synchronous RTL model's data-path circuit, data-path delays from STA, a resource information file, and latency constraint. To correspond to the synchronous RTL model's data-path circuit, in DFG, node n_m is categorized into each state, and node n_m is re-allocated a label that used operation. Between nodes are connected by edges that represent data dependence, resource sharing, and control dependence. Fig.2(a) shows an example of DFGs. We use STA to analyze data-path delays. The resource information file represents the range of each operation's delay. The proposed method gives high priority to operations that have long execution delay. Operation that takes long execution time can use mobility as much as possible. Latency constraint that represents the constraint from arrival of input signals to generation of output signals is computed by the product of clock cycle time and clock cycles.

3.3 Operation Delay Re-allocation

In the Case2, we eliminate control edges from a given DFG since it allows to change the execution timing of operations in non-critical paths. However, in the case that one operation's source register is shared as other operation's destination register in the same state, it may change the source register if the other operation completes early. In such a case, we need to keep control edges. Fig.4(a) shows the DFG that is eliminated control edges from the DFG in Fig.2(a). In this figure, source register of add0 is shared as destination register of mul1 at state2. Because of this, the control edge that exists between add0 to sub0 that comes executed after mul1 is kept. Next, we apply ASAP scheduling and ALAP scheduling and compute mobility in both Case1 and Case2. Fig.2(d) and Fig.4(b) represent Case1's mobility and Case2's mobility.

We execute operation delay re-allocation according to the flows in Fig.5. Fig.5(a) represents the Case1. First, we identify the critical path. We explore the number of resource

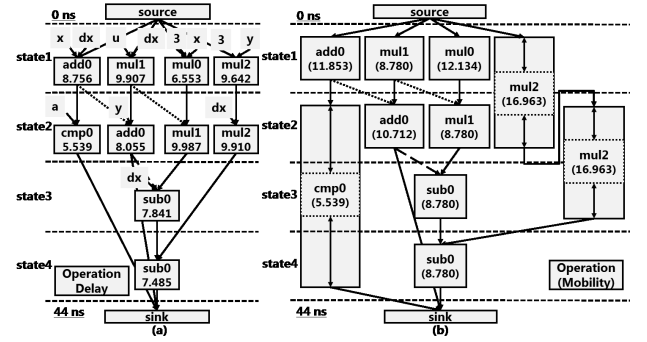


Figure 4: (a) DFG reduced control edges from Fig.2 (a), (b) Case2's mobility

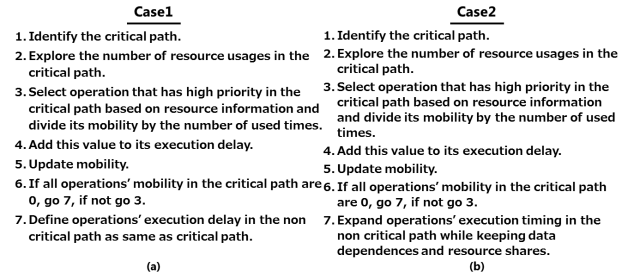


Figure 5: Re-allocation of execution delay (a) Case1, (b) Case2

usages in the critical path. Next, we select the operation in the critical path which has the highest priority based on the resource information file. Then, we divide the mobility by the number of resource usages of the corresponding resource. Next, we add this value to the execution delays of operation in the critical path that uses the same resource. Re-allocated execution delay corresponds to the delay of the state that belonging selected operation. Next, we update the mobility of all operations and select the operation that has the next highest priority and compute the execution delay. We repeat this processes until all mobility in the critical path become 0. When operation delay re-allocation for the critical path is finished, we set the execution delay of operating in non-critical paths based on the operation delay in the critical path. Fig.6 shows the operation delay re-allocation result for Case1.

Fig.5(b) represents the Case2. As same as Case1, first we re-allocate execution delay to operations in the critical path. However, there is a difference for the re-allocation of execution delay to operating in non-critical paths. In the Case2, there are less control edges than Case1. So we can use mobility more effectively. We may allocate more delay to operation in non-critical paths while keeping data dependences and resource sharing. Fig.7 shows the operation delay re-allocation result for Case2.

3.4 Outputs

Finally, we generate max delay constraints for all data-path based on re-allocated operation delays. They are classified paths, between registers, paths from primary inputs

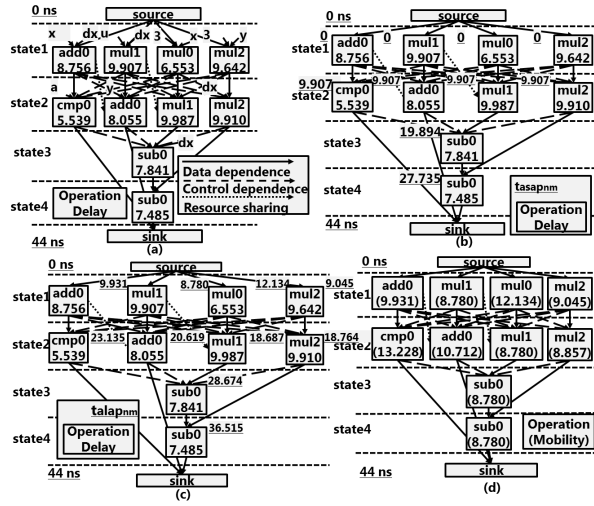


Figure 6: Operation delay re-allocation result for Case1

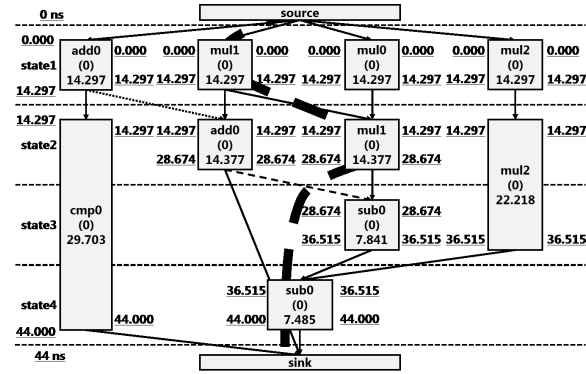


Figure 7: Operation delay re-allocation result for Case2

to registers, paths from registers to primary outputs, and paths from control module's latch to registers. Max delay constraints were generated using *set_max_delay* command as follows.

```
set_max_delay -from START -to END
               -through THROUGH DELAY
```

START and END represents path's start and end point. THROUGH represents multiplexer or functional unit's input or output port. DELAY represents the re-allocated execution delay. It may be required to change register write signals and multiplexer control signals in Case2. So, the proposed method generates a txt file which represents the change of control timing. Fig.8 shows the txt file. Designers are required to change asynchronous RTL model.

4. EXPERIMENTS

In the experiments, we synthesize differential equation solver (Diffeq) and ellipse wave filter (EWF) using the proposed design flow in Fig.3. We evaluate area, execution time, dynamic power consumption, and energy consumption of synchronous circuits (Sync), asynchronous circuits that the

```
[9] reg8 : ack7 -> ack8
    mux : in7 -> in8

[20] reg1 : ack10 -> ack11
    mux : in10 -> in11
```

Figure 8: txt output following Case2

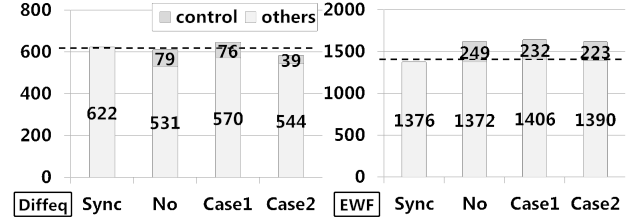


Figure 9: # of Logic Elements

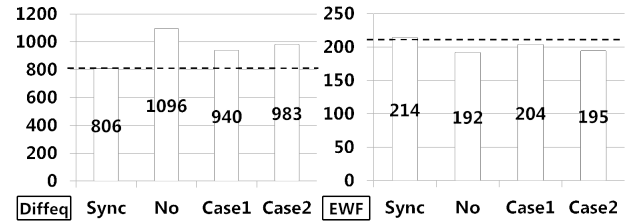


Figure 10: Execution time [ns]

proposed method is not applied (No), and asynchronous circuits using the proposed method (Case1 and Case2) to confirm the effect of the proposed method for reducing dynamic power consumption. For the experiments, we implemented the proposed method using Java and Eclipse.

We used Altera FPGA Cyclone IV (EPCE115F29C7). We used Quartus II 13.1 for synthesis. First we explore the fastest synchronous circuits satisfying timing constraints. Clock cycle time of Diffeq was 11ns and clock cycle time of EWF was 14ns. We generated asynchronous RTL models from synchronous RTL models. These models are modeled by Verilog HDL. The execution delay of operating used in the proposed method is defined by STA results for all data-paths of synchronous circuits. After that, we generated maximum delay constraints based on Case1 and Case2 under a given latency constraint that was computed by the product of clock cycle time and clock cycles. Finally, we synthesized asynchronous circuits using maximum delay constraints and adjusted delay elements repeatedly until all timing constraints for bundled-data implementation [5] were satisfied.

Fig.9 shows area. Area was evaluated from the report of Quartus II. In Diffeq, Case1 was increased about 6% and Case2 was reduced about 4%, compared with No. In EWF, Case1 was increased about 1% and Case2 was reduced about 1%, compared with No. In comparison with Sync, area was increased about 4% to 19%. This is because the overhead of the control circuits.

Fig.10 shows execution time that can be evaluated by simulating an arbitrary test pattern using ModelSim-Altera. In Diffeq, Case1 and Case2 were reduced about 14% and 11%, compared with No. In EWF, Case1 and Case2 were in-

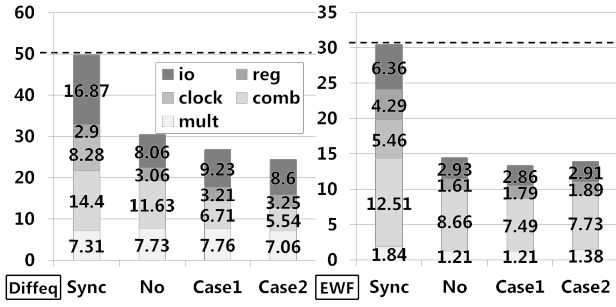


Figure 11: Dynamic power [mW]

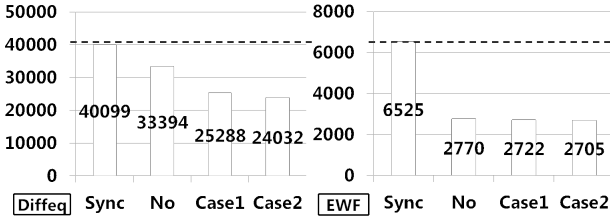


Figure 12: Energy [pJ]

creased about 6% and 1%, compared with No. Execution time was increased about 17% to 36% in DiffEq. In DiffEq, operations are repeatedly executed in a loop until a condition is satisfied. The idle phase of the control circuit between loop bodies could not be hidden. This is the reason why execution time is increased.

Fig.11 shows dynamic power that can be evaluated by PowerPlay Power Analyzer with a value change dump file obtained by simulation. In DiffEq, Case1 and Case2 were reduced about 12% and 20%, compared with No. In EWF, Case1 and Case2 were reduced about 7% and 4%, compared with No. Dynamic power is especially reduced in combinational logics due to the effect of the proposed method. In some cases where total area is increased, dynamic power could be reduced. This is considered as the reduction of dynamic power at routing resources.

Fig.12 shows energy consumption that was calculated by the product of execution time and dynamic power. In DiffEq, Case1 and Case2 were reduced about 24% and 28%, compared with No. In EWF, Case1 and Case2 were reduced about 2% and 3%, compared with No.

In the experiments, we could reduce dynamic power consumption more than asynchronous DiffEq (No) and EWF (No) using the proposed method. Especially, Case2 reduced more compared to Case1 due to the use of mobility more.

5. CONCLUSIONS

In this paper, we proposed a dynamic power optimization method for asynchronous circuits constraining operation delay using the mobility of operations. The proposed method generates maximum delay constraints to relax the execution time of operations that consume more power, while it tightens the execution time of operations that consume less power. We proposed two types of operation delay re-allocation methods. In the experiments, we have confirmed

that the proposed method reduced dynamic power consumption and energy consumption about 11% and 14% on average.

6. REFERENCES

- [1] J. Hansen and M. Singh. A fast branch-and-bound approach to high-level synthesis of asynchronous systems. *Asynchronous Circuits and Systems (ASYNC)*, IEEE International Symposium on, 10:107–116, May 2010.
- [2] Y. Huang and W. Shi. An optimized de-synchronization flow for power and performance optimization. *Systems and Informatics (ICSAI)*, International Conference on, 5:71–75, May 2012.
- [3] C. Jeong and S. Nowick. Block-level relaxation for timing-robust asynchronous circuits based on eager evaluation. *Asynchronous Circuits and Systems (ASYNC)*, IEEE International Symposium on, 10:107–116, May 2008.
- [4] L. Plana and S. Nowick. Architectural optimization for low-power nonpipelined asynchronous systems. *Very Large Scale Integration (VLSI) Systems*, IEEE Transactions on, 10:56–65, August 2002.
- [5] K. Takizawa, S. Hosaka, and H. Saito. A design support tool set for asynchronous circuits with bundled-data implementation on fpgas. *Field Programmable Logic and Applications (FPL)*, International Conference on, 4:1–4, September 2014.

Hybrid Tsunami Modeling Infrastructure: Tsunami Source Data and Bathymetry Editor

Sh. Takano, K. Hayashi

University of Aizu

Tsuruga, Ikki-Machi

Aizu-Wakamatsu, Japan

+81-242-37-2717

{s1200197, d8161103}@u-
aizu.ac.jp

A. Vazhenin

University of Aizu

Tsuruga, Ikki-Machi

Aizu-Wakamatsu, Japan

+81-242-37-2717

vazhenin@u-aizu.ac.jp

An. Marchuk

ICMMG SB RAS

pr. Akademika Lavrentjeva, 6

Novosibirsk, 630090, Russia

+7-383-33087-83

mag@omzg.sscs.ru

ABSTRACT

The important part of the tsunami science is focused on studying the considerable influence of natural geographical objects, like islands and coast bathymetry, on the tsunami waves. Currently, such investigations are mostly implementing by physical modeling allowing obtaining good results on impacting submarine barriers on tsunami wave propagation but actually very expensive. We are designing a system allowing numerical computer simulations for crucial coastal areas supporting so-called hybrid bathymetry that combines natural and artificial underwater objects as well as tools allowing the user to manipulate with them. The paper describes the main features the original Bathymetry and Tsunami Source Data Editor that allows tuning/editing bathymetric and tsunami source data by including/removing artificial barriers as well as specifying their placement, shapes and sizes.

Categories and Subject Descriptors

I.6.7 [Simulation and Modeling]: Simulation Support Systems – *Environments, Animation, Visual.*

General Terms

Algorithms, Design, Experimentation.

Keywords

Tsunami Modeling, Integration of Heterogeneous Software Components, Hybrid Bathymetry, Tsunami Source and Bathymetry Data Editor

1. INTRODUCTION

The important part of the tsunami science is focused on studying the considerable influence of natural geographical objects, like islands and coast bathymetry, on the tsunami waves. While the complete damage assessment for the Great Japanese Earthquake event is still underway, the immense impact of this tsunami raises questions about mitigating the

impact for such an event at different time scales, from real-time tsunami warning guidance to long-term hazard assessment. Assessing the use of real-time tsunami forecasting tools is an important part of this process. The necessity to embed modeling tools is shown for Earthquake and Tsunami Warning System for Natural Disaster Prevention and Defence Systems [1-2]. One such example known as “Matsushima effect” showed big influence of placement and sizes of natural geographical objects like islands and bathymetry on the tsunami wave parameters such as wave height and speed.

Currently, such investigations are mostly implementing by physical modeling in basins (several meters in length and less than one meter in depth) [3]. It was also shown that parameters of simulations could be transformed to natural conditions. These experiments showed good results on impacting submarine barriers on tsunami wave propagation but actually very expensive.

In addition to the physical modeling, we are designing a system allowing numerical computer simulations for crucial coastal areas supporting so-called hybrid bathymetry combining natural and artificial underwater objects as well as tools allowing the user manipulate with them [4-5] in order to make it more intensive. Accordingly, The aim of modeling process is in finding preliminary suitable number, sizes, and placement of submarine bathymetry objects in order to minimize the dangerous tsunami wave parameters (height and speed). In this way it may be possible to design and build a set of digital artificial objects (islands) that can be used to protect the coastal areas. In particular, such protection could be of extreme value in highly populated areas, as well as in industrial areas (e.g. nuclear plants, factories, airports, etc.).

The paper describes the main features the original Bathymetry and Tsunami Source Data Editor (TSB-editor) that allows tuning/editing bathymetric and tsunami source data by including/removing artificial barriers as well as specifying their placement, shapes and sizes. The rest of the paper is organized as follows. Section 2 explains the mathematical model for simulating wave propagation explaining data source sets needed for numerical modeling. In Sections 3, we show main features of the TSB-editor architecture and interface. Section 4 describes present results of tsunami modeling with hybrid bathymetry designed via the TSB-editor. Finally, we conclude with remarks and comments about future work.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.

Copyright 2015 University of Aizu Press.

2. TSUNAMI MODELING FEATURES

2.1 Theoretical Background

There are a number of algorithms and models developed for the tsunami risk mitigation covering phases of generation, propagation from the deep ocean to the coastal areas. The most known, accurate and widely used are TUNAMI [6] and MOST [7-8] packages calculating the long wave propagation in the ocean by the so-called shallow-water differential equations:

$$\begin{cases} H_t + (uH)_x + (vH)_y = 0, \\ u_t + uu_x + vv_y + gH_x = gD_x, \\ v_t + uv_x + vv_y + gH_y = gD_y, \end{cases} \quad (1)$$

where $H(x, y, t) = h(x, y, t) + D(x, y, t)$, h is the water surface displacement, D is depth, $u(x, y, t)$ and $v(x, y, t)$ are velocity components along the axis' x and y . g is acceleration of gravity. Accordingly, the tsunami propagation velocity does not depend on its length and is expressed by the so-called Lagrange formula $c = \sqrt{g(D + \eta)}$ [5,6] that plays the key role for the long-wave (tsunami) kinematics. The horizontal flow velocity depends on the wave amplitude and water depth

$$u = \eta \sqrt{\frac{g}{D}}. \quad (2)$$

The numerical algorithm is based on splitting the difference scheme, which approximates equations (1) by spatial directions as well as permits to set boundary conditions for a finite difference boundary value problem using a characteristic line method. The splitting method used for shallow wave equations (2) comprises a consecutive numerical solution of two one-dimensional systems of equations:

$$\begin{cases} V_t + UV_x = 0 \\ U_t + UU_x + gH_x = gD_x \\ H_t + (UH)_x = 0 \end{cases} \quad \begin{cases} V_t + VV_x + gH_x = gD_x \\ U_t + VU_x = 0 \\ H_t + (VH)_x = 0 \end{cases} \quad (3).$$

Eigenvalues of (3) are real and different, and the system can be

$$\begin{cases} V'_t + \lambda_1 V'_x = 0, \\ P'_t + \lambda_2 P'_x = 0, \\ Q'_t + \lambda_3 Q'_x = 0, \end{cases} \quad (4)$$

where $\lambda_1 = U$, $\lambda_{2,3} = U \pm \sqrt{gH}$ are eigenvalues, $V' = V'$, $P' = U + 2\sqrt{gH}$, $Q' = U - 2\sqrt{gH}$ are the Riemann invariants. The characteristic line method has been used to set the boundary conditions for the system. For the numerical solution of the system, the following finite difference scheme is used:

$$\begin{aligned} & \frac{\dot{W}_{i-1}^n - \dot{W}_i^n}{\Delta t} + A \frac{\dot{W}_{i-1}^n - \dot{W}_{i-1}^{n-1}}{2\Delta x} - A\Delta t \frac{A(\dot{W}_{i-1}^n - \dot{W}_i^n) - A(\dot{W}_i^n - \dot{W}_{i-1}^n)}{2\Delta x^2} = \\ & = \frac{\dot{F}_{i-1} - \dot{F}_i}{2\Delta x} - A\Delta t \frac{\dot{F}_{i-1} - 2\dot{F}_i + \dot{F}_{i-1}}{2\Delta x^2}, \end{aligned} \quad (5)$$

where

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \dot{W} = (V', P, Q), \dot{F} = (0, gD_x, gD_x).$$

The criterion of stability is $\Delta t \leq \Delta x / \sqrt{gH}$ requiring assignment of a smaller time step if a computational domain

contains deep-water areas. Obviously, the digital bathymetry D and initial water elevation (tsunami source) are the key elements influencing to tsunami waves. In our experiments we use the MOST package [7-8].

2.2 Bathymetry Data

Bathymetry is the measurement of the depth of water in oceans, rivers, or lakes [9-10]. Bathymetric maps look a lot like topographic maps, which use lines to show the shape and elevation of land features. On topographic maps, the lines connect points of equal elevation. On bathymetric maps, they connect points of equal depth. It can also indicate a seamount, or underwater mountain.

Figure 1 represents a way of coding the digitalized bathymetry and tsunami waves data. Depth values may be either negative or positive, but should all be understood to be negative. Elevations or wave heights are above sea level and considering to be positive. The gridded bathymetry is the two dimensional array that is used as a source data for modeling program. Additionally, it contains some information specifying longitude and latitude in order to associate these data with real geographic areas. Figure 1b shows a hybrid bathymetry obtained by recalculating the natural bathymetry and containing an artificial object. Shape, size and placement of this object are saved in a special file.

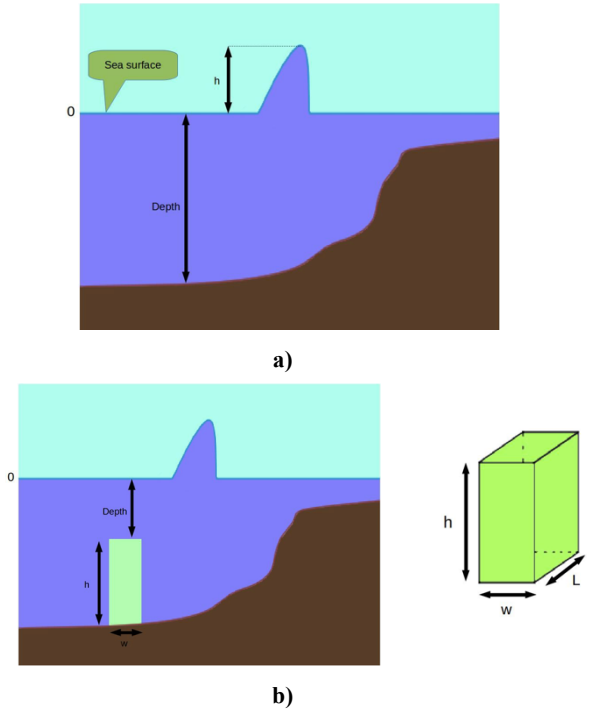


Figure 1. Natural and hybrid bathymetry

2.3 Tsunami Source Model

There is an initial water surface elevation as a result of numerical modeling of the elastic-plastic problem with seismic source with specified parameters. To study the ratio between the initial wave height and wave parameters near the coast, it is necessary to carry out a number of numerical experiments with specified initial parameters. Accordingly, the initial water surface displacement can be specified by ellipsoidal shape

$$H(i, j) = (1 + \cos(\pi \cdot \arg(i, j))) \cdot H_0$$

where H_0 is half of the water surface displacement at the central point (i_0, j_0) of the ellipse. The parameter $\arg(i, j)$ gives the ratio between the distance to the ellipse center and the distance to the ellipse border in this direction

$$\arg(i, j) = \left(\frac{(i - i_0) \cdot \cos(\beta) + (j - j_0) \cdot \sin(\beta)}{r_1} \right)^2 + \left(\frac{(j - j_0) \cdot \cos(\beta) - (i - i_0) \cdot \sin(\beta)}{r_2} \right)^2$$

Here r_1 , r_2 are the ellipse axis length and β is the long axis azimuth. Figure 2 shows the shape of the two meters height ellipsoidal source with the axis' ratio equal to two, and the water height distribution along the ellipse axis.

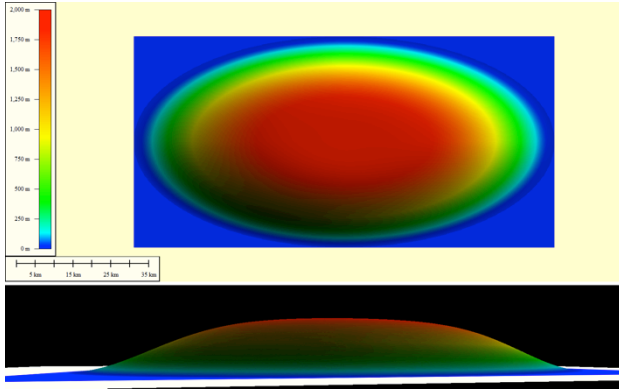


Figure 2. The shape and cross-section of a model ellipsoidal tsunami source

3. TSB-EDITOR ARCHITECTURE AND INTERFACE

TSB-editor was developed supporting object-oriented GUI-editing on bathymetric data as well as including/removing artificial objects with variable placement, shapes and size (Figure 3). As was pointed, above it manipulates with the gridded bathymetry as a two dimensional array that is used as a source data for modeling program. The editing process can be implemented using two windows named Controller and Viewer correspondingly (Figure 4). The Viewer window visualizes the state of the edit data, and Controller is used for specifying object parameters, control the editing mode, reading and saving bathymetry and tsunami source files, etc.

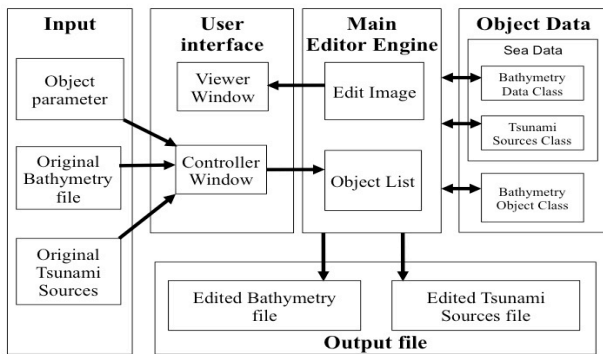


Figure 3. Main elements of the TSB-editor

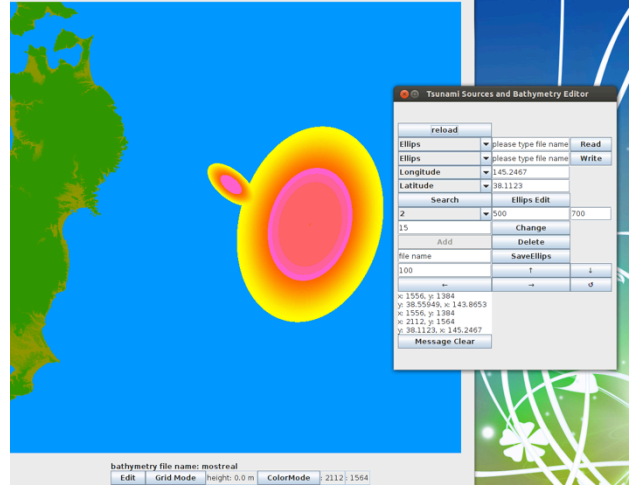


Figure 4. Interface of TSB-editor

An editing process is started from reading an bathymetry file and setup the editing parameters like array size, coordinates, etc. Then the user defines the object location by inputting directly its longitude and latitude values or getting position by clicking a picture on Viewer window. The user should specify height, width, angle and elevation/depth of an object that is generated by pressing the “ADD” button on the controller window. In any time the user can change any object by choosing it from a list, inputting new parameters, and fixing these changes by the “CHANGE” button. If the user would like to add a new object, user chose new object from list, a parameter is input and it's added.

The editor saves a file according to the edit mode (Tsunami sources or Bathymetry) including an edit history file. Every objects data is stored in a container of the vector type. The Object data has object number. This number is different respectively. The controller can search and edit every object data by object number. Objects data has object parameter and object's depth map (2-dimensional array data). Current version of the Editor is used the NOAA Bathymetry data. In future we are planning to extend it by other GIS data formats.

The editing is similar for both tsunami sources and bathymetry with a few differences. For bathymetry, the depth is expressed by a blue gradation, and an artificial object is also corresponding to this gradation. It should also be defined by its height and width. For tsunami sources, all sea surfaces have the same color, and the Tsunami waves height is expressed by a red colored gradation because of no relation to the sea depth. It requires specifying the side and lengthwise radiuses.

4. MODELING EXPERIMENTS

A set of numerical experiments was provided with hybrid bathymetry. Accordingly, the 2148x1074 knots gridded bathymetry was created for the Oppa Bay and the neighboring harbors (approximately 17 m). These data cover the geographical area from 141.41659° E to 141.75° E and from 38.5° N up to 38.6666° N. In Table 1, the results of ten numerical experiments are presented with the hybrid bathymetry of a parallelepiped shape barriers of different sizes and placement. Figure 5 shows a visual comparison of the

Tsunami wave behavior for natural (upper image) and hybrid bathymetry reflecting the case seven in the Table1.

Table 1. Source Data and Experimental Results

No	Object Place in °	Object Size, (km)	Object Climax, (m)	Max Wave (m)	Max Wave Place
0	None	None	None	3.10	None
1	38.577148 141.6129	9.443x 0.9443	-63	2.80	Dotted the coastline
2	38.577148 141.6129	4.721x 0.9443	-12	2.70	Dotted the coastline
3	38.577148 141.6129	9.443x 0.9443	-12	2.35	Intricate of South
4	38.577148 141.6129°	9.443x 0.9443	-63	2.80	Intricate of South
5	38.577148 141.5668	9.443x 0.9443	-35	2.75	Dotted the coastline
6	38.577148 141.566895	4.722x 0.9443	-35	2.75	Dotted the coastline
7	38.577148 141.566895	4.722x 0.9443	-7	2.50	Intricate of North
8	38.577148 141.566895	9.443x 0.9443	-7	2.40	Intricate of North

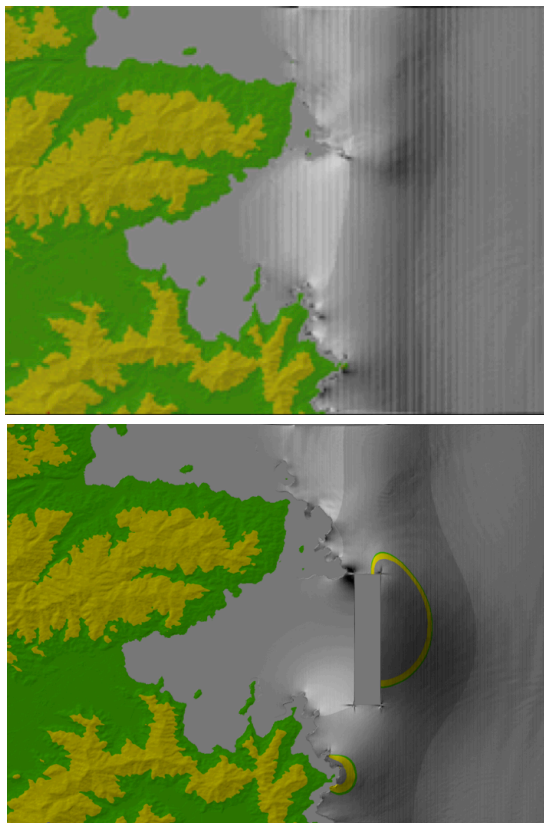


Figure 5. Visualization of Tsunami waves for natural (upper image) and Hybrid bathymetry of the Oppa bay

5. CONCLUSION

The current version of the Editor allows the user to manipulate with ellipsoidal shapes (for editing bathymetry and wave sources) described above as well as parallelepiped-like objects (for editing bathymetry) such as cubes, walls, etc. In the future we are planning to extend this set by pyramids, sphere, etc. The user can create composite objects using this set of elementary objects.

These results demonstrate a good possibility to realize this new type of tsunami modeling. As shown in figure, the tsunami wave behavior on hybrid bathymetry is significantly different from the natural one. This makes also possible to control the tsunami wave height by underwater artificial objects as well as provide studying features of the natural bathymetry.

6. ACKNOWLEDGMENTS

This work is supported by the Grand-In-Aid for Scientific Research (Basic), 2015-2018, Japan, Issue Number: 21500948.

7. REFERENCES

- [1] *The International 26th Tsunami Symposium (ITS2013)*, (Gosek, Turkey, September , 2013). <http://www.tsunami2013.org>
- [2] *The 25th (2015) International Offshore and Polar Engineering Conference*, (Kona, USA, June 21-26, 2015).
- [3] Fridman A. M., Alperovich L. S., L. Shemer, Marchuk An. G., at al. 2010. Tsunami wave suppression using submarine barriers, *Physics-Usppekhi*, 53, 8 (August, 2010), 809– 816. DOI: [10.3367/UFNe.0180.201008d.0843](https://doi.org/10.3367/UFNe.0180.201008d.0843).
- [4] Hayashi K., Vazhenin A., Marchuk An. 2014, Application Engines in VMVC-based Tsunami Modeling Environment, *Frontiers in Artificial Intelligence and Applications*, IOS Press, 265 (2014), 464-475. DOI: 10.3233/978-1-61499-434-3-464.
- [5] Marchuk An. , Hayashi K., Vazhenin A. 2015, Trans-boundary realization of the nested-grid algorithm for trans-pacific and regional tsunami modeling, *Bull. Nov. Comp. Center: Math. Model. in Geoph.*, NCC Publisher, 18 (2015), 35-47.
- [6] Shuto N., Imamura F., Yalciner A., Ozyurt, G. 1995, *TUNAMI N2: Tsunami Modeling Manual*, (1995). <http://tunamin2.ce.metu.edu.tr/>
- [7] Titov V. 1988, *Method for Numerical Modeling of Tsunami Taking into Account the Wave Transformation in Shallow Water*, Preprint #771, Computing center SD USSR Academy of Sciences, Novosibirsk, (1988).
- [8] Titov, V. 1989, Numerical modeling of tsunami propagation by using variable grid, In *Proceedings of the IUGG/IOC International Tsunami Symposium*, (Computing Center, Siberian Division, USSR Academy of Sciences, Novosibirsk, USSR, 1998), 46-51.
- [9] Smith, W. and Sandwell, D. 1997, Global seafloor topography from satellite altimetry and ship depth soundings, *Science* 277 (1997), 1956-1962.
- [10] Bathymetry Data (NOAA Bathymetry and Global Relief). <https://www.ngdc.noaa.gov/mgg/bathymetry/relief.htm>

Traveler Guide Assistant: Introducing an Application for an OpenStreetMap Based Travel Itinerary Construction

Alexander Baratynskiy
Peter the Great St. Petersburg Polytechnic
University
29 Polytechnicheskaya st.
195251 St. Petersburg Russia
baratynskiy@gmail.com

Evgeny Pyshkin
Peter the Great St. Petersburg Polytechnic
University
29 Polytechnicheskaya st.
195251 St. Petersburg Russia
pyshkin@icc.spbstu.ru

ABSTRACT

In this paper we describe an approach to the development of an application for user experience-centered travel itinerary construction. The developed prototype uses the OpenStreetMap API and provides the following major features: iterative tourist route construction; adding text, image and audio annotations; point of interest overview screen; access to recommendations for travelers.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous;
H.5.2 [Information Interfaces and Presentation]: User Interfaces—*human centered design*

General Terms

Design, Human Factors

Keywords

Human-centric computing, Information systems, Travel

1. INTRODUCTION

Being an immanent human activity since probably prehistoric times, traveling is an essential part of our life, one of the obvious ways to discover the outside world. Among different factors (including journey time, hotel accommodation, well-prepared equipment, etc.), the careful forethought of the journey makes the trip successful and coming up to the traveler's expectations. What kind of tools travelers (and guides as well) do usually use in order to prepare the tour itinerary? First, they use traditional paperback editions of numerous travel guides that contain plenty of map and city plan fragments with routes, points of interests together with some stories. Figure 1 shows an example that is taken from Dorling Kindersley travel guide for Paris [6].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT'15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

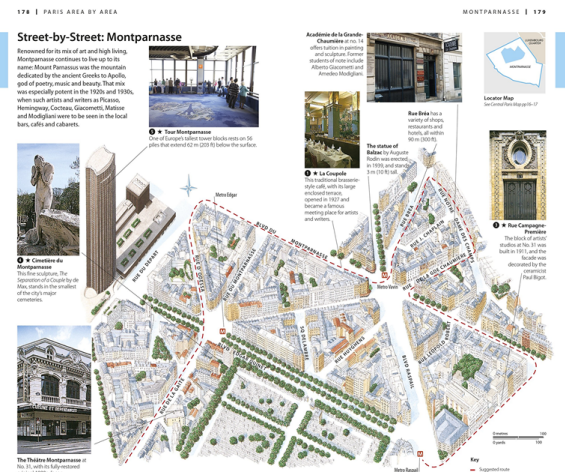


Figure 1: DK Eyewitness Travel Guide

Second, they consult different guides that are implemented as software applications (currently available on many mobile platforms). Third, they ask other people by using social networks and traveler forums like tripadvisor.com, expedia.com, etc. Finally, sometimes they construct their own itineraries, and this is the case that we address in our work above all things.

Thus, the general idea is to combine the user experience orientation and rich facilities of the present day information retrieval, processing and presentation tools for the domain of travel tour planning and construction and developing related services for the purposes of creating better infrastructure of big cities [8].

2. RELATED WORK

Tourism is the worldwide industry that demands huge amount of data delivery and processing. It is no wonder that nowadays numerous web resources are developed to allow people to plan their journeys and access different related services: reading travelers' reviews, comparing prices, reserving tickets, booking hotels and air flights, and so on.

Besides web resources there are research works and projects developed in the domain of information systems for tourism. In "Electronic Tour Guide for Android Mobile Platform with Multimedia Travel Book" [4] the services for creating travel

reporting in the form of a multimedia travel book are presented. Authors of “*Mobile Application for Guiding Tourist Activities: Tourist Assistant – TAIS*” [5] developed an application that generates recommendations collected on the base of other travelers’ experience and evaluations. There are also works focusing the problem of tourist route planning automation. An interesting example we can find in the work of Kachkaev and Wood from City University London presenting an approach for automated planning of leisure walks based on crowd-sourced photographic content [2].

Somehow, many current efforts are about delivering personalized solutions allowing travelers to leverage and to share their experience and to follow major scenarios we could learn on the base of study of traveler experience not without paying special attention to the mobile technology [3, 1].

Significant contribution is *Aurigo*, the project of Georgia Institute of Technology. This system seems to be one of the most advanced applications developed recently in the domain of interactive tools for travel planning [7]. The authors attempted to combine features for manual route construction with number of automated features.

3. MODELING A TRAVEL ITINERARY

Let us introduce an *annotated travel itinerary* concept. Unlike to the simple route which is often the same as a path (usually the shortest one) between two points on the map, the annotated travel itinerary is more complicated. We consider this as a set of points of interest (POI) representing different attractions like architectural sights, museums, historical places, monuments, view points, etc.

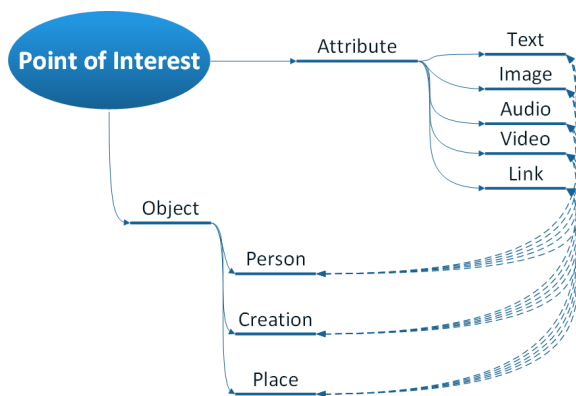


Figure 2: POI metaphor represented as a map of concepts

As Figure 2 illustrates, many related concepts can be linked to the central POI concept. In order to create a relevant annotation (which has to be able to be visualized in a sense), one should consider including such entities as description texts, images (photos, drawing, replicas, diagrams, etc.), multimedia objects (audio or video clips), web links, notes, citations, dates, timeline connections, information about related people, places or events, to cite a few.

Figure 3 shows a diversity of data to be visualized by using rich POI annotations (for example, the case of the Eiffel Tower illustrates well how many different allusions exist that a guide take into consideration in order to introduce the tourist attraction in the best way to the visitors).



Eiffel Tower

The Eiffel Tower is an iron lattice tower located on the Champ de Mars in Paris, France. It was named after the engineer Alexandre Gustave Eiffel, whose company designed and built the tower.

https://en.wikipedia.org/wiki/Eiffel_Tower

"I left Paris and even France because in the end, the Eiffel Tower annoyed me too much. Not only could you see it from wherever you went in the city, but you also found it everywhere, made in every material known to man, on sale in all the shop windows, an unavoidable and agonising nightmare."

Guy de Maupassant, La Vie Errante (1890)



Alexandre Gustave Eiffel

Alexandre Gustave Eiffel (15 December 1832 – 27 December 1923) was a French civil engineer and architect. A graduate of the prestigious École Centrale des Arts et Manufactures of France, he made his name with various bridges for the French railway network, most famously the Garabit viaduct. He is best known for the world-famous Eiffel Tower, built for the 1889 Universal Exposition in Paris, and his contribution to designing the Statue of Liberty in New York.

https://en.wikipedia.org/wiki/Gustave_Eiffel

Also,

Eiffel is an ISO-standardized, object-oriented programming language designed by Bertrand Meyer (an object-orientation proponent and author of Object-Oriented Software Construction) and Eiffel Software. The design of the language is closely connected with the Eiffel programming method.



Bertrand Meyer

Bertrand Meyer (born 1950) is a French academic, author, and consultant in the field of computer languages. He created the Eiffel programming language and the idea of design by contract.

https://en.wikipedia.org/wiki/Bertrand_Meyer

[https://en.wikipedia.org/wiki/Eiffel_\(programming_language\)](https://en.wikipedia.org/wiki/Eiffel_(programming_language))

Figure 3: Giving an idea of an annotation concept

The popular services like Google Play or AppStore give access to applications that in most cases implement the idea of book style travel guide providing access to the plenty of facts, but without rich facilities allowing users to be creative in planning their journey and to construct their own annotated routes. Particularly, if we use Google Custom Maps service, we are able to add POI markers together with some lines, shapes, etc. Our idea is to extend these facilities and to allow users to get more POI inspired information (including pictures, encyclopedic descriptions, external links, etc.) with respect to the particular map area and personal interests. We also pay attention to partial automation of itinerary construction with special focus on requirements provided by tourist guides.

Let’s consider what an amateur or a professional guide has to do in order to prepare a travel itinerary. One possible way is to rely on electronic cartographic services. Unfortunately, currently they don’t provide enough flexibility and user interface options allowing user to be able to construct reasonably complicated travel route (if this is not about finding the shortest path while getting directions).

So people often go back to traditional values and draw the route manually on the printed maps (having much flexibility in adding different kind of written annotations, but with few possibilities to share results of their work). Also, if one is not a professional tourist guide, he or she might need managing many different points of interest on the map.

To sum up, the process of the annotated travel itinerary construction is time and efforts consuming. Moreover, there are difficulties both to maintain the completed itinerary, and to implement the itineraries with variants. An ability to support variants is very important since there are many aspects conditioning variants: difference in travelers’ age or experience, possible changes in weather condition, unexpected closing of some sights, etc.

Thus, there are reasons to create an application implementing the introduced itinerary model and supporting major use cases that we believe useful for a traveler, and not only for a traveler alone, but also for an amateur or professional tourist guide who plans routes for his or her clients.

4. SOFTWARE PROTOTYPE

In this section we introduce the software prototype for personalized travel itinerary construction.

As we have mentioned in the previous section, the target user is twofold, since both travelers and tourist guides can be considered as users. Below major requirements for the software prototype are listed:

- An intuitive user-friendly interface;
- Ability to create the route itself and to add the annotations containing text information, images, audio, links to web sites, etc.;
- Ability to modify the created tourist route;
- Ability to search of POI-related information on the web.

In the following subsection we describe the major components related to the above mentioned scenarios.

4.1 Main Frame

The main frame (see Figure 4) contains the map itself and the tabbed panel with different tools that allow creating and modifying routes.

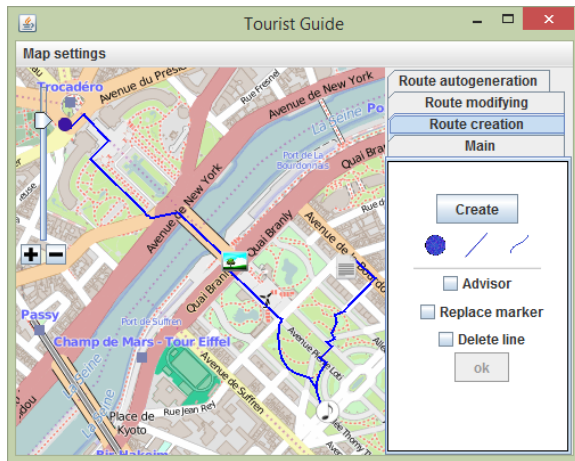


Figure 4: Main frame window with an example of constructed itinerary

4.1.1 Using OpenStreetMap as a Cartography E-Service

OpenStreetMap (OSM)¹ is a collaborative project to create a free editable world map. Currently OSM maps cover almost the whole world. The map is quite accurate; the accuracy depends on the users' activity in the region. The service is free to use, it doesn't have any restrictions to access its API. Also, OSM provides good map database. In our application the implementation uses *JMapView* component.

4.1.2 Annotated Itinerary Creation

With the help of the route creation panel user can create the annotated tourist route. This panel allows adding markers and lines, replacing markers and deleting lines. In

¹<http://www.openstreetmap.org>

the developed prototype a marker represents the view of annotation on the map. The icon of the marker depends on its annotation content. If the annotation contains an image, the marker would look like a photo. If the annotation contains an audio, the marker would look like the violin clef. It is important to mention that the constructed itinerary is modifiable.

4.2 Annotation Frame

As it was stated before, markers represent annotation views on the map. So, after clicking on the marker, the annotation frame opens. Images are loaded either from the local computer or from the Internet. The annotation frame user interface is shown in Figure 5.

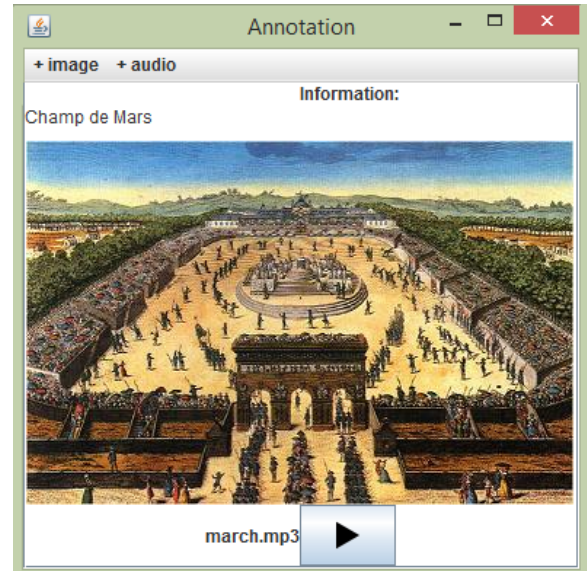


Figure 5: POI overview

4.3 POI Overview Window

The brief POI overview contains POI metadata. Currently they include POI address, name, related images and links (see Figure 6).

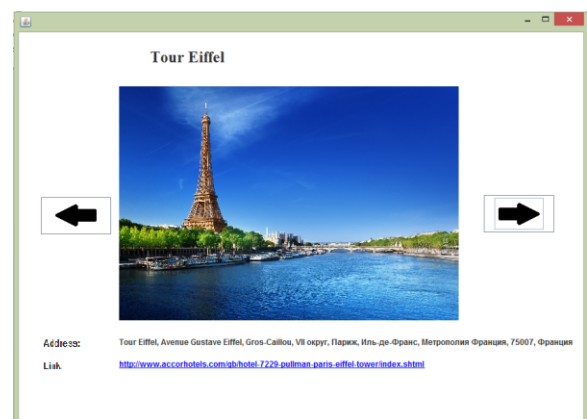


Figure 6: POI Overview Screen

To obtain the POI address by its latitude and longitude

values (got from the map) we use the methods of reverse geocoding based on *Nominatim*², a tool to search OSM data by name and address and to generate synthetic addresses of OSM points. Below is an HTTP-request example for extracting the POI address succeeded by the JSON-structure representing the delivered address information.

```
http://nominatim.openstreetmap.org/reverse?format=json&accept-language=en&lat=48.858&lon=2.2945&zoom=18&addressdetails=1
```

```
{
  "place_id": "47281851",
  "licence": "Data © OpenStreetMap ...",
  "osm_type": "way",
  "osm_id": "5013364",
  "lat": "48.8582602",
  "lon": "2.29449905431968",
  "display_name":
    "Eiffel Tower, Avenue Gustave Eiffel,
    Gros-Caillou, 7th Arrondissement, Paris,
    Ile-de-France,
    Metropolitan France,
    75007, France",
  "address": {
    "address29": "Eiffel Tower",
    "road": "Avenue Gustave Eiffel",
    "suburb": "Gros-Caillou",
    "city_district": "7th Arrondissement",
    "city": "Paris",
    "county": "Paris",
    "state": "Ile-de-France",
    "country": "France",
    "postcode": "75007",
    "country_code": "fr"
  }
}
```

The name is obtained from the POI address.

Images and links are obtained with the help of Google Web Search API.

4.4 Advisor

Users are able to select the area of search on the map in order to allow advisor component to suggest POIs considering including them to the itinerary. Figure 7 gives an idea of the advisor user interface.

5. CONCLUSION

In this short paper we described the software prototype for personalized interactive travel itinerary construction. In the future we plan to rethink the application architecture in order to provide support for the following requirements:

- Implementing software component to be executed on mobile devices;
- Integration with navigation facilities;
- Support for thematically oriented route construction;

²<http://wiki.openstreetmap.org/wiki/Nominatim>

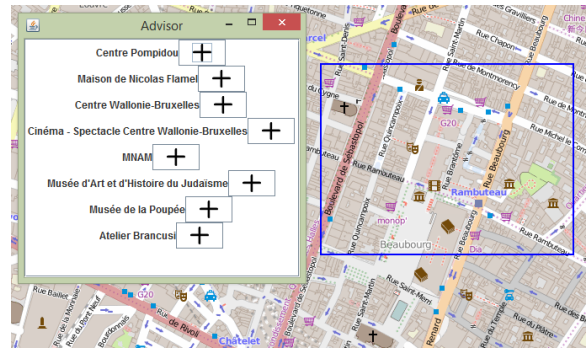


Figure 7: Advisor

- Integrations with the component for automated route generation.

The idea which differs significantly from similar applications mentioned before is to consider investigating the issue of using old maps available as images. A perspective to the ancient views accessible electronically may significantly extend the way to learn history while visiting tourist attractions all over the world.

6. REFERENCES

- [1] T.-D. Cao and N.-D. Tuan. Improving travel information access with semantic search application on mobile environment. In *Proceedings of the 9th International Conference on Advances in Mobile Computing and Multimedia*, MoMM '11, pages 95–102, New York, NY, USA, 2011. ACM.
- [2] A. Kachkaev and J. Wood. Automated planning of leisure walks based on crowd-sourced photographic content. *46th Annual Universities Transport Study Group Conference*, Newcastle, UK, 2014.
- [3] T. Y. Lim. Designing the next generation of mobile tourism application based on situation awareness. In *Network of Ergonomics Societies Conference (SEANES), 2012 Southeast Asian*, pages 1–7. IEEE, 2012.
- [4] V. Mladenovic, M. Lutovac, and M. Lutovac. Electronic tour guide for android mobile platform with multimedia travel book. In *Telecommunications Forum (TELFOR), 2012 20th*, pages 1460–1463, Nov 2012.
- [5] A. Smirnov, A. Kashevnik, N. Shilov, N. Teslya, and A. Shabaev. Mobile application for guiding tourist activities: tourist assistant-tais. In *Open Innovations Association (FRUCT16), 2014 16th Conference of*, pages 95–100. IEEE, 2014.
- [6] A. Tillier. *DK Eyewitness travel Guide: Paris*. Dorling Kindersley, 2014.
- [7] A. Yahi, A. Chassang, L. Raynaud, H. Duthil, and D. H. P. Chau. Aurigo: an interactive tour planner for personalized itineraries. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*, pages 275–285. ACM, 2015.
- [8] Y. Zheng, L. Capra, O. Wolfson, and H. Yang. Urban computing: concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(3):38, 2014.

Automated Leisure Walk Route Generation for an Interactive Travel Planner

Boris Skripal
Peter the Great St. Petersburg Polytechnic
University
29 Polytechnicheskaya st.
195251 St. Petersburg Russia
skipalboris@gmail.com

Evgeny Pyshkin
Peter the Great St. Petersburg Polytechnic
University
29 Polytechnicheskaya st.
195251 St. Petersburg Russia
pyshkin@icc.spbstu.ru

ABSTRACT

This paper presents a study of the state-of-the-art approaches for traveler assisting systems with particular attention paid to the problem of travel itinerary automated construction. We describe an algorithm for creating a route for a leisure walk to be included to the OpenStreetMaps based interactive travel planner. We discuss the current implementation and analyze further requirements from the viewpoint of better personalization of a tourist recommendation system.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous;
H.5.2 [Information Interfaces and Presentation]: User
Interfaces—*human centered design*

General Terms

Automation, Human Factors

Keywords

Human-centric computing, Information systems, Travel, Route generation

1. INTRODUCTION

Path finding is a common problem in many areas of information processing. It occurs while solving problems of getting directions, mapping a route for a leisure walk, finding an optimal topology of a digital circuit, tracing a software run and many others. The focus of this short paper is an application of route generation algorithms to the purposes of planning a traveler itinerary, probably the most obvious route construction related problem.

With current level of computer-assisted tools and methods, tourists expect more facilities than simply finding a shortest (or quickest) path or getting a direction. They require taking into account many competitive factors of route

construction, some of them (like sight attractiveness) aren't easily formalizable. The problems related to developing better personalized services for travelers are within the scope of the emerging domain of urban computing [13].

Tourists visiting some area for a certain period are unlikely able to visit every attraction. Effectively, they solve a kind of fuzzy optimization problem in order to select something more interesting to them personally. Tourists select the points of interest (POI) depending on their value from a certain point of view [11]. Hence, one of possible application of traveler advisory systems is to navigate the selection process by implementing the criteria representing a tourist attraction value by some formal schema.

In this work we describe an implementation of the algorithm supporting a travel itinerary generation which can be used as a component of an automated traveler recommendation and guide construction system.

2. PERSONALIZED SERVICES FOR TRAVELERS: STATE OF THE ART

Nowadays information services that can be used to support traveler needs include route and time planning, access to real-time information (alarms, traffic information, weather alerts, etc.), payment services, journey tracking via mobile and geo-navigation tools, access to relevant online resources, and transportation planning.

Today travelers require more than simply a nicely looking algorithm based on traditional salesman problem that isn't enough to leverage existing travel experience and to arrange planning with respect to requirements of delivering personalized cultural and historical information. Many research efforts are about developing some formalization of the traveler's degree of satisfaction. For example, in [4] the authors took steps toward better personalization while introducing a formal model for travel itinerary recommendation based on collaborative filtering and recommendations of other travelers with similar travel interests.

In [3] the authors proposed a near optimal daily route itinerary generation as an advanced version of known team orienteering problem. In contrast to this work, in our approach we try to work with an open list of POIs in order to detect the next POI to visit iteratively with having in mind the possibility to response to changes in traveler plans, environment conditioned changes nearly in real time.

The authors of [5] discuss a formal model with respect to such attributes as time, cost, distance and route infrastructure (including relief, signs, administrative restrictions,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

etc.). They represent the task as a multi-criteria decision making problem reducing the subjective influence in process of route planning. We dare to say that in our work we attempt to do the inverse: to increase the subjective influence while proposing an itinerary for both a traveler and a professional (or an amateur) travel guide.

The author of [6] lays emphasis on mentioning the difference between the tasks appearing in different traveling phases (before-visit, during-the-visit and after-visit) which could be periodically repeated. The above mentioned paper follows a conceptual technology-related discourse of Brown and Chalmers [1] where the authors argued over the importance of support for sharing visit experience with other travelers as well as the importance of post-visit retrospectives: *“Post-visiting is thus a powerful way of expanding the enjoyment of a tourist visit out beyond the visit itself”*.

We also have to mention two recent projects related to the scope of this research: *TAIS*, the mobile application for guiding tourist activity described in [7] and focused on step-by-step itinerary construction in response to user actions and movements with an interesting feature of collecting user impressions about visited POIs, and *Aurigo*, an interactive tour planner for personalized itineraries [12]. To a great extent, the latter is in the same direction as the project of ours. The authors described the idea of finding a balance between automated and purely manual approaches, together with the implementation of a tour planning system combining an algorithm for generating recommended itineraries with interactive visualized interface for better itinerary personalization. The authors pointed out three important aspects:

1. Algorithmic solutions for personalized itinerary generation often use approximation and heuristics.
2. While developing such algorithms and related tools, we have to learn from human behavior and experience.
3. The focus of such solutions is on creating technology which doesn't support only kind of time/cost/etc. optimization task but allows travelers to develop their own *memorable experience* (the latter term being borrowed from [10]).

3. AUTOMATIC TOURIST ROUTE GENERATION

Let us consider the problem of constructing a tourist route for a leisure walk within the limited period of time and/or tourist area region. The common steps required in order to construct a tourist route are the following (the simplified scenario):

1. Define a time slot, select the departure and destinations points. Define other constrains if required.
2. Explore potential locally accessible POIs for the current point on the route (initially the current point is departure point).
3. Evaluate POIs by using some formalized model for taking into account the degree of POI popularity, interestingness, price to visit, location, etc.
4. Rank the POI and add the best ranked POIs to the part of the route in progress.

Steps 2–4 are repeated until the destination point is reached.

However, there are many issues to be resolved before the general schema would work. First, we have to think about time limitation for the whole walk and about more or less equal distribution of POIs along the route¹.

3.1 The Model

The standard task of tourist route generation is formalized by Souffriau [8] as follows:

Assume there are N POIs.

For every POI: $x_{ij} = 1$, if a path between the POIs marked as i and j exists, otherwise $x_{ij} = 0$.

Each POI i has a score $S_i \geq 0$, where for a departure point $i = 1$, for the destination point $i = N$.

The shortest path points i to j requires time t_{ij} , the total score S_{total} has to be maximized under constrain of time limit T_{max} .

This model set boundary of the future route and determine criteria for the best tourist route that is the best suit of the tourist objects and the best path between this objects.

3.2 Exploring Locally Accessible POIs

Due to the fact that the POI score depends on its place along the route, each potentially accessible POI's score has to be recalculated at each iteration. In order to explore POIs locally we use the geometric model shown in Figure 1.

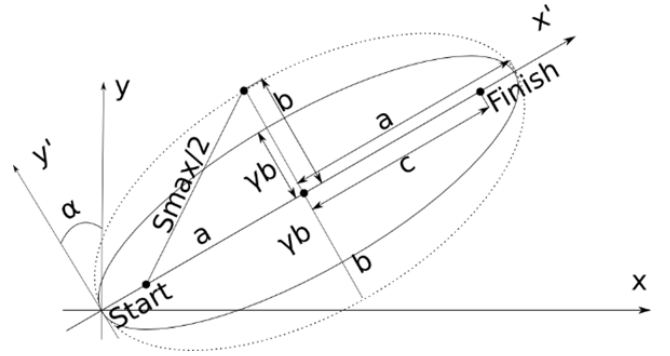


Figure 1: Exploring local area

This geometric model (Figure 1) has the following parameters: *Start*, *Finish* – the arrival and departure point within the route area, a – the semi-major axis of the search area, b – the semi-minor axis of the search area, c – half of the focal distance of the search area, $S_{max}/2$ – half of the maximum distance that can be covered for the remaining time, α – the search area angle of rotation, γ – the minimization coefficient of semi-minor axis.

Unlike to the model used in the earlier mentioned *Arigo* project, where the authors introduces a *Pop Radius* for exploring the POIs in the *circled* local area [12], we use an elliptic model for the local POI exploration area. The elliptic form of the local region, where the potential POIs to visit are explored, allows us to use ellipse particularity: the sum of the distances to the two focal points is constant for every point on the curve. It means that if this sum is equals to the distance to the most distant point a person can theoretically

¹The serious problem which is out of scope of this paper is that there could be different criteria of such an equal distribution.

reach within the given time and speed limits, each point outside the ellipse couldn't be reached. So, corresponding POIs must be excluded from the short list for further analysis. POIs located close to the border of the ellipse may also be excluded from the search, otherwise it could happen that the resulting route is too sparse: most time will be spent not to see the sights but to walk in between.

3.3 Adding POIs to the Search Scope

In current implementation, we used an extension of a gradient descent algorithm for searching POIs to be included to the route. The extension is as follows: searching a new POI to be included occurs in the area with a maximum distance between two objects in the buffer route. This modification makes possible to reduce the maximum distance between the tourist objects and to exclude from the consideration unreachable objects and therefore to improve performance. Then the selected objects are evaluated in order to find the object with the maximum score.

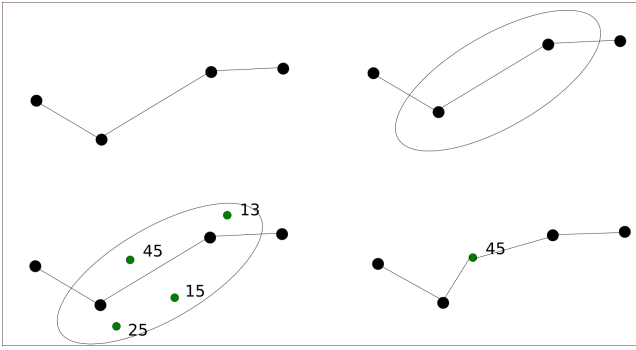


Figure 2: Iteration of adding the POI to the route

We also consider the greedy randomized adaptive search algorithm GRASP [9] as a good candidate to substitute the gradient descent stub in future implementation. One of the reasons is that GRASP uses randomized greedy heuristics in multistart-procedure and as a result generates different solutions even in equal conditions, and we believe it is rather interesting feature for a traveler route generation.

After adding a new POI to the route, we need calculate the path time.

If the path time is greater than the maximum limit, then the last added POI must be removed from the route, the process of finding POIs stops, otherwise we start new iteration.

The total time score is the sum of traveling time spent between POIs with addition of the time spent at each POI.

The last step is to minimize the traveling time between the POIs selected according to the above described procedure, and that means to solve the common task of finding a Hamiltonian path in a weighted undirected graph.

3.4 POI Evaluation

One of the most challenging problems is how to evaluate POIs and to obtain scores used in the route generation algorithm in order to have positive effect on creating interesting and personalized travel routes. Different characteristics of travel sights may compete with each other², they may de-

²For example, think of the popularity of a certain sight.

pend on the personal preferences, and also (sometimes) on the other objects.

One of the possible approach for POI attractiveness evaluation is using photos associated with tourist objects and posted and updated by users in in different social networks. If we analyze photos carefully, we can select the most interesting tourist objects and analyze variations of their popularity in “real time”. We are also able to extract other helpful information such as visit time, geo-coordinates, weather conditions, etc. [2].

After selecting photo series, we should create some area around every POI and analyze, how many photos are taken in the area, what is the period when most photos are taken, in order to select a period of minimum and maximum tourist activity.

In order to compute the score for each POI we use some type of linear convolution: each tourist object characteristic is measured at one-hundred units scale, from worst (0) to best (100). Here is the equation used to compute the score for some i -th POI:

$$\begin{cases} M_{totali} = \sum_{j=1}^{N_{metric}} k_j m_{ij} \\ m_{ij} = \frac{s_{ij}}{\max(s_{ij})} * 100; i = 1..N \\ \sum_{j=1}^{N_{metric}} k_j = 1 \\ 0 \leq m_{ij} \leq 100; i = 1..N; j = 1..N_{metric} \end{cases} \quad (1)$$

where M_{totali} being the final score for i -th POI, m_{ij} being the normalized score for j -th characteristic of the i -th POI, k_j being the weight of j -th characteristic, s_{ij} being the score of the j -th characteristic, N_{metric} being the number of characteristics used.

3.5 Software Prototype

Figure 3 illustrates the architecture of the software prototype developed in order to arrange experiments on generating tourist routes with further integration with the interactive travel planner system.

Figure 4 shows the application interface using OpenStreet-Map³ API in order to visualize the route on the map.

4. CONCLUSIONS AND FUTURE WORK

To improve the quality of generated tourist routes, we have to extend the list of criteria used to select and evaluate tourist POIs, supporting such factors as price, time windows, importance of tourist object in culture, seasoning, etc.

The latter observation leads us to think of generating a set of itinerary trees (instead of a set of itineraries) as proposed in [3].

We also have to pay attention to the whole multi-day journey planning where we have to take into account the whole set of itineraries in order to avoid undesired intersections or repeated routes, as well as to support wider range of possible constrains (including budget, weather conditions, sanitary concerns, respect to the age and disabilities, etc.)

With regard to the content related improvements, we believe that automatic tourist route generation algorithms have to be extended in order to support an extremely useful feature like constructing the thematically linked journeys.

“Popular” often means “interesting to see”. However at the same time it might mean that many other visitor come and we spent much time to wait for an entrance ticket.

³<http://www.openstreetmap.org>

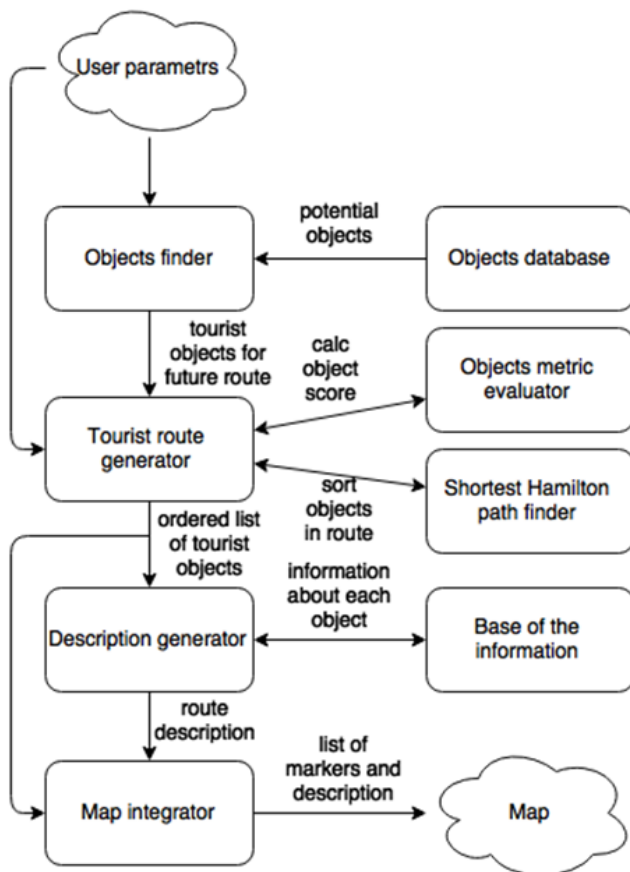


Figure 3: Architecture of the prototype.

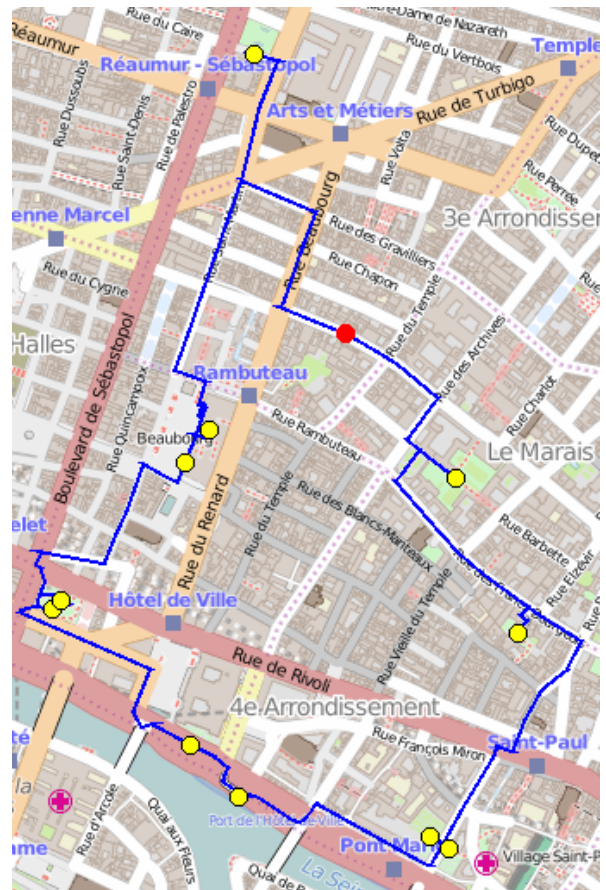


Figure 4: Example of auto-generated tourist route.

5. REFERENCES

- ## 5. REFERENCES
- [1] B. Brown and M. Chalmers. Tourism and mobile technology. In *ECSCW 2003*, pages 335–354. Springer, 2003.
 - [2] A. Kachkaev and J. Wood. Automated planning of leisure walks based on crowd-sourced photographic content. *46th Annual Universities Transport Study Group Conference, Newcastle, UK*, 2014.
 - [3] M. Kenteris, D. Gavalas, G. Pantziou, and C. Konstantopoulos. Near-optimal personalized daily itineraries for a mobile tourist guide. In *Computers and Communications (ISCC), 2010 IEEE Symposium on*, pages 862–864. IEEE, 2010.
 - [4] C. Li, J. Wang, L. Ma, and Q. Lu. Personalized travel itinerary recommendation service based on collaborative filtering and iec. In *Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference on*, pages 161–164, April 2010.
 - [5] Z. Li, J. Wang, Q. Yan, and L. Zhou. Route choice decision-marking analysis based on congestion charging. In *Multimedia Technology (ICMT), 2011 International Conference on*, pages 4073–4076, July 2011.
 - [6] T. Y. Lim. Designing the next generation of mobile tourism application based on situation awareness. In *Network of Ergonomics Societies Conference (SEANES), 2012 Southeast Asian*, pages 1–7. IEEE, 2012.
 - [7] A. Smirnov, A. Kashevnik, N. Shilov, N. Teslya, and A. Shabaev. Mobile application for guiding tourist activities: tourist assistant-tais. In *Open Innovations Association (FRUCT16), 2014 16th Conference of*, pages 95–100. IEEE, 2014.
 - [8] W. Souffriau. *Automated Tourist Decision Support*. Katholieke Universiteit Leuven, 2010.
 - [9] W. Souffriau, P. Vansteenwegen, G. V. Berghe, and D. Oudheusden. A greedy randomised adaptive search procedure for the team orienteering problem. In *proceedings of EU/MEeting*, 2008.
 - [10] V. W. S. Tung and J. B. Ritchie. Exploring the essence of memorable tourism experiences. *Annals of Tourism Research*, 38(4):1367–1386, 2011.
 - [11] P. Vansteenwegen. Planning in tourism and public transportation. *4OR*, 7(3):293–296, 2009.
 - [12] A. Yahi, A. Chassang, L. Raynaud, H. Duthil, and D. H. P. Chau. Aurigo: an interactive tour planner for personalized itineraries. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*, pages 275–285. ACM, 2015.
 - [13] Y. Zheng, L. Capra, O. Wolfson, and H. Yang. Urban computing: concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(3):38, 2014.

Human Response Delay in Balancing Virtual Pendulums with Overdamped Dynamics

Takashi Suzuki,¹ Shigeru Kanemoto,² Ihor Lubashevsky³

University of Aizu

Ikki-machi, Aizu-Wakamatsu, Fukushima 965-8560, Japan

¹m5181141@u-aizu.ac.jp, ²kanemoto@u-aizu.ac.jp, ³i-lubash@u-aizu.ac.jp

ABSTRACT

The results of our experiments aimed at estimating the distribution of the response delay time in human intermittent control over unstable mechanical systems are presented. A novel experimental paradigm: balancing an overdamped inverted pendulum was used. The created simulator of balancing a virtual pendulum makes the pendulum invisible when the angle between it and the upward position is less than 5° , which enabled us to measure directly the delay time. It is demonstrated, in particular, that (i) the response delay time may be treated as a random variable distributed within a wide interval and (ii) the response delay in human intermittent control may be determined by cumulative actions of two distinct mechanisms, automatic and international, endowing it with complex nonlinear properties.

Categories and Subject Descriptors

J.4 [Social and Behavioral Sciences]: psychology; H.1.2 [User/Machine Systems]: human factors

General Terms

Experiment

Keywords

Human intermittent control, response delay, pendulum balancing

1. INTRODUCTION

According to the modern paradigm, human control over unstable mechanical systems can be categorized as intermittent and this type of control being rather efficient on its own is a natural consequence of human physiology (see, e.g., [1]). With respect to human behavior the notion of intermittency implies that actions of a human operator in controlling a mechanical system form a sequence of alternate fragments of passive and active phases, with transitions between them

being event-driven. The latter means that the control is activated or halted when the discrepancy between the goal and the actual system state exceeds a certain threshold or falls behind it, respectively. Recently, we developed a concept of noise-driven control activation as a more advanced alternative to the conventional threshold-driven activation [2]. In this concept the transition from passive to active phases is probabilistic and reflects human perception and fuzzy evaluation of the current system state before making decision concerning the necessity of correcting the system dynamics. Broadly speaking, during the passive phase the operator accumulates the information about the system state, naturally, this process is not instantaneous but requires some time in addition to the physiological delay in human response. The cumulative effect of the two mechanisms, the accumulation of information about the system state and the physiological delay, can be described by some delay time τ_d . The found stochasticity of human intermittent control in experiments on balancing virtual stick [2] prompts us to expect that this delay time is not a fixed value but a random variable with a relatively wide distribution; this feature is crucial for modeling human intermittent control. To elucidate the probabilistic properties of this cumulative delay in human reactions in controlling unstable mechanical system we have conducted some experiments and the purpose of present work is to report the obtained results.

However, before discussing our results, some comments concerning the research of human response delay are worthwhile. Investigations of response delay in human reactions to various stimuli including visual ones has a relative long history (see, e.g., [3]). Usually in experiments on human visual perception the values of delay time $\tau_d \gtrsim 100$ ms are detected and they are unimodally distributed within a wide interval; the gamma or Weibull fits are often used to characterize the found results (see, e.g., [4]). During the last decade there has been accumulated some evidence that mental processes contribute substantially to the response delay and such factors as memory load and required attention are essential in this case (see, e.g., [5, 6] and references therein). Taking into account this fact and the possible existence of two mental systems of information processing (see a review [7]) we may expect the response delay time distribution to exhibit complex behavior especially in cases when it is related to decision-making in multifactorial processes like human intermittent control.

2. EXPERIMENTAL SETUP

As previously [2], in our investigations, we used a novel

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

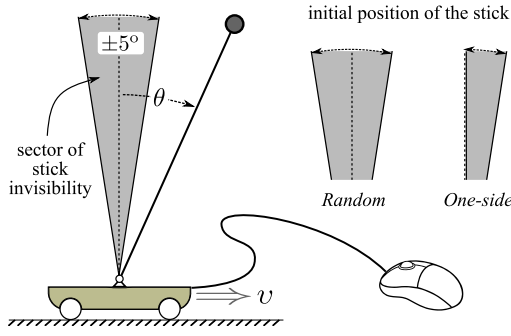


Figure 1: The used simulator of balancing a one-degree-of-freedom overdamped inverted pendulum (stick) with a region of invisibility.

experimental paradigm: balancing an overdamped inverted stick (Fig. 1) whose dynamics is governed by the equation

$$\tau \frac{d\theta}{dt} = \sin \theta - \frac{\tau}{l} \vartheta \cos \theta,$$

where the cart velocity ϑ is determined by operator actions, τ is the timescale of the stick motion, l is the stick length. The overdamping eliminates the effects of inertia, and therefore reduces the dimensionality of the system. The new feature of this balancing simulator is that the stick becomes invisible within the sector $\pm 5^\circ$ (Fig. 1).

Twelve subjects were instructed to maintain the upright position of the stick on the computer screen by moving the platform horizontally via the computer mouse. When the stick leaves this sector and becomes visible, a subject tries to restore its upright position. If, nevertheless, the stick falls it is returned immediately by computer into the invisibility sector and the balancing experiment is resumed. Two types of experimental set-up were studied, in the first case after falling, the stick is placed randomly inside the invisibility sector such that from which side the stick will appear is not predictable. In the second case the stick appears from one side only. The time span τ_d between the moments when the stick leaves the invisibility sector becoming visible and when a subjects starts cart (mouse) motion was fixed and regarded as the current value of the response delay time.

3. RESULTS AND DISCUSSION

The statistical data collected in this way are illustrated in Fig. 2. Their analysis enables us to draw the following conclusion.

- The response delay time characterizing the transition from passive to active phases in human intermittent control is a random variables changing in a wide interval from the lower value less than 100 ms up to upper value about 600 ms. It should be noted that the lower and upper values are typical delay times of human response to predictable and unpredictable events, respectively [1].
- The found distributions of response delay time for many subjects contain at least two different fragments represented individually in Fig. 2 by *Random* and *One-side* data, which argues for the hypothesis about the exis-

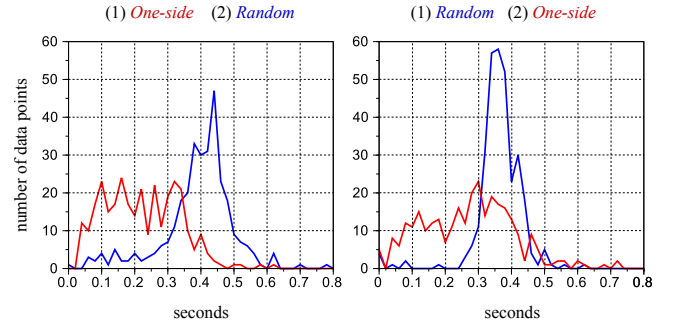


Figure 2: Two histograms of response delay time obtained based on the conducted experiments for different subjects within the *Random* and *One-side* stick appearance setup.

tence of two mental systems affecting together human decision-making.

- The individual difference in the response delay time distribution found for several subjects in experiments with predictable and unpredictable stick appearance demonstrates the substantial role of mental processes in evaluating the current state of system dynamics by human operators.

The found results are essential for constructing mathematical models of human intermittent control that operate directly with the human delay.

4. REFERENCES

- [1] Ian D. Loram, Henrik Gollee, Martin Lakie, and Peter J. Gawthrop. Human control of an inverted pendulum: is continuous control necessary? Is intermittent control effective? Is intermittent control physiological? *The Journal of Physiology*, 589(2):307–324, 2011.
- [2] Arkady Zgonnikov, Ihor Lubashevsky, Shigeru Kanemoto, Toru Miyazawa, and Takashi Suzuki. To react or not to react? Intrinsic stochasticity of human control in virtual stick balancing. *Journal of The Royal Society Interface*, 11:20140636, 2014.
- [3] Trisha Van Zandt. Analysis of Response Time Distributions. In Hal Pashle and John Wixted, editors, *Stevens' Handbook of Experimental Psychology, Volume 4, Methodology in Experimental Psychology*, pages 461–516. John Wiley & Sons, New York, 3rd ed., 2002.
- [4] Evan M. Palmer, Todd S. Horowitz, Antonio Torralba, and Jeremy M. Wolfe. What are the shapes of response time distributions in visual search? *Journal of Experimental Psychology: Human Perception and Performance*, 37(1):58–71, 2011.
- [5] Edward F Ester, Tiffany C Ho, Scott D Brown, and John T Serences. Variability in visual working memory ability limits the efficiency of perceptual decision making. *Journal of vision*, 14(4):2(1–12), 2014.
- [6] Benjamin Pearson, Julius Raškevičius, Paul M Bays, Yoni Pertzov, and Masud Husain. Working memory retrieval as a decision process. *Journal of Vision*, 14(2):2(1–15), 2014.
- [7] Jonathan St BT Evans. Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59:255–278, 2008.

Effect of Extended Visual Feedback on Human Information Processing in Virtual Stick Balancing

Takashi Suzuki
University of Aizu
Tsuruga, Ikki-machi
Aizu-Wakamatsu, Fukushima,
Japan
m5181141@u-aizu.ac.jp

Arkady Zgonnikov
University of Aizu
Tsuruga, Ikki-machi
Aizu-Wakamatsu, Fukushima,
Japan
arkady@u-aizu.ac.jp

Shigeru Kanemoto
University of Aizu
Tsuruga, Ikki-machi
Aizu-Wakamatsu, Fukushima,
Japan
kanemoto@u-aizu.ac.jp

Ihor Lubashevsky
University of Aizu
Tsuruga, Ikki-machi
Aizu-Wakamatsu, Fukushima,
Japan
i-lubash@u-aizu.ac.jp

ABSTRACT

Maintaining vertical position of an inverted pendulum is a simple balancing task, which is widely used to study human control behavior. Yet, much about this behavior remains poorly understood even in the context of simple virtual tasks. The purpose of this study was to investigate whether the control behavior of human operators depends on the type of visual feedback from the controlled system. We analyze the experimental data on human stick balancing on a computer screen. The previous studies reported detailed analysis of the task performance of human operators observing only the angular deviation of the stick from the vertical. In this study we augmented the information supplied to the operator by linear displacement of the upper tip of the stick from the reference point. This additional information was suggested to improve the performance of the operators. Surprisingly, the subjects not only exhibited better performance, but also supposedly employed structurally different control mechanisms in the linear displacement condition. The found results may have potential implications both for fundamental research aimed at investigating the basic properties of human control, and applied research on human factors.

Categories and Subject Descriptors

H.1.2 [User/Machine Systems]: Human information processing

Keywords

Stick balancing, Motor control, Intermittent control, Human

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

operator, Human-machine interaction

1. INTRODUCTION

Understanding how humans control unstable systems is a key issue in numerous applications, ranging from quiet standing to aircraft landing [7]. The task of stick balancing (Fig. 1) has been extensively used as an experimental paradigm for studying behavior of human operators controlling unstable systems [6, 5]. In different variations of this task studied in literature, human operators are provided with different information about the balanced stick. When balancing a stick on the fingertip, humans employ both visual and proprioceptive signals to control the stick (see e.g. review in Ref. [5]). In the paradigm of virtual stick balancing on a moving platform, typically only visual feedback on the system state is available for the operator [1, 6]. Although both the real and virtual balancing approaches often converge on the similar findings about the overall control properties (see e.g. [1]), little attention has been paid to the differences in the operators' behavior under different conditions of balancing task. Such differences, if any, may have profound impact on the methodology of experimental studies on human balancing behavior. Besides theoretical value, elucidating the effect of visual feedback on human operator performance can have potential implications for applied studies on human factors. For instance, it is long known that proper setup of visual displays can enhance the performance of human operators [7], which is particularly important in life-critical tasks such as airplane landing.

The purpose of this study was to investigate whether the control behavior of human operators depends on the type of visual feedback provided by the controlled system, using as an example the task of virtual balancing of overdamped stick. This task has been previously confirmed to provide a useful insight into properties of human control in diverse processes, e.g. quiet standing and car driving [5, 8].

Particularly, we analyze the data on human balancing of virtual stick obtained under two different conditions. In the first condition (which had been analyzed previously in details [8]), the subjects controlled the stick guided only by its

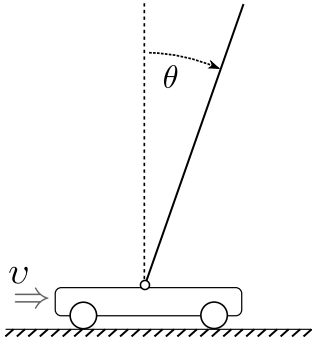


Figure 1: Inverted pendulum (stick) on a moving cart.

angular deviation from the vertical. The newly conducted experiments focused on the second condition, where the subjects not only observed the stick angle, but were also able to monitor the *linear displacement* of the upper tip of the stick with respect to a dynamic reference point. The two chosen conditions reflect the two classes of experimental paradigms employed in modern literature on human balance control. One of these classes comprises experimental setups where the subject only observes the angle of the balanced object (e.g. [2, 5]), and the other one focuses on the subjects matching the actual position of the upper tip of the virtual stick and the reference position of the upper tip when the stick is in vertical position (e.g. [1, 4]).

The experiments revealed that availability of the linear displacement in addition to the angular deviation information substantially improved the balancing performance of the subjects. The most surprising result, however, was that the additional visual feedback not only reduced the amplitude of the stick fluctuations under human control, but also resulted in the structurally different statistical distributions characterizing the control process. This suggests that human operators employ different control mechanisms depending on the particular type of information about the controlled system is available (linear or angular). Besides theoretical value, our findings may have potential implications for design of visual displays for human operators.

2. EXPERIMENTAL SETUP

This study analyzes human behavior during the task of balancing virtual overdamped¹ stick on the computer screen (Fig. 1). The goal of the task is to keep the stick upwards, whereas the cart position is controlled directly by a computer mouse, so the cart can be moved arbitrarily within the limits of the computer screen.

The physics of the overdamped stick is governed by the differential equation

$$\tau \dot{\theta} = \sin \theta - \frac{\tau}{l} v \cos \theta, \quad (1)$$

where θ is the stick angle, the cart velocity v is the control variable, $\tau > 0$ is the constant time scale of the stick motion, and $l > 0$ is the length of the stick.

¹The overdamping condition is essential, because it simplifies the analysis of the human behavior dynamics (see Ref. [8] for discussion)

In what follows we investigate two datasets. The first one has been obtained previously [8], and includes the data for ten subjects (six male, four female, aged 20 to 61), who performed the balancing task while observing only the cart and the stick on the computer screen. Each subject balanced the stick during three five-minute sessions, separated by two three-minute breaks. The kinetic parameters of the stick were set to $\tau = 0.3$, $l = 0.4$. The detailed description of the obtained data can be found in Ref. [8].

The second dataset has been obtained in the newly conducted experiment, which significantly differed from the previous one only in the type of visual information provided to subjects. In this new experiment, the subjects were shown not only the coupled cart-stick system, but also the mouse cursor on the computer screen. The subjects were specifically instructed to balance the stick by keeping the mouse cursor near the upper tip of the stick. We hypothesized that this would greatly simplify the task for the subjects, because even the smallest deviations of the stick from the vertical position can be easily detected and corrected in this case.² The eight volunteers (six male, two female) participated in the new experiment were aged from 23 to 63 years old, and four of them took part in the experiment reported in Ref. [8].

3. RESULTS AND DISCUSSION

The question we aimed to answer in this study is: Besides the expected quantitative difference, are there any fundamental, *qualitative* difference between the two conditions? For this purpose, we analyze the statistical distributions of the stick angle and cart velocity data (Figs. 2,3).

The distributions of the stick angle under two conditions were found to be markedly different (Fig. 2, left frame). Based on this fact, and given that the physics of the controlled system is the same in the two conditions, and only the type of the visual feedback differs, we conclude that human operators employ different control mechanisms when using the angular and linear visual information. At the same time, the distributions of the cart velocity in the “With cursor” and “No cursor” conditions exhibit the same form, which in turn indicates that the corrective movements executed by the operators have the same generating mechanisms.

Taking into account these considerations, and appealing to the hypothesis of discontinuous, or intermittent control [3, 5], we hypothesize that the key difference between the operators’ behavior in the two analyzed conditions lies in the domain of the *control activation* mechanism. To investigate this assumption further, we analyze the *action point* (AP) distributions (Fig. 3). In the context of the present work, an action point is a value of stick angle triggering the corrective action of the operator. Each AP corresponds to an instant when the operator begins to move the cart in response to the deviation of the stick from the vertical position. Thus, the AP distributions highlight the stick angle values which are likely to trigger reaction of human operator in each condition.

We found a pronounced difference in the AP distributions under two conditions, which indicates that different control

²By design of the simulator software, the horizontal position of the moving cart on the screen was the same as the mouse cursor position. Hence, if the stick deviated to the right of the cursor, simply adjusting the cursor position to the upper end of the stick would eliminate the deviation.

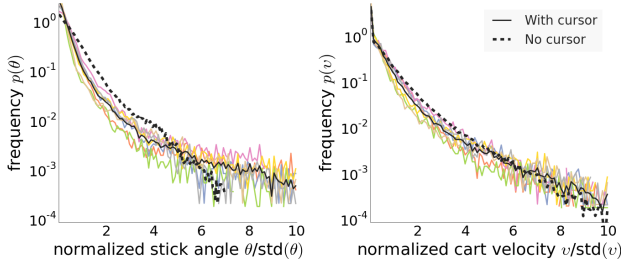


Figure 2: Distributions of stick angle and cart velocity exhibited by human operators in virtual stick balancing under the two tested conditions. Colored solid lines represent the individual distributions for each subject for the “With cursor” condition, whereas solid and dashed black lines show the average distributions for all the subjects in the “With cursor” and “No cursor” conditions, respectively.

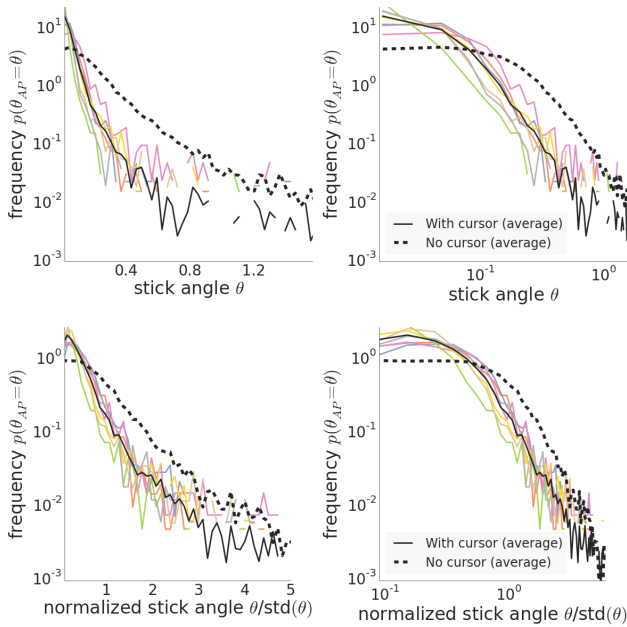


Figure 3: Distributions of action points (AP) exhibited by human operators in virtual stick balancing. The colored lines represent the individual distributions in the “With cursor” condition for each subject, whereas solid and dashed black lines show the average distributions for all the subjects in the “With cursor” and “No cursor” conditions. The top two frames illustrate the *scales* of the distributions in the two conditions. The bottom two frames (the stick angle values are normalized with respect to $\text{std}(\theta)$) demonstrate the structural difference between the two distributions.

activation mechanisms are employed by the operators provided with different types of visual feedback. If the information about linear deviation is available (“With cursor” condition), the AP distribution peaks at a small angle, and decreases much faster than in the case when only the angular deviation is supplied to the operator. Moreover, the log-log plot of the AP distributions (Fig. 3, bottom right plot) demonstrates that the tail part of the “With cursor” distribution may follow a power law (a straight line in the log-log scales), as opposed to the “No cursor” one, which can be approximated by the Laplace distribution [8].

Although our results highlight the difference in the control activation mechanisms under two tested conditions, this work leaves unanswered the question as to what these mechanisms are. At this point, we can only claim that humans’ visual perception of information on angular deviations differs from that of linear displacements, and that it results in different pattern of control activation (at least in the task of virtual stick balancing). However, to be able to speculate about the particular mechanisms, one should confront the data presented here to the various models of control activation, including threshold-based [3, 5] and noise-driven [8] models.

4. CONCLUSION

The present study investigates actions of human operators in a simple virtual balancing task, namely, maintaining vertical position of an overdamped inverted pendulum. We focus on the effect of visual feedback type on the performance of the operators, providing the subjects with two types of visual feedback — angular deviation and linear displacement. We analyze the previously collected and newly obtained experimental data. The analysis reveals that the statistical properties of the stick under human control under two considered conditions differ not only in scale, but also qualitatively, as highlighted by the distribution of the stick angle and action points (Figs. 2,3). This suggests that one must be aware of the potential differences in human control mechanisms even in the case when the difference between the experimental conditions is seemingly minor.

We conclude that while the control execution mechanisms supposedly do not change depending on the type of visual feedback, the *control activation* exhibited by the subjects are substantially different. One may even assume that humans’ processing of angles is governed by different cognitive mechanisms than that of linear displacements. However, the latter assumption is to be reinforced by appropriate modeling in the future studies. In any case, the very fact that control activation in human operators’ essentially depends on the type of visual feedback may have potential implications both for fundamental research aimed at investigating the basic properties of human control, and applied research on human factors.

5. REFERENCES

- [1] R. Bormann, J.-L. Cabrera, J. G. Milton, and C. W. Eurich. Visuomotor tracking on a computer screen—an experimental paradigm to study the dynamics of motor control. *Neurocomputing*, 58:517–523, 2004.
- [2] P. Foo, J. Kelso, and G. C. de Guzman. Functional stabilization of unstable fixed points: Human pole balancing using time-to-balance information. *Journal of*

- Experimental Psychology: Human Perception and Performance*, 26(4):1281, 2000.
- [3] P. Gawthrop, I. Loram, M. Lakie, and H. Gollee. Intermittent control: a computational theory of human control. *Biological cybernetics*, 104(1-2):31–51, 2011.
 - [4] I. D. Loram, M. Lakie, and P. J. Gawthrop. Visual control of stable and unstable loads: what is the feedback delay and extent of linear time-invariant control? *The Journal of Physiology*, 587(6):1343–1365, 2009.
 - [5] J. Milton. Intermittent Motor Control: The 'drift-and-act' Hypothesis. In *Progress in Motor Control*, pages 169–193. Springer, 2013.
 - [6] S. Suzuki, F. Harashima, and K. Furuta. Human Control Law and Brain Activity of Voluntary Motion by Utilizing a Balancing Task with an Inverted Pendulum. *Advances in Human-Computer Interaction*, 2010, 2010.
 - [7] C. D. Wickens and J. G. Hollands. *Engineering psychology and human performance*. Prentice Hall New Jersey, 1999.
 - [8] A. Zgonnikov, I. Lubashevsky, S. Kanemoto, T. Miyazawa, and T. Suzuki. To react or not to react? Intrinsic stochasticity of human control in virtual stick balancing. *Journal of the Royal Society Interface*, 11(99):20140636, 2014.

Is the Modern Theory of Stochastic Processes Complete? Example of Markovian Random Walks with Constant Non-Symmetric Diffusion Coefficients

Kosuke Hijikata,¹ Ihor Lubashevsky,² Alexander Vazhenin³

University of Aizu

Ikki-machi, Aizu-Wakamatsu, Fukushima 965-8560, Japan

¹)m5191113@u-aizu.ac.jp, ²)i-lubash@u-aizu.ac.jp, ³)vazhenin@u-aizu.ac.jp

ABSTRACT

A new type non-symmetric diffusion problem is considered and the corresponding Brownian motion implementing such diffusion processes is constructed. As a particular example, *random walks with internal causality* on a square lattice are studied in detail. By construction, one elementary step of a random walker on the lattice may consist of its two succeeding jumps to the nearest neighboring nodes along the x - and then y -axis or the y - and then x -axis ordered, e.g., clockwise. It is essential that the second fragment of elementary step is caused by the first one, meaning that the second fragment can arise only if the first one has been implemented, but not vice versa. In particular, if for some reasons the second fragment is blocked, the first one may be not affected, whereas if the first fragment is blocked, the second one cannot be implemented in any case. As demonstrated, on time scales much larger than the duration of one elementary step these random walks are characterized by a diffusion matrix with non-zero anti-symmetric component. The existence of this anti-symmetric component is also justified by numerical simulation.

Categories and Subject Descriptors

G.3 [Probability and Statistics]: stochastic processes

General Terms

Theory

Keywords

Stochastic process, diffusion matrix, boundary conditions

1. INTRODUCTION

The present paper poses a fundamental question about the completeness of the modern formalism of describing stochastic processes and, by way of example, the formalism of the

Fokker-Planck equations, or speaking more strictly, the forward Fokker-Planck equations is analyzed.

The Fokker-Planck equation (see, e.g., [1])

$$\partial_t G = \sum_{i=1}^N \partial_i \left\{ \sum_{j=1}^N \partial_j [D_{ij}(\mathbf{x}, t) G] - V_i(\mathbf{x}, t) G \right\} \quad (1)$$

subject to the initial condition

$$G(\mathbf{x}, t | \mathbf{x}_0, t_0) \big|_{t=t_0} = \delta(\mathbf{x} - \mathbf{x}_0), \quad (2)$$

where $\mathbf{x} = \{x_i\}_{i=1}^N \in \mathbb{Q} \subset \mathbb{R}^N$ and $t > t_0$, describes a wide class of Markovian random walks continuous in space and time for which the first and second moments of walker displacement are some finite space-continuous quantities. The matrix $\mathbf{D} = \|D_{ij}\|$ of diffusion coefficients and the velocity drift $\mathbf{V} = \{V_i\}$ in the “phase” space \mathbb{R}^N are introduced as

$$D_{ij}(\mathbf{x}, t) = \lim_{\tau \rightarrow 0} \frac{1}{2\tau} \langle (x'_i - x_i)(x'_j - x_j) \rangle_{x':(t+\tau|\mathbf{x}, t)}, \quad (3)$$

$$V_i(\mathbf{x}, t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle (x'_i - x_i) \rangle_{x':(t+\tau|\mathbf{x}, t)}. \quad (4)$$

Due to the form of the Fokker-Planck equation the diffusion coefficient matrix $\|D_{ij}\|$ must be symmetric, $D_{ij} = D_{ji}$, which follows from definition (3) as well.

Discrete random walks on lattices also admit this description on scales $t \gg \tau$, where τ is the characteristic time of the walker hopping to the neighboring lattice nodes. An example of symmetric (i.e. without regular drift, $\mathbf{V} = 0$) random walks on a square lattice is illustrated in Fig. 1: “diagram of transitions.” Within one elementary time step τ a walker hops to one of the nearest lattice nodes with the probability $p = \frac{1}{4}(1 - \epsilon)$ or to one of the next shell of nearest neighbors with the probability $q = \frac{1}{4}\epsilon$, here $0 < \epsilon < 1$ is a given parameter. For these random walks the diffusion matrix is of the diagonal form and can be characterized by one diffusion coefficient $D = (1 + \epsilon)a^2/(4\tau)$, i.e., $D_{xx} = D_{yy} = D$ and $D_{xy} = D_{yx} = 0$.

Appealing to the form of the Fokker-Planck equation (1) usually one draws a conclusion that the diffusion flux $\mathbf{J} = \{J_i\}$ is related to the distribution function G via the expression

$$J_i = - \sum_{j=1}^N \partial_j [D_{ij}(\mathbf{x}, t) G] + V_i(\mathbf{x}, t) G. \quad (5)$$

Then ascribing various physical properties to the medium boundary $\partial\mathbb{Q}$ the Fokker-Planck equation is subjected to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

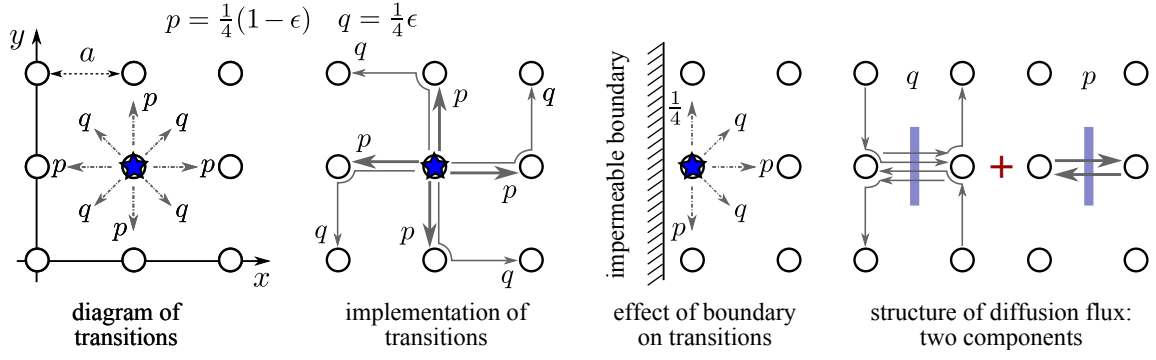


Figure 1: The analyzed random walks on the square lattice, from left to right, the diagram showing possible transitions of the walker within one elementary step and their probabilistic weights, spatial structure of these transitions, diagram and probabilistic weights of the walker near a impermeable boundary, the diagram illustrating the relationship between the diffusion flux and possible walker transitions.

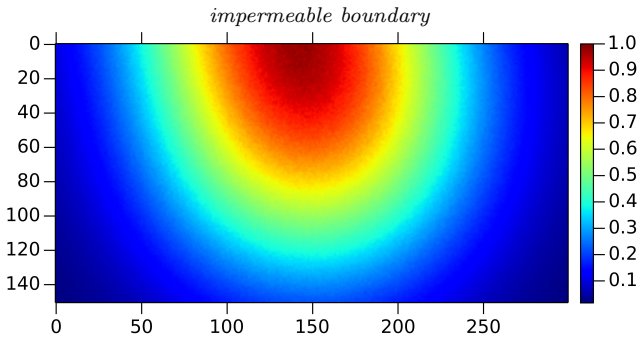


Figure 2: Distribution function normalized to its maximum. In numerical simulation the asymmetry parameter $\epsilon = 0.2$, the trajectory origin $\{x_0 = 150, y_0 = 50\}$, and the number of steps in one trajectory $N = 3y_0^2$ were used.

the corresponding boundary conditions (see, e.g., [1]). The latter completes the description of such stochastic processes in the framework of the Fokker-Planck equation. If it were so for the random walks illustrated in Fig. 1, at a impermeable boundary (third fragment counting from left in Fig. 1) the distribution function G would meet the boundary condition $\partial_x G = 0$ in the continuous approximation.

As far as the relationship between the Fokker-Planck equation (1) and the diffusion flux (5) is concerned, we note that the replacement $D_{ij} \Rightarrow D_{ij} + D_{ij}^a$, where D_{ij}^a is an asymmetric component, $D_{ij}^a = -D_{ji}^a$, does not change the form of the Fokker-Planck equation (1) but modifies substantially expression (5). The latter, in turn, changes the boundary condition, so, finally, contributes essentially to the description of stochastic process. Therefore the statement on the diffusion coefficient symmetry does not follow from the derivation of the Fokker-Planck equation but is an *addition assumption* that can be accepted for some physical reasons.

2. MODEL AND DISCUSSION

The example of random walks shown in Fig. 1 illustrates the stated above proposition, in particular, the diffusion coefficient matrix D_{ij} entering relationship (5) is of the form

containing antisymmetric component, namely,

$$D_{xx} = D_{yy} = \frac{(1 + \epsilon)a^2}{4\tau}, \quad D_{xy} = -D_{yx} = -\frac{\epsilon a^2}{2\tau}. \quad (6)$$

This fact must be reflected in the boundary conditions and, finally, cause the asymmetry of the distribution function for the diffusion problem in the region with the impermeable boundary with respect to the boundary point nearest to the origin of random walks. Numerical simulation justifies this statement (Fig. 2).

Concluding the obtained results, we pose a question about the completeness of describing stochastic processes in terms of the Fokker-Planck equation or stochastic differential equations. Indeed, this formalism ignores the *internal* structure of elementary steps, whereas the given example demonstrates the fact that particular spatial details of the walker motion within one elementary step can affect the macroscopic behavior of diffusion processes. Diffusion in magnetic field is also discussed as a characteristic example of physical systems where such phenomena can be pronounced. Besides it should be noted that the considered problem of non-symmetric diffusion coefficient matrix is partly related to the problem called non-symmetric diffusion dealing with a diffusion type stochastic processes governed by equations like

$$\partial_t G = \sum_{i,j=1}^N \partial_i [D_{ij}(\mathbf{x}) \partial_j G],$$

where the diffusion coefficient $D_{ij}(\mathbf{x})$ depends on the spatial coordinates \mathbf{x} and, so, its possible non-symmetry can be responsible for macroscopic effects (see, e.g., [2] and references therein).

3. REFERENCES

- [1] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer Verlag, 2009), 4th ed.
- [2] J.-D. Deuschel, T. Kumagai, Markov Chain Approximations to Nonsymmetric Diffusions with Bounded Coefficients, *Communications on Pure and Applied Mathematics*, 66:821–866, 2013.

Asymptotics of Psychometric Function of Gray Color Categorization

Marie Watanabe,¹ Ren Namae,² Ihor Lubashevsky³

University of Aizu

Ikki-machi, Aizu-Wakamatsu, Fukushima 965-8560, Japan

¹s1190150@gmail.com, ²s1210140@u-aizu.ac.jp, ³i-lubash@u-aizu.ac.jp

ABSTRACT

The results of conducted experiments on categorical perception of different shades of gray are reported. A special color generator was created for conducting the experiments on categorizing a random sequence of colors into two classes, light-gray and dark-gray. The collected data are analyzed based on constructing (i) the asymptotics of the corresponding psychometric functions and (ii) the mean decision time in categorizing a given shade of gray depending on the shade brightness (shade number). Conclusions about plausible mechanisms governing categorical perception, at least for the analyzed system, are drawn.

Categories and Subject Descriptors

J.4 [Social and Behavioral Sciences]: psychology; H.1.2 [User/Machine Systems]: human factors

General Terms

Experiment

Keywords

Categorical perception, psychometric function, asymptotics

1. INTRODUCTION

Categorical perception is a general term describing situations when we perceive our world in terms of the categories formed previously during our communication with the social environment. Our perceptions are warped such that differences between objects that belong to different categories are accentuated, and differences between objects that fall into the same category are deemphasized [1]. The reported research concerns categorization of colors. In spite of a vast amount of literature about various aspects of color categorization (see, e.g., [2]) even classification of possible mechanisms governing these processes is not understood well.

In previous pilot experiments on shape recognition near perception threshold [3] we found some arguments for the

hypothesis that this process is governed by a certain potential mechanism. The given conclusion is based on the analysis of the *asymptotics* of the psychometric functions; it seems to be a novel way to discriminating directly plausible mechanisms governing human recognition near perception thresholds. The purpose of the reported experiments is to verify this hypothesis for categorical perception being a more complex cognitive process involving mental processes in addition to pure physiological ones. By way of example, categorization of different shades of gray color was selected for investigation.

2. EXPERIMENTAL SETUP

Colors for categorization were generated as follows. A computer visualizes a window of size of 17×16 cm with a square \mathbb{S} of size of 11×11 cm placed at its center. Color inside this square is changed during experiments; the remaining window part is filled with a neutral gray. Each trial of shade categorization is implemented as follows. A random integer $I \in [0, 255]$ is generated and the area \mathbb{S} is filled with the gray color $G(I) := \text{RGB}(I, I, I)$. Then a subject has to classify the visualized gray color $G(I)$ according to his/her perception into two possible categories, “light gray” (LG) and “dark gray” (DG). A made choice is recorded via pressing one of two joystick buttons. Then a mosaic pattern of various shades of gray is visualized for 500 ms. This mosaic pattern is used to depress a possible interference between color perception in successive trials that can be caused by human iconic memory. After that a new number I is generated and the next trial starts. Four subjects, two female and two male students of age 21–22, were involved in these experiments. The experiments spanned 5 successive days, for each subject the total number of data records was 10 000. One day set comprised four blocks of 15 min experiments separated by 3 min rest.

3. RESULT AND DISCUSSION

The collected data have been used in constructing three functions. The first one is the psychometric function for the light-gray category, i.e., the probability $P_w(I)$ of choosing the light-gray category for the shade of gray, $G(I)$, specified by a given integer I . The second one, $P_b(I)$, is actually the same function for the dark-gray category. The third function, $T(I)$, is the mean duration time of making decision during one trial of categorization for a given shade $G(I)$ of gray. The results are illustrated in Fig. 1 for one of the four subjects.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWA/IT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

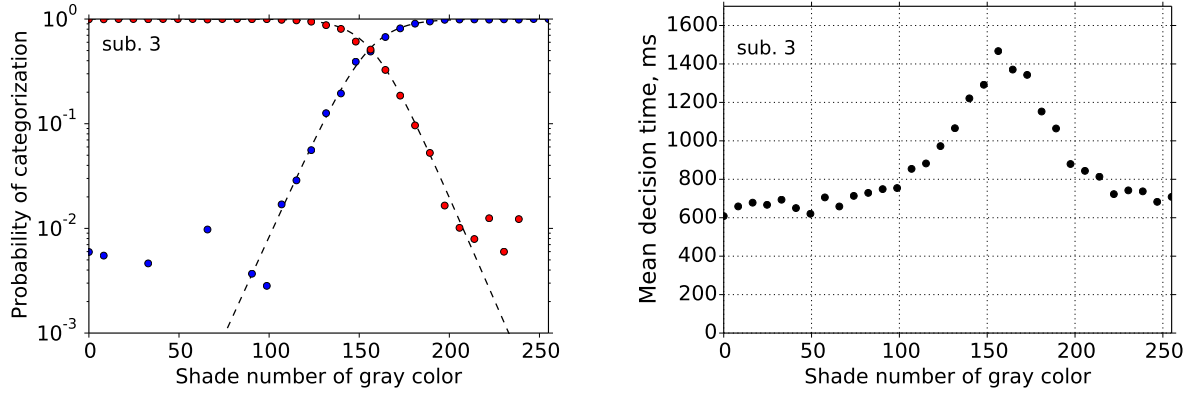


Figure 1: *Left frame:* the psychometric functions of categorization of shades of gray, red and blue points correspond to the light-gray and dark-gray categories, respectively, dashed lines represent fitting function (1). *Right frame:* the mean delay time in decision-making vs. the shade number of gray.

As seen in Fig. 1 (left frame), the asymptotics of the constructed functions $P_{w,g}(I)$ can be approximated by

$$P_{w,b}(I) \approx \frac{1}{2} \left\{ 1 + \tanh \left[\pm \frac{(I - I_{m:w,b})}{\delta I_{w,b}} \right] \right\}, \quad (1)$$

where $I_{m:w,b}$ is the center point of the crossover region and the value $\delta I_{w,b}$ characterizes its thickness. As seen in Fig. 1, the asymptotics of both the psychometric functions looks like exponential function of the shade number,

$$P_{w,b} \propto \exp(\pm I / \delta I_{w,p}),$$

what is the typical dependence for the model of random systems residing in a bimodal potential well near the equiprobable distribution. In this case the difference δU in the well minima leads to the asymmetry in the well populations quantified by their ratio proportional to $\tanh(\delta U/T)$ (here T is some constant). The found results can be regarded as some argumentation for the hypothesis that categorical perception, at least, in the analyzed case is described by a certain potential model, where the corresponding decision-making is regarded as some random process η in a potential relief $U(\eta|I)$ with two minima $U_w(I)$ and $U_b(I)$ determined by the shade number I as a control parameter.

The experimental data shown in Fig. 1 (left frame) contain two domains on both the sides of the crossover region that may be classified as domains of scattered data that are caused by some mechanism rather than lack of statistics in the collected data. In order to clarify a plausible mechanism that could be responsible for this anomalous behavior of these heavy tails of the psychometric functions, we analyzed the dependence of the mean time required for subjects to make decision about classifying a current shade of gray depending on the shade number I . The results are exemplified in Fig. 1 (right frame). For all the subjects these distributions exhibit remarkable peak located inside the crossover region. Outside this peak the values of the mean decision time T are located in the interval 500–700 ms, which can be regarded as the upper boundary of the human reaction delay time controlled by physiological processes of recognizing threshold events within their unpredictable appearance. The mean decision time in these peaks is about 1.5 s. To explain these values we pose a hypothesis that mental processes on their own rather than pure physiological mechanisms

contribute substantially to the decision making in color categorization under pronounced uncertainty, at least, in the analyzed case. As far as the scattered data domains in Fig. 1 are concerned, we explain their appearance by addressing this phenomenon to a relatively short time of decision when the right choice is rather evident, which can increase the probability of subjects pressing a wrong button accidentally.

4. CONCLUSION

Categorical perception, at least, of shades of gray, is governed by a potential mechanism of decision-making treated as a random process in a potential relief. The previous analysis of shape recognition [3] also revealed the same type asymptotics of psychometric functions. Noting that the cognitive perception and shape recognition should be governed by different mechanisms, the found universality allows us to pose an assumption that human decision-making under uncertainty is implemented via a common emergent mechanism. In this mechanism the uncertainty measure plays the role of a certain parameter aggregating in itself particular physiological details.

The characteristic time scale of decision-making during categorical perception, at least in the analyzed case, depends substantially on the uncertainty in classifying a current event; the higher the uncertainty, the longer the decision time. The obtained data enable use to relate this effect with considerable contribution of mental processes to categorization. In this feature categorization perception differs from recognition process near perception thresholds seemed to be governed by physiological mechanisms only.

5. REFERENCES

- [1] Robert L. Goldstone and Andrew T. Hendrickson. Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(1):69–78, 2010.
- [2] Henri Cohen and Claire Lefebvre, editors. *Handbook of categorization in cognitive science*. Elsevier, Amsterdam, 2005.
- [3] Minoru Kobayashi. Asymptotic Behavior of the Psychometric Function of Shape Recognition. Master's thesis, University of Aizu, 2014.

3D Automated Anatomically Correct Face Reconstruction: Facial Feature Points Positioning and Motion Analysis

Iurii Dorofeev
Peter the Great St. Petersburg
Polytechnic University
29 Polytechnicheskaya st.
195251 St. Petersburg Russia
dorofeevyury@gmail.com

Mikhail Glukhikh
Peter the Great St. Petersburg
Polytechnic University
29 Polytechnicheskaya st.
195251 St. Petersburg Russia
mikhail.glukhikh@gmail.com

ABSTRACT

This paper solves a task of facial feature points positioning in 3D space with the CLM algorithm and the structure from motion approach. In our implementation we use a video stream received through the web camera as an input. The person face position and feature points movements are determined using Viola and Johns algorithm. In order to get three-dimensional images, we register person's face movement as a result head rotations by a predefined angle which makes possible to reconstruct the 3D image.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous;
I.4.6 [Image Processing and Computer Vision]: Segmentation—*edge and feature detection*; I.5.4 [Pattern recognition]: Applications—*computer vision*

General Terms

Algorithms, Automation

Keywords

Facial feature points, Reconstruction 3D structure, Texture, Digital image processing

1. INTRODUCTION

Currently a problem of face feature points localization is an important aspect in the area of biometric applications (used in fingerprint identification, forensic science, anthropometric portrait compilation, etc.) Dynamic tracking of control points is used also in psychology to analyze motor activity which, in turn, could be a part of emotional state assessment techniques (facial action coding system, for example [6]).

In recent years, the application field of such systems widens significantly. Due to availability of powerful equipment and

tools there's more opportunities for automated approaches making possible to conduct experiments in nearly real-time, without a lengthy procedure of manual image marking. The main challenge is to achieve the accuracy of automatic feature points discovery comparable to the accuracy of manual image marking made by the experts. The next task is how to construct a 3D image model by using the data obtained as a result of the automatic stage of feature points detection.

Such 3D-reconstruction systems are used in numerous applications in medicine (e.g. for plastic surgery planning), in intelligent biometric security systems (e.g. for recognition and classification of persons), or in compute game industry (e.g. for animation of a character's face).

2. AUTOMATED FACIAL FEATURE POINTS POSITIONING

This work is focused on problems of automatic detection, analysis, and interpretation of the two-dimensional feature points array in order to construct a whole model of a face. Input data are the frames extracted from a video recording of the process of human head's rotation. The objective of the first stage (while processing the frames) is to detect the model's face detection. The second stage is aimed to recognize and highlight anatomical facial features. For the solution of the problem of finding the key points of the face, the most suitable algorithm 2: active appearance model (AAM) and active shape model (ASM). For our task, best suited ASM algorithm. The third stage is to determine the angle of face rotating in each frame and to get the texture. The final stage is to stack images into 3D model, by using the recognized facial features and calculated angles. Hence, the output is twofold: first, this is a sequence of feature points 3D coordinates, second, this is a face texture.

2.1 Face Localization

Face localization algorithm uses well-known approach of Viola and Jones [8]. The approach is based on the Haar wavelet which is a sequence of square-shape functions representing here (in a two-dimensional space) a pair of two adjacent rectangles: light and dark. In fact, the algorithm uses the primitives which are similar to Haar wavelet, but not exactly the same. That's why that functions are often called Haar functions, not Haar wavelets.

The process goes as follows. In the scanning window, the algorithm uses one of the primitives, superimposing it on the image window, then calculates the sum of pixel values for each area Haar function: $F = X - Y$, where X is the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

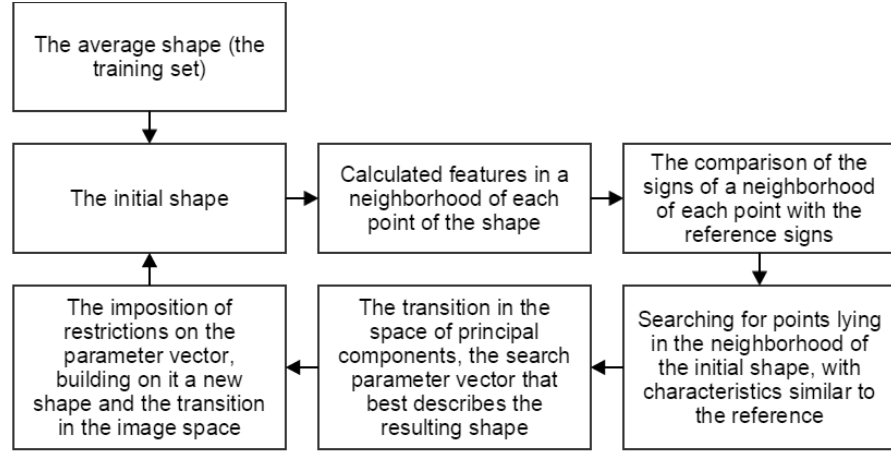


Figure 1: Scheme of the process of localization of feature points on the basis of ASM

amount of brightness of the light rectangle' points, Y is the amount of the dark rectangle' points. Haar functions describe the brightness difference in the image. Then, based on the fact that the nasal septum area is usually lighter than the area around the eyes, the algorithm gets the generalized characteristics of the image anisotropy [11].

2.2 Using Active Shape Models for Feature Points Localization

Active shape model (see [8, 9]) is used to determine a set of feature points as follows:

$$\hat{x} = \bar{x} + \Phi b$$

where \hat{x} is the desired shape vector ($n \times 2$ matrix, n is the number of feature points), \bar{x} is the mean shape vector, got with the use of the training sample images; b is the vector of parameters; Φ is the matrix of eigenvectors, where each eigenvector corresponds to the largest eigenvalue of the covariance matrix S , obtained according to the training sample as follows:

$$S = \frac{1}{n_{shapes} - 1} \sum_{i=1}^{n_{shapes}} (x_i - \bar{x})(x_i - \bar{x})^T$$

where n_{shape} is the number of shapes used in the training set.

Thus, the problem of determining the face shape (as a combination of its feature points) is reduced to determination of the parameter vector b , which is responsible for the variability of the desired shape relatively "the shape in average". During determination of the face feature points, the optimization tasks has being solved, in order to minimize the functional $d(x, T(\bar{x} + \Phi b))$, with d representing the distance, and x is a shape in question found as a result of face feature analysis. Figure 1 represents the feature points detection process. The described actions have to be repeated in a loop for different image scales (usually 4 scales are enough, starting with the smallest one).

2.3 3D Face Structure Reconstruction

In order to reconstruct the 3D face structure from a set

of 2D data, we use the approach known as *Structure from motion* [1, 3, 10, 7].

Inputs are the sets of 2D coordinates together with the information about the face rotation angles. Usually only two face projections are used to restore the structure: full face view (front view) and half face view (considering right and left face sides equal). In our approach we attempted to pay attention to the personality features and to take into consideration the fact that a human face isn't symmetric. Thus, to reconstruct the 3D structure, we use three sets of data: the first set corresponding to the head rotated by -20° around X -axis, the second set corresponding to the frontal face view, and the third set – to the head rotated by $+20^\circ$.

The results are represented as Euler angles. Thus, the rotation matrix is as follows:

$$M(\alpha, \beta, \gamma) =$$

$$\begin{pmatrix} \cos \alpha \cos \gamma - \sin \alpha \cos \beta \sin \gamma & -\cos \alpha \sin \gamma - \sin \alpha \cos \beta \cos \gamma & \sin \alpha \sin \beta \\ \sin \alpha \cos \gamma + \cos \alpha \cos \beta \sin \gamma & -\sin \alpha \sin \gamma \cos \alpha \cos \beta \cos \gamma & -\cos \alpha \sin \beta \\ \sin \beta \sin \gamma & \sin \beta \cos \gamma & \cos \beta \end{pmatrix}$$

The algorithm reconstructing the 3D face image performs the following steps:

1. Finding the maximum and minimum coordinates values for each of the three face projections. Calculate the geometric center of the face;
2. Computing relative coordinates (by using the geometric center determined at the previous step);
3. Multiplying the point coordinates by the rotation matrix (thus paying attention to the face projection);
4. Computing z coordinates for both turns as $z = \frac{p_2^1 \cos p - p_2^0}{\sin p}$, where p_2^1 and p_2^0 are position feature points before and after turn, p is the turn angle;
5. Group out points to three regions (as shown in Figure 2):
 - points from the left-hand area: 1–9, 18–22, 37–40;
 - central area: 28–36, 49–68;

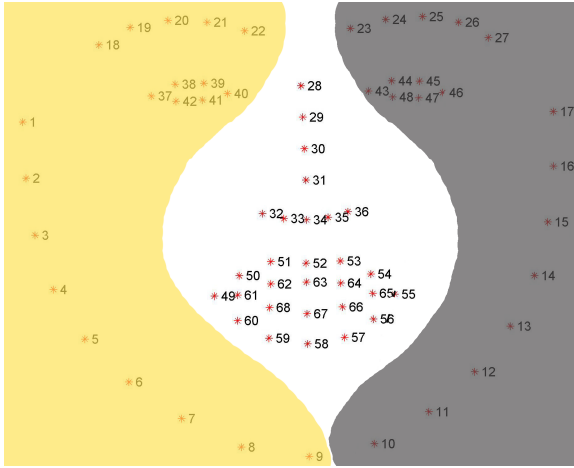


Figure 2: The distribution of the feature points on the face

- points from the right-hand area: 10–17, 23–27, 43–46.
6. For the left hand area, use z values computed for the left turn; for the right-hand area, using z values computed for the left turn; for the central area use the mean of z values for the left and right areas;
 7. For the coordinates X and Y using the full face projection coordinates;
 8. Going back to absolute coordinates.

2.4 Head Rotation Angle Determination

For determination of the head rotation angle we use the CLM Framework [4, 5]. For initial camera settings the following values are used:

- $cx = 0$ (the position of the optical center horizontally);
- $cy = 0$ (the position of the optical center vertically);
- $fx = 500$ (focal length horizontally);
- $fy = 500$ (focal length vertically).

The CLM Framework doesn't provide a method for computing the head rotation matrix. However, there is a method which determines the head position relative to the image center (fx, fy) . There is also a method determining how the camera is located relative to the face. By using these methods the solution of so-called PnP problem (perspective- n -point problem) is achieved with the help of OpenCV [2]. The rotation matrix can be obtained by multiplying the camera rotation value (relative to the face) by the head rotation value (relative to the image center). Figure 3 shows how the position of the coordinate axis is changed.

2.5 Texture Calculation

Assume the model's texture is a full face image. Since we don't need to reconstruct the full 3D scene but only the face, the only recognized face is cut from the full face image in order to get the texture.

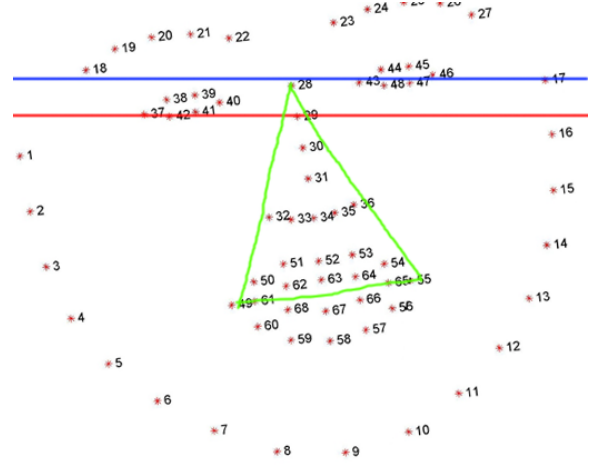


Figure 3: Texture rotation around of the center of on a face

Many editors work mostly with square textures, so we have to crop the image as a square shape:

For the reason that a face might be inclined, we have to rotate the texture by the desired angle. For this purpose we calculate the positional difference between the eye edges (points 37 and 46) and rotate the image until blue and red lines in Figure 3 coincide. The face center is assumed to be a triangle center with the vertices in points 28, 49, 55 (see green triangle in Figure 3).

2.6 Software Prototype

The software prototype is based on using the above mentioned CLM framework and OpenCV library. Figure 4 illustrates the program structure.

For our experiments we used a computer built-in web camera in order to obtain the video stream as input. The capture is made by using the OpenCV library.

Information is processed according to the stages described in the above subsections.

The program is written in C++ in Visual Studio 2012. Currently written version for Windows. Tested to work on Windows 8.1.

3. CONCLUSIONS AND FUTURE WORK

In this work, we developed a software application which allows to obtain 3D coordinates of the face and texture feature points in order to reconstruct a human face 3D model.

Currently the number of detected points is limited by 68. The accuracy can be increased if we increase the number of the points and if we pay attention to other areas (cheeks, forehead, cheekbones).

We arranged a set of experiments and revealed that in the case of significant head rotations (beginning from about 40°), the recognition results became much less accurate. At the moment the program has a very large size due to the use of OpenCV library. Should be removed from the library only the necessary functions that greatly reduce the size of files. Various disturbances of the face, such as (mustache, beard, glasses, etc.) can introduce errors found in the coordinates of the points, it is necessary to invent an algorithm that could determine the position of the interference on his face.

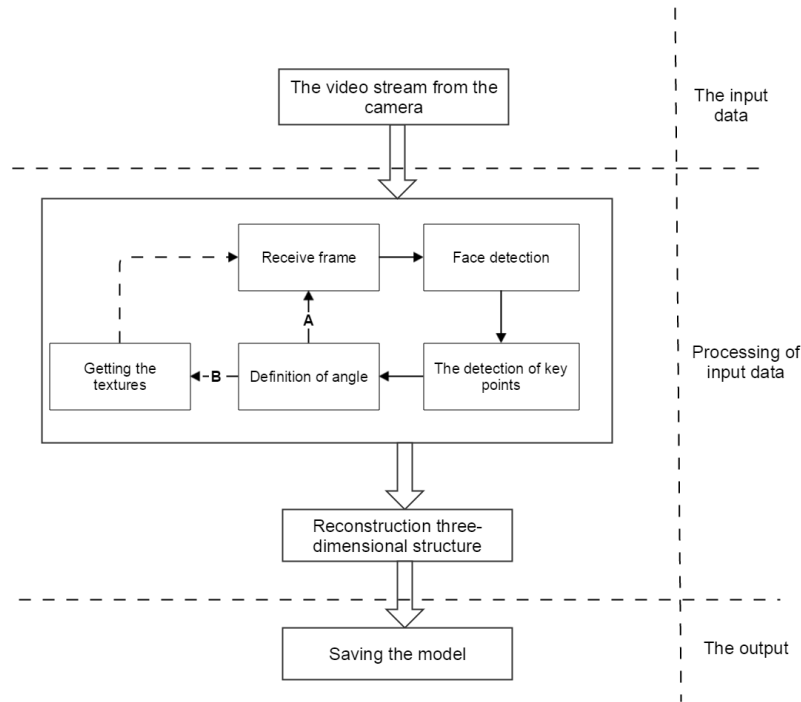


Figure 4: Program structure

4. ACKNOWLEDGMENT

Thanks to Evgeny Pyshkin for his helpful comments which significantly improved our contribution.

5. REFERENCES

- [1] Bundler: Structure from motion (sfm) for unordered image collections, <http://www.cs.cornell.edu/~snavely/bundler/>. [Online; accessed Jul 11, 2015].
- [2] Opencv (open source computer vision), <http://opencv.org/>. [Online; accessed Jul 11, 2015].
- [3] Structure from motion, <http://mi.eng.cam.ac.uk/~cipolla/publications/contributionToEditedBook/2008-SFM-chapters.pdf>. [Online; accessed Jul 11, 2015].
- [4] T. Baltrusaitis. Clm-framework, <https://github.com/TadasBaltrusaitis/CLM-framework>. [Online; accessed Jul 11, 2015].
- [5] T. Baltrusaitis, P. Robinson, and L.-P. Morency. Constrained local neural fields for robust facial landmark detection in the wild. In *Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on*, pages 354–361. IEEE, 2013.
- [6] P. Ekman and W. V. Friesen. Facial action coding system: a technique for the measurement of facial movement. 1978.
- [7] K. E. Ozden, K. Schindler, and L. Van Gool. Multibody structure-from-motion in practice. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(6):1134–1141, 2010.
- [8] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–511. IEEE, 2001.
- [9] P. Viola and M. J. Jones. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004.
- [10] C. Wu. Towards linear-time incremental structure from motion. In *3D Vision-3DV 2013, 2013 International Conference on*, pages 127–134. IEEE, 2013.
- [11] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International journal of computer vision*, 13(2):119–152, 1994.

Design of Automatic Speech Emotion Recognition System

Elena Dmitrieva
Peter the Great St. Petersburg Polytechnic
University
29 Polytechnicheskaya st.
195251 St. Petersburg Russia
ledmitr17@gmail.com

Kirill Nikitin
Peter the Great St. Petersburg Polytechnic
University
29 Polytechnicheskaya st.
195251 St. Petersburg Russia
execiter@mail.ru

ABSTRACT

In this paper we describe a speech emotion recognition system by using k nearest neighbor classifier of statistic features of prosodic contours. We survey major approaches to emotion recognition and argue for using an algorithm dealing with a selection of statistic features of the prosodic contours with further reduction feature space by using SFFS, PCA and LDA and classification provided by k-NN classifier. We tested the designed system by using different combinations of the mentioned algorithms in order to select the optimal combination.

Categories and Subject Descriptors

I.2.7 [Artificial Intelligence]: Natural Language Processing—*speech recognition and synthesis*; I.5 [Pattern Recognition]: Applications

General Terms

Algorithms

Keywords

Emotion recognition, speech processing, classification, feature extraction, feature preparation.

1. INTRODUCTION

Nowadays much attention is given to speech recognition systems used in many applications such as voice recognition in navigators, voice control in mobile devices, voice search systems, etc. In most implementations the speech is converted into the text form, and then processed by using natural language processing (NLP) technologies. However it's human to convey information not only by using words, but also with the help of emotions. There is a hypothesis that the quality of speech recognition process can be improved if the problem of recognizing emotions is taken into consideration. [16]

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

2. EMOTION RECOGNITION SYSTEM

An input for an emotion recognition system is a speech expected to contain emotions (emotional speech). The expected output is the classified emotion (we know that classification is the primary objective of any pattern recognition systems) [9]. The process consists of the following stages:

- Feature extraction component;
- Feature normalization;
- Feature preparation;
- Classification.

For emotion recognition system design Matlab environment is used because of its numerical computing orientation. For now to detect some features like pitch and formants PRAAT software is used. But in the future we plan to program all system parts on Matlab.

2.1 Feature Extraction

In fact, feature extraction is the most important and complicated step. The problem is that it is in a priori unknown what features should be extracted for efficient emotion recognition. That's why usually as many as possible features are being extracted in order to select the most informative of them during the further processing.

In our emotion recognition approach, we use 381 features described in table 1.

After the extraction the features are normalized by using their mean value and their standard deviation value as follows:

$$\hat{x} = \frac{x - \mu}{\sigma}$$

Short-Term Energy Features.

Short-Term energy is one of the most important features that gives good information about the emotion [9]. It can be calculated by:

$$E(n) = \sum_{m=n-N+1}^n [x(m)w(n-m)]^2$$

where $w(n-m)$ is the hamming window, n is the sample in analyzed window, and N is the window size.

Table 1: Extracted Features

Indices	Features
1 - 20	Mean, minimum, maximum, standard deviation, range, interquartile range, 30, 50, 90th percentile of ST energy and first derivative of ST-energy
21 - 30	Mean, minimum, maximum, standard deviation, range, interquartile range, 30, 50, 90th percentile of db-energy
31 - 50	Mean, minimum, maximum, standard deviation, range, interquartile range, 30, 50, 90th percentile of rising slopes of db-energy, falling slopes of db-energy
51 - 60	Mean, minimum, maximum, standard deviation, range, interquartile range, 30, 50, 90th percentile of first derivative of db-energy
61 - 70	Mean, minimum, maximum, standard deviation, range, interquartile range, 30, 50, 90th percentile of rising slopes of first derivative of db-energy
71 - 80	Mean, minimum, maximum, standard deviation, range, interquartile range, 30, 50, 90th percentile of falling slopes of first derivative of db-energy
79 - 86	Minimum, maximum, mean, range, standard deviation, interquartile range, 75 and 90th percentile of pitch
87 - 110	Minimum, maximum, mean, range, standard deviation, interquartile range, 75 and 90th percentile of rising slopes of pitch, falling slopes of pitch, plateaux at minima
111 - 118	Minimum, maximum, mean, range, standard deviation, interquartile range, 75 and 90th percentile of first derivative of pitch
119 - 142	Minimum, maximum, mean, range, standard deviation, interquartile range, 75 and 90th percentile of rising slopes, falling slopes of first derivative of pitch, plateaux at minima of first derivative of pitch
143 - 151	Jitter, RAP, PPQ5, DDP, Shimmer, APQ3, APQ5, APQ11, DDA
152 - 193	Maximum, minimum, mean, median, standard deviation, interquartile range, variance, skewness, 90th percentile of the 1st, 2nd and 3rd formants and BW of the 1st, 2nd and 3rd formants
194 - 201	Spectral energy between: 0 - 250, 0 - 600, 0 - 1000, 0 - 1500, 250 - 600, 600 - 1000, 1000 - 1500, 250 - 1000 Hz
202 - 211	Minimum, mean, range, median, standard deviation of voiced segments durations, unvoiced segments durations
212 - 213	Speaking rate, number of voiced segments
214 - 243	Mean, minimum, maximum, standard deviation, range, interquartile range of zero-crossing rate, spectral centroid, spectral rolloff, spectral flux, spectral crest factor, spectral flatness
250 - 303	Mean, minimum, maximum, standard deviation, range, interquartile range of 9 LPC-coefficients
304 - 381	Mean, minimum, maximum, standard deviation, range, interquartile range of 12 MFCC-coefficients

Pitch Features.

Pitch (or fundamental frequency) is connected to the possibility to distinct between male and female speeches. The pitch is determined by the frequency of vibration of the vocal cords. The pitch is individual for each person and depends on the structure of the vocal tract. There are many algorithms of pitch detection [9]: HPS, RAPT, AMDF, CPD, SIFT, etc. Pitch detection is not trivial problem [14], so, we use PRAAT software to minimize estimation error. PRAAT is a free software for analyzing, synthesizing, and manipulating speech [3].

Formants Features.

Unlike to pitch featuring a tone of voice, Formants characterize timbre of voice. Formants are characterized by the center frequency and bandwidth [6]. Formants frequencies and bandwidth can be calculated with the help of linear prediction analysis, but in our case we use the PRAAT software in order to minimize error.

Jitter and Shimmer Features.

Jitter and shimmer are measures of the cycle-to-cycle variations of the fundamental frequency and amplitude, which have been largely used for the description of pathological voice quality [8].

In this work we use local jitter, local shimmer, relative average perturbation (RAP), period perturbation quotient (PPQ), difference of differences of periods (DDP), amplitude perturbation quotient (APQ), difference of differences of amplitudes (DDA) calculated by PRAAT. These features are described in detail in [15].

Spectral Features.

Spectral features are calculated from Fast Fourier Transform (FFT) of every short-time frame of speech signal, except spectral energy which is calculated from the whole signal. In our work we use spectral features like spectral energy, spectral centroid, spectral rolloff, spectral flux, spectral crest factor, spectral flatness which are described in [11]:

Zero-crossing Rate.

Zero-crossing rate is the weighted average number of sign changes in each signal frame. It can be estimated by:

$$Z_n = \sum_{m=-\infty}^{\infty} 0.5 |sgn[x(m)] - sgn[x(m-1)]| w(n-m)$$

where

$$sgn(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$$

Speaking Rate.

Speaking rate is obtained by dividing number of voiced segments by number of all segments. It can be estimated by:

$$S = \frac{N_v}{N_{uv}}$$

LPC-coefficients.

Linear predictive coding (LPC) is based on the physical

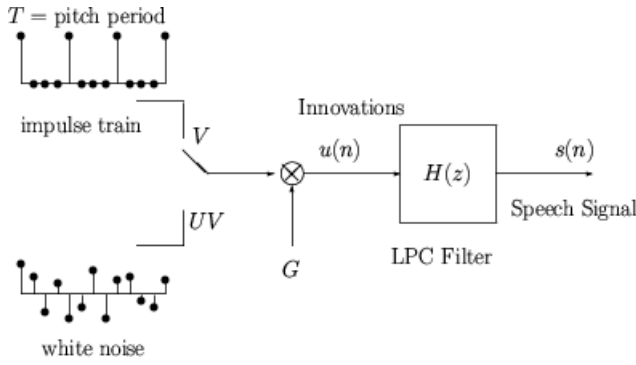


Figure 1: LPC model of speech signal

model of the human speech which is presented on fig. 1. The picture shows that speech can be modelled as the output of a linear, time-varying system excited by either quasi-periodic pulses (during voiced speech), or random noise (during unvoiced speech). The basic idea of linear predictive coding (LPC) is that a speech sample can be approximated as a linear combination of past speech samples. LPC-coefficients can be determined by minimizing the sum of the squared differences (over a finite interval) between the actual speech samples and the linearly predicted ones.

MFC-coefficients.

Mel-frequency cepstrum (MFC) is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency. So, mel-frequency cepstrum coefficients are describe an MFC.

2.2 Feature Preparation

After feature extraction we get the large set of features. If we try to classify emotions by using this set, we get very high error rate about $\sim 74\%$. So, let us feature set dimension. There are two known feature reduction methods:

- Selection of the most informative subset of features from the source set;
- Transformation of the source set to the new set, where result set is some function of the source set.

2.2.1 Feature Selection

We know the following hierarchy of feature selection algorithms:

- Optimal algorithms (exhaustive search, branch-and-bound, etc.);
- Suboptimal algorithms (sequential forward selection, plus-L minus-R search, sequential forward floating search, etc.).

To select the most informative set a sequential forward floating search (SFFS) was used which is suboptimal algorithm [13].

2.2.2 Feature Transformation

PCA (Principal Component Analysis) converts a set of features to a new set of linearly uncorrelated features called principal components.

LDA (Linear Discriminant Analysis) finds linear combination of features that separates object classes.

2.3 Review of Emotion Classification Methods

There are two major approaches of emotion classification [17]:

- Use of prosodic contours to classify emotions. It can be done by:
 - Artificial neural network (ANN)
 - Multichannel hidden Markov model (HMM)
 - Mixture of hidden Markov models
- Use of statistic features of prosodic contours. These methods are divided into two types:
 - With pdf modelling:
 - * Variations of Bayes classifier
 - * Parzen windows
 - Without pdf modelling:
 - * k-nearest neighbor (k-NN)
 - * Artificial neural network (ANN)
 - * Support vector machine (SVM)

In this work we use k nearest neighbor ($k-NN$) algorithm for the reason that it is relatively simple and allows extending the training data set.

2.4 Database

We use Berlin emotion database [4] in order to train and test the speech emotion recognition system. This database consists of seven basic emotions: anger, boredom, disgust, fear, happiness, sadness and neutral. The are simulated using 535 speech samples.

To fit the $k-NN$ classifier requirements we have to reduce the database: the reduced version consists of 322 emotional utterances.

2.5 Experiments

In our experiments different combinations of feature preparation algorithms were tested. For testing we use cross-validation [7]. In this method data set is divided into k (in our work $k = 10$) subsets and recognition procedure is repeated k times. Each time, one of the subsets is used as a testing set and other subsets are used as a training set. Recognition accuracy is computed as mean of k recognition accuracy results. Comparison of different combinations of feature preparation algorithms are shown in Table 2. It shows us, that without feature selection and preparation recognition quality is very low. Reason of that is what there are too many features, which duplicate and depend on each other. That is undesirable when k-NN classifier is used.

As you can see, even when we use the simplest feature extraction procedure, quality is improved sufficiently. The best result (accuracy of 81%) was reached with combination of SFSS + PCA + LDA which we consider a reasonably good value. Comparison with different researches results are shown in Table 3. Besides, best results are archived for anger (92%), and worst for happiness and fear (near 60%). More likely this is due to specialty of database used, emotion characteristics may also affect the result.

Table 2: Comparison of different combinations of algorithms

Algorithms	Accuracy
None	26%
SFFS	45%
PCA	63%
PCA + LDA	76%
SFFS + PCA	74%
SFFS + PCA + LDA	81%

Table 3: Comparison with other emotion recognition systems

Research group	Classification	Result
Shuller et al. [12]	k-NN	80.3%
Ayadia et al. [1]	HMM	77%
	HMM mixture	78.4%
	ANN	66.5%
	SVM	78.4%
Hendy, Farag [9]	PNN	84%
	LVQNN	71%
	BPNN	74%
Kotti, Paterno [10]	k-NN	75.5%
	Gaussian SVM	83.6%
	Linear SVM	85.6%
Busso et al. [5]	k-NN	83.5%
Our system	k-NN	81%

It is also worth noticing, that recognition quality depends on speaker language: best quality is achieved when both - learning and testing are executed using same language. So, when testing with polish emotion database we get accuracy just 15%.

3. CONCLUSION

In this work emotion recognition system was designed. Feature set consists of 381 features was extracted. The best result was reached with SFFS + PCA + LDA feature reduction algorithms combination and classification by k-NN classifier - 81%.

Very high recognition accuracy was reached, but it probably can be a little higher.

Despite rather good accuracy value we achieved in our experiments, there is space for further improvements in the algorithms of feature preparation and selection. We believe interesting to investigate possibilities to use support vector machines, deep learning strategy, HMM mixture [2] as a model which could improve the classification accuracy and speed.

4. REFERENCES

- [1] M. E. Ayadia, M. S. Kamel, and F. Karray. Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, 44(3):572–587, 2002.
- [2] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [3] P. Boersma and V. van Heuven. Speak and unspeak with praat. *Glott International*, 5:341–347, 2001.
- [4] F. Burkhardt, A. Paeschke, M. Rolfes, W. Sendlmeier, and B. Weiss. A database of german emotional speech. In *Proceedings of Interspeech*, pages 1517–1520, 2005.
- [5] C. Busso, M. Bulut, and S. S. Narayanan. *Toward effective automatic recognition systems of emotion in speech*. Oxford University Press, 2012.
- [6] A. de Cheveigne. Formant bandwidth affects the identification of competing vowels. In *International conference on phonetic sciences*, pages 2093–2096, 1999.
- [7] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Macmillan Publishers Limited, 1998.
- [8] M. Farrus, J. Hernando, and P. Ejarque. Jitter and shimmer measurements for speaker recognition. *Interspeech*, pages 778–781, 2007.
- [9] N. A. Hendy and H. Farag. Emotion recognition using neural network: A comparative study. *World Academy of Science, Engineering and Technology*, 7(3):1149–1155, 2013.
- [10] M. Kotti and F. Paterno. Speaker-independent emotion recognition exploiting a psychologically-inspired binary cascade classification schema. *International Journal of Speech Technology*, 15(2):131–150, 2012.
- [11] G. Peeters. A large set of audio features for sound description (similarity and classification) in the cuidado project. Tech. rep., IRCAM, 2004.
- [12] B. Schuller, M. Lang, and G. Rigoll. Automatic emotion recognition by the speech signal. In *6th World Multiconference on Systemics, Cybernetics and Informatics*, pages 367–372, 2002.
- [13] P. Somol, P. Pudil, J. Novovicova, and P. Paclik. Adaptive floating search methods in feature selection. *Pattern Recognition Letters*, 15(11):1119–1125, 1999.
- [14] X. Sun. A pitch determination algorithm based on subharmonic-to-harmonic ratio. In *the 6th International Conference of Spoken Language Processing*, pages 676–679, 2000.
- [15] J. P. Teixeira, C. Oliveira, and C. Lopes. Vocal acoustic analysis – jitter, shimmer and hnr parameters. *Procedia Technology*, 9:1112–1122, 2013.
- [16] M. S. Unluturk, K. Oguz, and C. Atay. Emotion recognition using neural networks. In *Proceedings of the 10th WSEAS International Conference on Neural Networks*, pages 82–85, 2009.
- [17] D. Ververidis and C. Kotropoulos. Emotional speech recognition: Resources, features, and methods. *Speech Communication*, 48(9):1162–1181, 2006.

Using Dynamic Predicate Logic for Pronominal Anaphora Resolution in Russian Texts

Nikita Gerasimov
St. Petersburg State University
7-9, Universitetskaya nab.
St. Petersburg, Russia
n@tariel.ru

Evgeny Pyshkin
Peter the Great St. Petersburg
Polytechnic University
29 Polytechnicheskaya st.
St. Petersburg, Russia
pyshkin@icc.spbstu.ru

ABSTRACT

This study is focused on semantically-linked words detection in natural language written documents with particular attention paid to pronominal anaphora resolution in Russian language. The objective of this study is to investigate whether the approach based on using dynamic predicate logic (DPL) is appropriate to formalize a Russian pronominal anaphora and to resolve anaphoric connections and antecedents in Russian texts. As a result of the experiments that we arranged, we realized that currently the DPL model and related algorithms are not adequate to be directly applied to anaphora resolution in Russian texts due to high computational complexity and low anaphora detection accuracy.

Categories and Subject Descriptors

I.2.4 [Artificial Intelligence]: Natural Language Processing

General Terms

Theory

Keywords

Pronominal anaphora, Predicate logic, PLA, DPL, NLP

1. INTRODUCTION

In linguistics, *anaphora* is the use of a language expression that can be interpreted correctly by taking into consideration its dependency on another expression (the latter being its *antecedent* or *postcedent* depending on whether it is used after or before the referred sentence). For AI systems such references are the very problem on the score of understanding anaphoras and antecedents as different objects. This case impairs quality of knowledge extraction and AI processing. Anaphora identification is one of the complex tasks in

the domain of semantic analysis and natural language processing (NLP). The problem is particularly challenging for the current human-centric computer systems that use many different ways of human oriented interaction and require support for NLP.

Anaphora resolution could be considered as a special case of a general problem of recognition of referencing one semantic object by using different words and implicit connections. For example, learning cognitive synonyms can significantly improve the quality of web search and query expansion [12, 7]. A particularly difficult case of anaphora resolution is the case in which the reference word is a pronoun (pronominal anaphora).

Despite since about 1970s we know many approaches and algorithms developed for pronominal anaphora resolution (see the exhaustive study of Ruslan Mitkov [10], for example), the problem isn't totally fixed even for the English language [3, 6]. In contrast to English, developing algorithms for Russian language anaphora resolution seems to be still an emerging area. Among recent works we could cite the *An@phora* system [8] based on machine learning techniques and knowledge engineering formalisms as well a rule based approach described in [9].

In this short paper we report an attempt to use the Paul Dekker's approach described in [2] for pronominal anaphora resolution in Russian texts. We also report some preliminary results of testing our implementation against a corpus of Russian texts.

2. DYNAMIC PREDICATE LOGIC

Dynamic predicate logic (DPL) is a dynamic semantic interpretation of the first-order logic language for the purposes of defining a compositional theory of discourse semantics [5].

2.1 DPL in Brief

Lets us illustrate the difference between the first-order logic and the DPL by an example. Look at the the following sentence:

Maria borrowed the textbook from her professor.

By using the first-order logic we can represent the semantic of this sentence by the following construction:

$$\exists x[B(x) \wedge \exists y(P(y, x) \wedge T(m, y, x))] \quad (1)$$

where $B(x)$ is the textbook, $P(y, x)$ represents the fact that professor y owns the book x , and $T(m, y, x)$ represents the fact that Maria m borrowed x from y .

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

The anaphoric connection of *her* to *Maria* is represented correctly since all constructs in the conjunction from the equation (1) have the same scope (or, by using terms from [5] they are bound by the same existential quantifier [...]).

Let's add the following phrases:

The textbook(x) was full of comments and remarks. Surely, he(y) was reading it(x) very attentively.

If we analyze the latter discourse independently of the above introduced sentence, it would be hard to resolve cross-sentential anaphoric references from *he* to *the professor*, as well as from *it* (the textbook) to that exact textbook borrowed by *Maria*.

Under the DPL model, the values x and y are, in a sense, "saved" in order to expand the search of suitable values for the whole domain D within the framework of the model $M = \langle D, I \rangle$, D being the domain of individuals under discussion, while I – an interpreting function.

Thus, DPL is a kind of first-order logic where the following rules are introduced:

THEOREM 1. *Egli theorem*

$$(\exists x \varphi \wedge \psi) \leftrightarrow \exists x(\varphi \wedge \psi)$$

THEOREM 2. *Egli corollary*

$$(\exists x \varphi \rightarrow \psi) \leftrightarrow \forall x(\varphi \rightarrow \psi)$$

In order to see how the theorems (1–2) could be used, let's consider the following classical whistler-examples:

A Kid is Going home. He is Whistling.

A Kid who is Going home is Whistling.

By applying the above rule, both examples are being translated into the following predicate logic formula:

$$(\exists x(Kx \wedge Gx) \wedge Wx) \leftrightarrow \exists x((Kx \wedge Gx) \wedge Wx) \quad (2)$$

Moreover, unlike to classic logic, DPL offers idempotence and commutativity properties: $\phi \wedge \psi \leftrightarrow \psi \wedge \phi$ or $\phi \leftrightarrow \psi \vee \phi$ is false if the interpretation changes from ϕ to ψ . Again, if the first part of $\psi \vee \psi$ has the interpretation which is different from the second part's one, the whole formula yields false.

2.2 Predicate Logic for Anaphora Resolution Language

PLA (Predicate logic for anaphora) resolution language includes the following entities:

1. Relational constants R^n
2. Individual constants $c \in C$
3. Variables $x \in V$
4. Pronoun variables p_i
5. Terms $t \in (c, x, p_i)$
6. Formulas $\phi \in (R_n t_1 \dots t_n, \neg x, \exists x \phi, \phi \wedge \phi)$

Constants and *variables* are being interpreted according to the classic first-order logic. *Pronoun* p_i interpretation depends on the context.

PLA pronouns represent functions choosing the correct antecedent. Antecedent candidates are represented by the list of *terms* where i -th pronoun selects i -th existence quantifier.

2.3 Anaphora Resolution Process By Examples

Let us introduce a couple of examples to show how the anaphora is resolved. Here is the first one:

A Kid Walks down the park ($\exists x(Kx \wedge Wx)$).

There is also a Dog ($\exists y Dy$).

It Frightens him and he Chases it ($Fp_1 p_2 \wedge Cp_2 p_1$).

Sentences are transformed into the following formula:

$$(\exists x(Kx \wedge Wx) \wedge \exists y Dy) \wedge (Fp_1 p_2 \wedge Cp_2 p_1) \quad (3)$$

or (in reduced form):

$$\exists y \exists x(((Kx \wedge Wx) \wedge Dy) \wedge (Fyx \wedge Cxy)) \quad (4)$$

where: Kx being the term "a Kid exists", Wx being the term " x Walks", Dy being the term "a Dog exists", $Fp_1 p_2$ being the term " p_1 Frightens p_2 ", $Cp_2 p_1$ being the term " p_2 Chases p_1 ".

The equivalence can be easily proved. The formula (3) requires y and x to be d (dog), and k (kid) to $Cx, Wx, Dy, Fp_1 p_2, Cp_2 p_1$ to be true. Thus, dk is antecedent queue.

The formula (4) produces the same result: first, the system searches the value of x such as both Kx and Wx are true. Further, y values are being searched to get Dx true. Then the final queue is dk .

The final queue fulfills the first proposition: the kid x frightens the dog y and the kid x chases the dog y . If the formula yields true, the queue dk is a correct antecedent queue.

Here is one more complex example:

Once there was a Queen ($\exists x Qx$).

Her Son Fell in Love with a frog ($\exists y(Sy \wedge \exists z(Fz \wedge Lyz))$).

The prince Kissed it and she got Mad ($Kp_1 p_2 \wedge Mp_3$).

The first proposition is true for every queen q . The second one is true for every son (i.e. prince) s who fell in love with a frog f . The resulting pair is sf . The queue sfq is generated after processing the first two propositions. The resulting queue is inserted into the formula so as the last proposition looks like the following: $Ksf \wedge Mq$.

Hence, the whole transformation is as follows:

$$((\exists x Qx \wedge \exists y(Sy \wedge \exists z(Fz \wedge Lyz))) \wedge (Kp_1 p_2 \wedge Mp_3)) \leftrightarrow \exists y \exists z \exists x((Qx \wedge (Sy \wedge (Fz \wedge Lyz))) \wedge (Kyz \wedge Mx))$$

3. ANAPHORA RESOLVING ALGORITHM AND ITS IMPLEMENTATION

3.1 Transformation Algorithm

THEOREM 3. If \hat{x} is a sequence $x_1 \dots x_n$, ϕ and ψ is closed, ϕ doesn't contain \hat{y} , and ψ doesn't contain \hat{x} :

$$(\exists \hat{x} \phi \wedge \exists \hat{y} \psi) = \exists \hat{y} \exists \hat{x}(\phi \wedge [\hat{x}/p_i] \psi)$$

where variables x_i are free for p_i insertions in ψ

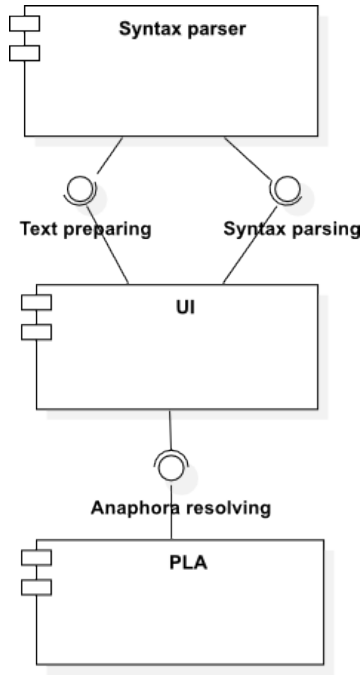


Figure 1: System structure

The formulas ϕ and ψ are closed. Also they don't contain "active" existence quantifiers and locally-resolved pronouns.

Theorem (3) shows that $\exists \hat{x}$ from the left part affects on the right part, if pronouns p_i replaced by x_i are also quantified. After $\exists \hat{x}$ coverage has enlarged, $\exists \hat{y}$ action state has to be checked, so $\exists \hat{y}$ coverage is also enlarged. The nested formulas ϕ and ψ must be closed before resolving the pronouns.

If the number of pronouns is more than of existence quantifiers, the pronoun selects an antecedent from the previous proposition.

Theorem (3) can be used for converting a PLA formula to a classic first-order logic formula. The algorithm returns a first-order logic formula ϕ by using the function $[\phi]^?$ returning the effect $[\psi]^!$ on ψ by ϕ . The algorithm uses the rules as presented in definition 1:

Definition 1.

$$\begin{aligned} &[(\phi \wedge \psi)]^? \rightarrow ([\phi]^? \wedge [\psi]^?); \\ &([\exists \hat{x} \phi]^! \wedge [\exists \hat{x} \psi]^!) \rightarrow [\exists \hat{y} \exists \hat{x} (\phi \wedge [\hat{x}/p_i] \psi)]^! \\ &[\exists x \phi]^? \rightarrow \exists x [\phi]^?; \exists x [\phi]^! \rightarrow [\exists x \phi]^! \\ &[\neg \phi]^? \rightarrow \neg [\phi]^?; \neg [\phi]^! \rightarrow [\neg \phi]^! \\ &[Rt_1 \dots t_m]^? \rightarrow [Rt_1 \dots t_m]^! \end{aligned}$$

3.2 Implementation for Russian Texts

The algorithm described in the above section gets a syntactically processed text as input. System component structure is presented in Figure 1. After the PLA subsystem has got syntax trees for the processed sentences (see Figure 2), building a DPL becomes possible.

To proceed with the input text syntax analysis we use the component described in our earlier work [4]. For morphological analysis of the Russian texts the *Mystem*¹ analyzer

¹<https://tech.yandex.ru/mystem/>

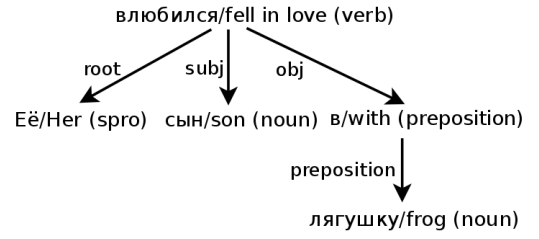


Figure 2: Syntax tree example

is used. Model required for the *MaltParser* [11] is prepared by using the National Corpus of Russian Language².

We implemented the user interface in two forms: CLI and a Java web application developed by using the *Play*³.

Figure 3 represents the user web interface and illustrates how the Russian text equivalent to the above studied *queen-prince-frog* example is processed and resolved.

3.3 Anaphora Resolution Subsystem

The anaphora resolution component is implemented as an independent subsystem (which could be deployed on a separate server) communicating via *Apache Thrift*. This subsystem gets *CoNLL-X* [1] syntactically coded sentences and transforms them to the tree suitable to be represented as a DPL formula, in which the nodes are words, and the edges are their syntactic dependencies. Figure 2 shows an example of the syntax tree generated for the above mentioned sentence "Her son fell in love with a frog". In the process of building a DPL formula both an object noun (e.g. "a frog" in our example) and a subject noun (e.g. "son") are translated into the existence quantified variable with the terms $\exists ySy$ and $\exists zFz$ respectively. Object and subject pronouns are translated into the pronoun variables p_i . The predicates are represented by the terms with the above introduced quantified variables (e.g. "y fell in love with z" to $Lyzy$).

The output data array is as follows: DPL formula, classic first-order formula, resolved sentence.

3.4 Experiments

To arrange the experiments, we used three sets of manually tagged sentences selected from the following sources:

1. 60 sentences from *Syntagrus*;
2. 60 sentences from "Monday Begins on Saturday", the novel by Boris and Arcady Strugatsky;
3. 60 light-syntax structure sentences similar to the above mentioned examples.

As a result of applying the DPL based approach for anaphora resolution in Russian language texts we get the following accuracy values: *Syntagrus* – 9,6%; "Monday Begins on Saturday" – 6,4%; Light-syntax structures – 44,8%.

4. CONCLUSION

Despite the general appropriateness of the DPL to the problem of anaphora resolution, the results we achieved for the Russian texts are rather discouraging. That's why we

²<http://ruscorpora.ru/en/>

³<https://www.playframework.com/>

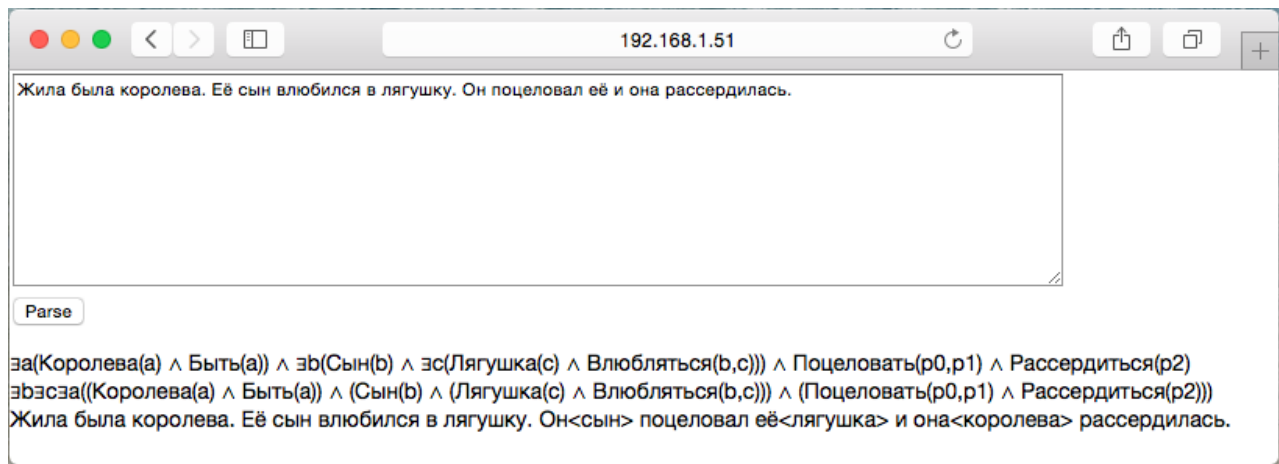


Figure 3: PLA resolution component web interface

have to conclude that both the algorithm and its implementation require further study in order to pay more attention to obvious particularities of Russian language.

Leastwise, we have to mention the following major drawbacks:

1. The model doesn't fit perfectly the task of anaphora resolution within the context of Russian text processing: many important structural linguistics properties are missing e.g. grammar case, gender, adjectives, adverbs;
2. The computational complexity of the process of building DPL formulas will be increasing significantly if we take into account more (currently missing) semantic and grammatical information which could be potentially presented in syntax trees.

To sum up, in its current state the DPL model seems to have no advantages against other known techniques in order to describe adequately the Russian language structures and to achieve satisfactory levels of anaphora resolution accuracy in Russian texts. Word position and appearance order only don't serve anaphora resolution process well. Possible efforts to take into consideration such categories as genre, number, grammatical case, object properties expressed by adjectives and adverbs, etc., might lead to higher complexity of the model without guarantee of better anaphora resolution results.

At present time machine learning techniques like [8] gain rather considerable results and look interesting for further researches and using on practice.

5. REFERENCES

- [1] S. Buchholz and E. Marsi. Conll-x shared task on multilingual dependency parsing. In *Proceedings of the Tenth Conference on Computational Natural Language Learning*, pages 149–164. Association for Computational Linguistics, 2006.
- [2] P. Dekker. *Dynamic Semantics*. Studies in Linguistics and Philosophy. Springer Netherlands, 2012.
- [3] A. Ekbal, S. Saha, O. Uryupina, and M. Poesio. Multiobjective simulated annealing based approach for feature selection in anaphora resolution. In *Proceedings of the 8th International Conference on Anaphora Processing and Applications, DAARC'11*, pages 47–58, Berlin, Heidelberg, 2011. Springer-Verlag.
- [4] N. Gerasimov, M. Mozgovoy, and A. Lagunov. Semantic sentence structure search engine. In *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, Warsaw, Poland, September 7-10, 2014.*, pages 255–259, 2014.
- [5] J. Groenendijk and M. Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14(1):39–100, 1991.
- [6] K. Karthikeyan and V. Karthikeyani. Understanding text using anaphora resolution. In *Pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013 International Conference on*, pages 346–350, Feb 2013.
- [7] V. Klyuev and Y. Haralambous. A query expansion technique using the ewc semantic relatedness measure. *Informatica: An International Journal of Computing and Informatics*, 35(4):401–406, 2011.
- [8] A. Kutuzov and M. Ionov. The impact of morphology processing quality on automated anaphora resolution for Russian. In *Computational Linguistics and Intellectual Technologies: papers from the Annual conference "Dialogue"*, pages 232–241. Moscow, RGGU, 2014.
- [9] M. Malkovskiy, A. Starostin, and I. Shilov. Method of pronoun anaphora resolution in parallel with syntactic analysis. *Perspective innovations in science, education, production and transport proceedings*, 11(4):41–49, 2013.
- [10] R. Mitkov and W. W. Sb. Anaphora Resolution: The State Of The Art. Technical report, 1999.
- [11] J. Nivre, J. Hall, and J. Nilsson. Maltparser: A data-driven parser-generator for dependency parsing. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, volume 6, pages 2216–2219, 2006.
- [12] E. Pyshkin and A. Kuznetsov. Approaches for web search user interfaces. *Journal of Convergence*, 1(1), 2010.

A Program Complex for Learning of Optimal Control Problems

Andrey Klyuenkov
Saint-Peterburg State University
7-9, Universitetskaya nab.
Saint-Petersburg, 199034, Russia
andr_esk@mail.ru

ABSTRACT

In the present paper the task is to develop a learning program complex on optimal control problems. For solving optimal control problems the adaptive algorithm of R. Gabasov is used. The structure of a program complex for continuous models is proposed. The solution is divided into two stages: reducing the optimal control problem to an interval linear programming problem (ILPP), and applying the adaptive method to the ILPP. The developed complex is implemented in MATLAB environment. Stabilization of three-mass oscillatory system with minimal fuel consumption problem is used for testing. The optimal control is computed, and results are presented graphically.

Categories and Subject Descriptors

I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, and Search — *control theory*

General Terms

Theory

Keywords

Optimal control problem, interval linear programming problem, learning

1. INTRODUCTION

In many control problems we have to find the optimal control. Often this problem is quite difficult, however for controllable systems the solution can be found with the adaptive method by R. Gabbasov [1, 2, 3]. Teaching students this method and its program implementation gives them a way to solve optimal control problem and opens a new viewpoint towards the problems. Furthermore this contains unconventional approach to the linear programming problem.

The task was to develop a learning complex on optimal control problems.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

2. STRUCTURE OF THE COMPLEX

Problem statement. The adaptive method can be applied to differential as well as to discrete systems. Consider a system of ordinary differential equations with a control parameter and given initial conditions

$$\dot{x} = A(t)x + b(t)u, \quad (1)$$

$$x(t_*) = x_0, \quad (2)$$

$$t \in [t_*, t^*] = T, \quad t_* < t^* < +\infty,$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^1$, $A(t)$ is a $(n \times n)$ -matrix, $b(t)$ is a vector function whose elements are piecewise continuous functions for $t \in T$.

The function $u(t)$, $t \in T$ denotes a *discrete control* (with the quantization step h), if

$$u(t) = u(t_* + kh) = u_{kh}, \quad t \in [t_* + kh, t_* + (k+1)h] \quad (3)$$

for every $k = \overline{0, N-1}$. Here $h = (t^* - t_*)/N$ is the quantization step and N is a natural number. Besides, $|u(t)| \leq L$ for $t \in T$, $L > 0$. Denote the set of decomposition points as $T_u = \{t_*, t_* + h, \dots, t^* - h\}$.

Let's formulate a terminal control problem: To find the solution of the initial value problem (IVP) (1), (2) with piecewise constant control, that maximizes the objective function

$$c^T x(t^*) \rightarrow \max \quad (4)$$

provided that

$$Hx(t^*) = g, \quad (5)$$

where $g \in \mathbb{R}^m$ and $\text{rank } H = m < n$.

The problem is solved with the adaptive Gabasov method. The method is applied in two stages. *The first stage* is reducing of the optimal control problem to an *interval linear program problem* (ILPP). *The second stage* is solving the obtained ILPP with the adaptive method.

The first stage can be presented as following. The general solution of the IVP (1),(2) at time $t = t^*$ is

$$x(t^*, t_*, x_0) = Y(t^*) \left[Y^{-1}(t_*)x_0 + \int_{t_*}^{t^*} Y^{-1}(\tau)b(\tau)u(\tau)d\tau \right], \quad (6)$$

where $Y(t)$ is the fundamental matrix of the corresponding homogeneous system. The substitution of (6) to (4) gives an equivalent representation

$$\int_{t_*}^{t^*} c^T Y(t^*)Y^{-1}(\tau)b(\tau)u(\tau)d\tau \rightarrow \max. \quad (7)$$

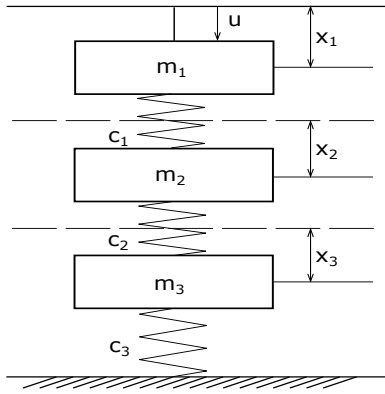


Figure 1: Oscillatory system

Consider the dual system with initial conditions

$$\dot{z} = -A^T(t)z, \quad (8)$$

$$z(t^*) = c, \quad (9)$$

where c and t^* comes from the conditions (4), (5). Its solution can be written as $z_c(t) = Z(t)Z^{-1}(t^*)c$, where $Z(t)$ is the fundamental matrix of the system (8). The main property of the fundamental matrix is $Z^T(t) = Y^{-1}(t)$. With use of this property it's easy to find that

$$z_c^T(t) = c^T (Z^{-1})^T(t^*)Z^{-1}(t) = c^T Y(t^*)Y^{-1}(t). \quad (10)$$

Considering (10) the condition (7) becomes

$$\int_{t_*}^{t^*} z_c^T(\tau)b(\tau)u(\tau)d\tau \rightarrow \max. \quad (11)$$

Now consider the terminal condition (5). Using (6) it can be represented as

$$\int_{t_*}^{t^*} HY(t^*)Y^{-1}(\tau)b(\tau)u(\tau)d\tau = g_0, \quad (12)$$

where $g_0 = g - HY(t^*)Y^{-1}(t_*)x_0$ is a vector of size m .

Taking into account that we search for control in the form (3), we can rewrite (11) and (6) as

$$\sum_{t \in T_u} c_k u(t) \rightarrow \max, \quad (13)$$

$$\sum_{t \in T_u} d_k u(t) = g_0, \quad (14)$$

respectively, where

$$d_k = \int_t^{t+h} HY(t^*)Y^{-1}(\tau)b(\tau)d\tau, \quad c_k = \int_t^{t+h} z_c^T(\tau)b(\tau)d\tau,$$

for $t = t_* + kh$, $k = \overline{0, N-1}$.

Let's denote

$$U = (u_0, u_1, \dots, u_{kh}, \dots, u_{(N-1)h})^T,$$

$$C = (c_0, c_1, \dots, c_{N-1}), \quad D = (d_0, d_1, \dots, d_{N-1}).$$

Finally the problem (13), (14) can be represented in the form

$$\begin{cases} C^T U \rightarrow \max, \\ DU \leq g_0, \\ |u_k| \leq L, \quad k = \overline{1, N}. \end{cases} \quad (15)$$

We give a short outline of the second stage following [4]. It's important since the package of applied programs, that implement the construction of optimal control, was written on its base.

Often it is impossible to get exact results of measuring on practice, in which case it's unnecessary to find the exact solution. The adaptive method can find approximate (suboptimal) solutions [1, 4].

Let's show its general scheme by the example of a general ILPP

$$\varphi(x) = c^T x \rightarrow \max, \quad b_* \leq Ax \leq b^*, \quad d_* \leq x \leq d^*. \quad (16)$$

We need to find an ε -optimal solution x^ε , $\varepsilon \geq 0$, using an a priori known information about the solution of the problem (16). To find the solution we construct a sequence of approximations $\{x^i\}$, $i = 1, 2, \dots$, which are plans of the problem (16). A basic instrument of the adaptive method are *supports*. Thereby a sequence of approximations and a sequence of supports $\{K_{op}^i\}$, $i = 1, 2, \dots$ are constructed. At every iteration of the method a new pair is constructed

$$\{x^k, K_{op}^k\} \rightarrow \{x^{k+1}, K_{op}^{k+1}\}.$$

The method consists of two phases. At the first phase on base of the known information about the problem (16) solution a first plan $\{x^1, K_{op}^1\}$ is constructed. The construction of following approximations is the second phase of the adaptive method.

In the adaptive method at every step it's possible to find the target value deviation, named suboptimality estimation and denoted $\beta(x^k, K_{op}^k)$. The method is based on the principle of suboptimality reduction:

$$\beta(x^{k+1}, K_{op}^{k+1}) \leq \beta(x^k, K_{op}^k).$$

The estimation of suboptimality allows the representation

$$\beta(x^k, K_{op}^k) = \mu(x^k) + \mu(K_{op}^k),$$

so two separate procedures can be done:

- 1) plan update $x^k \rightarrow x^{k+1}$, and
- 2) support update $K_{op}^k \rightarrow K_{op}^{k+1}$.

It is proven [4] that the adaptive method of finite number of steps converge to optimal solution.

The method was implemented in MATLAB environment and is suitable for solving various optimal control problems.

Note that the developed algorithm has some drawbacks. This method can converge slow enough if dimension of the system will be high. Also in the last iteration the rate of convergence reduced, especially for small values of ε .

3. EXAMPLE

As an example we consider the problem of fixed time stabilization of a three-mass oscillatory system with minimal fuel consumption (Fig. 1).

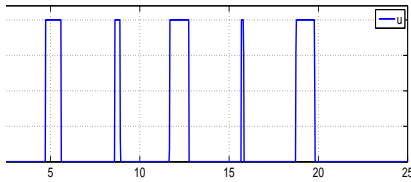


Figure 2: Optimal control

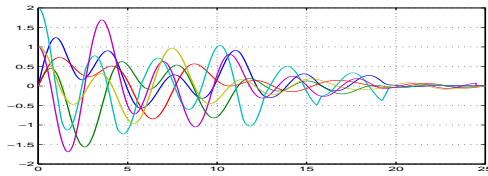


Figure 3: State variables. Transient

A mathematical model of the problem is

$$\begin{cases} \dot{x}_1 = x_4, \\ \dot{x}_2 = x_5, \\ \dot{x}_3 = x_6, \\ \dot{x}_4 = -2.33x_1 - x_2 + 0.33x_3 + u, \\ \dot{x}_5 = -0.88x_1 - 2.06x_2 - 0.88x_3, \\ \dot{x}_6 = 0.25x_1 - 0.75x_2 - 1.75x_3, \end{cases}$$

$$x_1(0) = x_2(0) = x_3(0) = 0, \quad x_4(0) = 2, \quad x_5(0) = 1, \quad x_6(0) = 1, \\ x_1(25) = x_2(25) = x_3(25) = x_4(25) = x_5(25) = x_6(25) = 0,$$

$$\int_0^{25} u(t)dt \rightarrow \min, \quad 0 \leq u(t) \leq 1, \quad t \in [0, 25],$$

where $x_1 = x_1(t)$, $x_2 = x_2(t)$, $x_3 = x_3(t)$ are deviations from the equilibrium position of the first, second and third masses, respectively, and $u = u(t)$ is the control parameter.

This optimal control problem with the quantization period $h = 0.025$ is equivalent to a linear programming problem of dimension 6×1000 . The adaptive method was used to construct the optimal control. The problem is solved in 141 iterations (Figures 2, 3).

4. CONCLUSION

A learning software system on optimal control problems is described. The stages of the complex are presented, the software package implementing the method is written. It provides the tools for visualization of the the adaptive Gabasov method work, and for optimal control construction for arbitrary mathematical models.

The software package is used and developed at a special seminar of the Department of Mathematical Modelling of Economical Systems, Faculty of Applied Mathematics and Control Processes, Saint-Petersburg State University.

In the next stage to expand the methodological basis of the program complex we are planning to use the results and examples of works [5–13].

5. REFERENCES

- [1] R. Gabasov and F. M. Kirillova. *Constructive Optimization Methods. P. 2. Control Problems*. Minsk, USSR, 1977. (in Russian)
- [2] N. V. Balashevich, R. Gabasov, and F. M. Kirillova. Numerical methods of programmed and positional optimization for linear systems. *Journal of Computational Mathematics and Mathematical Physics*, 40(6):838–859, 2000. (in Russian)
- [3] R. Gabasov, F. M. Kirillova, and E. A. Rugitskaya. Dampening and stabilization of equilibrium with great initial disturbance. *Intelligence of Academy of Science. Theory and Systems of Control*, No 1: 29–38, 2001. (in Russian)
- [4] V. V. Al'sevich, R. Gabasov and V.S. Glushenkov. *Optimization of Linear Economic Models. Static Problems*. Belarus State University Press, Minsk, 2000. (in Russian)
- [5] I. Solovyeva. Positional optimization in a certain problem of multiprogrammed control. In *Proceedings of the 11-th International Conference on Humans and Computers*, pages 359–393, University of Aizu Press, Japan, November 2008.
- [6] N. V. Smirnov and I. V. Solovyeva. Application of a closed-loop optimization method for multiprogram stabilization of bilinear systems. *Vestnik of St. Petersburg University. Ser. 10: Applied Mathematics. Informatics. Control Processes*, (3):251–259, September 2009 (in Russian).
- [7] A. Zgonnikov. The optimum feedback synthesis for a certain nonlinear mechanical system. In *Proceedings of the 12-th International Conference on Humans and Computers*, pages 235–239, University of Aizu Press, Japan, December 2009.
- [8] M. Smirnov. Suppression of Bounded Exogenous Disturbances Act on a Sea-going Ship. In *Proceedings of the 13-th International Conference on Humans and Computers*, pages 114–116, University of Aizu Press, Japan, December 2010.
- [9] M. Fedorova. Computer Modeling of the Astatic Stabilization System of Sea-going Ship Course. In *Proceedings of the 13-th International Conference on Humans and Computers*, pages 117–120, University of Aizu Press, Japan, December 2010.
- [10] N. Smirnov. Multiprogram Control for Dynamic Systems: a Point of View. In *Proceedings of the Joint International Conference on Human-Centered Computer Environments (Aizu-Wakamatsu, Japan, Duesseldorf, Germany, HCCE 2012)*, pages 106–113, ACM, New York, NY, March 2012. DOI=10.1145/2160749.2160773.
- [11] N. V. Smirnov, T. E. Smirnova, Multiprogram digital control for bilinear object. *2014 International Conference on Computer Technologies in Physical and Engineering Applications (ICCTPEA)*. Editor: E.I. Veremey, pages 168–169, St. Petersburg State University, Russia, June–July 2014.
- [12] N. V. Smirnov, T. E. Smirnova and Ya. A. Shakhov. Stabilization of a given set of equilibrium states of nonlinear systems. *Journal of Computer and Systems Sciences International*, 51(2):169–175, April 2012.
- [13] N. V. Smirnov and T. Ye. Smirnova. The synthesis of multi-programme controls in bilinear systems. *Journal of Applied Mathematics and Mechanics*, 64(6):891–894, December 2000.

Study of Patterns in the Hyperlink Structure of Large Sites

Ivan Blekanov

Saint-Petersburg State University
7-9 Universitetskaya Naberezhnaya,
St. Petersburg, 199034, Russian
+7 921 339 53 43

i.blekanov@gmail.com

Sergei Sergeev

Saint-Petersburg State University
7-9 Universitetskaya Naberezhnaya,
St. Petersburg, 199034, Russian
+7 921 381 23 62

slsergeev@yandex.ru

Evgenii Klemeshov

Saint-Petersburg State University
7-9 Universitetskaya Naberezhnaya,
St. Petersburg, 199034, Russian
+7 961 808 63 52

zhklem14@gmail.com

ABSTRACT

This paper presents experimental results on websites of four universities investigated with the aim of drawing up lists of pages and links. Variants of functions approximating the experimental data are considered. The resulting approximation estimates indicate that the approximating functions proposed are of high precision.

Categories and Subject Descriptors

G.2.2 [Discrete Mathematics]: Graph Theory – *graph algorithms*.

General Terms

Measurement, Experimentation.

Keywords

Web-site, approximation, hyperlinked structure of the site, webometrics, web-graph.

1. INTRODUCTION

Today, you can hardly come across an organization that does not have its own website. In a sense, a website is the face of an organization and the quality of the site is of great importance. The quality of a website consists of many components – page count, page design, page content, and site structure.

A website structure is usually presented in the form of a directed graph whose nodes are documents and whose edges are links connecting these documents [1]. Many researchers today study the global structural characteristics of information networks on the Web. For example, Broder A. and Kumar F. [1] represented large website communities in the form of a strongly connected graph component, components In, Out and Tubes; [2]-[5] deal with distribution of external links and citation index of various university sites; authors in [6] introduce site connectivity characteristics; [7] is dedicated to the study of the method of automatic classification of links and

pages by their characteristics; and others.

This paper aims at studying a website structure. More precisely, it tries to identify the kind of functional dependence that exists between the page count of a site and the internal link count of that site.

2. HYPOTHESIS VERIFICATION

2.1 Problem statement

Special crawler [8] was used to solve this problem. In scanning the site, the robot creates two lists: a list of found pages (web graph nodes) and a list of links connecting the found pages (web graph edges).

An act of finding e links will be considered as a step of the scanning algorithm. That is, $E_i = e \cdot i$ links are found after i steps. Let us denote with $v_i = v(E_i)$ the number of pages found after i steps. We denote the number of all the links and the number of all pages found in the site with e_0 and $v(e_0) = v_0$ respectively. The search robot can find not only the number of pages and links in the site, but also get a graph of the $v(e)$ function.

Obviously, $v \leq e$. Here, $v_0 < e_0$ (it is very rare when $v_0 = e_0$ for websites). It is also obvious that increase in $v(e)$ is gradually slowed down. This is due to the fact that with increasing number of found pages, the likelihood that the next link will point to a page already found increases.

Let us try and choose an analytic function approximately describing experimental function $v(E)$, with the following properties:

1. Passes through the origin in the plane (e, v) .
2. Increases with an increase in e_0 .
3. Derivative of the function decreases.
4. The function has a simple form and is dependent on a small number of parameters.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.

Copyright 2015 University of Aizu Press.

2.2 Experiment

Let's compare the function with experimental set $V_i, E_i (i = \overline{1, N})$, where $N = \left\lceil \frac{e_0}{e} \right\rceil$. We investigated the websites of four universities and obtained the following sets (Table 1).

To assess the approximation quality, we will use average relative error:

$$\Delta = \frac{1}{N} \sum_{i=1}^N \frac{|V_i - v_i|}{v_i}.$$

Table 1. Universities sites studied

University	URL	Total pages	Total links
Saint-Petersburg State University	www.spbu.ru	41183	2664000
Moscow State University	www.msu.ru	47832	1891000
The University of Aizu	www.u-aizu.ac.jp	4161	49900
The University of Tokyo	www.u-tokyo.ac.jp	> 17000	240000

Next, we consider three possible approximating functions:

$$v^{(1)} = \alpha \cdot E^\beta,$$

$$v^{(2)} = \frac{\alpha \cdot E}{E + \beta},$$

$$v^{(3)} = \alpha \cdot (\ln(1 + E))^\beta.$$

Let us examine them one by one:

- 1) We take the logarithm of the equation:

$$\ln v^{(1)} = \ln \alpha + \beta \cdot \ln E.$$

This gives a system of linear equations

$$x + A_i \cdot y = B_i, \quad i = \overline{1, N},$$

where $x = \ln \alpha$, $y = \beta$, $A_i = \ln E_i$, $B_i = \ln V_i$.

Its solution with method of least squares:

$$\alpha = \exp \frac{C_1 C_4 - C_2 C_3}{N \cdot (C_1^2 - C_3)}, \quad \beta = \frac{C_1 C_2 - C_4}{C_1^2 - C_3},$$

where

$$C_1 = \sum_{i=1}^N A_i, C_2 = \sum_{i=1}^N B_i, C_3 = \sum_{i=1}^N A_i^2, C_4 = \sum_{i=1}^N A_i B_i.$$

- 2) System of linear equations:

$$\alpha \cdot E_i - \beta \cdot V_i = V_i \cdot E_i, \quad i = \overline{1, N}.$$

Its solution with method of least squares:

$$\alpha = \frac{a_3 \cdot a_4 - a_2 \cdot a_5}{a_1 \cdot a_4 - a_2^2}, \quad \beta = \frac{a_2 \cdot a_3 - a_1 \cdot a_5}{a_1 \cdot a_4 - a_2^2}$$

Where

$$a_1 = \sum_{i=1}^N E_i^2, \quad a_2 = \sum_{i=1}^N E_i V_i, \quad a_3 = \sum_{i=1}^N E_i^2 V_i, \\ a_4 = \sum_{i=1}^N V_i^2, \quad a_5 = \sum_{i=1}^N V_i^2 E_i.$$

- 3) We take the logarithm of $v^{(3)}$:

$$\ln v^{(3)} = \ln \alpha + \beta \cdot \ln \ln(1 + E)$$

We get a system that almost coincides with the first case. Its solution will be the same, except that in the first case $A_i = \ln E_i$, while in this case,

$$A_i = \ln \ln(1 + E_i).$$

2.3 Results of the experiment

Figures 1, 2, 3 and 4 present the graphs of functions $v^{(1)}, v^{(2)}, v^{(3)}$ and V , for each of the universities listed.

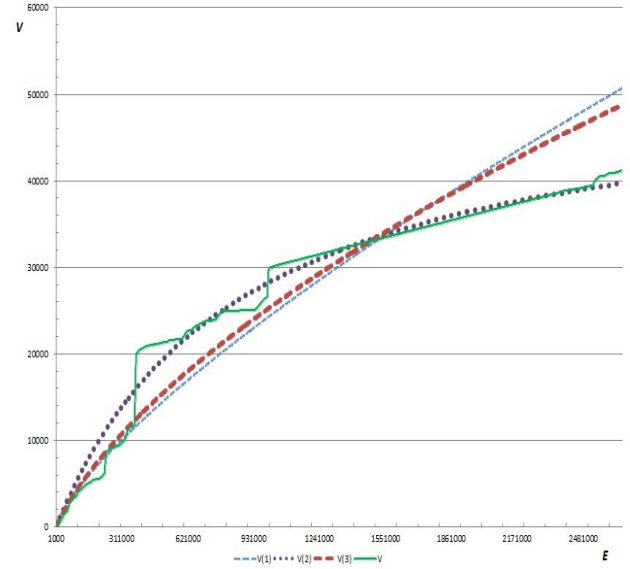


Figure 1. Functions $v^{(1)}, v^{(2)}, v^{(3)}$ and V for Saint-Petersburg State University

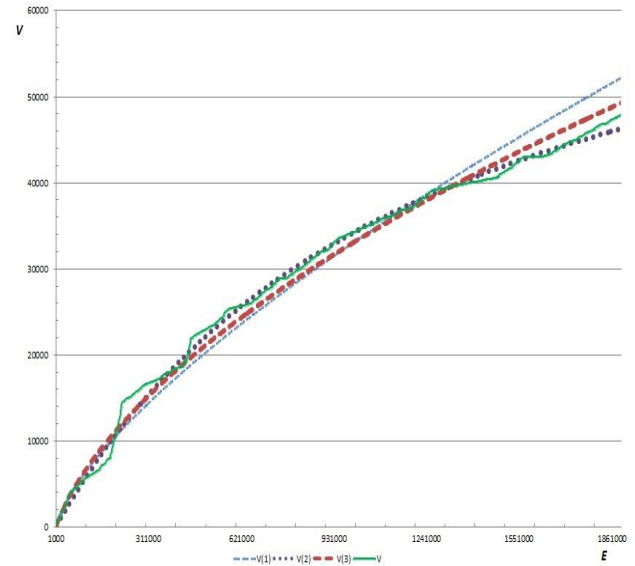


Figure 2. Functions $v^{(1)}, v^{(2)}, v^{(3)}$ and V for Moscow State University

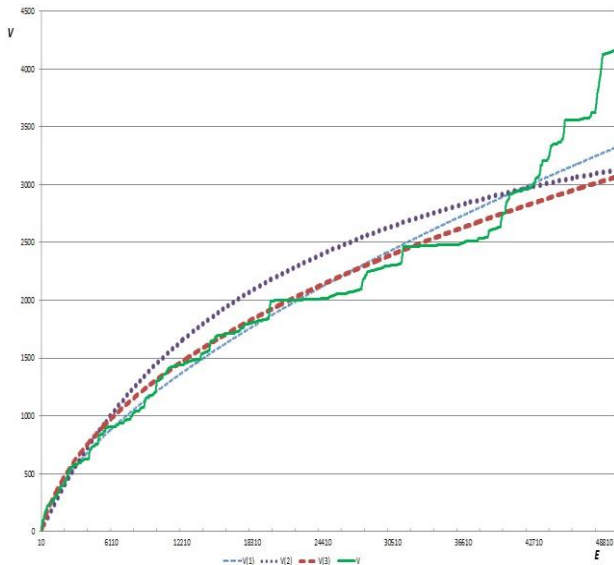


Figure 3. Functions $v^{(1)}, v^{(2)}, v^{(3)}$ and V for the University of Aizu

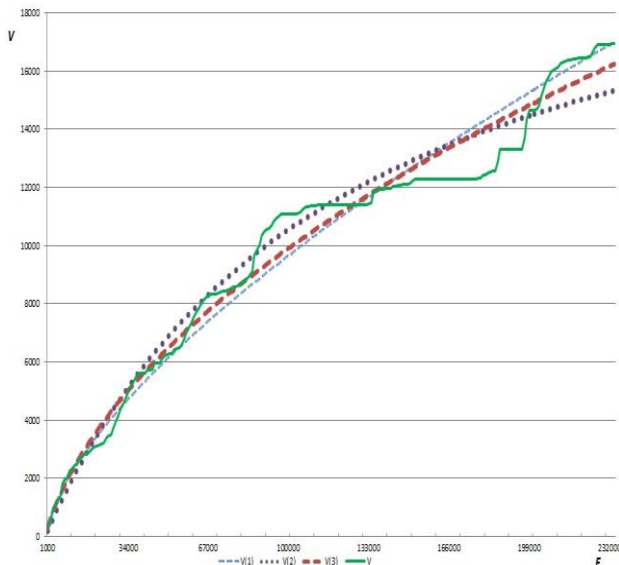


Figure 4. Functions $v^{(1)}, v^{(2)}, v^{(3)}$ and V for the University of Tokyo

Table 2 shows the relative errors of each of the formulas for each of the universities.

Table 2. Average relative error of functions $v^{(1)}, v^{(2)}, v^{(3)}$ for each university

University	$v^{(1)}$	$v^{(2)}$	$v^{(3)}$
Saint-Petersburg State University	0,166	0,141	0,099

Moscow State University	0,092	0,037	0,014
The University of Aizu	0.167	0.209	0.229
The University of Tokyo	0.068	0.076	0.062

3. CONCLUSION

The paper considered the problem of finding a function that approximates the experimental graph of dependence of web page count on link count. Three approximations were proposed. It was revealed that the best approximations are obtained for power and linear fractional function. The functions proposed can be used to study the parameters of sites and their clustering.

4. ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, grant № 15-01-06105

5. REFERENCES

- [1] Broder, A., R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins and J. Wiener, 2000. Graph structure in the Web: Experiments and models. In WWW9 (Vol. 33, #1–6), Elsevier Science, pp: 309–320.
- [2] Thelwall, M., 2013. Webometrics and Social Web Research Methods. University of Wolverhampton, pp: 8–39.
- [3] Thelwall, M., Zuccala, A. 2008. A university-centred European Union link analysis. *Scientometrics*, 75(3), 407–420.
- [4] Thelwall, M., Wilkinson, D., Musgrove, P. B., 2005. National and international university departmental web site interlinking, part 2: Link patterns. *Scientometrics*, 64(2): 187–208
- [5] Pechnikov, A. and A. Nwohiri, 2012. Webometric analysis of Nigerian university websites. *Webology*. Vol. 9, Num. 1, June. (<http://www.webology.org/2012/v9n1/a95.html>).
- [6] Blekanov, I.S., S.L. Sergeev and A.I. Maksimov, 2014. Analysis of the topology of large Web segments using Broder's bow-tie model. *Life Science Journal*, Vol. 11: 258–261.
- [7] Kenekayoro, P., K. Buckley, M. Thelwall, 2014. Automatic classification of academic web page types. *Journal Scientometrics*, Vol. 101(2): 1015–1026.
- [8] Blekanov I., S. Sergeev and I. Martynenko, 2012. Construction of subject-oriented web crawlers using a generalized kernel. *Scientific and technical bulletins of St. Petersburg State Polytechnic University*. St. Petersburg State Polytechnic University, # 5 (157). pp: 9–15.

Hierarchical clustering of large text datasets using Locality-Sensitive Hashing

Vasilii Korelin

Saint-Petersburg State University
7-9 Universitetskaya Naberezhnaya,
St. Petersburg, 199034, Russia
+7 911 266 12 99
vn.korelin@gmail.com

Ivan Blekanov

Saint-Petersburg State University
7-9 Universitetskaya Naberezhnaya,
St. Petersburg, 199034, Russia
+7 921 339 53 43
i.blekanov@gmail.com

ABSTRACT

In this paper, we present a hierarchical clustering algorithm of the large text datasets using Locality-Sensitive Hashing (LSH). The main idea of the LSH is to “hash” items several times, in such a way that similar items are more likely to be hashed to the same bucket than dissimilar are. The main drawback of the conventional hierarchical algorithms is a large time complexity (e.g. Single Linkage method has time complexity of $O(n^2)$). Proposed algorithm reduces the time complexity to $O(Pn)$. Here, P represents the maximum number of items going to the single bucket. P is a small constant as compared to n for the large number of buckets.

Clustering results of the hierarchical clustering algorithm, that uses LSH, are similar to the clustering results of the classical single linkage method. The main advantage of the hierarchical clustering algorithm, that uses LSH, is a significant increase in speed for large datasets clustering in comparison with classical algorithms.

Categories and Subject Descriptors

H.2.8 [Database Applications]: Data mining; H.3.3 [Information Search and Retrieval]: Clustering.

General Terms

Algorithms, Performance, Experimentation, Languages.

Keywords

Hierarchical clustering, Locality-Sensitive Hashing, Minhashing, Shingling.

1. INTRODUCTION

For today clustering of the large text datasets (e.g. clustering of the web pages) is one of the urgent data mining issues. Conventional clustering algorithms allow creating clusters with some accuracy, F-measure and etc. but when it comes to the clustering of the larger datasets (high-resolution pictures, fingerprints or web pages) the vast majority of algorithms have poor speed performance [1]. For example, despite the fact that

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.

Copyright 2015 University of Aizu Press.

Single Linkage algorithm allows detecting clusters in arbitrary shapes, it has a large time complexity of $O(n^2)$ where n is number of objects. Such time complexity is inappropriate for big data clustering. To avoid the dimensional issue new clustering algorithm that uses LSH method was proposed. LSH reduces the dimensionality of high-dimensional data. It hashes input items so that similar items map to the same “buckets” with high probability (the number of buckets being much smaller than the universe of possible input items) [2].

2. DIMENSIONAL REDUCTION FOR MINING MASSIVE DATASETS

There are many ways to represent documents as sets for the purpose of identifying lexically similar documents. The most effective way is to construct from the document the set of short strings that appear within it (shingles). Further, for document comparison Minhashing method is used that allows comparing multidimensional sets. These conversions must be performed before applying fast hierarchical algorithm with LSH using. Below are descriptions of Shingling, Minhashing, LSH methods and new proposed algorithm. Figure 1 presents the process of finding similar items in large documents collection for every document.

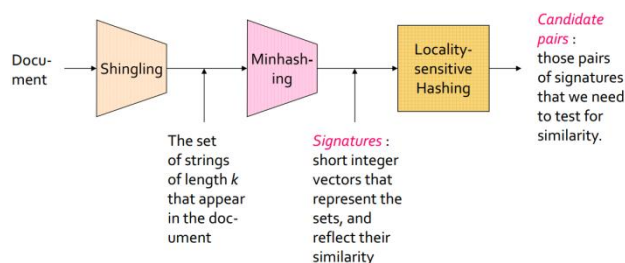


Figure 1. Dimension reduction of documents datasets

2.1 Shingling

A document represents a string of characters. Shingling is a process that creates sets of k -shingles. Define a k -shingle for a document to be any substring of length k found within the document [3].

For example, $k=2$, $doc="a b c a b"$. Then after shingling next sets can be obtained: $\{a, b\}$, $\{b, c\}$, $\{c, a\}$.

Similar documents to each other will have a lot of equal shingles.

2.2 Jaccard Similarity

In this paper to determine the similarity of two sets Jaccard similarity is used. The Jaccard similarity of sets S and T is the ratio of the size of the intersection of S and T to the size of their union [4].

$$Sim(C_1, C_2) = \frac{|C_1 \cap C_2|}{|C_1 \cup C_2|}$$

Figure 2 depicts two sets C_1 and C_2 . There are 3 elements in their intersection and a total of eight elements that appear in S or T or both.

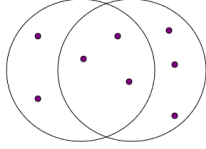


Figure 2. Two sets with Jaccard similarity 3/8

Jaccard similarity of two sets equals:

$$Sim(C_1, C_2) = \frac{3}{8}$$

Every document can be represented as set of k -shingles. Therefore the sets of documents can be represented as quite sparse boolean matrix. Rows are elements of the universal set (e.g. the set of all k -shingles). Columns correspond to the documents. 1 in row e and column S if and only if e is a member of S . Column similarity is the Jaccard similarity of the sets of their rows with 1. Figure 3 depicts the boolean matrix that represents two documents.

	C_1	C_2
a	1	1
b	1	0
c	0	1
d	0	0

Figure 3. Boolean matrix representing two documents

$$Sim(C_1, C_2) = \frac{1}{3} - \text{similarity of two documents.}$$

2.3 Minhashing

Typical for the large number of documents boolean matrix is quite sparse. Therefore its further processing will be very time consuming. To solve this problem boolean matrix can be compressed to the signature matrix M in such a way that we can still deduce the similarity of the underlying sets from their compressed versions. This technique is called "Minhashing".

To minhash a set represented by a column of the characteristic matrix, a permutation of the rows should be picked. The minhash function $h(c)$ of any column is the number of the first row, in the permuted order, in which the column has a 1 [5].

Therefore the number of random permutations determines the number of the minhash functions. For example we can use 100 random permutations to create 100 signatures for every column of the boolean matrix.

Signatures can be stores in the special signature matrix M . The rows of the signature matrix are minhash values and columns

correspond to the documents. Figure 4 presents the example of the signature matrix M creation by using 3 minhash functions (3 random permutations).

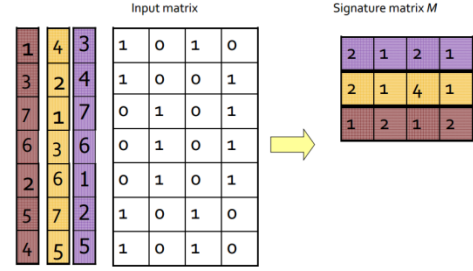


Figure 4. Minhashing

The probability (over all permutations of the rows) that $h(C_1) = h(C_2)$ is the same as $Sim(C_1, C_2)$. The similarity of signatures is the fraction of the minhash functions in which they agree. Thus, the expected similarity of two signatures equals the Jaccard similarity of the columns or sets that the signatures represent. And the longer the signatures, the smaller will be the expected error [5].

This important feature allows compressing large sparse boolean matrixes to the signature matrixes with short defined number of rows with preserving similarity between rows. Thus, every document can be represented as a vector and its number of elements equals the number of minhash functions.

In spite of boolean matrix is compressed it still may be impossible to find the pairs with greatest similarity efficiently. The reason is that the number of pairs of documents may be too large, even if there are not too many documents.

2.4 Locality-Sensitive Hashing

General idea: Generate from the collection of all elements (signatures in our example) a small list of candidate pairs: pairs of elements whose similarity must be evaluated. For example, for signature matrix every column should be hashed several times and columns with equal hash values should be placed to the same bucket. Candidate pairs are those that hash at least once to the same bucket. To compare similarity of two sets threshold t should be picked ($t < 1$). A pair of documents is considered to be similar only if their signatures agree in at least fraction t of the rows [6].

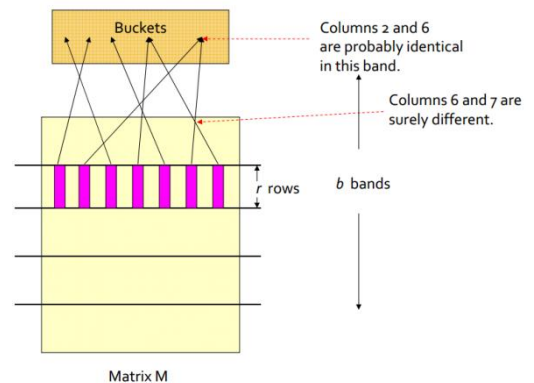


Figure 5. Hash functions for one band

An effective way to choose the hashings is to divide the signature matrix M into b bands consisting of r rows each (Figure 5). For each band, hash function takes vectors of r integers (the portion of one column within that band) and hashes them to some large number of buckets. Same hash function can be used for all the bands, but for each band there should be a separate bucket array, so columns with the same vector in different bands will not hash to the same bucket. Those columns that at least once were hashed to the same bucket are considered as candidate pairs. To catch most similar pairs, but few non similar pairs, b and r should be tuned attentively.

3. HIERARCHICAL CLUSTERING USING LSH

Proposed algorithm that exploits the hash tables generated by LSH. This algorithm outputs clustering results that approximate those obtained by the single linkage method [7]. The following is a detailed description of the algorithm.

Preconditions:

1. $t < 1$ – threshold that determines the Jaccard similarity of 2 documents.
2. r – the initial value of the rows in each band. It should depend on number of signatures in signature matrix.
3. r_{min} – the minimum value of the rows of signatures in each band.
4. Δ - parameter is used for r reduction.
5. Each document is view as single cluster.

Steps:

Step 1: For each band, hash vectors of r integers to the buckets. In the i -th hash table, column d (document) is stored in the bucket with index $h_i(d)$. However, if another column belonging to the same cluster as d has already been saved in the very bucket, d is not stored in it.

Step 2: For each column d , from the set of columns that enter the same bucket as d in at least one hash table, find columns whose distances from d are less than t .

Step 3: The pairs of clusters, each of which corresponds to a pair of columns obtained in Step 2, are connected (Figure 6).

Step 4: If $r \leq r_{min}$, algorithm terminates. Otherwise, advance to Step 5.

Step 5: $r = r - \Delta$. Advance to Step 1.

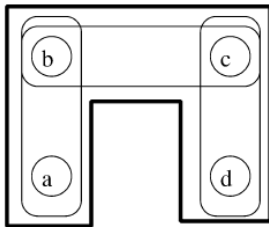


Figure 6. Merging of clusters

4. CLUSTERING EVALUATION

Reuters 21578, test collection of documents, was used to evaluate quality and performance of hierarchical clustering algorithm with LSH. New algorithm was compared with classical Single Linkage method. Table 1 presents that quality of clustering (accuracy and F-measure) for 1000 and 10000 number of documents.

Table 1. Comparison of the clustering quality

Algorithm	Documents count	Accuracy	F-measure
Single-Link	1000	75%	68%
	10000	79%	71%
Single-Link + LSH	1000	72%	73%
	10000	72%	74%

Clustering results show that accuracy and F-measure of two algorithms are similar.

Figure 7 presents dependence of the execution time of the number of input documents.

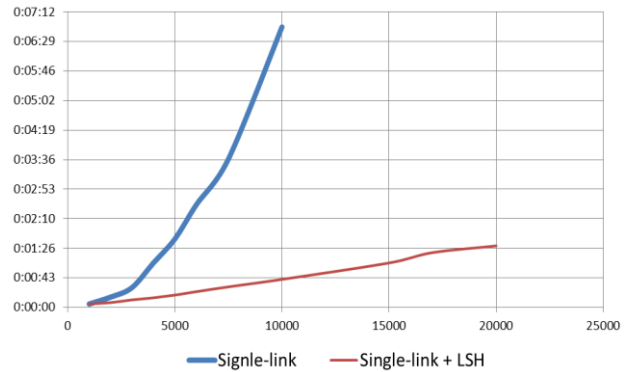


Figure 7. Comparison of the execution time

Analyzing the graphs, one can conclude that the clustering algorithm using LSH is much faster than the algorithm without LSH. Execution time of the algorithm with LSH increases linearly with an increasing number of documents.

5. CONCLUSIONS

In this paper we have proposed fast hierarchical clustering algorithm that uses LSH algorithm (LSH used for reducing the dimensionality of high-dimensional data). Developed algorithm is optimally suited for the massive text datasets clustering. Proposed algorithm was tested on Reuters collection of documents and showed reasonable accuracy and F-measure in comparison with classical clustering algorithms but in speed new algorithm is much superior to them. Fast developed clustering algorithm that uses LSH can be modified and used for clustering and analysis in such areas as medicine, criminalistics, sociology, etc. [2].

6. ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, grant № 15-01-06105.

7. REFERENCES

- [1] A. Ene, S. Im, and B. Moseley. Fast clustering using MapReduce. In KDD, pages 681–689, 2011.
- [2] J. Buhler. Efficient large-scale sequence comparison by locality-sensitive hashing. *Bioinformatics*, 17(5):419–428, 2001.
- [3] Broder, A.Z. Identifying and Filtering Near-Duplicate Documents. In proceedings of the 11th Annual Symposium on Combinatorial Pattern Matching, pp. 1-10, 2000.
- [4] Jatsada Singthongchai and Suphakit Niwattanakul, "A Method for Measuring Keywords Similarity by Applying Jaccard's, N-Gram and Vector Space," Lecture Notes on Information Theory, Vol.1, No.4, pp. 159-164, Dec. 2013. doi: 10.12720/Init.1.4.159-164.
- [5] CHUM, O., PERDOCH, M., AND MATAS, J. 2009. Geometric minhashing: Finding a (thick) needle in a haystack. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 17–24
- [6] A. Gionis, P. Indyk, and R. Motwani. Similarity search in high dimensions via hashing. In VLDB, 1999.
- [7] Koga, Hisashi, Tetsuo Ishibashi, Toshinori Watanabe. 2007. Fast agglomerative hierarchical clustering algorithm using Locality-Sensitive Hashing. *Knowledge and Information Systems* 12.1, 2007.

Developing a Mobile Application for Wine Amateurs

Anton Kiselev
Peter the Great St.Petersburg Polytechnic
University
29 Polytechnicheskaya st.
St. Petersburg 195251 Russia
anton.kiselev.94@inbox.ru

Andrey Kuznetsov
Motorola Solutions, Inc.
St. Petersburg Software Center
12 Sedova st
St. Petersburg 192019 Russia
andrei.kuznetsov@motorolasolutions.com

ABSTRACT

In this paper, we describe a motivation to develop an Android mobile application for a wine lover. We introduce a simple map of concepts used for analysis and modeling of the subject domain. We also briefly describe the application architecture and the interfaces developed to support major use cases that a wine amateur may require in order to store and manage information about wines.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous

General Terms

Design, Human Factors

Keywords

Software, Semantic network, Domain ontology, Mobile applications, User interface, Android

1. INTRODUCTION

A wine is complex phenomenon related to the aspects linked by many different domains: biology, oenology, history, cooking, medicine, economics, etc. Professional sommeliers aren't the only ones who like to learn about wines, to discover their special characteristics and to organize their knowledge and wine tasting experience. There are wine amateurs interested not only in tasting wines but also in discovering wine inspired knowledge.

Currently numerous wine information and wine cellar management applications are available ranging from a simple wine notebook to data organizers going to the extreme depth of details. There are solutions for professional experts, educators, sellers and customers as well as many for amateur wine lovers, however "this is a market that's still emerging and in flux" [8]. Existing applications provide features

which allow storing information about wine properties, accessing numerous wine rankings, or organizing user personalized tasting notes and evaluations.

If one carefully thinks of what is the information that describes a wine, he or she could be surprised to discover so many attributes and classification properties that one has to take into consideration. That's why before collecting requirements for software that would help to manage most of related information (and to add some novelty compare to existing applications), one has to study the subject domain quite attentively. One way of such a study is to try to define a map of concepts which could serve as a foundation for developing use cases and for designing an application.

The next step is to study, in what extent a *mobile* application would fit such requirements better in order to leverage existing sommeliers' experience and to provide rationale for including new features which seem to be hardly implementable without using a mobile platform.

Among others we could cite the following features that customers expect to have in such applications:

- Ability to search for wine information via identifiers like barcodes or labels;
- Ability to link to the geographic, encyclopedic and web resources providing more detailed information about the wines;
- Ability to sort wine directories and creating own wine collections;
- Ability to view and edit wine charts;
- Ability to support advises and tools assisting wine tasting (specialized note editors, ranking schema, access to database);
- Integration with social networks to enable users to share their assessments and reviews.

In this work we describe an Android based software prototype featuring some of the above mentioned requirements partially.

2. RELATED WORK

We discovered few works related to the wine business and culture focusing aspect of wine informatics and possible applications of information technology to benefits of better wine information flow.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWA/IT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

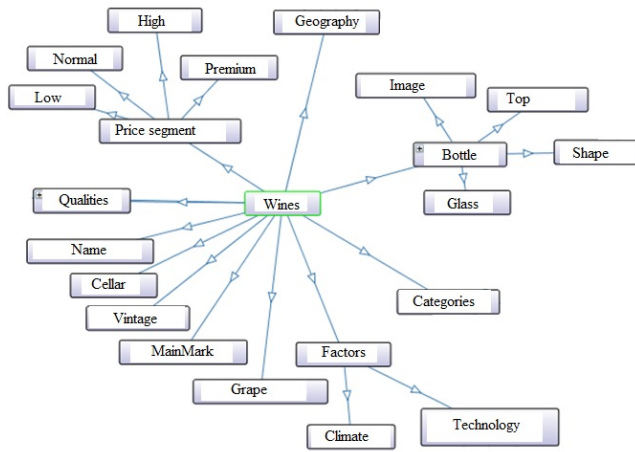


Figure 1: Wine: Graph of major concepts

Thus, in [5] the authors used the wine experience metaphor as an illustration of the process of developing knowledge representation, retrieval and distribution. They use a task of defining a wine ontology as a “test ground” of theoretical ontology engineering models. Despite using the wine related story as an example, the authors elaborated a kind of the very carefully thought ontology representing most of important entities and relations needed to represent and distribute wine related knowledge (at bottom, ontologies are being developed to enable knowledge sharing and reuse [7]).

An attempt of defining a fuzzy ontology to represent obvious and unusual links like computing the wine score as a product of scores reflecting whether the wine goes well with game or with pork or suits a dinner with friends or even one with wine experts shows well what are the challenges in creating relevant “wineinformatics” system in such a really uncertain domain.

The term “wineinformatics” was introduced in the work [6] where the authors described how techniques of data mining and machine learning might be helpful in researching large number of existing wine reviews and data sheets in systematic way to benefits of wine makers, distributors and consumers.

We found also a couple of wine ontologies where the wine domain is used again like an example. We could mention the example of using the OWL web ontology language [4] which in turn was derived from the DAML Wine ontology [3] with substantial changes, with particular respect to the region based relations.

3. DEVELOPING THE APPLICATION

Major entities related to the wine subject domain are shown in Figure 1 in the form of a map of concepts. In fact, this is also the step towards wine ontology definition. However the discussion on possible improvements of existing wine ontologies is out the scope of this paper.

The map of concepts shown in Figure 1 is used to define the entities, their attributes and relationships, and over all, as a foundational concept model for our application. Every entity is linked to another entity. For example, entity ‘Wine’ contains the next attributes: name, vintage, grap, bottle, e t.c. So the communication between concepts allow to define

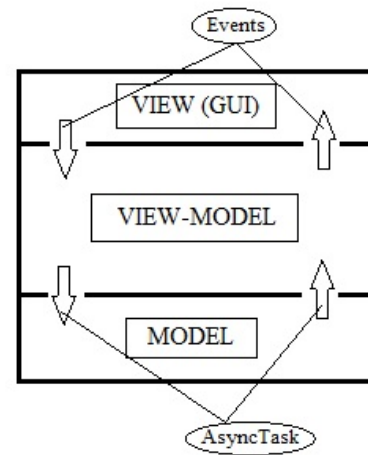


Figure 2: Application architecture

the relationship between entity and attribute. Wine testing holds on wine qualities, and catalogues may be sorted by major characteristics.

In this work we discuss the special class of software – mobile applications. From the viewpoint of customization the mobile application is targeted to the concrete user (unlike web sites) and may use features of the whole device (e.g. device camera, multimedia facilities, touch screen, etc). Application architecture is designed by the well known *MVVM* (Model-View-Viewmodel) template [2].

The architecture of the developed application is shown in Figure 2.

The application concept model forms the *Model*, while the *View-model*, constructed over the model, serves all GUI related application logic, handles the *View* events and includes the *Model*’s links. The *View* forms the user interface prototypes with the help of XML descriptions.

Let us mention that it is important to use a separate thread for data model conversion. For this purpose the framework provides an *AsyncTask* class enabling the use of the UI thread in such a way that background operations are performed in a separate thread, and their results are published on the UI thread without having to manipulate threads and/or handlers [1].

Some examples of the application interface design are presented in Figure 3. These examples show a selection of scenarios supported by the current version including creating wine collections, attaching photos taken by using the device camera, allowing links to geographic maps, and some facilities of diagramming. To reduce the number of created windows we used the library component *PagerSlidingTabStrips*.

The application deals with user wine collections as with catalogues. Each catalogue may be sorted by name, vintage, producer or by wine type.

Actually, we included two basic methods for wine evaluation, but there is space for further improvements.

One specific feature is an ability to have access to the geographical maps allowing users to add markers describing geographical origin of the wine. We also consider to support geographical information as it is mentioned in [9].

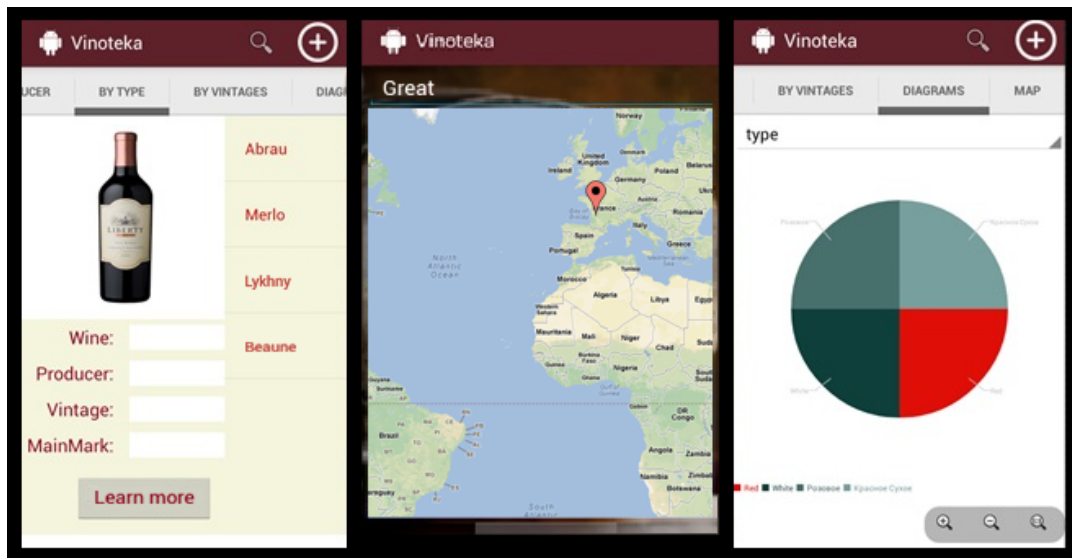


Figure 3: Application interface

4. CONCLUSION AND FUTURE WORK

In its current state the developed mobile application can be addressed to the wine lovers organizing their own local wine cellar information service with the following features:

- Add/Remove wine records and view wine information;
- Organizing wine catalogues and wine search by name;
- Access to the related web site(s);
- Add/show anchors to the geographic map;
- Create/view simple diagrams and graphs base on wine rankings;
- Using device photo camera to get images and to link them to the wine metadata;
- The developed prototype supports two tasting systems: 100 point scale and 20 point.

One of the possible future steps is to overcome limitations conditioned by the current architecture allowing to deploy only local one user application. Changing the architecture to the client-server model will allow us not only to leverage existing experience of one concrete user but to share experience, notes, rankings and other wine information by many users and to access common wine metadata from the server-side database.

With respect to the analysis of existing applications surveyed in [8], currently we couldn't say that our prototype exceeds the analogues but serves as a starting point for further development and for introducing new features such as sommelier training mode, the possibility to compare end-user's wine notes to those of wine experts and others. Let us mention that in order to support such features we have to redesign the application by using a client-server architecture.

5. REFERENCES

- [1] AsyncTask, <http://developer.android.com/reference/android/os/AsyncTask.html>. [Online; accessed Jul 1, 2015].
- [2] Mvvm, <https://ru.wikipedia.org/wiki/Model-View-ViewModel>. [Online; accessed Jul 10, 2015].
- [3] Ontolingua daml wine ontology, <http://ontolingua.stanford.edu/doc/chimaera/ontologies/wines.daml>. [Online; accessed Jul 7, 2015].
- [4] Owl wine ontology, www.w3.org/TR/owl-guide/wine.rdf. [Online; accessed Jul 7, 2015].
- [5] C. Carlsson, M. Brunelli, and J. Mezei. Fuzzy ontologies and knowledge mobilisation: Turning amateurs into wine connoisseurs. In *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on*, pages 1–7, July 2010.
- [6] B. Chen, C. Rhodes, A. Crawford, and L. Hambuchen. Wineinformatics: Applying data mining on wine sensory reviews processed by the computational wine wheel. In *Data Mining Workshop (ICDMW), 2014 IEEE International Conference on*, pages 142–149, Dec 2014.
- [7] T. R. Gruber. Toward principles for the design of ontologies used for knowledge sharing? *International journal of human-computer studies*, 43(5):907–928, 1995.
- [8] C. Null. Pocket vino: We review 5 apps for managing your wine cellar, March 2015. [Online: <http://www.techhive.com/article/2893433/pocket-vino-we-review-5-apps-for-managing-your-wine-cellar.html>; posted Mar 5, 2015; accessed Jul 10, 2015].
- [9] S. Shanmuganathan, P. Sallis, L. Pavesi, and M. Munoz. Computational intelligence and geo-informatics in viticulture. In *Modeling Simulation, 2008. AICMS 08. Second Asia International Conference on*, pages 480–485, May 2008.

Multicriteria Regulation of Investments in the Economy of the Russian Federation

Alexander Popkov
Saint-Peterburg State University
199034, Universitetskaya nab., 7-9
Saint-Petersburg, Russia
alexandr.popkoff@gmail.com

ABSTRACT

In the present paper the problem of optimal distribution of investments to the sectors of the economics is considered. A dynamic input-output model for describing the development of multi-commodity economics is used. On the base of the available input-output tables (Federal State Statistics Service, Rosstat) a system of differential equations with a control vector in the right-hand side was constructed. There were two goals of regulation: to maximize the GDP and to minimize the value of investments. To solve the multicriteria task the regulation problem was reduced to a linear programming problem. Four different choice principles were considered. For each of them the optimal plan of investments was found.

Categories and Subject Descriptors

I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, and Search — *control theory*

General Terms

Theory

Keywords

Optimal control, multicriteria problems, linear programming, input-output model, economy

1. INTRODUCTION

The theoretical foundations of the input-output model were provided by works of Nobel laureates in Economics W. W. Leontief and L. V. Kantorovich [1, 2]. Currently, the input-output model is one of the internationally acknowledged scientific instrument for analysis of regional economic systems and macroeconomic trends in these systems. The International Input-Output Association [3], which brings together scientists concerned with the theory and practice of application of input-output models, has existed and has been

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWA/IT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

actively functioning for 25 years. In Russian Federation, the statistics on changes in the coefficients of input-output tables are published annually by Rosstat [4].

In this work the dynamic input-output model [5] is used for describing development of a multicommodity economy. The problem of distribution of investments in sectors of economy of the Russian Federation is considered. Data for this model for 2004 – 2006 years are obtained from the official site of Rosstat [4]. The proposed model contains a control vector — quarterly investments in sectors of the economy. The investment process has two objectives: to maximize the GDP and to minimize value of investments. To solve this multicriteria task we use the approach of the multi-objective optimization theory [6].

2. PROBLEM STATEMENT

The dynamic of the main indicators of Russian economy in 2006 can be described by the following linear system of differential equations with control vector [5, 7]:

$$\begin{aligned} \dot{I} &= D_{2006}I + Q_{2006}u, \\ 0 &\leq u_i \leq L_i, \quad i = \overline{1, n}, \\ I(0) &= I_{2006}, \end{aligned} \quad (1)$$

where the first $n - 1$ components of I are the volumes of outputs by sectors in terms of money, I_n — GDP. In this paper there are three aggregated sectors: industry, life support and infrastructure. Data from other sectors are included in the process of reduction, u — n -dimensional vector of investment in the industry, D_{2006}, Q_{2006} — $(n \times n)$ — matrices, obtained using the input-output tables for 2006 [4], I_{2006} — output and GDP vector for 2006:

$$I(2006) = \begin{pmatrix} 13.784 \cdot 10^{12} \\ 5.439 \cdot 10^{12} \\ 14.500 \cdot 10^{12} \\ 21.619 \cdot 10^{12} \end{pmatrix}.$$

Consider the regulation process for one year: $t \in [0, 1]$. Let investment is flowing in the economy on a quarterly basis. Therefore the control parameters can be represented as

$$u_i(t) = u_i^j, \quad (j - 1)/4 < t \leq j/4, \quad \forall i = \overline{1, 4}. \quad (2)$$

Now we introduce two target functionals:

- maximization of GDP by the end of the year:
 $I_4(1) \rightarrow \max,$
- minimization of expenses: $\int_0^1 (\sum_{i=1}^4 u_i) dt \rightarrow \min.$

Let us assume these objectives are equivalent.

3. MULTICRITERIA OPTIMIZATION

By writing the solution of differential equations (1) in the form of Cauchy, at $t = 1$:

$$I(1) = e^D \left(I(0) + \int_0^1 e^{-Dt} Q u dt \right),$$

and by replacing the control u using representation (2), the problem can be reduced to a linear programming problem with two objective functions:

$$\begin{aligned} f_1(x) &= c_1^T x \rightarrow \max_u, \\ f_2(x) &= c_2^T x \rightarrow \min_u, \\ 0 &\leq x_i \leq L_i, \quad i = \overline{1, 4n}, \end{aligned} \quad (3)$$

where $x_{4(i-1)+j} = u_i^j$.

At first we will find the set of admissible plans as the locus that satisfying to restrictions of a task and designate this set by symbol X^* . Obviously that in this task this set is represented by the 16-dimensional parallelepiped.

Next we will carry out the procedure of natural normalization for the functions considered:

$$g_i(x) = \frac{f_i(x) - \min_{x \in X^*} f_i(x)}{\max_{x \in X^*} f_i(x) - \min_{x \in X^*} f_i(x)}, \quad i = 1, 2.$$

Now the values of both functions are measured in dimensionless sizes, their minima are equal to 0, maxima — to 1. Because of the first functional is need to tend to the maximum, and the second to a minimum, we consider auxiliary functions $h_1(x)$ and $h_2(x)$:

$$h_1(x) = g_1(x), \quad h_2(x) = 1 - g_2(x).$$

Now we will enter the set

$$Y^* = \{Y = (y_1, y_2) \mid y_1 = h_1(x), y_2 = h_2(x), x \in X^*\}.$$

Let's say that Y' is more preferable than a vector Y'' ($Y' \succ_Y Y''$), if $y'_1 \geq y''_1$ and $y'_2 \geq y''_2$, and one of these inequalities is strong. We will call Y' non-dominated on the set Y^* , if in this set there is no such vector Y , that $Y' \succ_Y Y$. We denote the set of non-dominated vectors (the set of Pareto-optimal solutions) as $P(Y^*)$.

Consider the vectors $Y', Y'' \in P(Y^*)$. Let $|y'_1 - y''_1| = w_1$, $|y'_2 - y''_2| = w_2$. We introduce coefficients of the relative importance:

$$\theta_1 = \frac{w_1}{w_1 + w_2}, \quad \theta_2 = \frac{w_2}{w_1 + w_2}.$$

Further we will enter a set of relative important vectors:

$$V(Y^*) = \{Y' \in P(Y^*) \mid \theta_1 \geq \theta_1^*, \theta_2 \geq \theta_2^*, \forall Y'' \in P(Y^*)\},$$

θ_1^*, θ_2^* are advance set coefficients.

Obviously that $V(Y^*) \subseteq P(Y^*) \subseteq Y^*$. We will look for the solution of a problem on a set $V(Y^*)$.

Algorithm has some drawbacks. For solution is needed to find set X^* . It's very difficult or impossible even for linear restrictions of the form $Ax = b$. Therefore used only box restrictions $0 \leq x_i \leq L_i$.

4. PRINCIPLE OF A CHOICE

To find the optimal vector $Y^0 = (y_1^0, y_2^0)$ we will use the following principles of choice:

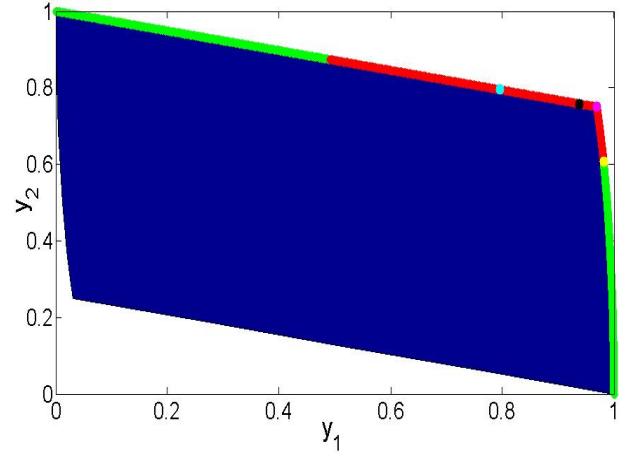


Figure 1: The space of values of target functions

1. the principle of dominant result :
 $\max(y_1^0, y_2^0) = \max_{V(Y^*)} \max(y_1, y_2)$
2. the principle of aggregated efficiency:
 $y_1^0 + y_2^0 = \max_{V(Y^*)} (y_1 + y_2)$
3. the principle of guaranteed results:
 $\min(y_1^0, y_2^0) = \max_{V(Y^*)} \min(y_1, y_2)$
4. the principle of least deviation:
 $\|1 - Y^0\| = \min \|1 - Y\|, \forall Y \in V(Y^*)$.

5. NUMERICAL REALIZATION

To solve this problem, the authors implemented a program in MATLAB covering the following methodology requirements:

1. Reduction of input-output tables from the 16×16 to 4×4 .
2. Construction of a system of differential equations (calculation of matrix D , Q and vector I).
3. Reduction of regulation problem to a linear programming problem.
4. Alignment of target functions.
5. Construction sets $Y^*, P(Y^*)$ and $V(Y^*)$.
6. Finding the optimal vector Y^0 for each principle of choice.
7. The calculation of the optimal control u^0 and optimized functional values for each vector Y^0 .

6. RESULTS

The following parameters values were taken for modeling:

$$L_i = 10^{12}, \quad \theta_1^* = 0.03, \quad \theta_2 = 0.2045.$$

Fig. 1 represents a target functions space. Blue color denotes the set Y^* , green color is used for $P(Y^*)$, red — for $V(Y^*)$. The optimal vector Y^0 for the 1st principle of choice is illustrated by yellow dot, for the 2nd — by purple dot, for the 3rd — by blue dot, for the 4th — by black dot.

	1st qr.	2nd qr.	3rd qr.	4th qr.
u_1^0	0	0	0	0
u_2^0	$2.5 \cdot 10^{11}$	0	0	0
u_3^0	$2.5 \cdot 10^{11}$	$0.725 \cdot 10^{11}$	0	0
u_4^0	$2.5 \cdot 10^{11}$	$2.5 \cdot 10^{11}$	$2.5 \cdot 10^{11}$	$2.5 \cdot 10^{11}$

Table 1: Optimal control for the principle of dominant results

The following values of the control vector and optimized functionals were obtained for optimal vectors Y^0 :

$$I_4^0(1) = 23.336 \text{ trillion roubles}$$

$$\int_0^1 \left(\sum_{i=1}^4 u_i^0 \right) dt = 1.57 \text{ trillion roubles}$$

	1st qr.	2nd qr.	3rd qr.	4th qr.
u_1^0	0	0	0	0
u_2^0	0	0	0	0
u_3^0	0	0	0	0
u_4^0	$2.5 \cdot 10^{11}$	$2.5 \cdot 10^{11}$	$2.5 \cdot 10^{11}$	$2.5 \cdot 10^{11}$

Table 2: Optimal control for the principle of aggregated efficiency

$$I_4^0(1) = 23.332 \text{ trillion roubles}$$

$$\int_0^1 \left(\sum_{i=1}^4 u_i^0 \right) dt = 1 \text{ trillion roubles}$$

	1st qr.	2nd qr.	3rd qr.	4th qr.
u_1^0	0	0	0	0
u_2^0	0	0	0	0
u_3^0	0	0	0	0
u_4^0	$2.5 \cdot 10^{11}$	$2.5 \cdot 10^{11}$	$2.5 \cdot 10^{11}$	$0.7 \cdot 10^{11}$

Table 3: Optimal control for the principle of guaranteed result

$$I_4^0(1) = 23.268 \text{ trillion roubles}$$

$$\int_0^1 \left(\sum_{i=1}^4 u_i^0 \right) dt = 0.82 \text{ trillion roubles}$$

$$I_4^0(1) = 23.320 \text{ trillion roubles}$$

$$\int_0^1 \left(\sum_{i=1}^4 u_i^0 \right) dt = 0.9675 \text{ trillion roubles}$$

7. CONCLUSIONS

In this work two goals of regulation of the economy trends are considered. According to the available input-output tables the system of differential equations with a control parameter vector is constructed. Within the framework of the

	1st qr.	2nd qr.	3rd qr.	4th qr.
u_1^0	0	0	0	0
u_2^0	0	0	0	0
u_3^0	0	0	0	0
u_4^0	$2.5 \cdot 10^{11}$	$2.5 \cdot 10^{11}$	$2.5 \cdot 10^{11}$	$2.175 \cdot 10^{11}$

Table 4: Optimal control of the principle of least deviation

mathematical model the problem of optimal distribution of investments in the sectors of the economy was solved. The optimal regulation problem was reduced to the linear programming problem, because for problems of this class the approach of multi-criteria optimization is well developed. Four principles of a choice are used. For each of them the optimal plan of investments is created. The final result depends strongly on the selection of choice principle. It is impossible to select the best principle from the set of all kinds of principles. This choice depends on the character of the problem and objectives.

8. REFERENCES

- [1] W. W. Leontief. *Essays in Economics. Theory, research, facts and policy*. Politizdat, Moscow, 1990. (in Russian)
- [2] L. V. Kantorovich. *Mathematical methods of organizing and planning production*. Publ. House of Leningrad State University, Leningrad, 1939. (in Russian)
- [3] International Input-Output Association (IIOA) [Electronic resource]. *The official website of the IIOA*. URL: <http://www.iioa.org/> (date of reference: 17.05.2015).
- [4] Federal State Statistics Service, Rosstat [Electronic resource]. *The official website of the Rosstat*. URL: http://www.gks.ru/free_doc/new_site/vvp/tab-zatr-vip.htm (date of reference: 17.05.2015).
- [5] V. P. Peresada, N. V. Smirnov and T. E. Smirnova. Development control of a multicommodity economy based on the dynamical input-output model. *Vestnik of St. Petersburg University. Ser. 10: Applied Mathematics. Informatics. Control Processes*, (4):119–132, December 2014.
- [6] V. D. Nogin. *Making decisions in many criteria*. Fizmatlit, Moscow, 2002. (in Russian)
- [7] A. S. Popkov. Identification of dynamic input-output model for the Russian economy and the optimal investment distribution. *Control Processes and Stability*, 2(1):696–701, April 2015. (in Russian)

Modeling Technical Systems with smartIflow for Safety Related Tasks

Philipp Hönig
Hochschule Ulm
University of Applied Sciences
Institute of Computer Science
hoenig@hs-ulm.de

Rüdiger Lunde
Hochschule Ulm
University of Applied Sciences
Institute of Computer Science
r.lunde@hs-ulm.de

Florian Holzapfel
TU München
Institute of Flight System
Dynamics
florian.holzapfel@tum.de

ABSTRACT

SmartIflow (State Machines for Automation of Reliability-related Tasks using Information FLOWs) is a modeling formalism for automating safety related tasks. It considers systems on a quite high level of abstraction without losing too much predictive power. Since existing approaches often lack in integration into the development process, we present a new method that allows safety engineers to model in their familiar environment. Simulink, Simscape, and Stateflow offer an optimal platform to visualize and edit smartIflow models. This paper describes the graphical notation of smartIflow components in Simulink and the translation into the original language. The practicability of our new graphical modeling approach is shown by an example system that we have modeled successfully.

Categories and Subject Descriptors

B.8.1 [Hardware]: Reliability, Testing, and Fault-Tolerance;
I.6.2 [Computing Methodologies]: Simulation Languages

General Terms

Algorithms, Reliability, Languages

Keywords

Model-Based Safety Analysis, smartIflow, AltaRica, HiP-HOPS, Simulink, Simscape, Stateflow

1. INTRODUCTION

Assistance systems in cars are primarily designed to improve safety and comfort. However, a failure in such systems can lead to critical situations like abrupt unintentional steering movements or brake actions. Such systems are called safety-critical [4], since an error can lead to high material damage or even to loss of life. Therefore, it is an essential task to ensure system safety even in critical situations. This

is achieved by applying several analysis tasks during the development process. Fault Tree Analysis (FTA) and Failure Modes and Effects Analysis (FMEA) are probably the most famous methods for safety assessment. The problem with these traditional analysis techniques is the only partly available automation, which makes the analysis time consuming and also error-prone (especially for more complex systems). Model-based safety analysis (MBSA) as defined in [3] tries to overcome this problem by using a formalized and computer understandable model of the system to automate the analysis process. There are already a lot of approaches to MBSA available whose underlying models are quite different. Distinguishing criteria are for instance the connection modeling, the incorporated knowledge and especially the level of abstraction. The main problem behind the existing approaches is the level of abstraction. Very detailed models will result in great modeling effort and consequently to high computational effort while using highly abstracted models will lead to incomplete results which means that some effects can't be revealed. Another problem of existing approaches is the integration into the development process. Most of the available analysis tools require models in special formalism independent of the models for simulation. Engineers will have to use a separate tool and spend much time in keeping the models consistent.

In this paper we present the modeling approach of smartIflow which was first described in [2]. It supports reasoning about failures and their effects through the whole product cycle. We briefly describe the concepts behind smartIflow and show how systems models can be composed graphically using Simscape and Stateflow.

This paper is organized as follows. Hereafter, we briefly analyze two existing approaches and their basic limitations. In Section 3 the concepts behind our modeling formalism are described. Graphical modeling will be explained in Section 4. In Section 5 the results of our experiments with an example system are shown. Finally a short conclusion will be given.

2. RELATED WORK

The MathWorks Simulink¹ and Simscape² are de facto standard for simulation of technical systems. In Simulink models are composed of a set of mathematical blocks such as integrators, derivatives or basic arithmetic operations that are connected via directed connections. The resulting block

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

¹<http://www.mathworks.com/products/simulink/>

²<http://www.mathworks.com/products/simscape/>

diagram is equal to the mathematical model of the system under analysis. Simscape is an extension to Simulink, which enables physical modeling of multidomain systems. Models are composed of blocks representing the components in the real system. Simscape already provides a set of predefined blocks from various domains such as electric motors, resistors, or valves. Additionally, custom blocks can be created. The connections in Simscape are bidirectional. Thus the structure of the models matches the structure of the real system. Simulink and Simscape are optimal tools for simulation, but not for getting widespread information about system safety. The level of abstraction in particular of Simscape is too low. Both simulation tools use continuous solvers. Thus analyzing all possible failure situations would lead to state-space explosion. Furthermore, models in Simulink and Simscape only support deterministic behavior. To reach an appropriate level of completeness with acceptable computational efforts, one will often need more abstraction. The levels of abstraction of approaches to MBSA are quite different.

HiP-HOPS (Hierarchically Performed Hazard Origin & Propagation Studies) [6] is a modeling formalism using an extremely high level of abstraction. Models are based on Simulink block diagrams wherein a fault model is assigned to each block. The fault model is composed of a set of expressions which specify how a component responds to internal or failures created by other components. Fault modeling is quite limited since there are no state variables and failure propagation is restricted to one failure per connection. The directed connections also lead to the problem that situations in which the flow direction reverses can't be handled. In addition, the structure of both, the real system and the model is quite different. Multiple failure propagation channels can only be handled by creating multiple connections between components.

In contrast to HiP-HOPS, AltaRica [1] is a modeling formalism whose level of abstraction is somewhere between Simscape and HiP-HOPS. In principle, components in AltaRica are represented by nodes comprising variables, events, transitions, and equations to describe the behavior. Flow variables are used to exchange information between components, while state variables can be used to represent the operational- or failure-modes of a component. Events are used to represent state changes. Transitions consist of an event name that induces the state change, a guard and a list of state updates. A guard is the condition which must be fulfilled to enable a transition. As described in [5] there are well known limitations in the first AltaRica approach. AltaRica Data-Flow and AltaRica 3.0 [7] are the successor of AltaRica. AltaRica Data-Flow tries to reduce the computational effort by updating variables by using a fixed order for value propagation. However, no bidirectional flows and no looped systems are allowed. AltaRica 3.0 handles looped systems and bidirectional flows [7]. However, modeling physical flows is still very limited (there are no built-in elements for flow direction calculation) and events can only be triggered externally but not by components within system.

3. THE SMARTIFLOW FORMALISM

3.1 Basic Concept

In principle the main concepts behind AltaRica and smartIfflow are quite similar. In fact, both approaches follow the

principle of Discrete Event Systems (DES). SmartIfflow regards components as finite state machines, which means that each component consists of a set of state variables and the corresponding values. In contrast to AltaRica, events that updates component states are generated based on signal changes initiated by other components. The behavior of a component is specified in terms of value changes of state variables as reaction on events, modifications of the network structure, and property publication through the network. Each component has a set of typed ports over which the components are linked. SmartIfflow allows both directed and undirected connections. These connections are used to propagate abstracted physical flow information and additional data through the network. The propagated information is specified in terms of properties and depends on the current system state since the behavior of a component is state-dependent. Properties are key-value pairs and they are used to abstract from the physical values since we are not interested in specific signal values. Properties for instance can indicate a flow in a system that deviates from the expected value. Besides property propagation, the behavior description of a component also allows the modification of the network structure. For flow direction determination, smartIfflow has some built-in components like sources, drains or bipolar sources. One simulation step is comprised of three sub steps. First, network structure and property values at specific flow variables are updated based on the current values of the state variables. Then, physical flow directions are calculated and property values are propagated through the network. Finally, the values of the state variables are updated based on events produced by changed property values.

3.2 Language

The following listing shows an excerpt of a valve modeled in the smartIfflow modeling language.

```

1  class Valve {
2      Ports:
3          LogicalInput in1;
4          Fluidal p1;
5          Fluidal p2;
6      States:
7          Enum[Position1, Closed] s = Closed;
8          Enum[StuckAtPosition1, Ok, StuckAtClosed]
9              fm = Ok;
10     Transitions:
11         when(in1.val==Closed && fm==Ok &&
12             s == Position1) {
13             s = Closed;
14         }
15         when(in1.val==Position1 && fm==Ok &&
16             s == Closed) {
17             s = Position1;
18         }
19     Behavior:
20         if(fm==Ok && s==Position1) {
21             connect(p1, p2, [flow.val=controlled]);
22         }
23         if(fm==StuckAtPosition1) {
24             connect(p1, p2);
25         }
26         ...
27     }
    
```

The modeling language is component- and object-oriented. For each system component a separate class is created. Similar classes only need to be created once, because classes can be instantiated several times. Model components can, of course, be nested to create hierarchical structures. A component class comprises five *sections*, namely *Components*,

Ports, States, Transitions and Behavior. The section *Components* is used to instantiate subcomponents. Each subcomponent can be accessed with a unique identifier. Connection points that are used to connect components are called ports. All ports of a component are declared in the section *Ports*. A port is specified by a type and identifier that must be unique within a component. The states and failure modes of a component are specified by state variables in section *States*. Each variable has a number of potential values and also a default value. The section *Transitions* describes how a component changes its state. Transitions are specified by a when-clause followed by a set of state assignments. Each when-clause consists of a logical expression that specifies the condition under which the state updates are performed. A logical expression is composed of state comparisons and comparisons between port properties. A state update is realized by assigning a new value to a state variable. In case of non-deterministic transitions, a number of new values, separated by the keyword *or* can be assigned. As already said, the behavior is state-dependent. Therefore section *Behavior* consists of a set of conditional branches that specify the behavior in the different states. The conditions consist of a logical combination of state comparisons. The behavior of a component is specified by a set of actions namely *set* and *connect*. While *set* only puts a number of properties to a port, the *connect*-function links two ports and additionally a set of properties can be added. By omitting the if-clause, default actions can be created that are performed unconditionally.

4. GRAPHICAL MODELING

4.1 Motivation

Safety engineers are often not familiar with modeling languages such as AltaRica or smartIfflow. Therefore creating system models in textual mode even with tool support is very complicated, tedious, time consuming and also error prone. Well known techniques from code editors like syntax highlighting, auto completion or code templates support the user in coding. However, it will be still hard to get an overview about the component instances and how they are interconnected. Another problem arises when specifying the transitions in smartIfflow since the textual representation is very hard to understand in case of more complex components. A way to get rid of this problem is to use graphical modeling. Basically there are different ways to enable graphical composition for smartIfflow models. One possibility could be for example to build a modeling tool from scratch or to use a framework like eclipse graphical modeling framework (GMF)³. Obviously this will lead to an optimal platform for modeling systems, however the engineers will then have to use another tool besides their standard simulation tools. Furthermore, this will result in additional effort since the models have to be created twice even though the structure of both models is quite similar. In addition, the engineers will have to spend a lot of time in keeping the models consistent. Since MatheWorks simulation tools Simulink and Simscape are widely used in industry, we decided to use them as platform for modeling technical systems in the smartIfflow formalism.

³<http://www.eclipse.org/modeling/gmp/>

4.2 Matlab Integration

Since the language of smartIfflow is component-oriented and object-oriented, we utilize the possibility of creating custom libraries in Simulink. This means that a component only needs to be modeled once and multiple instances can be created by drag and drop. Components are connected as usual. Figure 1 shows the structure of a valve modeled with the smartIfflow formalism in Simulink.

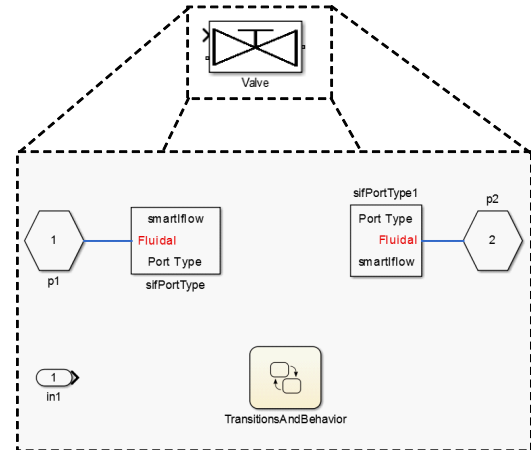


Figure 1: Component structure of a valve

Basically, each smartIfflow class is represented by a Simulink subsystem. SmartIfflow allows you to create hierarchical models. Thus each component can include several component instances. Ports are represented by Simulink *Inputs* respectively *Outputs* in case of directed connections and in case of undirected connections by Simscape physical ports (*PMIOPort*). In addition, a type must be specified for physical ports. This is realized by connecting a special component called *sifPortType* to a port. The type can be specified in the mask of the component. In case of our example component, there are two ports of type *fluidal* (*p1*, *p2*) and one input port (*in1*). When it comes to the representation of transitions and the behavior of a component, Stateflow seems to be a perfect candidate. Stateflow is a toolbox in Matlab for state diagram modeling. As shown in Figure 1, each component may contain exactly one Stateflow chart. This chart is used to represent the transitions and behavior of a component. As shown in figure 2, the Stateflow chart consists of two superstates, namely *Transitions* and *Behavior*. The state *Transitions* is used to define the state variables and, as the name implies, to specify the state transitions. Therefore, the substates of *Transitions* represent the various state variables of a component and their substates represent the corresponding potential values. Substates with dashed borders indicate a parallel decomposition, while solid borders indicate exclusive state decomposition. Default transitions (transitions with no source object) are used to specify the initial value of a state variable. Transitions between state values are labeled with conditions which specify, under which conditions the value change can take place. Non-deterministic transitions can be modeled by creating transitions that have the same source state and label and end at different destination states. The second superstate *Behavior* is used to define the state-dependent behavior. Each behavioral description is modeled by a sin-

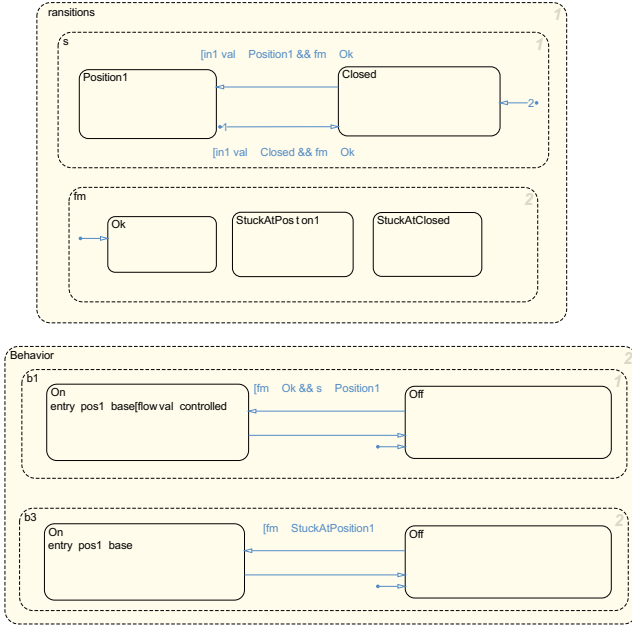


Figure 2: Transitions and behavior of a valve

gle state which contains two substates, namely *On* and *Off*. The transition from state *Off* to *On* specifies the condition under which the behavior is executed while the actual behavior is modeled as entry action inside state *On*. The behavior is described in a simplified syntax. For example, $pos1=base[flow.val=controlled]$ connects the ports *pos1* and *base* and publishes the property $flow.val=controlled$. Whenever the condition of the Off-On-Transition is not fulfilled, the unconditional On-Off-transition will change the state of the behavior to *Off*. The modeling in Simulink has the same expressive power as the smartIfflow language. Therefore, a system modeled with the smartIfflow formalism in Simulink is semantically equal to the model in raw smartIfflow classes. Only the kind of representation differs.

4.3 Conversion Algorithm

Since we need a clear interface between the graphical representation and the simulator, Simulink models are converted in the modeling language of smartIfflow. Matlab provides a set of functions to access the blocks, their parameters and the structure of a Simulink model. The function *find_system* finds all blocks in a model and returns a handle to them. With function *get_param* and a block handle the properties of a block such as instance name, the name of the parent block or the port connectivity can be accessed. Stateflow charts can be accessed with function *find*. This function can be parameterized to analyze the composition of a chart and the transitions between the states. Since smartIfflow supports hierarchical models, the conversion algorithm starts at top level and goes recursively down to the blocks at the lowest level. All gathered information about a component is stored in an object, which is then finally written to a smartIfflow class.

Our graphical conversion tool can be used to translate a Simulink model into raw smartIfflow classes. The tool helps to select the input model and an output folder where the resulting classes are stored. To use the converter, it suf-

fices to add the path to the conversion tool to the Matlab search path and enter *ConverterGUI* in the Matlab Command Window.

5. EXPERIMENTAL VALIDATION

Figure 3 shows an example system that we have successfully modeled in Simulink with the smartIfflow formalism. The example shows a fuel system [8] that could be deployed in the real aircrafts. Main purpose of the system is to support the two engines with fuel that comes from the wing tanks. There are different ways of supporting the engines with fuel, such as crossover, where both engines are supplied by one tank (left/right). The model consists of over

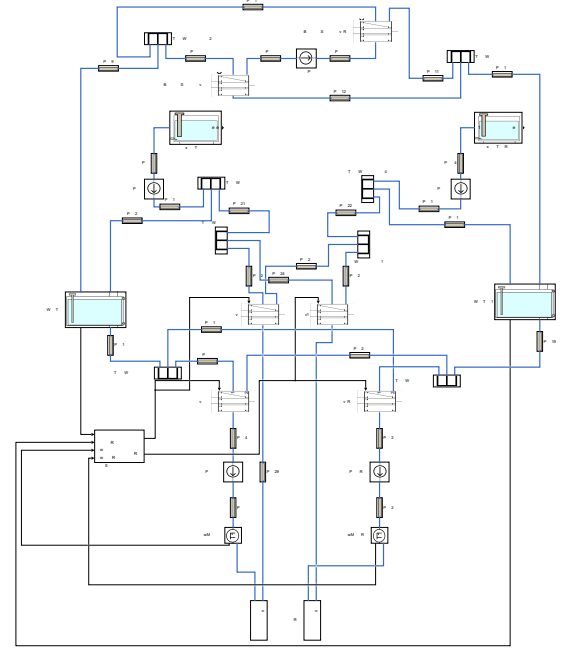


Figure 3: Example fuel system

60 components such as valves, pipes and controller. The conversion of the model into the smartIfflow language took about 1.2 seconds on a workstation with an Intel Core i5 at 3.3 GHz running Windows 8.1 and Matlab 2014b.

Maucher discussed in [5] different ways of modeling the fuel system in AltaRica 1. The experiments have shown that the example system can be modeled only with massive simplifications since the formalism behind AltaRica has too many limitations (e.g. missing flow direction determination and incompatibility with looped systems). In addition, the lack of tool support for graphical model composition limits the practicality for bigger systems.

Thanks to the Simulink integration, modeling systems in HiP-HOPS is in principle quite easy and intuitive, even for more complex systems. The aircraft fuel systems consists of several pipes that allow flow in both directions. To handle such situations in HiP-HOPS, additional connections are necessary since HiP-HOPS uses unidirectional connections for information exchange between the components. The restrictions of HiP-HOPS (e.g. no state-dependent behavior) again lead to an extremely simplified system model.

6. CONCLUSION

In this work we showed the main concepts of smartIfFlow. SmartIfFlow tries to overcome the problems behind existing approaches, which mainly originates from the choice of the level of abstraction, by treating the connections between the components as information channels. This concept in combination with flow direction determination will help to handle situations with unforeseen flow direction changes. The Simulink integration allows even non-computer scientists to model systems quite easily. The component-oriented perspective of Simulink provides an optimal platform for modeling in smartIfFlow. Even complex transitions, can be specified straightforward in Stateflow. The experiments with the example system have shown that even large systems can be modeled in a simple and clear way.

7. REFERENCES

- [1] André Arnold, Gérard Point, Alain Griffault, and Antoine Rauzy. The altarica formalism for describing concurrent systems. *Fundam. Inf.*, 40(2,3):109–124, August 1999.
- [2] Philipp Hönig and Rüdiger Lunde. A new modeling approach for automated safety analysis based on information flows. In *25th International Workshop on Principles of Diagnosis (DX14)*, 2014.
- [3] Anjali Joshi. *Behavioral Fault Modeling and Model Composition for Model-based Safety Analysis*. PhD thesis, Minneapolis, MN, USA, 2008. AAI3324436.
- [4] J.C. Knight. Safety critical systems: challenges and directions. In *Software Engineering, 2002. ICSE 2002. Proceedings of the 24rd International Conference on*, pages 547–550, May 2002.
- [5] Daniel Maucher. Model-based safety analysis with altarica.
- [6] Yiannis Papadopoulos and JohnA. McDermid. Hierarchically performed hazard origin and propagation studies. In *Computer Safety, Reliability and Security*, volume 1698 of *Lecture Notes in Computer Science*, pages 139–152. Springer Berlin Heidelberg, 1999.
- [7] Tatiana Prosvirnova, Pierre-Antoine Brameret, and Antoine Rauzy. Model-based safety assessment: The altarica 3.0 project. *INSIGHT*, 16(4):24–25, 2013.
- [8] Neal Andrew Snooke and Mark H. Lee. Qualitative order of magnitude energy-flow-based failure modes and effects analysis. *CoRR*, abs/1402.0581, 2014.

Fragments Video Protection Mechanism for Video on Demand Server

S. Sevryukov

Faculty of Applied Mathematics and
Control Processes, St. Petersburg
State University, St. Petersburg,
Russia
7-9, Universitetskaya nab.,
St. Petersburg, 199034
+7-911-910-94-88
s.sevryukov@spbu.ru

A. Matlash

Faculty of Applied Mathematics and
Control Processes, St. Petersburg
State University, St. Petersburg,
Russia
7-9, Universitetskaya nab.,
St. Petersburg, 199034
+7-931-002-14-61
matlashallex32@gmail.com

ABSTRACT

Article describes the mechanism which provides access control to video fragments broadcast by the video stream server. Paper present architecture of the developed video materials provision system and principal components of the system. Approaches which can be applied in case of implementation of the above mechanism taking into account use as the broadcasting server of the commercial solution Wowza Streaming Server are described. Own mechanism of fragments video protection presented in this article was designed on the basis of one

Categories and Subject Descriptors

Categories and subject descriptors: C.2.4 [Computer-communication networks] : Distributed Systems---Distributed applications; H.5.1 [Information interfaces and presentation] : Multimedia Information Systems---Video; K.6.5 [Management of computing and information systems] : Security and Protection---Authentication

General Terms

Design, security

Keywords

Streaming media, security, video on demand

1. INTRODUCTION

As technologies in the sphere of media content broadcasting advanced, the need for protection of transmitted data became more urgent. Such a mechanism must allow to do a number of things, such as controlled access to the provided materials, users authentication, copy protection, etc. [2] The need for data

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

protection may be caused by a number of factors depending on a specific tasks. One can cite such examples as the need to protect copyright for the provided materials, control of the provided materials on a commercial basis (subscriptions), etc.

This article describes one of possible architectural solutions making it possible to control and monitor access to the fragments of video materials broadcast with the help of Web-technologies.

It should be noted that the implemented facilities are intended for controlling access to the video fragments broadcast in a VOD mode [1]. This article does not consider so-called streaming broadcasting.

2. ARCHITECTURE DESCRIPTION OF THE VIDEO MATERIALS PROVISION SYSTEM

The system of video materials provision for users includes the following components.

2.1 Broadcasting server

One of the principal components of the system.

The server provides facilities for delivering video to end users enabling them to view it "on the fly". Video is provided by creating so-called broadcasting points: links via which it is possible to obtain necessary data.

Besides, the broadcasting server is involved in the procedure of access control to the provided video fragments. For that purpose, Wowza Media Streaming Server [3] proprietary software is used as a solution

2.2 Data storage

A component for storing video and subtitling intended for broadcasting.

2.3 Web server

A component on the basis of Node.js technologies providing user interface for the procedures of users registration, authentication, authorization and provision of links to video

materials. Besides, the web server is involved in access control alongside with the broadcasting server.

2.4 Database

A document-oriented MongoDB database for storing user data (emails and hash passwords), video materials metadata (names and links) and so-called access scenarios (see below for further discussion).

2.5 User interface

The main component is a HTML5-based [4] Dash.js video player [5] that can handle the data flows provided by the broadcasting server, including the MPEG-DASH adaptive broadcasting technology (a more detailed description of the video provision protocols, as well as the description of adaptive broadcasting technology can be found in [6]).

3. VIDEO FRAGMENTS ACCESS CONTROL SYSTEM

There are several methods of carrying out such control. [7] Among them are such approaches as splitting the file being broadcasted into smaller fragments and creating a separate broadcasting point (and a proprietary broadcasting link) for each of them with subsequent provision of a certain links set to a certain user, as well as, in case the media server is provided with certain functionalities, the control logic implementation exclusively at the software level. [8]

In view of the fact that the Wowza Media Streaming Server commercial solution is used in the developed video materials provision system for which the task of an access control system is being implemented, as well as in view of the solution functionalities, we chose the second method.

The functionality provided by the broadcasting server is the capability of software processing of the received requests for broadcasting with the utilization of server-side API [9], as well as the availability of methods permitting to use a set of certain parameters in the process of broadcasting implementation (in particular, it is possible to specify the start time and the video length during its broadcasting). Those methods were used for the implementation of a proprietary module for the broadcasting server that expands its basic functionality.

It should be noted that the developed solution focused on access control implementation is applicable not only to the Wowza Media Streaming Server but to any broadcasting server subject to the support of the above functionalities by the latter.

3.1 Implementation

Based on the above, the task of implementing access control of the provided video fragments boils down to the task of shaping the "scenarios" of access rights and their application in the process of the broadcasting procedure realization.

By "scenarios" we mean a set of certain parameters applied to broadcasting, in particular the parameter determining the timepoint of the video file from which broadcasting will be started, as well as the length of the video file fragment in relation to the starting point (in seconds).

These scenarios are stored in the database, a certain key is assigned to each of them. Scenario structure example: {id_scenario: 'some_id', key: 'some_key', settings: {start_time: 0, duration: 300}}. This scenario allows to view video since the beginning duration of 5 minutes.

The diagram of process receiving a video fragment according to the rule of access is provided in a figure 1

A user sends request to the web server (1). After the procedure of authentication (2-3), he gets a list of accessible video and chooses one of them (4). Web server sends database request for receiving a scenario key which determinates access rights for this user to the selected video (5).

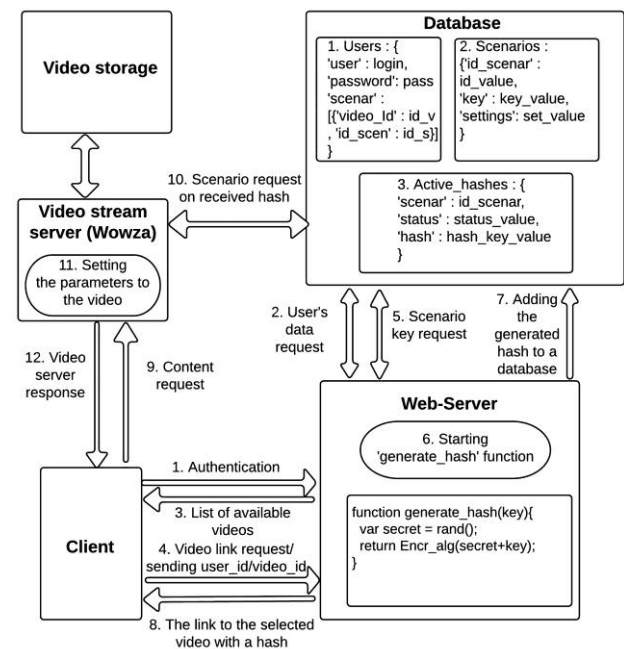


Figure 1. Process of receiving a video fragment according to the access rule

After that random string is generated in a web server, then one of encryption algorithms is applied to this string and the scenario key (6) (SHA-1, for example) [10]. The calculated hash is added to a collection 'Active_hashes' (7). User receives the link of the selected video and a hash (8) and sends a request to video stream server (9).

Video stream server receives request and searches this hash in the 'Active_hashes' collection (10), and, in case the hash is found, receives the scenario with a set of access parameters (11) and uses them for broadcasting video stream to user (12). Then the record which contain the hash is deleted from database. Thus, a hash can be used only once. The web server creates new hash for each new request. If outsider will intercept this hash, he won't be able to obtain data in a case when the user already sent request to stream server.

4. CONCLUSION

The goal of this paper was to describe one of methods making it possible to control and monitor access to the fragments of video materials broadcast. In the course of work principal components

of the developed system of provision of video materials were described. Some of possible approaches which can be applied to the solution of the task access control to video fragments taking into account use in system of provision of video records of the commercial software of Wowza Streaming Server are described. On the basis of the method consisting in use of the functional capabilities, namely Server-Side API given by the server of a broadcasting own module in the Java was realized. This module realizes processing of incoming requests and applies a set of the parameters, such as start time and duration of video broadcast to broadcasting depending on the certain hash-key. The received results are planned to use further in the developed system of video materials provision.

5. REFERENCES

- [1] D. Cymbalák, F. Jakab M. Michalko, Advanced multimedia solution for interactive teaching via whiteboards over IP networks Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Software Engineering (JSSE), November Edition, 2012
- [2] E. Lin, G. Cook, P. Salama, and E. Delp, An overview of security issues in streaming video, Proceedings of the International Conference on Information Technology: Coding and Computing, pp. 345-348, Las Vegas, 2001
- [3] Meyn A., Browser to browser media streaming with HTML5, Kongens Lyngby, 2012
- [4] Wowza Media Systems, Inc. Wowza Media Server 3 – Overview
<http://www.wowza.com/resources/WowzaMediaServer3/Overview.pdf>
- [5] Dash.js player. <https://github.com/Dash-Industry-Forum/dash.js/>
- [6] Ali Begen , Tankut Akgul, Mark Baugher, Watching Video over the Web: Part 1: Streaming Protocols, IEEE Internet Computing, v.15 n.2, p.54-63, March 2011, DOI= [10.1109/MIC.2010.155](https://doi.org/10.1109/MIC.2010.155)
- [7] Michalko, M., Bobko, T., & Fogaš, P. Protected streaming of video content to mobile devices. Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Telecommunications (JSAT), 2014 Edition, Vol. 4, No. 14
- [8] S. Sevryukov, Y. Griva, A. Matlash, Review implementations of access control system to video fragments. The XLVI annual international conference control Processes and Stability (CPS'15), Saint-Petersburg, 2015
- [9] Wowza Media Systems. (2006-2014). Wowza Media Server 3 - Server-Side API
https://www.wowza.com/resources/WowzaMediaServer_ServerSideAPI.pdf
- [10] Eastlake, D. 3rd and P. Jones, "US Secure Hash Algorithm 1 (SHA1)", [RFC 3174](https://tools.ietf.org/html/rfc3174), September 2001

Towards Self-Learning AI for the Videogame of Tennis

Akane Yamada
The University of Aizu
Tsuruga, Ikki-machi,
Aizuwakamatsu, Fukushima, Japan
+81-242-37-2664
m5191102@u-aizu.ac.jp

Maxim Mozgovoy
The University of Aizu
Tsuruga, Ikki-machi,
Aizuwakamatsu, Fukushima, Japan
+81-242-37-2664
mozgovoy@u-aizu.ac.jp

ABSTRACT

While behavior patterns of AI-controlled videogame characters are typically designed manually, self-learning AI systems possess unique attractive features. They can be used to create distinct character personalities of different skill levels, and “train your own character” can be a major user-end feature of a game. In this paper we present our work-in-progress attempt to develop a self-learning AI for the game of tennis. We show how an AI system can learn behavior patterns from user behavior, and act accordingly in similar game situations.

Categories and Subject Descriptors

I.2.1 [Artificial Intelligence]: Applications and Expert Systems – *games*.

General Terms

Algorithms, Experimentation

Keywords

Case-based reasoning, learning by observation, behavior capture.

1. INTRODUCTION

Artificial intelligence systems that control game characters are essential to most computer games, ranging from simple arcade games like Pacman to modern virtual game worlds, inhabited with computer-controlled non-player characters (NPCs). Typically, books on game AI suggest a variety of methods that use handcrafted rules, decision trees, goal hierarchies, etc. [1] These well-established solutions, however, do not help game designers to create diverse and realistic behavior of NPCs, as each behavior pattern has to be designed by hand.

One possible way to remedy the situation is to use learning by observation techniques. We propose the following analogy: in mid-90s, the technology of *motion capture* revolutionized game animation, since the designers could use motion patterns obtained from real people instead of inferior hand-drawn

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

animation. Similarly, handcrafted behavior cannot replicate the whole complexity of realistic NPC reasoning, so it can be reasonable to take the same approach: instead of designing behavior patterns, one can build a “behavioral knowledgebase” by observing human-controlled game characters.

Unfortunately, behavior acquisition is not a straightforward process, and it cannot be applied in a universal way to any computer game. However, for certain types of games this approach is feasible and not exceedingly complex. In this paper, we will describe our work-in-progress effort to create such a self-learning character for a video game of tennis. We will show that the character is able to learn a variety of complex strategies, and we will discuss our preliminary results on character acting.

2. TENNIS GAME ENGINE

As a vehicle for the experiments, we use a custom-designed tennis game engine, developed with Unity3D [2] (see Figure 1).



Figure 1. Tennis game engine.

Each player in the game can perform one of the two actions: run to the specified location and shot the ball into the specified point on the court. The game physics is accurate, so, for example, if a player tries to send a ball coming at a high speed to a nearby point on the opponent’s side of a court, a backspin shot will be performed, so the ball will fly high over the net.

3. AI LEARNING MODE

To learn practical tennis behavior patterns, we played a series of five mini-games, approximately 1-1.5 min each. During a single

game, each player typically performs 100-150 actions. Every game session was recorded to a file for subsequent processing.

Before learning behavior patterns, we stripped recording fragments containing undesirable actions (made due to human errors). In learning mode, our AI system processes actions of a specified player (so it learns from one player in the given game). When an action occurs, the system extracts observable features of the current game situation, and stores the resulting <situation, action> pair in a game graph (see Figure 2). In our current experiments we actually build a family of game graphs for the different sets of observable features. The most accurate graph uses 27 features, while the most abstract game representation relies on 15 features only. These principles are explained in more detail in our earlier work [3].

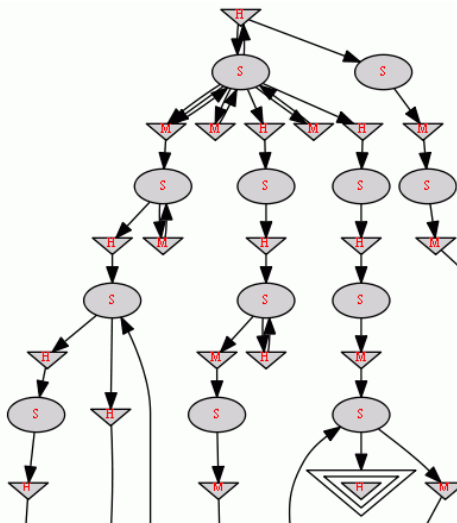


Figure 2. A fragment of a game graph (visualized with AT&T GraphViz).

This representation preserves the specific behavioral patterns of an observed player, since all the actions are stored in the graph.

Even at this stage, our system can be useful for the purpose of game analysis. Game graph provides a “bird’s-eye view” on a game, and can reveal various game peculiarities, such as repeatability (whether two different game sessions share a significant fraction of identical situations), the differences in acting styles of distinct players, average duration of a game episode between subsequent ball servings, etc. We believe this type of instrument can be also applied in real conditions, where digitalized recordings of actual tennis games are used instead of video game data.

4. AI ACTING MODE

In AI acting mode the learned game graph is used to support decision making process. When a player needs to act, our AI subsystem polls the graph to find the matching game situation, and uses a weighted random choice to select one of the attached actions. The procedure also takes into account the previously performed action, and first checks whether the next game

situation in the action chain matches the current game situation, so the next action in the chain can be applied.

If the search yields no result, the procedure is repeated for a less detailed graph (learned with fewer observable attributes). The algorithm fails only if no search returns a matching action.

Our preliminary experiment show that the current agent is able to reproduce typical game patterns such as serving and responding to an opponent’s serve. However, the agent often fails to reproduce longer strategies, involving over 7-10 consecutive actions. We are actively working on these issues; previously we successfully addressed similar problems in a game of boxing [4].

One of primary challenges of tennis is caused by the necessity to determine reasonable decision making points. Human player can act at any time, but as stated above, a typical observed acting rate is around 100 actions per minute. Therefore, our AI system should also act with certain non-uniform frequency, reflecting typical tennis gameplay strategy. In the current implementation the AI system acts every 10 frames, which causes frequent change of behavior patterns. It should be noted, though, that this problem is relevant to any tennis AI system regardless of the underlying approach.

5. CONCLUSION

Game AI is a very specialized topic, since computer games set unusual requirements for AI, such as human-likeness, unpredictability, skill level adjustments, and *fun factor*, a core of any entertainment. Self-learning AI systems have a potential to address these challenges, since they can directly learn from human opponents, and thus exhibit distinct human-like behaviors of different skill levels.

In this paper, we briefly discussed our work-in-progress AI systems for the game of tennis. While our AI did not reach a production-quality stage yet, it demonstrated the ability to learn and repeat simple behavioral patterns of human players. We plan to achieve the desired quality of acting within this year.

6. REFERENCES

- [1] Millington, I., Funge, J. Artificial Intelligence for Games, 2nd Ed. CRC Press, 2012, 896 p.
- [2] Unity 3D Game Engine. Project website: <http://unity3d.com> [Accessed: 20 July 2015].
- [3] Mozgovoy, M., Umarov, I. Behavior Capture with Acting Graph: a Knowledgebase for a Game AI System. *Lecture Notes in Computer Science*, 2011, vol. 7108, pp. 68-77.
- [4] Mozgovoy, M., Umarov, I. Building a Believable and Effective Agent for a 3D Boxing Simulation Game. *Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology*, vol. 3, Chengdu, China, 2010, pp. 14-18.

Realistic Ball Motion Model for a Tennis Videogame

Lopukhov Andrey

National University of Science and Technology MISiS
Moscow, Ul. Dmitiya Ulyanova 13\1, 52
+79261204135

andrew_lopukhov@mail.ru

Serikov Alexander

National University of Science and Technology MISiS
Moscow, Ul. Profsoyuznaya 83\2
+79258554854

ivanmitrafanych@gmail.com

ABSTRACT

In this paper we speak about the modern simulation games, tennis in particular. We give reasons for using complex mathematical models for providing more realistic physics. The majority of forces influencing the ball flight are described, thus forming a system of motion equations, which possible solution with numeric methods we illustrate. A respective optimization problem is stated and solved with gradient projection. In conclusion areas of further study are suggested.

General Terms

Modeling

Keywords

Videogame, mathematical model, physical model

1. INTRODUCTION

Videogame industry is one of the leading consumers of modern computer technologies. In fact, virtually every game uses complex mathematical apparatus. In contemporary simulator games an average player demands high degree of realism. To achieve this, one has to take into account real world laws of physics when developing the mathematical model. In tennis, for instance, initial ball spinning effects its motion severely so respective laws of motion should be considered. Applying angular velocity to the ball allows the player perform a variety of different hits, e. g. using topspin with maximum force one can send the ball to the court, whereas just a flat hit would have caused an out. To implement such opportunities one have to consider Magnus force and some others, which will be described below.

In our model we try consider all forces that have a great influence on the ball movement so that the best user experience can be provided. In this paper we also introduce methods for calculating initial ball hit conditions, which correspond with the motion model.

2. TENNIS BALL MOTION

Equations of motion can be obtained from Newton's laws of motion. For tennis ball we have three main forces [1][2]:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

gravity, air resistance and Magnus force.

While ball moves in the air, it has air resistance. This drag force can be as large as gravitation force and cannot be neglected.

In tennis and other ball sports angular velocity of ball is very important. If symmetrical object (such as a tennis ball) translates in air or liquid it experiences only drag force. But when it has angular velocity it receives force perpendicular to its travelling direction. This fact is called Magnus effect.

From articles [1][2], we take following information about forces.

For flat hits (with zero angular velocity) the force of air resistance is proportional to the square of tennis ball speed:

$$F_d = C_d A d \frac{v^2}{2}, \quad (1)$$

where $A = \pi r^2$ is the cross-sectional area of the tennis ball, r is ball radius, $d = 1.21 \frac{kg}{m^3}$ is air density, v is ball speed and $C_d \approx 0.55$ is drag coefficient.

For Magnus force Rod Cross gives following formula:

$$F_m = C A d \frac{v^2}{2}, \quad (2)$$

where

$$C_l = \frac{1}{2 + \frac{v}{\omega r}}, \quad (3)$$

where ω is angular velocity of the ball.

Ball spin change the drag coefficient. A good fit to the experimental data is given by

$$C_d = 0.55 \left[\frac{1}{22.5 + 4.2 \left(\frac{v}{|\omega| r} \right)^{0.4}} \right]^{2.5}, \quad (4)$$

which indicates that C_d can rise by about 50%.

In this work, we consider two-dimensional ball movement that is projected to three-dimensional space. In such case, we can describe angular velocity as scalar ($\omega > 0$ when Magnus force is directed down and $\omega < 0$ in other case). Using Newton's laws of motion and equations (1)-(4), we have the following system of ordinary differential equations:

$$\left\{ \begin{array}{l} \frac{dx}{dt} = v_x \\ \frac{dy}{dt} = v_y \\ \frac{dv_x}{dt} = -k_d v_x \sqrt{v_x^2 + v_y^2} + \\ \quad + \text{sign}(\omega) k_m v_y \frac{|\omega| r \sqrt{v_x^2 + v_y^2}}{2|\omega| r + \sqrt{v_x^2 + v_y^2}}, \\ \frac{dv_y}{dt} = -g - k_d v_y \sqrt{v_x^2 + v_y^2} - \\ \quad - \text{sign}(\omega) k_m v_x \frac{|\omega| r \sqrt{v_x^2 + v_y^2}}{2|\omega| r + \sqrt{v_x^2 + v_y^2}} \end{array} \right. \quad (5)$$

where $k_m = \frac{d\pi r^2}{2m}$, $k_d = C_d \frac{d\pi r^2}{2m}$, m is mass of ball and C_d is given by (4) if $\omega \neq 0$ or $C_d = 0.55$ otherwise.

Assuming that initial conditions for this system are given, we can treat it as a Cauchy problem. For solving this problem we can use any numeric method.

Modern game engines provide us with opportunity to perform computations between frames with great ease. That is the main reason we considered using one-step methods with fixed time step. Another important criteria of a proper numeric method is its simplicity, conventionality and computational ease. Thus the best fitting methods are Euler's method[4] and the family of second-order Runge-Kutta methods with two stages[5]. These methods obviously calculate only the approximation of a real process. We decided to state that the absolute value of the error must not exceed 0.1 meter per each coordinate during the motion process.

Euler's method fails to meet the precision criteria (for comparison we used Runge-Kutta methods with 4 stages and 1/15 or 1/30 time step). From Runge-Kutta family we tested midpoint method[5] and Heun's method[3]. Both succeeded to meet the criteria, varying in precision depending on the initial conditions. Heun's method appeared to fit out purposes the best.

For videogames, 30 frames per second is the sufficient frequency. Our Heun's method implementation in Unity3D computes the solution much faster than 1/30 of a second, thus making it affordable. Further evaluations that are described in par. 3 take time of 2-3 frames, taking in consideration other important computations running simultaneously.

3. INITIAL CONDITION CALCULATION

In tennis player chooses target at court and tries to perform a hit that will accelerate the ball towards the target. An ideal method should evaluate initial conditions (velocity vector

$\vec{v}_0 = (v_{x0}; v_{y0})$ and angular velocity value ω) that satisfy the following requirements:

- trajectory of ball goes through target point $(X_{target}; 0)$;
- ball doesn't hit net, which is segment $(X_{net}; 0) - (X_{net}; H_{net})$;
- time of flight is as minimum as possible;
- $v_{min} \leq v \leq v_{max}$, $\alpha_{min} \leq \alpha \leq \alpha_{max}$, $\omega_{min} \leq \omega \leq \omega_{max}$, where $v = |\vec{v}_0|$, $\text{tg}(\alpha) = v_{y0} / v_{x0}$. The minimum and maximum values represent the player's skill, equipment, game situation, etc.

As far as there are restrictions for angle of the starting velocity, we search initial conditions as three values: absolute value of ball velocity v , its angle relative to the ground α and angular velocity ω .

We introduce implicit function $\bar{x}(v, \alpha, \omega)$ that evaluates the x coordinate of ball bounce point. This function can be find as a numeric solution of the system (5).

This system cannot be solved analytically thus an explicit equation for trajectory $y = f(x)$ or $x = g(y)$ cannot be found. Despite this fact we can assume that function \bar{x} has following properties:

- is monotone function;
- is strictly increasing function of v for any fixed α and ω ;
- is concave function of α for $\alpha_{min} \leq \alpha \leq \alpha_{max}$ with other arguments fixed;
- is concave function of ω for $\omega_{min} \leq \omega \leq \omega_{max}$ with other arguments fixed.

3.1 Nested searches

In this method, we assume that for minimizing fly time maximizing initial velocity is enough. This is not strictly right, though provides us with a decent approximation.

Working with this idea, we can search for a solution by performing nested searches of each argument. The outer one is a binary search of velocity, while the nested ones are ternary searches of angular velocity and angle.

This solution can find initial conditions for trajectory that goes through target point, however this does not guaranty that the ball will not hit the net. This problem can be solved by iterating through a set of such solutions. Since the real range of angles is relatively small, we can find solution for different fixed angles and then choose the best one.

3.2 Optimization problem

The method suggested above can give us a solution that is very performance costly and does not fulfill conditions stated before.

So we consider the following optimization problem:

$$\left\{ \begin{array}{l} F(v, \alpha, \omega) = \bar{\tau}(v, \alpha, \omega) + p_{target}(v, \alpha, \omega) + p_{net}(v, \alpha, \omega) \rightarrow \min \\ v_{min} \leq v \leq v_{max} \\ \alpha_{min} \leq \alpha \leq \alpha_{max} \\ \omega_{min} \leq \omega \leq \omega_{max} \end{array} \right. \quad (6)$$

where $\bar{\tau}(\nu, \alpha, \omega)$ is implicit function evaluating flight time, $p_{target}(\nu, \alpha, \omega)$ and $p_{net}(\nu, \alpha, \omega)$ are penalties for missing the target and hitting the net.

We defined $p_{target}(\nu, \alpha, \omega)$ and $p_{net}(\nu, \alpha, \omega)$ with the following formulas:

$$p_{target}(\nu, \alpha, \omega) = \begin{cases} 0, & |\bar{x}(\nu, \alpha, \omega) - X_{target}| < \varepsilon \\ R_{target} (\bar{x}(\nu, \alpha, \omega) - X_{target})^2, & \text{otherwise} \end{cases}$$

$$p_{net}(\nu, \alpha, \omega) = \begin{cases} 0, & \bar{y}_{net}(\nu, \alpha, \omega) \geq H_{net} \\ 2R_{net}, & \bar{x}(\nu, \alpha, \omega) < X_{net} \\ R_{net}(1 + H_{net} - \bar{y}_{net}(\nu, \alpha, \omega)), & \text{otherwise} \end{cases}$$

where $\bar{y}_{net}(\nu, \alpha, \omega)$ is implicit function that evaluates the y coordinate of ball trajectory at X_{net} (see Figure 1), R_{target} and R_{net} are penalty coefficients.

We use gradient projection method to solve the optimization problem (6). This solution is very sensitive for the starting point. For choosing it we can use brute force or method described in subsection 3.1.

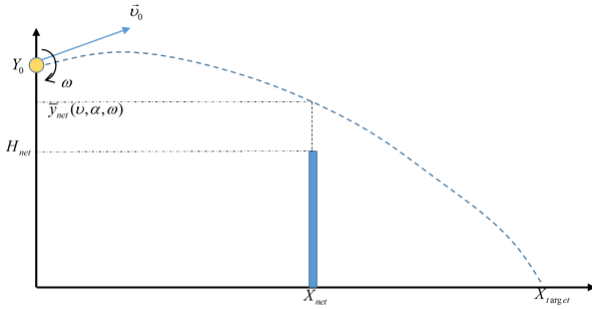


Figure 1

4. RESULTS

We implemented model and methods described above in C# for Unity3D engine. The ball flight of this implementation is more realistic than built-in Unity physics engine solution as it considers Magnus effect and its influence on air resistance.

The initial conditions calculating algorithm is accurate in terms of hitting target.

5. CONCLUSION

We have presented model of realistic tennis ball motion. This model can be used not only for tennis, but also for another ball game with only backspin and topspin rotation type.

Our algorithm for calculating initial conditions of this model is nearly optimal in terms of precision and evaluation performance.

In future work, we plan to optimize initial conditions calculation method to make possible calculation of different hits during one frame.

6. REFERENCES

- [1] R. Cross and C. Lindsey, Technical Tennis, Racquet Tech Publishing, Vista, CA, USA (2005).
- [2] R. Cross, Aerodynamics in the classroom and at the ball park, Am. J. Phys. 80, 289-297 (2012)
- [3] E. Süli, D. Mayers, An Introduction to Numerical Analysis, Cambridge University Press, ISBN 0-521-00794-1 (2003)
- [4] Lakoba, Taras I., Simple Euler method and its modifications (2012)
- [5] Cellier, F, Kofman, E., Continuous System Simulation, Springer Verlag ISBN 0-387-26102-8 (2006)

An Evaluation of Classification Accuracy in a Multilayer Perceptron

Hiroaki Yui
The University of Aizu
Database Systems Laboratory
Ikki-machi, Aizu-Wakamatsu, Fukushima
965-8560
hiroakiyui@gmail.com

Subhash Bhalla
The University of Aizu
Database Systems Laboratory
Ikki-machi, Aizu-Wakamatsu, Fukushima
965-8560
bhalla@u-aizu.ac.jp

ABSTRACT

In data mining, there are a number of algorithms for classification, association rule, clustering and regression. In this article, we classify the data sets into each instance by a Multilayer Perceptron (MLP), which is one of artificial neural networks. The algorithm precisely classifies the data sets as an instance. We evaluate the classification accuracies by changing various hidden layers in the MLP. Also we compare the classification accuracies by the MLP and the other classifiers.

Categories and Subject Descriptors

H.2.8 [Database Applications]: Data mining

General Terms

Machine Learning

Keywords

Data mining, Neural network, Perceptron, Neuron, Sigmoid function, C4.5, Machine learning, Support vector machines

1. INTRODUCTION

An artificial neural networks (ANNs) [1] are applied for a number of pattern recognitions and predictions such as image recognition, handwriting recognition, speech recognition, weather prediction, financial stock prediction and so on. The ANNs are statistical learning models, based on a model of biological neural structures connected with neurons.

The ANNs have a long history in computer science over 50 years. The first computational model for neural network, based on mathematics and neuroscience, was created by Warren McCulloch and Walter Pittsin in 1943 [2]. Rosenblatt developed an algorithm applied for the perceptron in 1959 [3]. The first perceptron has only two layers in a neural network.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

Nowadays, the ANNs present two kind of layer systems such as a single-layer perceptron and a multilayer perceptron (MLP) [4]. The single-layer perceptron is a simple neural network, which consists of a single layer of output nodes. The MLP is organized as a set of hidden layers of artificial neurons. These hidden layers have many neurons, which send and receive messages from multiple sources for computation of neural network. Recently, the ANNs improve the pattern recognition and classification, adding to artificial algorithms, developed by University of Toronto in 2009. We call new type of neural network a deep learning [5].

In this article, we discuss only about the MLP. We evaluate the instances by changing various hidden layers in the MLP. Also we compare the classification accuracies by the MLP and J48, Naive Bayes, AdaBoost and Support vector machines (SVM) [6, 7].

1. J4.8 is a decision tree algorithm and an implementation of the C4.5 algorithm in the weka data mining tool.
2. Naive Bayes classifiers are based on the Bayes' Theorem.
3. AdaBoost is binary classification in a machine learning meta-algorithm
4. Support vector machines are a supervised algorithm and one of classification models.

2. MULTILAYER PERCEPTRON

The perceptron in the MLP is organized by many layers between inputs and outputs. Each layer involves a number of neurons. These neurons connect each other. We called these connections synapses.

2.1 Perceptron

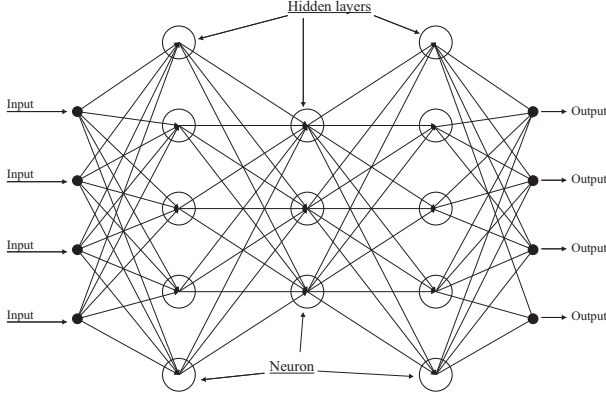
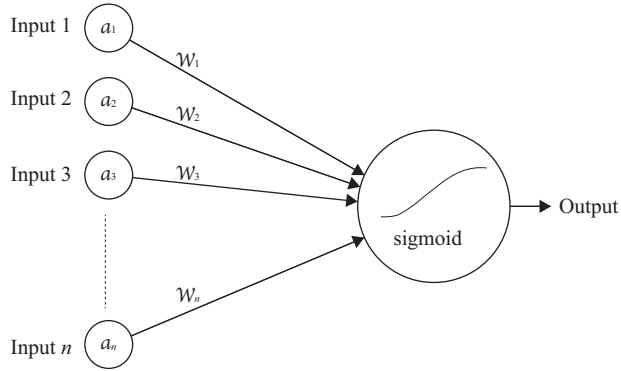
Figure 1 describes the MLP, including 13 neurons, 3 hidden layers, 5 inputs and 5 outputs, From the left layer in this figure, we call these layers *sensory*, *association* and *response*.

2.2 Neuron

The neuron receives n messages a_1, a_2, \dots, a_n with the weights w_1, w_2, \dots, w_n on synapses in a unit of neuron (Figure 2). These messages are sent into one output S computing the sigmoid function.

Table 1: Comparison of the number of hidden layers and neurons in the MLP. The first column corresponds the number of neurons in sensory, association and response.

Number of neurons	Correctly Classified Instances (%)	Mean absolute error	Root mean squared error	Root relative squared error (%)
(3, 3, 0)	75.3906	0.3041	0.4192	87.9540
(5, 5, 0)	75.1302	0.2983	0.4281	89.8090
(3, 3, 3)	76.1719	0.4463	0.3078	87.5931
(3, 3, 5)	75.7813	0.3092	0.4162	87.3101
(3, 5, 5)	76.0417	0.3129	0.4192	87.9519
(5, 5, 5)	73.8281	0.3125	0.4302	90.2485
(5, 5, 10)	75.0000	0.3051	0.4220	88.5258
(10, 10, 10)	74.4792	0.3036	0.4355	91.3773
(20, 20, 20)	74.4792	0.2978	0.4274	89.6779


Figure 1: Description of the Multilayer perceptron. The MLP includes 13 neurons, 3 hidden layers and 5 inputs and 5 outputs.

Figure 2: Description of the Neuron.

The formula for a sum of the incoming messages is described below:

$$S_j = \sum_{i=1}^n w_i a_i, \quad (1)$$

where w_i is the weight for input unit i , a_i is the activation value of the unit i , and n the number of these units.

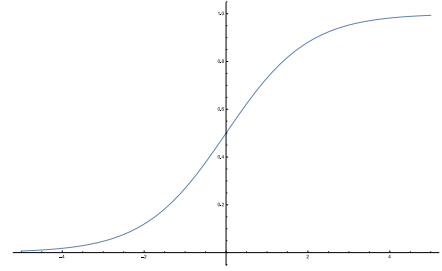
2.3 Sigmoid

Each neuron receives multiple messages and come out. The neuron computes these multiple messages with a sigmoid function. The sigmoid function is based on biological

neuron model. The function is described as below:

$$\text{simoid}(x) = \frac{1}{1 + e^{-x}}, \quad (2)$$

where x is the input in the neural network. Figure 3 shows the sigmoid curve as below:


Figure 3: Sigmoid function.

3. EXPERIMENTS AND DISCUSSION

3.1 Environment and Equipment

In this section, we use a data mining tool, which is called Weka 3.6.12 [8, 9], including classification, clustering and prediction and association rule generation. We run Weka for evaluating the data sets on a machine with a 2.66GHz Intel Core 2 Duo processor and 8G-byte memory running under MacOS 10.10.4. For the experiments, we use a Pima Indians Diabetes data set, provided by National Institute of Diabetes and Digestive and Kidney Diseases [10]. The data set is organized by numerical values. The number of instances, attributes and classes are 768, 8 and 2, respectively.

3.2 Experiments

We evaluate the data set by changing the number of hidden layers and neurons. In addition, we compare the MLP with other classifiers such as J4.8, Naive Bayes, AdaBoost and SVM. The experiments are used in weather data set provided by Weka. In the experiment, we use C4.5, AdaBoost.M1 and Sequential Minimal Optimization (SMO) instead of J4.8, AdaBoost and SVM, respectively. The Weka implements a variant of C4.5 called J4.8 [11]. AdaBoost.M1 gives higher accuracy [12]. In Weka, SMO indicates with SVM. Also we describe statistical terms below:

- *Correctly Classified Instances* is the actual correct classification.

Table 2: Comparison of MLP and other classifiers.

Layers	Correctly Classified Instances (%)	Mean absolute error	Root mean squared error	Root relative squared error (%)
MLP (3, 3, 3)	76.1719	0.4463	0.3078	87.5931
J4.8	73.8281	0.3158	0.4463	93.6293
Naive Bayes	76.3021	0.2841	0.4168	87.4349
AdaBoost	74.3490	0.3127	0.4178	87.6631
SVM	77.3438	0.2266	0.4760	99.8620

- *Mean absolute error* (MAE) measures the average magnitude of the errors in a set of forecasts. The MAE formula is described below:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| = \frac{1}{n} \sum_{i=1}^n |e_i|,$$

where $e_i = |f_i - y_i|$.

- *Root mean squared error* (RMSE) is the standard deviation of the model prediction error. The RMSE formula is described below:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (f_i - y_i)^2}{n}},$$

where the predicted values on the test instances are f_1, \dots, f_n ; the actual values are y_1, y_2, \dots, y_n .

3.3 Results and Discussion

In the MLP, Table 1 shows the classified accuracy instances. The classified instances are not proportion to the number of layers and neurons. We have the best classified accuracy instance when the number of neurons in sensory, association and response are 3,3 and 3, respectively. The result is 76.1716%. Even though the number of layers are the same in the perceptron, the results are different.

Table 2 shows that comparison of the MLP and other classifiers. We choose the best MLP accuracy instance in Table 1, comparing with other classifiers. From the results, we know that the MLP is not the best result from other classifiers' experiments.

4. CONCLUSIONS

In future, we would like to evaluate more data sets in the multilayer perceptron and deep learning algorithm.

5. REFERENCES

- [1] J. E. Dayhoff and J. M. DeLeo, "Artificial neural networks", *Cancer*, Vol. 91, No. 8, pp. 1615–1635, Apr. 2001.
- [2] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity", *The bulletin of mathematical biophysics*, Vol. 5, No. 4, pp. 115–133, Dec. 1943.
- [3] F. Rosenblatt, "A comparison of several perceptron models", *Self-Organizing Systems*, Spartan Books, Washington
- [4] R. Kruse, C. Borgelt, F. Klawonn and et. al., "Multi-Layer Perceptrons", *Computational Intelligence*, pp. 47–81, 2013.
- [5] Y. LeCun, Y. Bengio and G. E. Hinton, "Deep Learning", *Nature*, Vol. 521, pp. 436–444, May 2015.
- [6] H. Byun and S. Lee, "Applications of Support Vector Machines for Pattern Recognition: A Survey", *SVM '02 Proceedings of the First International Workshop on Pattern Recognition with Support Vector Machines*, pp. 213–236.
- [7] Ruggieri, S, "Efficient C4.5 [classification algorithm]", *Knowledge and Data Engineering, IEEE Transactions on*, Vol. 14, No. 2, pp. 438–444, Mar. 2002.
- [8] M. Hall, E. Frank, G. Holme and et. al., "The WEKA data mining software: an update", *ACM SIGKDD Explorations Newsletter*, Vol. 11, No. 1, pp. 10–18, Jun. 2009.
- [9] Weka Tool, available from <http://www.cs.waikato.ac.nz/ml/weka/>.
- [10] Pima Indians Diabetes dataset. Available from: <http://archive.ics.uci.edu/ml/machine-learning-databases/pima-indians-diabetes/pima-indians-diabetes.data>. Accessed: 1st of May, 2008.
- [11] P. O. Gislason, J. A. Benediktsson and J. R. Sveinsson, "Random Forests for land cover classification", *Pattern Recognition Letters*, Vol. 27, No. 4, pp. 294–300, Mar. 2006.
- [12] Y. Keok and L. Hwee Tou Ng, "An empirical evaluation of knowledge sources and learning algorithms for word sense disambiguation", *EMNLP '02 Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, Vol. 10, pp 41–48, doi:10.3115/1118693.1118699
- [13] C. Jean, "Assessing agreement on classification tasks: The kappa statistic", *Computational Linguistics*, Vol. 22, No. 2, pp. 249–254, 1996.

Compact Algorithm of Extrema Count in the Interference Measurement of Transparent Film's Thickness

Alexander V. Boldyrev
Peter the Great St. Petersburg
Polytechnic University
aleksander.boldyrev@
gmail.com

Igor V. Gonchar,
Aleksey S. Ivanov
The National Mineral
Resources University
unityry@gmail.com,
ivaleks58@gmail.com

Alexander B. Fedortsov
The National Mineral
Resources University
borisovitch-
f@yandex.ru

ABSTRACT

In this paper, we describe the solution of the problem of automation of quick count of thickness method developed by Oyama and Mori, and the implementation of an efficient algorithm for the automatic calculation of the interference extremums.

Categories and Subject Descriptors

D.3.2 [Programming Languages]: Language Constructs and Features – graphical structures, *control structures*.

General Terms

Algorithms, Experimentation, Measurement.

Keywords

Thin film, measurement, ADC, algorithm, contactless methods, automation

1. INTRODUCTION

To control the thickness of the dielectric layers (films), which are transparent in the visible or infrared region, it is convenient to use laser interference methods. One of the methods developed for measuring the thickness of the so-called "thick" films (10-1000 μm) is a laser-interferometric method proposed by Oyama and Mori [1]. The film thickness can be determined by counting the number of extrema m , occur when the angle of incidence of the laser beam θ on the film in a predetermined angular range from θ_1 to θ_2 . The refractive index n of the film material must be known.

$$d = \frac{m\lambda}{2(\sqrt{n^2 - \sin^2 \theta_2} - \sqrt{n^2 - \sin^2 \theta_1})}$$

The method proposed by Oyama and Mori was realized with the help of a pantograph, on which the laser and photoelectric detector were placed on adjacent shoulders. The surface of the controlled sample was at the axis of the hinge connecting the

shoulders. The laser beam, which is incident to the previously specified point on the sample surface, will hit on the photoelectric detector at any angle of incidence on the sample. In the opinion of Japanese researchers the main disadvantages of this implementation method is the long duration of the measurement (10 seconds) and the lack of precision in a number of applications.

In the works [2-5] optical-mechanical schemes have been proposed, which allowed increasing the rate of change of the angle of incidence of the laser beam at a given point on the surface of the sample significantly and reducing the time of a single measurement from 10 seconds to 0.02 seconds. One of them [2, 3] contains a rotating flat mirror and two fixed ellipsoid mirrors. The laser beam, which is reflected by a flat rotating mirror from one focus of an ellipsoid, after reflection from the elliptical surface, hits its second focus, which coincides with a predetermined point on the sample surface. Changing of the angle of incidence of the laser beam on the sample is set by rotating a plane mirror. Subsequently, the ellipsoid mirrors have been replaced with spherical [4] or with spherical lenses [5]. When calculating the thickness according to the formula (1), the main task is to count the number of extrema m . To automate the calculation and reduce the time of receipt of the thicknesses we have applied digital signal processing and specially designed software.

2. ALGORITHM

2.1. Description of the algorithm

The chain of receiving and processing information about the interference in the sample is shown on Figure 1.

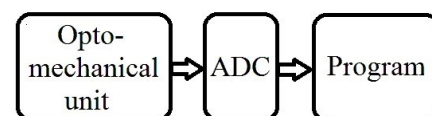


Figure 1. The signal processing chain to obtain values of thicknesses of the films studied.

Signal containing information about the interference in the sample is converted by a photodetector into an electric and goes from the output of the optical-mechanical unit performing a scan of the laser beam on the surface of the film, to the analog-to-digital converter (ADC).

Once converted to digital form from the output of ADC we get the flow of information in an array of data. The data, which

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

contains information about the interference is coming from the optical-mechanical unit with a periodicity of 50 times per second, because, conventionally the whole data stream can be divided into 50 frames. Every frame contains information about the interference pattern in the film. The frames are different one from another because of different disturbing factors - noise superimposed on the useful interference signal. For getting the finite values of the thickness of the analyzed film, each frame is processed on the fly by software. The result of processing - the value of the film thickness, is displayed graphically on the screen of a hardware and software complex.

Some of the information contained in the data stream does not carry information about the interference, so, to reduce the amount of processed data we implement the ability of selection and manual adjustment of the useful part of the information flow. This will allow us to get rid of redundant data and we increase the precision of calculations the number of local extremes without loading the CPU with too much information. Next, the selected data in the one-dimensional array enters the main processing unit - the cycle (with a nominal delay) with a count supplied by all elements of the array. For convenience, the program counts the required number of iterations carried out in the main body of the program automatically.

Previously [7] we have proposed a version of the software processing of the signal of the optical-mechanical unit. The same way in [8, 9], we proposed to combine the methods of optical measurements to repeatedly increase the precision. The current version of the data processing program has rather simple implementation of the algorithm. This will allow us, if necessary, to implement the extrema detection algorithm in the form of a separate electronic circuit and the microcontroller for embedded ADC.

The basic idea of the new algorithm is a selection of the incoming stream from the ADC at each iteration of the cycle of the three elements of the data array, making the detector differences and their subsequent comparison and verification of conditions of extreme detection.

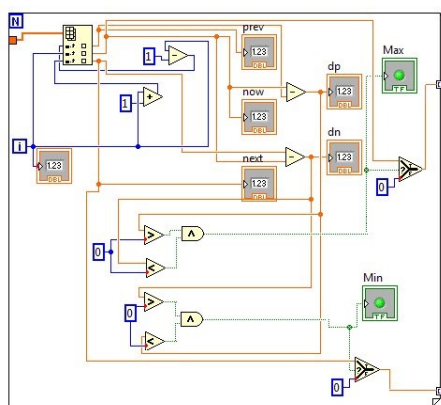


Figure 2. Algorithm for the detection of interference extremes.

Consider the bulk of software implementation of the interference extremes detection unit (Fig. 2): On each loop

iteration via subdivide of index selection elements from the incoming data stream, we extract three elements. For convenience, let us call them the previous, current and next element. After that, we create and calculate the intermediate differences between selected items (On the picture marked, respectively, dp и dn). It is used to set up a sign comparison of the differences, followed by verification of the interference extremes detecting conditions. Under these conditions, to produce the differences we form the following data stream and write it to an array of amplitude values for all elements in the special way. When the “different sign” conditions specified by us for detecting peaks are met (in a similar way to the lows), the current element for this iteration is written in the formed array and, when failed to fulfill the conditions, zero is placed in the array. Thus, now the array generates information about the minima and maxima of interference. Further processing does not represent any difficulties and can be solved in any convenient way. We, in turn, form, for solution of the problem of counting interference extrema from the resulting array, additional array containing zeros and ones, putting ones in an element place in the array, if it is the interference extremum and zero - otherwise. Next, the finding of the number of the interference extrema is a trivial task. Due to the implementation of the proposed extrema detection algorithm, film thickness can be determined with a precision of half-cycle interference.

2.2. Problems, solved for the practical implementation of the algorithm

After digitizing of the signal from the optical-mechanical unit, we receive digital data array that carries the information of the amplitude envelope of the output signal taken from the photodetector unit. The array includes detailed information about the interference in the measurement sample. Information from the opto-mechanical unit is supplied periodically with a frequency of 50 times per second (50 fps), and only a tenth of the time of each frame contains the necessary information. The rest of the analog information from the optical-mechanical unit does not bear information about interference and to free up CPU time for useful for us computations, we must get rid of it. In order to implement calculation of thickness, we solved the problem of the selection of that frame, which carries the interference pattern formed by the sample. For clarity, in the data processing part of the program we provided graphic display of data from the ADC.

The interferences superimpose on the interference pattern formed by the test sample superimpose, and these interferences are distorting information about the interference. One of the confounding factors - the unevenness of rotation of the engine, carrying out a scan of the laser beam. The second factor - unevenness of the samples of the oscillator of analog-digital converter (ADC). The third factor, the most significant - is electromagnetic interference arriving at the input of the ADC together with the signal from the photodetector carrying useful information. All three of the above factors affect the measurement precision of the proposed method of calculation of the digital interference maxima. To deal with them we had to complicate the algorithm of the software processing of measurement results, and namely: to allocate a certain margin

on the leading and trailing edges of the frame carrying the interference pattern. So the algorithm developed by us of extremum detecting has been further complicated and includes a cut-off unit of redundant information, which does not carry the information about interference, with the possibility of manual adjustment of the allocated frame block. Furthermore, to eliminate errors in extrema detection caused by noise imposed on the signal, we allocate additional subarray of data from the stream received from the ADC by sampling every n -th element (in this case, tenth). Additional processing of the subarray level allowed us to avoid errors associated with noise, the frequency of which affect the information within one half-cycle of the interference pattern.

In addition, we applied the algorithm for calculating the interference extrema, which allows us to solve the problem of determining the thickness of the film by the method of Oyama-Mori with the precision up to half period of the interference pattern, which is coming from the sample. We have following conditions: the wavelength of a helium-neon laser, $\lambda = 628$ nm, the refractive index of the polyethylene sample $n = 1,54$, the sampling frequency of used by us twelve bit ADC NI-5122 and 300 kHz and sweep range of the laser beam 35° (from 25° to 60°), provided by the optical-mechanical unit. When detecting 17 of the interference maxima, the measured thickness of the selected sample of the film was $26.49 \mu\text{m} \pm 0,76 \mu\text{m}$. In case of counting the whole number of periods of interference, the variation of thicknesses would be $\pm 1,51 \mu\text{m}$.

2.3. The precision of the measurement method

When increasing film thickness within a range of angles of the sweep of the laser beam, increasing number of interference periods will be placed and each period will fit inside the smaller angular range sweep of the laser beam. Since precision is characterized by taking into account one half period of interference, increasing the thickness of the sample, relative tolerance range of calculated thickness will decrease (at constant absolute tolerance range). This will increase the precision of determining the relative thicknesses. However, when increasing the amount of interference extrema in the given range of the laser beam sweep, the amplitude span between maximum and minimum interference would be reduced due to scattering effects in the real non-plain-parallel sample. In addition, the current peak values of the interference and noise while reducing the useful amplitude swing of the electrical signal will have a more significant impact on the analyzed signal, hence the risk of error of detecting amount of extrema increases. These factors will reduce the precision of the method of measurement of thick films. Using our optical-mechanical unit, we were able to measure the maximum film thickness of $105.95 \mu\text{m}$.

Limitation for small thicknesses can be determined from the fact that within a range of sweep of the beam fit one interference period. Then, as referred to the article's parameters of the measuring system and the refractive index of the sample, the minimum thickness measured by unit was 1.51

$\mu\text{m} \pm 0,76 \mu\text{m}$. It is understood that the relative accuracy of the measurement in this case is low.

A compromise solution will be an artificial restriction of the range of measured thicknesses both above and below, all other constant parameters of the system, such as: the range of the beam sweep, carried out by opto-mechanical unit, the wavelength of the used laser and the refractive index of the sample.

2.4. Advantages of the algorithm

The software module has a simple structure, which allows us to implement it in simple programmable complexes. Local extremes are counted precisely: in most raw data the results of program extrema counting is identical with the calculation of the waveform manually. Compact software system coupled with the ease of implementation does not require large amounts of CPU computing time and allows us to implement his work on a single chip with integrated ADC.

2.5. Disadvantages of the algorithm

The relative simplicity of the proposed algorithm is both an advantage and a disadvantage, because of the influence on the result of counting of the thickness of noise and disturbances superimposed on the useful signal. These disturbances manifest themselves and cause minor error in extrema detection that are not handled with the proposed algorithm. About 20% of the values of the thickness are differ from the bulk of values of thicknesses on half-period interference. Increasing the accuracy of counting of the interference extrema requires further complicating of data processing algorithms, and as well usage of statistical processing of the values of thickness;

3. ACKNOWLEDGMENTS

Our thanks to National Instruments for allowing us to use their licensed product LabView 8.5 that they had developed. In addition, we thank The National Mineral Resources University for providing us with all the hardware equipment and with all the experimental tests, where they used modern facilities and some software and hardware equipment made by their best specialists.

4. REFERENCES

- [1] Oyama, D. Mori, «Optical method of measurement of uniform thicknesses of solid and liquid films in range of around $0,01 \dots 1$ mm». Research instruments, № 10, 1987 year, pp. 70 – 75.
- [2] A.B. Fedortsov, D.G. Letenko, Yu.V. Churkin, I.A. Torchinski, A.S. Ivanov «A fast operating laser device for measuring the thicknesses of transparent solid and liquid films». Pub.AIP, "Review of Scientific Instruments", 1992 year, ed. 63, № 7, pp. 3579-3582.
- [3] L.M. Tsintsiper, A.B. Fedortsov, D.G. Letenko, «Device for the research of the kinetics of spreading and evaporation of liquid films in the realtime». Pub. «Science», Instruments and Experimental Techniques, № 1, 1996 year, pp. 154-1575

- [4] A.B. Fedortsov, «Device for non-destructive measurement of the thickness of the dielectric and semiconductor films». RF utility patent № 2102702, 1998 year
- [5] A.S. Ivanov, V.V. Manuchov, A.B. Fedortsov, Yu.V. Churkin, «High-speed control device for the angular dependence of the reflectance of the laser beam». High Education Institutes Proceedings, Device design. Pub. ITMO, ed. 54, № 3, 2011 year, pp. 61 – 64.
- [6] A.B. Fedortsov A.S. Ivanov, I.V. Gonchar, V.V. Manuhov. «Device for non-destructive measurement of the thickness of the dielectric and semiconductor films». Utility patent № 120-90. Registered in the State Register of utility models of the Russian Federation in September 20, 2012
- [7] I.V. Gonchar, A.S. Ivanov, A.B. Fedortsov, Yu. V. Churkin. Certificate of state registration of the electronic computer, № 2012611152, «The program for data processing of program-apparat complex "Automated high-speed laser interferometer for measuring the thickness of transparent solid and liquid films».
- [8] I.V. Gonchar, A.S. Ivanov, A.B. Fedortsov. Multiple lazer-interferometrical improving of the precision of the method of measurement of thickness of "thick" non-metallic films. SPb: // Scientific and technical sheets of St. Petersburg State Polytechnic University, a series of physical and mathematical sciences», 2012 year, № 3, pp. 48–56.
- [9] I.V. Gonchar, A.S. Ivanov, A.B. Fedortsov. Development of the method of Oyama-Mori to control the thickness of the optically transparent dielectric layers. // SPb: Proceedings of XIII International conference "Physics of Dielectrics" 2014, Volume 2, pp. 74 - 76. Ed. RSPU. Herzen.

OD-Matrix Estimation for Urban Traffic Area Control

Anastasiya Raevskaya^{*}
Saint-Petersburg state university
198504, Universitetskii prospekt 35, Petergof,
Saint-Petersburg, Russia
a.p.raevskaya@yandex.ru

Alexander Krylatov[†]
Saint-Petersburg state university
198504, Universitetskii prospekt 35, Petergof,
Saint-Petersburg, Russia
aykrylatov@yandex.ru

ABSTRACT

Congestion, accidents, greenhouse gas emission and others seem to become unsolvable challenges for all levels of management in modern worldwide large cities. The increasing dynamics of motorization requires development of innovative methodological tools and technical devices to cope with problems appeared on the road networks. Primarily, control system for urban traffic area has to be created to support decision makers via operating with big transportation data. The input for a such system is a volume of travel demand between origins and destinations – OD-matrix. The present work is devoted to the problem of OD-matrix estimation. The original technique of plate scanning sensors location is offered. The developed method has been tested on the experimental parameters of the Saint-Petersburg road network.

1. INTRODUCTION

The problems of estimation and reconstruction of trip matrices are difficult and actual problems in transportation researches. In general, trip matrices estimation and reconstruction are different problems and their solutions can be unequal [1]. One of the first mathematical models of a trip matrix estimation developed in the end of XX century was formulated as a bi-level program [2]. Despite numerous publications, this problem is still pressing scientific issue that require further research [8]. Among the recent results in this area the paper [3] should be mentioned, since in this article authors consider the problem of trip matrix and path flow reconstruction and estimation by combining data from plate scanning and link flow observation. A detailed comparative analysis of the three methods of trip matrix estimation (a method of linear programming, Bayesian approach, the

method of time-varying network tomography) was made in [4]. The actual size of transportation networks are very large and it is reasonable to formulate the problem of minimizing the number of sensors [5, 6]. In [7] a new model based on Total Demand Scale is developed.

The important practical implications on an efficient estimation of traffic flows by using plate scanning sensors was made by [3]. However, the problem of the optimal sensors location in terms of maximum observation of traffic flows has been never considered. This paper deals with a such formulation and, simultaneously, proposes a model of plate scanning sensors location, that ensures the effective application of the Castillo method [3].

2. THE OPTIMAL PLATE SCANNING SENSORS LOCATION

We assume that the plate scanning sensor could identify the plate of a vehicle on the lane where the sensor has been installed. Consider the transportation network presented by digraph $G = (N, A)$, N — set of vertices, A — set of arcs. We introduce the following notation: W — set of OD-pairs, $w \in W$; K^w — set of routes between OD-pair w , $w \in W$; q_a and c_a — number of sensors and number of lanes on the arc $a \in A$ corresponding; $q = (\dots, q_a, \dots)$; A_k — set of arcs belonging to the route $k \in K^w$, $A_k \subset A$; \bar{f}_k — *a priori* traffic flow through route $k \in K^w$. The probability of identification of a vehicle on the arc $a \in A$ with c_a lanes is $\frac{q_a}{c_a}$, $\forall a \in A$.

We assume that the random events of plate identification on a sequence of arcs are independent. The goal could be formulated as follows:

$$\max_q z(q) = \max_q \sum_{r \in R} \sum_{s \in S} \sum_{k \in K^{rs}} \left(\prod_{a \in A_k} \frac{q_a}{c_a} \right) \bar{f}_k, \quad (1)$$

where R is for origins and S — for destinations.

Thus we maximize the probability of fixing traffic flows throughout the route. At the same time, the condition that at least one arc in each path is scanned, could be given by the following inequality:

$$\sum_{a \in A_k} q_a \geq 1 \quad \forall k \in K^{rs}. \quad (2)$$

The existence of at least one scanned arc, belonging to the route k_1 and not belonging to any other route k_2 , is guar-

^{*}postgraduate student of applied mathematics and control processes faculty

[†]Cand. Sci., assistant of applied mathematics and control processes faculty

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

anted by the following constraint:

$$\sum_{a \in A} q_a \delta_a^{k_1 k_2} \geq 1 \quad \forall k_1, k_2 \in K^{rs} : k_1 \neq k_2, \quad (3)$$

$$\delta_a^{k_1 k_2} = \begin{cases} 1, & \text{if } a \in k_1, a \notin k_2, \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

Consider the budget constraint, where Q — the number of available sensors:

$$\sum_{a \in A} q_a \leq Q. \quad (5)$$

The number of sensors on the arc should not exceed the number of lanes:

$$0 \leq q_a \leq c_a, \quad \forall a \in A.$$

So, we formulated the integer program on a limited set of solutions. The solution of this problem exists when the set of possible solutions is not empty.

3. COMPUTATIONAL EXPERIMENT

Consider a multi-lane traffic network of Saint-Petersburg city center. We define nine areas of origin-destination (fig. 1), nine routes between pairs of origins and destinations and, hence, 21 arcs (fig. 2).

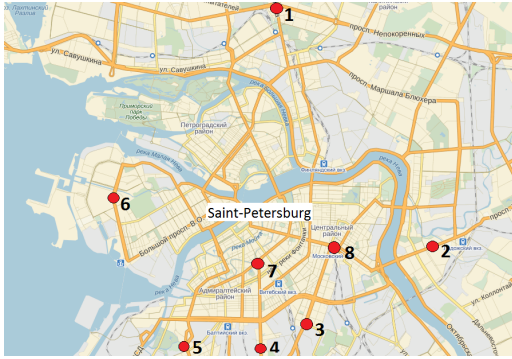


Figure 1: Origins-destinations

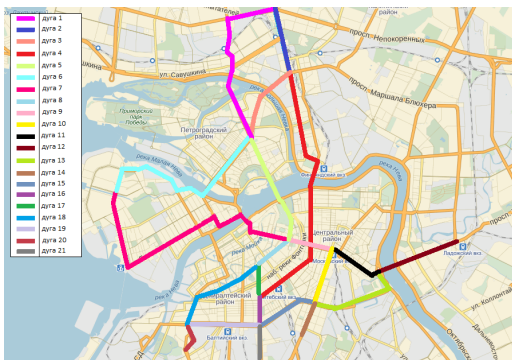


Figure 2: Links

We define the number of lanes c_a and *a priori* traffic flows \bar{f}_k on the available routes. Then we calculate optimal plate scanning sensors location for the different budget constraints

varying budget values from 0 to 50. Since there is the optimization problem with nonlinear functional and linear constraints, we employ *fmincon*-function in a software environment MatLab. The dependence between the observed flows value and the number of sensors located on the network are presented on the fig. 3.

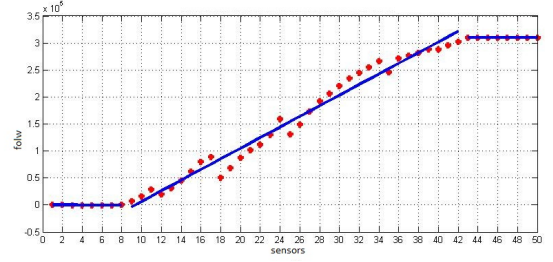


Figure 3: Total flow

Consider the different combinations of constraints — separately (5), then (5) and (2), and then all constraints (2) — (5). The dependence between the value of the total observed flow and the number of sensors on a network for different combinations of constraints does not change. Therefore, this analysis was performed for each of considered routes. If we analyze the dependence between the number of sensors on the network and the value of observed flows then we recognize that routes with significant flows and small number of arcs are observed primarily. By increasing the amount of constraints large flows on routes with lots of arcs and lanes begin to occur later (fig. 4).

Number of sensors on the network	The proportion of the observed flow $f(1)$, one constraint.	The proportion of the observed flow $f(1)$, two constraints.	The proportion of the observed flow $f(1)$, all constraints.
5	0	0	0
10	0	0	0
15	1,00	0	0
20	1,00	1,00	0
25	1,00	1,00	1,00
30	1,00	1,00	1,00
35	1,00	1,00	1,00
40	1,00	1,00	1,00
45	1,00	1,00	1,00
50	1,00	1,00	1,00

Figure 4: Combinations of restrictions

Conclusion

The article is devoted to the problem of optimal location of traffic plate scanning sensors on the network. A brief review of the literature is carried out and the method of Castillo is recognized. To increase the efficiency of the Castillo method, an original model of optimal location of plate scanning sensors on the network of general topology is developed. The developed model allows to maximize probability of identification of the most significant traffic flows throughout the corresponding routes. We obtained the following theoretical and practical results:

- a deterministic model of optimal plate scanning location on the network of general topology;
- an algorithm for solving deterministic problem of the sensors location;

- implementation of developed algorithm on the transportation network of St. Petersburg.

4. REFERENCES

- [1] Hazelton M. Inference for origin-destination matrices: estimation, prediction and reconstruction // *Transportation Research Part B*. 2001. No 35. P. 667–676.
- [2] Yang H., Sasaki T., Iida Y., Asakura Y. Estimation of origin-destination matrices from link traffic counts on congested networks // *Transportation Research Part B*. 1992. No 26 (6). P. 417–434.
- [3] Castillo E., Menedez J. M., Jimenez P. Trip matrix and path flow reconstruction and estimation based on plate scanning and link observations // *Transportation Research Part B*. 2008. No 42. P. 455–481.
- [4] Medina A., Taft N., Salamatian K., Bhattacharyya S., Diot C. Traffic matrix estimation: existing techniques and new directions // *Computer Communication Review – Proceedings of the 2002 SIGCOMM conference*. 2002. No 32. P. 161–174.
- [5] Bianco L., Cerrone C., Cerulli R., Gentili M. Locating sensors to observe network arc flows: exact and heuristic approaches // *Computers and Operation Research*. 2014. No 46. P. 12–22.
- [6] Minguez R., Sanchez-Cambronero S., Castillo E., Jimenez P. Optimal traffic plate scanning location for OD trip matrix and route estimation in road networks // *Transportation Research Part B*. 2010. No 44. P. 282–298.
- [7] Bierlaire M. The total demand scale: a new measure of quality for static and dynamic origin-destination trip tables // *Transportation Research Part B*. 2002. No 36. P. 282–298.
- [8] Zakharov V., Krylatov A. OD-matrix estimation based on plate scanning // *2014 International Conference on Computer Technologies in Physical and Engineering Applications (ICCTPEA)*. 2014. P. 209–210.

Developing Google Chrome Extensions: Case Study

Kentaro Murai

Software Engineering Lab, University of Aizu
Aizu-Wakamatsu City,
Fukushima-ken, 965-8580, Japan
s1210250@u-aizu.ac.jp

Vitaly Klyuev

Software Engineering Lab, University of Aizu
Aizu-Wakamatsu City,
Fukushima-ken, 965-8580, Japan
vkluev@u-aizu.ac.jp

ABSTRACT

There are many Web browsers on the Internet. Although any OS has a pre-installed browser, some people use another. They choose more functional and convenient browsers. A browser extension is a key feature which enhances its functionality. In this paper, we review publications on browser extensions and fundamental issues on the Google Chrome extension.

Categories and Subject Descriptors

H.3.5 [Online Information Services]: Web-based services

General Terms

Languages, Algorithms

Keywords

browser extension, Google Chrome, Context Menus

1. INTRODUCTION

A browser extension is a small software that extends its capability by adding a certain function. The value of browser extensions is growing with the growing number of webapps. Webapps use browsers as a UI. Almost all browsers have their extensions. A new version of Microsoft Windows – Windows 10 – will finally have a browser that supports extensions.

Recently, smart device apps are gaining popularity. Browser extensions will attract many web developers in the near future. There are some app markets for smart devices such as App Store and Google Play. There are several markets that distribute browser extensions such as Chrome Web Store. As Microsoft launches Windows 10 with a new browser, marked by extension functionality, it gets more value.

2. BROWSER EXTENSIONS

A browser extension adds a new value to the browser. There are many browser extensions. For example, a certain browser extension enables us to look for a word or text selected on the web page in an online dictionary. The browser extension is also called plug-in or add-on. The Table 1 shows worldwide browser

share in the past year and information about extensions.

Table 1. List of widely known browsers


product	YTD share [7]	support extensions?	extension made of
Google Chrome	48 %	Yes	HTML, CSS, JavaScript [4]
Internet Explorer	19.44 %	No	-
Mozilla Firefox	16.8 %	Yes	HTML, CSS, JavaScript, XUL [5]
Safari	10.52 %	Yes	HTML, CSS, JavaScript [6]
Opera	1.53 %	Yes	mostly the same as Google Chrome [1]
Microsoft Edge	-	to-be	-

3. GOOGLE CHROME EXTENSIONS

3.1 Extension Types

Google Chrome is one of the famous web browsers. It has many extensions in the Chrome Web Store [3]. Let us consider a Google Chrome extension as an example because of its largest share of the browser. Google Chrome extensions are classified into several types of extensions. The description of these types are given further.

3.1.1 Browser Action

Browser Action is one of the extensions. Its icon () locates next to 3-lined button (see Figure 1). One can click the icon to have the extension doing something.

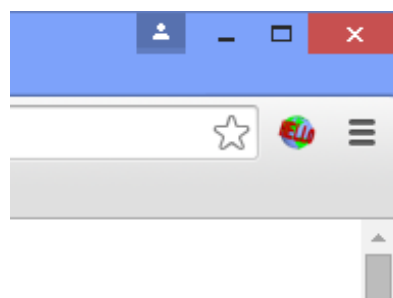



Figure 1. Browser Action

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.

Copyright 2015 University of Aizu Press.

3.1.2 Page Action

Page Action icon () can be found next to the star icon on the address bar (see Figure 2). Unlike Browser Action icon, the Page Action icon is not always visible. It is visible under the specific condition. While the icon is visible, the extension can work. For example, the web page, which the browser displaying, belongs to a specific domain. To create a Page Action, we can use chrome.pageAction API provided by Chrome.

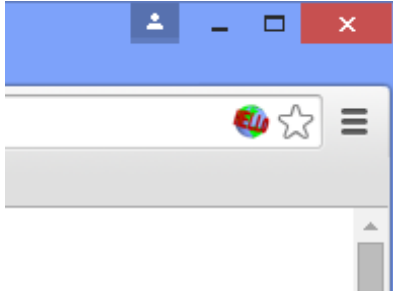


Figure 2. Page Action

3.1.3 Content Scripts

Content Scripts run JavaScript code automatically on a web page. They work in the background, thus they do not have an icon on the UI of Google Chrome.

3.1.4 Context Menus

Context Menus insert items to the context menu. A context menu is a small menu which is displayed near the cursor in response to the mouse right-click. To add items to the context menu, we can use the chrome.contextMenus API.

3.2 Example of the Basic Extension

We illustrate the features of Context Menus with an example. We created an extension by referring to the programming learning site [1]. This illustrative example can search on the University of Aizu Official Website [8] for the selected text. This example is able to search on online shopping sites, SNS, online dictionaries. This extension runs on a Google Chrome. The extension consists of only three files – manifest.json, background.js, and iconimg_16x16.png.

manifest.json is a manifest file. They describe the extension and associates its resources (such as background.js) with the extension. It is written in JSON format. JSON can be considered as an associative array. It is similar to an object in JavaScript. The content of manifest.json is presented below.

```
{
  "name": "u aizuSearch",
  "version": "0.1",
  "manifest_version": 2,
```

```
  "description": "Search for selected text on the
  University of Aizu Official Site.",
```

```
  "permissions": [
    "tabs", "http://*/*", "contextMenus"
  ],
  "icons": {
    "16": "iconimg_16x16.png"
  },
  "background": {
    "scripts": ["background.js"]
  }
}
```

In manifest.json, there are some properties regarding the extension such as "name", "version", and "permissions". This information is required for manifest.json to use a function of Google Chrome, declarations of permission.

background.js is a JavaScript file. It describes what the extension does. The discussed extension has a very simple behavior. The background.js file is given below.

```
chrome.contextMenus.create({
  "title": "Search for '%s' on the University of
  Aizu Official Site",
  "type": "normal",
  "contexts": ["selection"],
  "onclick":
    function(info) {
      var url =
        'http://www.u aizu.ac.jp/e search.html?q=' +
        encodeURIComponent(info.selectionText);

      chrome.tabs.create({url: url});
    }
})
```

We explain how background.js behaves. chrome.contextMenus API is used in the background.js. When the user selects a text on the web page and right clicks it, Google Chrome displays the context menu including the item of the extension (see Figure 3). Then, the extension can get the selected text in the parameter "selectionText" via the object "info". After that, the extension assembles URL that searches the university site and opens a new tab with it. Finally the user gets a search result page opened in a new tab in this way.

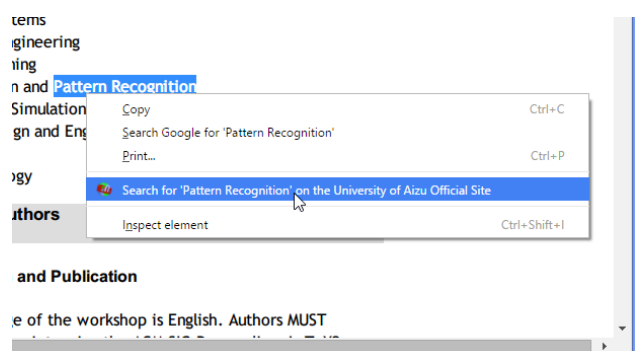


Figure 3. Context Menu example

iconimg_16x16.png is an image file used as an icon (see Figure 4). If an extension is either the Browser Action or Page Action, the image will appear on the UI of Google Chrome. Figure 4 is reprinted from the material on official website of Google Chrome extension for developers [4].



Figure 4. Image used as icon

3.3 How to Create Extensions

An extension needs a JSON file, JavaScript, HTML, and CSS if necessary. A creator for the Google Chrome extension should be familiar with these languages. An experienced web developer can easily create extensions. One can refer to the reference manual for Google Chrome extension [4]. We can begin to create it with the text editor and Google Chrome for free.

4. CONCLUSION

Creating browser extensions is not so difficult. If one is unfamiliar with browser extensions, the best strategy is to check web resources such as Chrome Web Store [3] and others. Many web browsers are cross-platform. Browser extensions are cross-platform too. It can improve productivity of web developers.

5. REFERENCES

- [1] Dev.Opera, Opera extensions documentation, <https://dev.opera.com/extensions/> [Accessed on August 18, 2015]
- [2] Dotinstall, Introduction to the Google Chrome extension, http://dotinstall.com/lessons/basic_chrome_v2 [Accessed on August 18, 2015]
- [3] Google Chrome, Chrome Web Store, <https://chrome.google.com/webstore> [Accessed on August 18, 2015]
- [4] Google Chrome, What are extensions?, <https://developer.chrome.com/extensions> [Accessed on August 18, 2015]
- [5] Mozilla Developer Network, Add-ons, <https://developer.mozilla.org/en-US/Add-ons> [Accessed on August 18, 2015]
- [6] Safari Developer Library, Extensions Overview, <https://developer.apple.com/library/safari/documentation/Tools/Conceptual/SafariExtensionGuide/ExtensionsOverview/ExtensionsOverview.html> [Accessed on August 18, 2015]
- [7] StatCounter, Top 5 Desktop, Tablet & Console Browsers from July 2014 to July 2015, <http://gs.statcounter.com/#browser-ww-monthly-201407-201507-bar> [Accessed on August 18, 2015]
- [8] University of Aizu Official Website (English), <http://www.u-aizu.ac.jp/e-index.html> [Accessed on August 18, 2015]

3D Interface for Easy Operation of Disaster Management Robot Teleportation

Venushka Thisara Dharmasiri

Informatics Institute of Technology

57, Ramakrishna Road

Colombo-06, Sri Lanka

00600

p.dharmasiri@my.westminster.ac.uk

Thilak Chamindra

Informatics Institute of Technology

57, Ramakrishna Road

Colombo-06, Sri Lanka

00600

chaminh@westminster.ac.uk

ABSTRACT

It is difficult to understand the surrounding around a robot from local cameras in a disaster situation since the disaster environment could change rapidly and it could cause a large mental load for human operators in that kind of environment. Human operator needs to have a clear understanding about the height and the distance to control a robot. In addition to that this paper proposes to be developed using a multicore computer with low computational time with low memory consumption. Main contribution is made using none of precisely calibrated cameras, blue screen, special sensor device or special hardware. This paper discusses computer vision based 3D simulation application for a disaster management robot. We also performed an experiment using Flea2 camera dataset. This has got seven hundred frames for a view. All the frames captured under medium illumination level. All the frames we achieved are 3D views of every frame. Mainly our proposed approach is capable of taking multiple cameras feeds (2D feeds) and project a 3D cam feed to the computer screen. This identifies and matches common pixels of images and those matched images are used for 3D video processing and manipulation of 3D stream based on controllers.

Categories and Subject Descriptors

1.4.1 [Digitization and Image Capture]: Imaging geometry

1.4.6 [Segmentation]: Edge and feature detection

General Terms

Algorithms, Performance

Keywords

3D Interface, RANSAC, SGM, Fronto parrel, Harris detector, SIFT, FLANN matcher, Epipolar geometry

1. INTRODUCTION

The existing system is based on the mobile robot. Cameras are mounted at various angles. Ground control and mobile robot communicates in a blind way. Ground control has one screen for one camera view. Sometimes a large screen is divided into views. The human operator can monitor all camera feeds and take decisions according to the environment further; in general situations robots are mounted with two cameras. One camera is a self- rotating camera.

Human operators are capable of viewing more than one screen and rotate the camera and take decisions within a very limited time period. But in a disaster environment, time is the most valuable factor. Most of the times in disaster situations, the environment changes dynamically & rapidly. Then controlling a mobile robot would be a challenging task. In that kind of a situation the human stress & fatigue will increase and it may cause large mental load for a human operator. This problem occurs when viewing various video feeds at the same time. If human operator had a clear understanding about the height and distance to control a robot, the impact on the operator wouldn't be the same. To overcome that issue our proposed solution is to develop a 3D Interface for the disaster management robot.

3D applications can be divided into hardware oriented & software oriented. Each approach has strong points and defects.

Most of the Hardware based approaches are based on sensors. [1, 2, 3, 4]. Following limitations can be identified in hardware based approaches. Uses special sensor nodes which are expensive and sensor integration, fusion takes time. On the other hand, software based approaches are based on vision based algorithms. SLAM concept [5] which is used for constructing the 3D depth map of an unknown environment or update a map within the known environment. Sato [6] has used multiple fish eye cameras to create bird eye to increase the scope of the view. Some approaches are not good under various lighting conditions [7]. Strictly calibrated, fixed position and focuses are used by Matsuyama[8]. This method is difficult to extend in physical resources.

SPS-StFI method [9] uses semi global block matching to construct a semi dense depth map on the reference image. This method says that a pixel is connected within the stereo field, if the semi global block matching does not return an estimate for that pixel. Defined unary cost of a depth in a pixel is the average of the costs of the flow & stereo matching. Under this method they developed two cost functions for stereo & motion pair if does not return an estimate for that pixel.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.

Copyright 2015 University of Aizu Press.

2. SYSTEM OVERVIEW

To develop the 3D algorithm, we use photos which are taken from 3D objects and convert those into 2D environment. This paper finds the relationships between those images (2D images) and recreates the 3D scene back.

This system uses a faster implementation. After acquiring a frame of two images, the system finds the correspondence between right & left images. To do that the proposed approach is key point based matching method. Detecting key points and extract the key points are more important. To extract the key points, scale and meaningful neighborhood is considered.

To extract precise image correspondence, Scale invariant Feature tracker (SIFT) is a well-known method for image matching [10]. Each extracted key point has a descriptor value. These descriptor values are matched using fast approximate nearest neighbor search method. (FLANN) most of the key points have a nearest descriptor value. By using FLANN matcher these key points can be identified. It helps to remove false matches.

Changing octaves count increases the computation time and key points. To get the correct octaves count, image size and resolution needs to be taken. The proposed solution's octaves count is kept as three. Increasing the octave number causes detection of both small and large size features of an image.

Images are taken around the scene. According to matched points, image rectification is applied. Most of the researchers use camera insentric parameters and exentric parameters based on the perspective transformation to image rectification. This paper approach is based on uncelebrated cameras. In order to perform it, fundamental matrix [11, 12] is calculated which is 3X3 matrix which determines corresponding points of images. x and x' are corresponding points in stereo image pair. Fx describes a line on the corresponding point x' on the other image. That means, for all pairs of corresponding points holds as follows

$$x'^T F x = 0 \text{ ----- (1)}$$

To calculate the fundamental matrix, RANSAC algorithm is used. RANSAC algorithm takes inliers and do not take outliers as fundamental points and takes optimal fitting points. Which consider equals or less than eight best points to calculate the fundamental matrix. This ensures that selected points are not affected by noise.

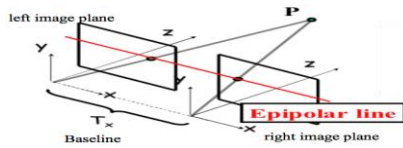


Figure 1: After rectifying two images

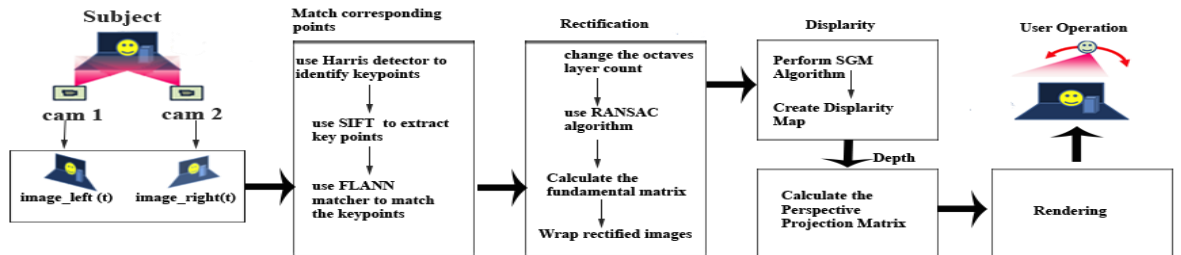


Figure 2: Scematic diagram of 3D interface for disaster management robot teleportation

After finding the corresponding points of images, perspective transformation is applied to output images according to points of input images which is called as rectification.

The image transformation is used under epipolar geometry. This ensures that number of geometric relations between the 2D points and their projection onto the 3D image that lead to constraint's between the image points.

By using rectified images, disparity map is created. Disparity refers to the distance between two corresponding points of the common object images. Every pixel contains the disparity /distance value for that pixel. In this disparity map the brighter shades represent more shifts and lesser distance from the point of view (camera). The darker shades represent a lesser shift and therefore greater distance from the camera.

The Semi-Global Matching stereo method (SGM) [18] is one of the most frequently used depth cost estimation technique for block-matching because it is lightweight and easy to implement, which is a great advantage in real-time applications with limited memory and computing power. Which helps perform correlation methods to correction methods simplicity which assumes that all pixels within the window have the same distance from the camera. For an example, Slanted surfaces and abrupt changes, as caused by depth discontinuities, which helps to avoid result in wrongly including non- corresponding image parts to make a disparity map calculate.

For Right Camera,

$$cx = f \frac{x}{z} \quad y_l = f \frac{y}{z} \text{ ----- (2)}$$

For Left Camera,

$$cx' = f \frac{x - Tx}{z} \quad y_r = f \frac{y}{z} \text{ ----- (3)}$$

Perspective transformation matrix from uncelebrated cameras is a research filed. We used a solution which is based on fronto parall assumption.

cx, cy takes as the principal point. x, y are image coordinates of input images. cx is the image center of x & cy is the image center of y . f is the focal length of camera. Tx is the difference between two cameras or translation between two cameras.

Left camera y and right camera y has the same coordinate. By changing cx & cx' It is possible find the depth of pixel. In order to find the depth of pixel perspective matrix is used (Q).

$$Q = \begin{bmatrix} 1 & 0 & 0 & -cx \\ 0 & 1 & 0 & -cy \\ 0 & 0 & 0 & f \\ 0 & 0 & -1/Tx & (cx - cx')/Tx \end{bmatrix} \text{ ----- (4)}$$

3. EXPERIMENTS

Table 1: Specification sheet one

Machine	Apple MacBook Pro
CPU	2.4 GHz Intel Core i5
Memory	8 GB 1600 MHz DDR3
Graphics	Intel Iris 1536 MB
Data set	technique of [13], published in [14, 15](Art, Books, Dolls, Laundry, Moebius, Reindeer, Computer, Drumsticks, Dwarves)

This section presents performance evaluation. In order to perform evaluation, larger width pixels and height pixels image are reduced by 1, 1.5, 2.0, 2.5... etc. Take the width axis and measure the algorithms computational time. Small images such as 32X 24 types cannot be taken for experimental use which does not help to identify the points. Small images have fewer details which may cause the lack of matching points.

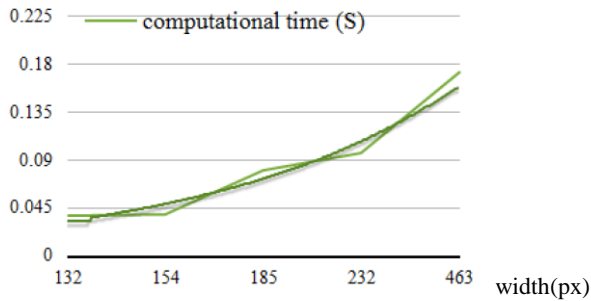


Figure 3: Specification dataset one computational time

The image points matching key points change the octaves count (pyramid layer) from which the key point has been extracted. Which Can be changed and obtain the best octave count. Changing the octaves count may cause high computational time and outliers. Outlier algorithms then take much time to take inliers by considering the facts experimentally. Octaves count is kept under the value of three. Less octaves count causes low matching points. Few matching points directly effects to the 3D model of the system. Low matching points are gradated incomplete 3D scene.

Table 2: By changing octaves count of a stereo frames

Octaves count	Computational time (S)	3D effect
2	0.108	Very bad
2.1 - 2.2	0.104 – 0.105	Very bad
2.3 - 2.8	0.105-0.108	Fair
2.9 - 3	0.107	Good
3.1 - 3.9	0.102-0.1107	Bad
4 - 4.2	0.11	Fair
4.3 - 4.4	0.106-0.109000	Bad
4.5 -5	0.109 -0. 111	Fair
5-10	0.107-0.113	Very bad



Figure 4: illumination levels

An algorithm is checked using various light conditions illumination level 1, illumination level 2 and illumination level 3. Light condition directly affects the accuracy of the 3D model.

Changing the light condition causes the reduction of good matching points of the images. It causes low computational time.

Low lighting conditions are tested using experimented data set. Under illumination 1 to 3 with exposure level 0. Tested images (figure 4)

The algorithm is checked against the light intensity values.

The Computational times of above illuminations are as follows,

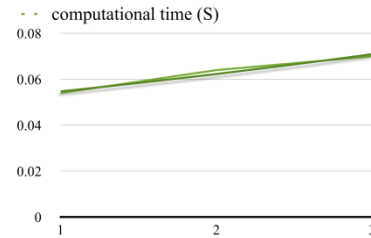


Figure 5: Computational time of figure 4

Table 3: Specification sheet two

Machine	Apple MacBook Pro
CPU	2.4 GHz Intel Core i5
Memory	8 GB 1600 MHz DDR3
Graphics	Intel Iris 1536 MB
Camera	3×PointgrayFlea2 IEEE1394b
Image size	320×240QVGA,RGB888Color
Camera separation	From 10 cm to 50 cm, optional
Camera separation	From 2 m to 4 m, optional

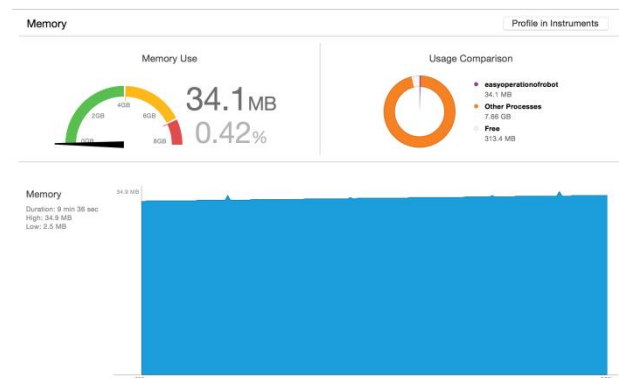


Figure 6: Memory consumption of specification sheet two

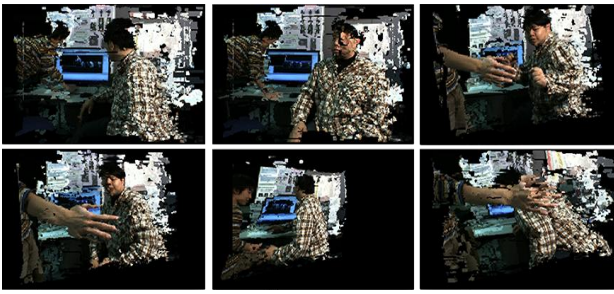


Figure 7: 3D Effect of Specification sheet two

4. CONCLUSION

In this proposed solution, in order to obtain the 3D model baseline of the stereo should be measure by the user. In a robotic environment, it can be planted to know the length. So baseline can be obtained easily.

Most of the real world images are non-linearly deformed when compared with those who captured at a different time or in a different viewpoint. Therefore in order to increase the accuracy of proposed solution it needs to have strict matching. In order to perform, the dense based matching is more suitable instead of using key point based matching. Daisy descriptor is one of the leading dense descriptors which matches images accurately and within less time when compared to SIFT, SURF descriptors.

SIFT uses Key points; Key points are good for analyzing the correspondence of many scenes which are significantly different. Scenes with a small displacement or with a tiny difference can use Dense SIFT or Daisy. Further development needs dense analysis rather than sparse analysis most of the cases. In a disaster situation, cameras can be damaged. The known base line length may change under various circumstances. The factorization method [16, 17] doesn't need the baseline length of cameras. That method needs multiple views (greater than two views) and it assumes affine space instead of the projection space. By using the factorization method, it is possible to calculate the depth of each matched points. Further, above approaches helps to develop a good 3D interface with low memory consumption.

5. ACKNOWLEDGEMENTS

We would like to thank Professor Kaitaro nurse & Associative professor Yuchi Yaguchi (University of Aizu, Japan) wonderful cooperation and helping us dataset and valuable comments throughout the research.

6. REFERENCES

- [1] Kuntze, H., Frey, C. W., Tchouhenkov, I., Staehle, B., Rome, E., Pfeiffer, K., Wenzel, A. & Wollenstein, J. Seneka - sensor network with mobile robots for disaster management. Homeland
- [2] Joochim, C. & Roth, H. Development of a 3D mapping using 2D/3D sensors for mobile robot locomotion. Technologies for Practical Robot Applications, 2008. Tepra 2008. IEEE International Conference on, 10-11 Nov. 2008 2008. 100-105.
- [3] Bruemmer, D. J., Boring, R. L., Few, D. A., Marble, J. L. & Walton, M. C. "I call shotgun!": an evaluation of mixed-initiative control for novice users of a search and rescue robot. Systems, Man and Cybernetics,
- [4] 2004 IEEE International Conference on, 10- 13 Oct. 2004 2004. 2847-2852 vol.3.

- [5] Baker, M., Casey, R., Keyes, B. & Yanco, H. A. Improved interfaces for human- robot interaction in urban search and rescue. Systems, Man and Cybernetics, 2004 IEEE International Conference on, 10-13 Oct. 2004 2004. 2960-2965 vol.3.
- [6] Cui, Y., Schuon, S., Thrun, S., Stricker, D. & Theobalt, C. 2013. Algorithms for 3D shape scanning with a depth camera. IEEE Trans Pattern Anal Mach Intell, 35, 1039-50.
- [7] Sato, T., Moro, A., Sugahara, A., Tasaki, T., Yamashita, A. & Asama, H. Spatio-temporal bird's-eye view images using multiple fish-eye cameras. System Integration (SII), 2013 IEEE/SICE International Symposium on, 15-17 Dec. 2013 2013. 753-758.
- [8] Xiaojun, W., Takizawa, O. & Matsuyama, T. Parallel Pipeline Volume Intersection for Real-Time 3D Shape Reconstruction on a PC Cluster. Computer Vision Systems, 2006 ICVS '06. IEEE International Conference on, 04-07 Jan. 2006 2006. 4-4.
- [9] Matsuyama, T. Exploitation of 3D video technologies. Informatics Research for Development of Knowledge Society Infrastructure, 2004. ICKS 2004. International Conference on, 1-2 March 2004 2004. 7-14.
- [10] K. Yamaguchi, D. Mcallester and R. Urtasun: Efficient Joint Segmentation, Occlusion Labeling, Stereo and Flow Estimation. ECCV 2014.
- [11] D. Lowe, "Distinctive image features from scale invariant keypoints," IJCV, vol. 60, no. 2, pp. 91-110, 2004.
- [12] Q.T. Luong and T. Vieville, "Canonic Representations for the Geometries of Multiple Projective Views," Computer Vision and Image Understanding, vol. 64, no. 2, pp. 193-229, 1996.
- [13] Bartoli, A. & Sturm, P. 2004. Nonlinear estimation of the fundamental matrix with minimal parameters. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 26, 426-432.
- [14] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003), volume 1, pages 195-202, Madison, WI, June 2003.
- [15] D. Scharstein and C. Pal. Learning conditional random fields for stereo. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007), Minneapolis, MN, June 2007.
- [16] H. Hirschmüller and D. Scharstein. Evaluation of cost functions for stereo matching. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007), Minneapolis, MN, June 2007.
- [17] Triggs, B. Factorization methods for projective structure and motion. Computer Vision and Pattern Recognition, 1996. Proceedings CVPR '96, 1996 IEEE Computer Society Conference on, 18-20 Jun 1996 1996. 845-851.
- [18] Vidal, R., Soatto, S. & Sastry, S. S. A factorization method for 3D multi-body motion estimation and segmentation. Proceedings of the Annual Allerton Conference ON Communication Control and Computing, 2002. Citeseer, 1626-1635.
- [19] Hirschmuller, H. Accurate and efficient stereo processing by semi-global matching and mutual information. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 20-25 June 2005 2005. 807-814 vol. 2.

Entity Resolution using Co-occurrence Graph and Continuous Learning

Anoop Kumar Pandey
International Institute of Information Technology
Bangalore
26/C Electronics City
Bangalore, India - 560100
anoopmis@gmail.com

Srinath Srinivasa
International Institute of Information Technology
Bangalore
26/C Electronics City
Bangalore, India - 560100
sri@iiitb.ac.in

ABSTRACT

In a community setting, *Utilitarian Knowledge* or “Knowledge that works” are routinely diffused through social media interactions. The aggregation of this knowledge is a divergent process, where common knowledge gets segregated into several local worlds of utilitarian knowledge. If the community as a whole is coherent, these different worlds end up denoting different aspects of the community’s dynamics. To capture and represent this knowledge, several data models have been proposed. One of the model organizes concepts (atomic or simpler elements) in a hierarchy namely concept hierarchy (“is-a”) in which concepts are added manually at the most appropriate level inside the hierarchy. To minimize manual intervention in entity resolution, this paper proposes entity resolution based on co-occurrence graph and continuous learning, thereby eliminating the bottleneck of manual concept entry. While traditional Supervised Learning methods require sufficient training data before hand which is not available in a community setting at start, Continuous Learning method could be useful which can acquire new behaviours and can evolve as the community data evolves.

Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous;
D.2.8 [Web Science]: Metrics—*performance measures*

General Terms

Algorithms, Performance

Keywords

Knowledge Base, Co-occurrence Graph, Utilitarian Knowledge, Continuous Learning

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

1. INTRODUCTION

In on-line communities, several netizens interact and exchange or share large amount of knowledge among themselves. The knowledge, they share, comes in two flavours: Encyclopaedic or Informational Knowledge & Utilitarian Knowledge that can be put to use. For maintaining Encyclopaedic knowledge many readymade ontologies like DBpedia are available, however there is no such knowledge base to capture Utilitarian Knowledge. To capture and represent the same, a data model called Many Worlds on a Frame (MWF)[4] was proposed. The data model contains several concepts organized in hierarchies. Concept in general, could refer to all the terminologies and vocabulary of a particular domain which is used to describe it. The definition of concepts and relationship between the concepts is typically captured using a simple structure called concept tree that captures two kinds of relationships: ‘is-a’ and ‘is-in’. For instance, the concepts ‘Java’ **is-a** ‘Programming Language’ depicts ‘is-a’ relationship while “Bangalore” **is-in** “Karnataka” depicts containment. These hierarchies contain many concepts beyond the encyclopaedic concepts which are relevant to that domain. Therefore in such kind of knowledge base, concepts need to be added manually and hence poses a bottleneck in evolution of the knowledge base. A need for an entity resolution system, that could, given a concept, label it with proper ‘is-a’ parent thereby placing it appropriately in the concept hierarchy, could help the knowledge base to grow with minimal manual intervention. While conventional Supervised Learning methods require sufficient training data before hand which is not available in a community setting at start, Continuous Learning method could be useful which can acquire new behaviours and can evolve as the community data evolves.

2. APPROACH

We formulate the problem of finding ‘is-a’ parent in community knowledge base as a class labelling task. Considering co-occurrences to be the only observed facts in the community posts, we wish to generate a signature for each class (is-a label) using co-occurrence graph. While the co-occurrences between terms in a document represent the associations between terms, signature for a class ‘C’ represents a vector with co-occurring classes as dimensions and their average co-occurring frequency with ‘C’ as the value along that dimension. The knowledge base, would then predict ‘is-a’ parent for an unknown concept by comparing its co-occurrence neighbours with the class signatures using cosine similar-

ity and shall recommend the label having highest similarity. The signatures will be improvised upon user feedback and as when a new instance of that class is added to knowledge base.

2.1 Definitions

- **Co-occurrence Graph:**[2] A graph data structure that maintains a weighted set of co-occurrences of terms across the corpus. This graph approach is related to word clustering methods, where co-occurrences between words can be obtained on the basis of grammatical or collocational relations[5]. Formally we define the undirected co-occurrence graph G as

$$G = (T, P, w)$$

where, T is the set of all terms in the given corpus. P is the set of all pair-wise co-occurrence between terms in T . The function $w : P \rightarrow \mathbb{N}$ is the corresponding pair-wise co-occurrence count.

- **Continuous Learning:** Continuous Learning [3] is the process of constant improvement with no fixed end and the final goal is improvement itself. A continual learning algorithm is characterized by the following properties: (i) It should be autonomous. It must behave in its environment and be able to assign credits to desirable and undesirable behaviours. (ii) These behaviours should be capable of spanning for arbitrary periods of time i.e. the duration of a behaviour is not determined before hand. (iii) Continual Learning algorithms should acquire new behaviours only when useful.

2.2 Algorithm for Signature Generation for a class C

Before the start of algorithm, presence of co-occurrence graph, concept & containment hierarchy is presumed in the system. A class(or label) typically represents a “is-a” parent while instance represent all concepts that inherit that class. For example “Person” is a class while “Anoop” is an instance of “Person” class. Hierarchically ‘Anoop’ is-a ‘Person’.

Algorithm for generating the signature is depicted in Algorithm 1

2.3 Algorithm for ‘is-a’ parent determination for an unknown concept

Prologue: This algorithm doesn’t resolve the class of the unknown entity at an instant when a text segment is posted in a community website, rather it is resolved at a later stage subject to the condition given in Algorithm 2.

Algorithm for finding ‘is-a’ parent of an unknown entity is depicted in Algorithm 2

3. RESULTS

Dataset: Seekha (<http://www.facebook.com/seekhain>) is a academic networking portal. We used the dataset of 11K Seekha concepts and co-occurrence graph for our experiment. The dataset was already labelled.

Experiment and Results: From the co-occurrence graph, we created signatures for 20 main classes. The graph wasn’t dense enough to create more signatures. Then we took 500

Algorithm 1 Algorithm for signature generation for a class

Find all instances of class C from the concepts hierarchy and store them in an array ‘ N ’.

From the co-occurrence graph find all the co-occurring concepts corresponding to each element of ‘ N ’ and store them in a 2-D associative array ‘ NC ’ such that key refers to a concept in ‘ N ’ and the value refers to an array which holds all co-occurring concepts corresponding to the key concept.

From the concept hierarchy we find the “is-a” parent of the co-occurring concepts in ‘ NC ’ and store it in a 2-D array ‘ SC ’ such that key refers to a concept in ‘ N ’ and the value refers to a multi-set array which holds “is-a” parent of all co-occurring concepts corresponding to the key concept. Find the union of multi-set array ‘ SC ’ and find the signature as

$$\vec{S} = \sum_{k=1}^X j_k * CC_k$$

where CC_k is the concept class in the union of multiset ‘ SC ’, X being the total number of distinct classes CC_k and coefficient

$$j_k = \frac{\text{total no of occurrences of } CC_k}{\text{total no of instances of } C}$$

To narrow down the components of the signature \vec{S} use 80-20 rule [1]. For applying this rule, order the components in descending order of their coefficient and select top u components CC_i for which the sum of co-efficient is 80% of the sum of co-efficient of all the ‘ n ’ elements.

$$\sum_{i=1}^u i = 0.8 * \sum_{i=1}^n i$$

Algorithm 2 Algorithm to find ‘is-a’ parent

Prologue: Presence of Co-occurrence Graph, Class Signature Database created using section 2.2 and Concept & Containment Hierarchy

Algorithm:

When a text segment is posted in a community website, extract concepts using a concept extraction algorithm and populate them in the co-occurrence graph.

Any new entity, not present in the knowledge base, is labelled as “Unknown” in the ‘is-a’ hierarchy.

Over time many of the unknown concepts in ‘is-a’ hierarchy are manually labelled randomly.

When an unknown entity has enough labelled neighbours (threshold can be heuristically set. For e.g. at least 5 and 50% of all neighbours) in the co-occurrence graph, we create a candidate vector composed of ‘is-a’(label) of the co-occurring neighbours of the unknown entity.

This candidate vector is compared against the signature database using cosine similarity to determine the most matched label.

Upon resolution of unknown entity to a class, signature of that class is updated using algorithm 1

If an unknown entity is resolved incorrectly, the threshold of labelled neighbours around an unknown entity is adjusted. The threshold for no. of instances and no. of labelled neighbours around the instances that participate in signature generation for a class, is also adjusted.

instances(concepts) assuming them as unknown that belonged to these classes, and determined their ‘is-a’ parent algorithmically. We were able to get 62% accuracy (percentage of concepts correctly labelled algorithmically) in class identification against the actual class for the instances.

4. CONCLUSION AND FUTURE WORK

We used the co-occurrence graph in a continuous learning environment to generate signature for a class (label). The signature database is later used to resolve unknown entities entered into the knowledge base. The signature for a class is updated when a new instance of that class is added to the system. We didn’t use traditional supervised learning method since in a community setting, we didn’t have training data in advance. The knowledge base evolves over time in community setting. Also the data churn in a community is tremendous. Introduction of new class can be easily accommodated in a continuous learning environment.

5. REFERENCES

- [1] D. M. Powers. Applications and explanations of zipf’s law. In *Proceedings of the joint conferences on new methods in language processing and computational natural language learning*, pages 151–160. Association for Computational Linguistics, 1998.
- [2] A. R. Rachakonda, S. Srinivasa, S. Kulkarni, and M. Srinivasan. Mining analytic semantics from unstructured text. Technical report, Technical report, 2012.
- [3] M. B. Ring. *Continual Learning in Reinforcement Environments*. PhD thesis, University of Texas at Austin, 1994.
- [4] S. Srinivasa. Aggregating operational knowledge in community settings. In *On the Move to Meaningful Internet Systems: OTM 2012*, pages 789–796. Springer, 2012.
- [5] D. Widdows and B. Dorow. A graph model for unsupervised lexical acquisition. In *Proceedings of the 19th international conference on Computational linguistics- Volume 1*, pages 1–7. Association for Computational Linguistics, 2002.

Data Collection for Investigation of Reliable Reviews

Jun Kikuchi

Software Engineering Lab, University of Aizu
Tsuruga, Ikki-machi, Aizu-Wakamatsu City,
Fukushima, 965-8580 Japan
(+81)242-37-2603
s1190017@u-aizu.ac.jp

Vitaly Klyuev

Software Engineering Lab, University of Aizu
Tsuruga, Ikki-machi, Aizu-Wakamatsu City,
Fukushima, 965-8580 Japan
(+81)242-37-2603
vkluev@u-aizu.ac.jp

ABSTRACT

Nowadays the Internet is growing quickly. People use for many different purposes, mainly for online shopping. However, there is a large number of websites, which offer similar products and features. It makes a choice for customers difficult. As a result, they refer to online reviews. The main purpose of this research is to improve a reliability of a review. Increasing reliability is useful and helpful for customer's decision, and it can also filter out spam reviews. We are detecting a pattern by morphological analysis. Obtaining a reliable review pattern could be used to collect useful reviews automatically. Results of this study can be applied to any kinds of reviews and opinions.

Categories and Subject Descriptors

H.3.3[Information Storage and Retrieval]: Information Search and Retrieval – *information filtering*

General Terms

Algorithms, Experimentation, Languages, Theory

Keywords

pattern, text mining, opinion extraction

1. INTRODUCTION

Nowadays, many people access to the Internet for many different purposes. The main purpose is to obtain useful and helpful information about products, places, restaurant, and etc. There are a lot of websites and web applications which provide this information, and most people rely on them as important tool because it is easy to exchange opinion between individuals. In addition, many people search for products and buy them online, so reviews are important and necessary because they provide customers with an easy way to evaluate a product. Changing a style of purchasing makes review websites and applications to analyze user features more accurate. These review sites and applications, however, do not cover all kinds of products, services, or places which customers want to know in detail. Moreover, these review websites and applications utilize different evaluation approaches. Another problem is that we are

not certain that these reviews are reliable and precisely.

To evaluate a product, we need a measurement and an evaluation model. For example, people could not rely on a certain review which is provided by a few users because reviews reflect a personal opinion. This research collected certain amount of opinions and reviews expressed by people who experienced or used products or services before to seek some features of useful reviews.

Our study goal is to create a valuable and reliable review evaluations system in different fields. We are planning to detect standardized patterns of these reviews. The first step is creating and developing a review corpus. We will find features of valuable review by part of speech (POS) tagging. We will also make these reviews up-to-date and current because people always need actual information to help in their decision making processes. One solution is social networking sites (SNS). The SNS provides customers with an opportunity to exchange information easily by any kinds of devices. The SNS are creating different kinds of information quickly in large quantities. Another merit of using this kind of data is that these opinions are very frank and clear, so product evaluation is more reliable. An application which collects and analyzes big data for gathering better reviews is necessary and helpful in the current circumstance.

2. RELATED WORK

2.1 Opinion Mining

Opinion mining applied to the review analysis deals with types of reviews and problems in this area. Reviews have many kinds of writing style and a large variety of words, so opinion mining concentrates on the automatic analysis of the reviews in natural language.

Generally, opinion mining organizes and obtains an entity, aspect, and opinion orientation. An entity is a target object such as a product that has been evaluated by someone. An aspect mentions a component or a sub-component of the entity. An opinion orientation expresses positive, negative or neutral attitude in an opinion sentence. These three components are essential to analyze any kind of reviews. This basic organization of reviews has been used to solve a lot of problems of the review analysis such as sentence subjectivity, sentiment classification, lexical expansion, and etc. [1]

Reference [2] used a support vector regression (SVR) based outlier detectors which designed by collecting reviews on the Internet. The SVR detector predicts a score of hotel ranking to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.

Copyright 2015 University of Aizu Press.

improve a reliability of hotel ranking. This research did morphological analysis to collect an opinion dictionary using part of speech (POS) pattern. This dictionary focused on patterns of adverb and intransitive adverb. Therefore, a feature value assigned each opinion word in this dictionary. As a result, the SVR detector witch trained by the opinion dictionary also has a feature of TF-IDF. These features help to analyze reviews on the Internet and to measure reliability.

3. APPROACH

3.1 Implementation

We analyzed several kinds of reviews on the Internet, and investigated critical and important information that a better review has. Accordingly we collected reviews from the Amazon.co.jp website. Amazon offers many kinds of products. In this study, our objection is to find out features of useful reviews, so we created a review corpus from gathering reviews. We gathered 169 reviews on the Sporting Goods from Amazon.co.jp. In addition to, it has implemented a rating system that asks customers for a feedback on a review. We also used this rating system for judging whether a review is useful or not.

Amazon provides developers with the APIs to obtain information for products, so we used it to obtain title, reviews, products' name, and reviews' rating to analyze patterns or template for good reviews. All reviews are only in Japanese. We analyzed opinion sentences with MeCab and TermExtract [3] in order to look at POS's features from collected reviews. Based on this POS tagging, we investigate a reliable and useful review.

3.2 Data Collection

We divided all gathered reviews into two groups. One includes opinion sentence, and another one consists of non-opinion sentences. We defined a difference between a review and an opinion. An opinion should include a target and subjective expression. An opinion target must be related a product or this aspects, and subjective expression are personal feelings, views or beliefs. In this study, opinion sentences are defined by two main types: regular opinion and comparative opinion. A regular opinion commonly expresses an opinion on any aspects of products. On the other hand, a comparative opinion expresses a relation of similarities or differences between two or more products or aspects. All opinion sentences commonly include positive, negative, or neutral sentences, emotion, attitude, etc. Table 1 summarizes collected Sporting Goods reviews. It shows two data. One is a number of reviews in Sporting Goods category. Second data shows numbers after calculating an average of positive feedback numbers and an average of all feedbacks number. A rate of reviews must be evaluated by more than 5 customers to obtain more accurate template. In this case, about 24 customers evaluated a review, and about 22 customers think a review is useful on average. Table 2 shows summarizing data on two types of sentences: opinion sentences or non-opinion sentences.

Table 1. Statistics of Sporting Goods Review Data from Amazon

Collection	Number of Reviews	Number of Positive Feedbacks / Number of All Feedbacks
Sporting Goods	169	22.54/24.51

Table 2. Statistics of Sporting Goods Review for each Review Group

Collection	Number of Sentences	Number of Words
All Sentences	839	27193
Opinion Sentences	532	17299
Non-Opinion Sentences	307	9894

3.3 Collected Data Analysis

All collected reviews are rated as a useful review by Amazon customers, so it is critical to look at any features in these collected reviews. Opinion sentences are more informative. They are focused on the product. Therefore, we analyzed only collected opinion sentences. As a first step of the analysis, we did POS tagging to all opinion sentences with MeCab and TermExtract. Important information of a review is opinion targets, subjectivity and emotions. A noun could indicate an opinion target or an aspect, and an adjective also could indicate emotions and subjectivity. Because of these reasons, we concentrated on two POSs, which are noun and adjective. In this study, these two word lists are created from collected opinion sentences. Table 3 shows the result after gathering on two POSs words.

Using these created POSs lists, we automatically collected sentences which include two POSs as a keyword from all collected reviews. This second step of the analysis identifies how these two POSs relate to important opinion sentences. Moreover, we investigated what kinds of words are necessary for a review corpus. Table 4 summarizes the result of the analysis. It shows a number of collecting sentences as an opinion sentence (OS) by each word lists and a number of correct opinion sentences. Correct opinion sentences include a manually collection of opinion sentences, and have at least one word in POSs lists. Accuracy is calculated by dividing a number of correct opinion sentences by a number of all collected sentences.

Table 3. Statistics of Nouns and Adjectives

Collection	Number of Words
Nouns	277
Adjectives	33

Table 4. Statistics after each POS collection

Collection	# OS	# Correct OS	Accuracy
All Noun	751	437	58.20
Top 20 Noun	563	315	55.95
Top 10 Noun	486	267	54.94
All Adj	469	284	60.55
Top 20 Adj	400	263	65.75
Top 10 Adj	365	242	66.30
Top 10 Noun AND Top 10 Adj	229	141	61.57
Top 10 Noun OR Top 10 Adj	622	367	59.00

4. DISCUSSION

Collected data on noun and adjectives show us two points. One point is that noun keywords have too many varieties and less duplication to obtain better collection of opinion sentences. In this case, a noun list has many words, so we obtain automatically correct opinion sentences with a high rate. Each noun collection, however, is not better accuracy than adjective collections. Adjective keywords could not gather much number of opinion sentences, but the collection of top 10 adjective keywords result on a highest accuracy. In this search area, adjective always treat as a special keyword because an opinion sentence mainly expresses based on adjective words. Therefore, an adjective word is an important issue to achieve a better pattern of reviews, so we need a deeper analysis about adjectives. Moreover, combination of nouns and adjectives shows an interesting and better result. Collected opinion

sentences that use top 10 noun keywords or top 10 adjective keywords have a good number of correct opinion sentences. It means that we need to inspect opinion sentences in details in order to know a relation between noun keywords and adjective keywords.

5. CONCLUSION

We collected a Japanese corpus of reviews from Amazon.co.jp. We conducted a preliminary analysis of every reviews manually and automatically to find the common features or patterns. We concentrated on analysis of opinion sentences. An analysis of opinion sentences in the way of inspecting adjectives and nouns shows relation between two POSs and opinion sentences. This data is necessary information for creating a better review corpus. However, using two POS words picks many non-opinion sentences up, and it missed a lots opinion sentences as well.

Therefore, this analysis needs deeper research to seek a feature or pattern. To increase the accuracy in obtaining correct opinion sentences, we need to analyze details of nouns, adjectives, and relations of two types of words. Moreover, it is necessary to study that these keywords could achieve effectively opinion sentences in different categories.

6. REFERENCES

- [1] Liu, B. and Zhang, L. 2012. A survey of opinion mining and sentiment analysis. *Mining Text Data*, 415-463.
- [2] Hsieh, H. Klyuev, V. Zhao, Q. and Wu, S. 2014. SVR-based outlier detection and its application to hotel ranking. *In Proc. of the 2014 IEEE 6th International Conference on Awareness Science and Technology (iCAST)*, 1-6.
- [3] Miyashita, M. and Klyuev, V. 2014. TermExtract: Accuracy of Compound Noun Detection in Japanese. *Future Information Technology*, 189-194.

Personal Data Leakage: Android Case Study

Yoichi Saito
University of Aizu
Tsuruga, Ikki-machi Aizu-Wakamatsu,
Fukushima, Japan 965-8580
+81-242-37-2603
s1210196@u-aizu.ac.jp

Vitaly Klyuev
University of Aizu
Tsuruga, Ikki-machi Aizu-Wakamatsu,
Fukushima, Japan 965-8580
+81-242-37-2603
vkluev@u-aizu.ac.jp

ABSTRACT

Android has gained popularity explosively in these days. Android has many security problems. Because of this, various incidents have occurred. A typical example is the leakage of personal information. This incident occurred in the Skype application. Android security features are important to prevent the leak of sensitive information. This study characterizes security issues in Android OS.

Categories and Subject Descriptors

K.6.5 [Management of Computing and Information Systems]: Security and Protection.

General Terms

Security

Keywords

Android Security, Information personalization

1. INTRODUCTION

It is difficult to resolve all possible security issues while developing Android apps because the cost of development should be low. However, if the incident with the application occurs, it is more difficult to recover from losses [1]. Security risks are spreading on the Internet so developers should pay more attention to the problem. In this study, we characterize key security mechanisms.

2. ANDROID SECURITY MECHANISMS AND COMPONENTS

2.1 Android Security Mechanisms

Android security architecture is composed of certificates, user ID, permission and access permission to the files. By combining these elements, Android OS presents the secure structures as follows [2]:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.

Copyright 2015 University of Aizu Press.

- Applications cannot access reciprocal data directly.

- Applications can access determinate parts within the functions of Android.

2.1.1 Cryptography

Android disk encryption is based on dm-crypt, which is a key feature that works at the block device layer. Because of this, encryption works with Embedded MultiMediaCard and similar flash devices that present themselves to the kernel as block devices. Encryption is not possible with YAFFS, which talks directly to a raw NAND flash chip [3].

2.1.2 IPC

Binder and Messenger are classes. They provide the very useful interface. They are preferred for Android IPC. A developer should not design the interface that requires the special permission because objects of the binder and Messenger are not declared within the manifest. As a result, the developer can't change the permission of them. They inherit the permission of the Service or Activity so their permissions need to be set appropriately [4].

2.1.3 SandBox

Android applications follow the rules SandBox framework an application. If it needs to access the resources and data that SandBox doesn't provide, it must declare the access to OS. This declaration is set in the AndroidManifest.xml by the permission [2].

2.2 Components

Components are the elements that compose the application.

- Activity

It provides the function of the user interface [2].

- Service

It doesn't have the user interface. It keeps processing in the background even if other applications are started [2].

- BroadcastReceiver

It receives the broadcast from OS or the other application [2].

- ContentProvider

Thanks for ContentProvider, an application can abstract the entity of the data when it exposes the data outside [2].

3. Data Leakage from the Telephone Directory

Some useful mechanisms are prepared for developers but there are issues to concern. In this section, we characterize these issues.

3.1 Problem Description

When the users exchange the telephone numbers, there is a way that the telephone number is exported to the QR code by the application and the camera of the receiver takes the QR code and gets the telephone number. When this application is installed, “reading the data of the telephone directory” permission is displayed. The user does not expect that this application has the Internet access.

This application may collect data away safe it to the data based on the Internet. This database is the source for data leakage [2].

3.2 How to Get the Personal Information

If the third party analyzes this application, they may know that it is possible to access the ContentProvider. He or she who is malicious develops the application to steal the personal information by using fragilities. Suppose it is an application to download the wallpapers. When this application is installed, “Network Communication/Complete Internet Access” permission is displayed naturally. The user installs this application without any suspicion. But this application does not only download the wallpapers but also pick up the personal information and send them to the server. In Android OS, the application process is defended against the other application, so the other processes do not know how data is send. Any security software cannot detect such a movement [2].

3.3 The Matter to Be Considered

The user has the personal information stolen by the application. To solve this problem, the contentprovider that accesses the personal information should set the appropriate access permission. [2].

A contentprovider is declared in <provider> element of the AndroidManifest.xml file. We characterize two distributions below:

3.3.1 *android:enabled distribution*

This distribution assigns whether this contentprovider is used or not. If the value of “true” is set, this contentprovider can be used. If the value of “false” is set, this contentprovider can’t be used. The default value is “true” [2].

3.3.2 *android:exported distribution*

This distribution controls the setting of the contentprovider extension and assigns whether this contentprovider can be used by the other applications. If the value of “true” is set, all applications can use this contentprovider. If the value of “false” is set, the other application can’t use it, and this application itself or the application that has a same user ID by shareUserId assignment only access this contentprovider [2].

In general, the declarations of the access permission of the other components are done in the same way.

4. CONCLUSION

Android devices are very popular nowadays. They are under attacks from the Internet. Various tools are developed to defend them. Access permission to files is used for android security as well as PC security. In addition, an Android security architecture has the AndroidManifest.xml file. The appropriate definitions in this file make the application secure. However, the attention of developers is more important because they cannot resist attacks unless they notice security risks. Every developer should clearly understand that security techniques have limitations.

5. REFERENCES

- [1] 情報セキュリティ読本 (Information Security Reader, in Japanese). Information-technology Promotion Agency, Jikkyou Publisher, Japan, 2012.
- [2] Taniguchi G., 安全なアプリケーションを作るために (Development of the Secure Application, in Japanese). Impress Japan, Tao Software Corporation, 2011.
- [3] Encryption. Retrieved August 26 , 2015 from Android Open Source Project:
<http://source.android.com/devices/tech/security/encryption>
- [4] Security. Retrieved August 26 , 2015 from Android Open Source Project:
<https://source.android.com/devices/tech/security/>

A Consideration of Query Improvement to the SMOKA Astronomical Archive System

Yilang Wu
University of Aizu
Aizu-wakamatsu
Fukushima, Japan, 965-0085
d8152103@u-aizu.ac.jp

Wanming Chu
University of Aizu
Aizu-wakamatsu
Fukushima, Japan, 965-0085
w-chu@u-aizu.ac.jp

Subhash Bhalla
University of Aizu
Aizu-wakamatsu
Fukushima, Japan, 965-0085
bhalla@u-aizu.ac.jp

ABSTRACT

The Web technology has brought the significant changes to astronomical archive services. In this study, we investigate the limitations in data search and query of the SMOKA (Subaru-Mitaka-Okayama-Kiso-Archive) astronomical archive system, which heavily depends on Java Server Pages (JSP) and Servlet web technology. By discussing the new trends of the software systems for astronomy, we propose to utilize the emerging or popular web technologies and also the novel query interfaces to improve the query service in SMOKA.

Keywords

Astronomical Data Archive, SMOKA, Query, Web Interface

1. INTRODUCTION

The Subaru-Mitaka-Okayama-Kiso-Archive (SMOKA) is a public science archive system which provides access to the data of the Subaru Telescope, the 188 cm telescope at Okayama Astrophysical Observatory, and the 105 cm Schmidt telescope at Kiso Observatory/University of Tokyo [1]. The normalization of Flexible Image Transport System (FITS) keywords among various instruments of the SMOKA associated Telescope makes it easy to construct and maintain the database. As a successor of the MOKA3 [2], the SMOKA is mainly constructed using Web technology such as Java Servlet and JavaServer Pages, which is fit to the component-based development for the multi-tiered system architecture. The UI-tier of SMOKA provides a well-defined interface to access the data. The middle-tier is in charge of retrieving the data from the back-end database-tier. Such component-based approach and multi-tiered architecture derived a number of separated search functions (see Table 1).

However, such system framework and architecture is insufficient for query needs and further development nowadays, because of its separated search interfaces and duplicated search functions. The domain specific users [3] such as amateur or professional astronomers prefer 1) easily understandable user interface, for instance with the style of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IWAIT '15, Oct. 8 – 10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

Table 1: SMOKA Search Functions and Interfaces

Search Types	Archives					Mixed With	Search Functions
	S	O	K	M	H		
Simple Search	○	○	○	○	○	—	Search data from a list of object names
Advanced Search	○	○	○	○	○	SUP, KWF	Data from various search constraints.
SUP Search	○	○	○	○	○	—	Search raw data and astrometric calibrated data.
KCD (pin-point) search (KCD)	○	○	○	○	○	—	Search frames whose fields contain the search coordinates.
Calendar Search	○	○	○	○	○	—	Search data from a calendar (observation date) with weather data.
MB Search	○	○	○	○	○	SUP, KCC, KCD	Search data which may contain minor bodies in the solar system.
All Keywords Search	○	○	×	×	○	—	Search data by FITS Header Keywords values.
Full-Text Search	○	○	○	○	○	—	Search any words in FITS headers (including History and Comment).
Overlapped Area Search (Old & New)	○	○	○	○	○	SUP, KCD	Search areas(s) observed multiple times.
Summit Log Search	○	×	×	×	×	—	Search any words in Summit Logs.
Number of Frame Search	○	○	○	○	○	—	Search number of frames and IDs from object coordinates.
Description	S (Subaru), O (OAO), K (Kiso), M (MITSuME), H (HHO) ○ (Supported), × (Not Supported)						

QBO (Query-By-Object) [4]; 2) facilitating work-flow between query steps, such as the support of multi stage query [5]; 3) reusing or adapting existing services or interfaces as much as possible, and eliminating the duplicated features.

In this study, we investigate the system limitations in SMOKA and analyze the prominent software systems which have been widely used in recent popular astronomical data archives, and propose to utilize the emerging web technology and interface to improve the query service in SMOKA.

2. BACKGROUND

On the top of the emerging web technologies, the once monolithic and file-server oriented web servers are evolving into easily programmable server applications capable to cope with the complex interactions made by new generation of browsers [6]. Accessing astronomical data archives, schedule, control and monitor observatories, and in particular robotic telescopes, supervising data reduction pipelines, all are capabilities that can now be implemented in a JavaScript web application.

The emerging high-end web technology with open and well documented interface is considered to be the sustainable approach to further advance many scientific and astronomical data resources and systems. It is a trend that many astronomical data archives are merging together into larger data repository with more advanced search and query interface.

The CDS portal¹ is such a mash-up application of various web based searching services for astronomical archive data. It aims at providing a single entry point to search and access the different CDS services (such as the Simbad VizieR, Aladin, X-Match) and facilitating the work-flow between the services. Nevertheless, it opens rich programming interfaces for developers, including various Java libraries, Unix/Linux clients, XML Web services, jQuery Widget scripts; furthermore, it also targets the smartphone OS (Android and iOS), HTML5 and WebGL developers. Various sources code of CDS Portal components is available on public (for instance via Github²) with detailed technical documents, and it has formed a large open source communities, connecting with the world wide users and developers.

3. DISCUSSION

The JSP, which is used by SMOKA is better to be treated only as a presentation (or representation) technology, not as the central structuring technology of the application. The problems of developing a pure JSP web-app without framework are endless; for instance there are limited abilities to usefully re-use code across pages, difficult error handling, limited ability to implemented caching strategies, tedious validation logic, and constant battle to ensure HTTP GETs. Thus the SMOKA adopted the JSP together with Servlet to overcome that limitations, and provide a component-based, platform-independent method for building Web-based applications, without the performance limitations of Common Gateway Interface (CGI)³ programs. The Servlets is advanced in the following features: 1) portability across operating systems and across web servers; 2) harnessing the full power of the core Java APIs, such as networking and URL access, multi-threading, image manipulation, data compression, JDBC, object serialization, internationalization; 3) efficiency and endurance in memory manipulation; 4) supporting safe programming since it inherits Java's strong type safety, exception-handling mechanism; 5) elegant with clean code and models; 6) tightly integrated with the server. However the design in Servlet is difficult and slows down the application; developers have to write complex business logic, making the application difficult to understand; the maintenance of Java Runtime Environment on sever side results in extra work to administrators.

There is increasing demand of easy to learn and start with framework and technology to build data-intensive real-time applications running across distributed devices. Along with the development of HTML and HTTP, JavaScript, the browser-side programming language has become a lot more powerful. The Node.js⁴ is such popular technology, with coders everywhere using it to create APIs and build a new matrix of interoperability across the Internet. Furthermore, according to the experiment in [7], Node.js performs much better than the traditional technology such as PHP in high concurrency situation, no matter in benchmark tests or scenario tests. PHP handles small requests well, but struggles with large requests. Besides, Node.js prefers to be used in

the IO-intensive situation, not compute-intensive sites; but that is enough to support the query to the astronomical big data. On the other hand Python-Web is also not suitable for the compute-intensive website.

4. PROPOSAL

A series of novel web technologies, such as Node.js with C++ *add-ons*, SVG⁵ and HTML5 web based visualizing tools, MongoDB for complex data structures have been proposed for astronomy in paper [6]. Based on the aforementioned technologies, we propose to upgrade the software system in SMOKA by integrating the the existing but independent search functions (see Figure 1) into a simple query interface such as the QBO and multi stage query [5], in which users can make complex query with guidance in steps. And to fully utilize the advanced new web technologies, the open and well-documented programming interface for data access is highly suggested to connect the world wide developers.

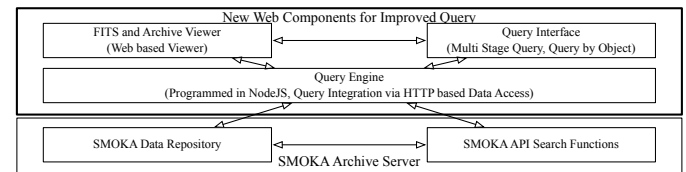


Figure 1: Proposed System Architecture

5. SUMMARY AND FUTURE WORK

To ease the system development and maintenance and to improve user experience for SMOKA, it is time to embrace the new generation of Web technology and more open community. We plan to build new query application by introducing the simple query interface like QBO, multi-stage query work-flow, and also eliminating the duplicated functions.

6. REFERENCES

- [1] H. Baba, N. Yasuda, S.-I. Ichikawa, M. Yagi, N. Iwamoto, T. Takata, T. Horaguchi, M. Taga, M. Watanabe, T. Ozawa, and M. Hamabe, "Development of the Subaru-Mitaka-Okayama-Kiso Archive System," *Adass Xi*, vol. 281, pp. 298–+, 2002.
- [2] T. Horaguchi, E. Nishihara, M. Yoshida, K. Aoki, T. Ito, M. Watanabe, S. Ichikawa, T. Takata, S. Yoshida, and M. Hamabe, "An astronomical data archive system with a java-based user interface," *Publications of the Astronomical Society of Japan*, vol. 51, no. 5, pp. 693–701, 1999.
- [3] A. Madaan and W. Chu, "Handling domain specific document repositories for application of query languages," in *Databases in Networked Information Systems* (A. Madaan, S. Kikuchi, and S. Bhalla, eds.), vol. 8381 of *Lecture Notes in Computer Science*, pp. 152–167, Springer International Publishing, 2014.
- [4] A. Yasir, M. Kumara Swamy, P. Krishna Reddy, and S. Bhalla, "Enhanced query-by-object approach for information requirement elicitation in large databases," in *Big Data Analytics* (S. Srinivasa and V. Bhatnagar, eds.), vol. 7678 of *Lecture Notes in Computer Science*, pp. 26–41, Springer Berlin Heidelberg, 2012.
- [5] Y. Wu and W. Chu, "Query languages for domain specific information from ptf astronomical repository," in *Databases in Networked Information Systems*, pp. 237–243, Springer, 2015.
- [6] P. Sprimont, D. Ricci, and L. Nicastro, "New web technologies for astronomy," in *Revista Mexicana de Astronomia y Astrofisica Conference Series*, vol. 45, p. 75, 2014.
- [7] K. Lei, Y. Ma, and Z. Tan, "Performance comparison and evaluation of web development technologies in php, python, and node.js," in *Computational Science and Engineering (CSE), 2014 IEEE 17th International Conference on*, pp. 661–668, IEEE, 2014.

⁵Scalable Vector Graphics

¹<http://cdsportal.u-strasbg.fr>

²<http://github.com>

³The Common Gateway Interface (CGI) is a standard for interfacing external applications with information servers, such as HTTP or Web servers.

⁴<http://nodejs.org>

The Time Synchronization Model of Sensors and Reference-Transponders for the Aircraft Navigation System

Dmitriy Novopashin
Saint-Petersburg State University
7-9, Universitetskaya nab.,
St. Petersburg, 199034, Russia
veshchiy@yandex.ru

ABSTRACT

We consider a system of navigation for the aircraft consisting of sensors that use technology MLAT. The main purpose of the paper is to solve applied problems of sensors time synchronization using reference-transponder within the group of sensors. One of the most important conditions for using this technology is bringing a group of sensors to a single timeline concerning the timing-pulse generator frequency of one of the sensors. Additionally, determination of the conditions is needed for a finding of solid zero-point timing within this group of sensors. Also applied problems of time synchronization of several groups of sensors are solved.

Categories and Subject Descriptors

J.2 [Computer Applications]: Physical sciences and engineering – *aerospace, engineering*.

General Terms

Theory.

Keywords

Time synchronization model; sensors; reference-transponders; MLAT; aircraft navigation system.

1. INTRODUCTION

In previous works we have dealt with the issue of aircraft maintenance in terms of software for processing and filtering of the signals received on the side of the controller [1]. At this time, we consider the problem which arises at an earlier stage – between the reception of the signal and processing by the observer. We consider a system located in the territory of difficult terrain, working on technology Multilateration (MLAT) [2].

One of the most important conditions for this technology is time synchronization of all sensors. Each sensor calculates the timing of receiving or sending signals based on the frequency of its internal timing-pulse generator.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IWAIT'15, Oct. 8–10, 2015, Aizu-Wakamatsu, Japan.
Copyright 2015 University of Aizu Press.

Frequency of timing-pulse generator is not a static value, it can have an error in the direction of increasing and decreasing or alternately one or the other. This may occur due to overheating of the sensor at the sun in summer or overcooling in winter, also the reason may be the aging of the generator or the deterioration of the equipment interacting with the generator. Furthermore, each sensor is activated asynchronously with the other and has its own starting counter value generator.

Thus the problem of synchronization involves bringing time of all sensors to a single timeline and, in some cases, finding a solid zero point.

2. DESCRIPTION OF MODEL

Let us consider a situation where for the coordination of sensors is used reference-transponder – a special device that sending and receiving the test messages and have access to satellite time.

A group of one reference-transponder and N sensors that are within reach of the transponder's signal and sufficiently frequently receive the synchronization signal will be called the domain. The model of calculation the internal time has the form:

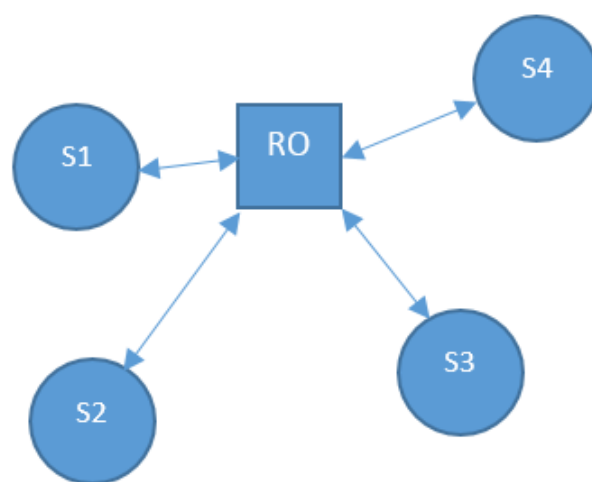


Figure 1: Example of domain. RO is reference-transponder, S1, S2, S3, S4 are sensors.

$$N_s = N_0 + f_s(T_s - T_0 + S),$$

where N_0 is the counter of timing-pulse generator clock at a time T_0 , T_0 is the time at some point, conditional zero, T_s is the time at the moment of emission timing signal, S is the time during which the signal reaches the sensor (temporal distance), f_s is a conversion function from seconds of real time to sensor timing-pulse generator cycles, inverse of timing-pulse generator frequency, N_s is the value of generator counter at the moment of receiving a synchronization signal.

Based on the geometrical position reference-transponder and sensors, temporal distance of i -th sensor is calculated by the formula [3,4]:

$$S_i = \sqrt{(x_i - x_r)^2 + (y_i - y_r)^2 + (z_i - z_r)^2} / V_{light},$$

where x_r, y_r, z_r are coordinates of reference-transponder, x_i, y_i, z_i are coordinates of i -th sensor.

The values of N_0 and T_0 are abstract and used in the model only for convenience, as in the subsequent calculations used only intervals ΔN and ΔT .

3. SYNCHRONIZATION MODEL OF SENSORS INSIDE A DOMAIN

For synchronization of one domain we will consider two consecutive signals from the reference-transponder on i -th sensor:

$$N_{S_i}^{(1)} = N_0 + f_i(T_r^{(1)} - T_0 + S_i),$$

and

$$N_{S_i}^{(2)} = N_0 + f_i(T_r^{(2)} - T_0 + S_i),$$

then

$$\Delta N_{S_i}^{(2)-(1)} = f_i \Delta T_r^{(2)-(1)}.$$

Since the time between signals $\Delta T = T_r^{(2)-(1)}$ is equally for all sensors then we obtain:

$$\frac{\Delta N_{S_i}^{(2)-(1)}}{f_i} = \frac{\Delta N_{S_j}^{(2)-(1)}}{f_j},$$

for any i -th and j -th of sensors in the domain.

Therefore it is possible to scale a timeline concerning to the frequency i -th sensor for all sensors of the domain:

$$f_j = \frac{\Delta N_{S_j}^{(2)-(1)}}{\Delta N_{S_i}^{(2)-(1)}} f_i.$$

By adding to the domain, which synchronized concerning to the i -th sensor, one reference-transponder and sending at the same time synchronization signals from each transponders, we obtain:

$$N_{S_i}^{ro_1} = N_0 + f_i(T^{ro_1} - T_0 + S_{1,1}),$$

$$N_{S_i}^{ro_2} = N_0 + f_i(T^{ro_2} - T_0 + S_{1,2}),$$

$$N_{S_j}^{ro_1} = N_0 + \alpha f_i(T^{ro_1} - T_0 + S_{2,1}),$$

$$N_{S_j}^{ro_2} = N_0 + \alpha f_i(T^{ro_2} - T_0 + S_{2,2}),$$

where $\alpha = f_j / f_i$.

Let present the $N_{S_i}^{ro_2} - N_{S_i}^{ro_1}$ and $N_{S_j}^{ro_2} - N_{S_j}^{ro_1}$, considering difference $T^{ro_2} - T^{ro_1}$ as one variable ΔT :

$$N_{S_i}^{ro_2} - N_{S_i}^{ro_1} = f_i(\Delta T + S_{1,2} - S_{1,1}),$$

$$N_{S_j}^{ro_2} - N_{S_j}^{ro_1} = \alpha f_i(\Delta T + S_{2,2} - S_{2,1}).$$

Therefore we can uniquely express f_i :

$$f_i = \frac{\Delta N_{S_j} - \alpha \Delta N_{S_i}}{\alpha(S_{2,2} - S_{1,2} + S_{1,1} - S_{2,1})}$$

4. SYNCHRONIZATION MODEL OF DOMAINS

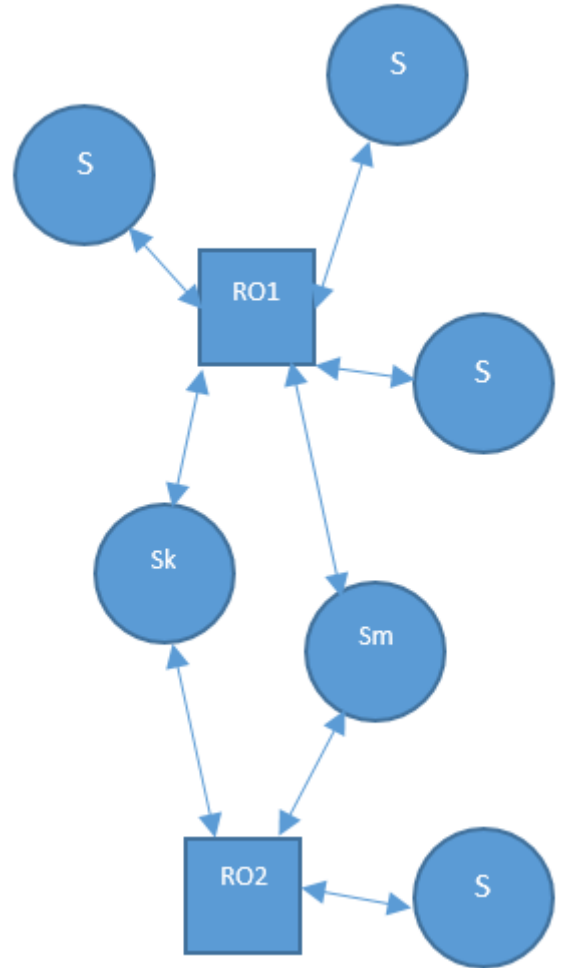


Figure 2: Example layout of two domains.

A more complex case of placing sensors is the situation when two domains share one common k -th sensor. Let us consider the

domains as two groups – the domain of reference-transponder 1 (ro_1) and the domain of reference-transponder 2 (ro_2). Let us find the equation concerning the emission time synchronization signals for each sensor in the domains, which will look like:

$$N_i^{ro_1} = N_0 + f_i(T^{ro_1} - T_0 + S_{i,ro_1})$$

or

$$N_j^{ro_2} = N_0 + f_j(T^{ro_2} - T_0 + S_{j,ro_2}).$$

Synchronizing the sensors in each domain concerning to the k-th sensor we obtain:

$$N_i^{ro_1} = N_0 + \alpha_{i,k} f_k(T^{ro_1} - T_0 + S_{i,ro_1})$$

or

$$N_j^{ro_2} = N_0 + \alpha_{j,k} f_k(T^{ro_2} - T_0 + S_{j,ro_2}).$$

Thus, the time of both sensors can be scaled to a common timeline concerning to the frequency of k-th sensor. It does not matter by which sensor's frequency of generator time was synchronized within each domain.

Adding to a model with two domains one more common m-th sensor we obtain the sub-domain that consists of two sensors and two reference-transponders on the intersection of domains coverage area. This will allow due to the sensors synchronization algorithm within the domain with two reference-transponders uniquely express the frequency of k-th or m-th sensor and thus bind to the solid zero point sub-domain, and through it both domains.

5. CONCLUSIONS

In such a way it can be argued that having a system of N domains, intersecting by one sensor, is enough to place in the intersection area of adjacent reference-transponders one additional sensor, frequently enough receiving a time synchronization signal from these two transponders, that in the acquired sub-domain it can be uniquely calculate the frequency of synchronizing sensor and thereby bind the system to a common time scale with solid zero point.

Should be noted that the algorithms developed in this article and in [1] in the future will become a part of the intellectual system support aircraft. It is planned to use them to identify the

trajectory of the aircraft in various modes. In addition, to improve the quality of this identification is expected to use the methods of mathematical theory of control [5, 6] and the multi-program control [7-9].

6. REFERENCES

- [1] Novopashin D.V. 2014. Statistical data processing of the secondary radiolocation using orthogonal Chebyshev polynomials. *Control Processes and Stability*. Vol. 1, No. 1, 352-356.
- [2] Multilateration & ADS-B Executive Reference Guide. URL: <http://www.multilateration.com/>.
- [3] Malikov M.F. 1949. *Fundamentals of metrology*. Kommerpibor, Moscow (in Russian).
- [4] Potapov, A. A. 2005. *Fractals in radiophysics and radiolocation. The topology of the sample*. University Book, Moscow (in Russian).
- [5] Andreev, Yu. N. 1976. *Control for n-dimensional linear objects*. Nauka, Moscow (in Russian).
- [6] Solis-Daun, J., Suarez, R. and Alvarez-Ramirez, J. 2000. Global stabilization of nonlinear systems with inputs subject to magnitude and rate bounds: a parametric optimization approach. *SIAM J. Control Optim.* Vol. 39, No 3, 682-706.
- [7] Zubov, V. I. Synthesis of multiprogram stable controls. *Reports Academy of Sciences USSR*. Vol. 318, No 2, 274-277 (in Russian).
- [8] Smirnov, N. 2012. Multiprogram Control for Dynamic Systems: a Point of View. *Proceedings of the Joint Intern. Conference on Human-Centered Computer Environments* (Aizu-Wakamatsu, Japan, Duesseldorf, Germany, March 8-13, 2012), HCCE 2012. ACM, New York, NY, 106-113. DOI=10.1145/2160749.2160773.
- [9] Smirnov, N. V., Smirnova T. E., Shakhov Ya. A. 2012. Stabilization of a given set of equilibrium states of nonlinear systems. *Journal of Computer and Systems Sciences International*. Vol. 51, No. 2, 169-175.

Author Index



A

Akhin, Marat	5
Ando, Hiromasa	8, 10

B

Baratynskiy, Alexander	25
Bhalla, Subhash	84, 109
Blekanov, Ivan	58, 61
Boldyrev, Alexander V.	87

C

Chamindra, Thilak	97
Chu, Wanming	109

D

Dharmasiri, Venushka	97
Dmitrieva, Elena	47
Dorofeev, Iurii	43

E

Eleshevich, Andrei	5
--------------------	---

F

Fedortsov, Alexander B.	87
-------------------------	----

G

Gerasimov, Nikita	51
Glukhikh, Mikhail	43
Gonchar, Igor V.	87

H

Hayashi, Kensaku	21
Hijikata, Kosuke	39
Holzapfel, Florian	71
Hönig, Philipp	71
Hosaka, Shunya	16

I

Ivanov, Alexey S.	87
-------------------	----

K

Kanemoto, Shigeru	33, 35
Khassina, Eugenia M.	1

Author Index (cont.)



K

Kikuchi, Jun	104
Kiselev, Anton	65
Klemeshov, Evgenii	58
Klyuenkov, Andrey	55
Klyuev, Vitaly	XI, 94, 104, 107
Korelin, Vasilii	61
Korotkova, Nelli	13
Krylatov, Alexander	91
Kuznetsov, Andrey	65

L

Lomov, Andrei A.	1
Lopukhov, Andrey	81
Lubashevsky, Ihor	8, 10, 33, 35, 39, 41
Lunde, Rüdiger	71

M

Marchuk, Andrey	21
Matlash, Aleksandr N.	76
Mozgovoy, Maxim	79
Murai, Kentaro	94

N

Namae, Ren	41
Nikitin, Kirill	47
Novopashin, Dmitriy	111

O

Oka, Ryuichi	X
--------------	---

P

Pandey, Anoop Kumar	101
Popkov, Alexander	68
Pyshkin, Evgeny	XI, 25, 29, 51

R

Raevskaya, Anastasiya	91
-----------------------	----

S

Saito, Hiroshi	16
Saito, Yoichi	107
Serikov, Alexander	81
Sevryukov, Sergey Y.	76
Sergeev, Sergei	58

Author Index (cont.)



S

Skripal, Boris	29
Srinivasa, Srinath	101
Suzuki, Takashi	33, 35

T

Takano, Shunsuke	21
------------------	----

V

Vazhenin, Alexander	21, 39
Veremey, Evgeny	13

W

Watanabe, Marie	41
Wu, Yilang	109

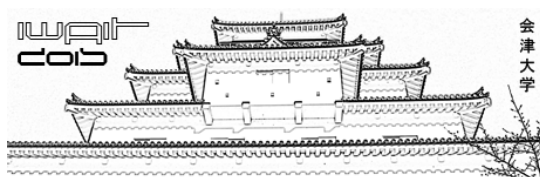
Y

Yamada, Akane	79
Yamauchi, Ryoji	8, 10
Yui, Hiroaki	84

Z

Zgonnikov, Arkady	8, 10
-------------------	-------

Proceedings of the International Workshop on
Applications in Information Technology (IWAIT-2015)



THE UNIVERSITY OF AIZU
Tsuruga, Ikki-machi, Aizu-Wakamatsu City
Fukushima, 965-8580 Japan
phone: +81-242-37-2521 fax: +81-242-37-2531



***Published by
the University of Aizu Press***

ISBN 978-4-900721-03-6