

# TCSW2013

The 1<sup>st</sup> International Tsukuba Computer Science Workshop

Theme:

“Human Centric Computing and Related Technology of Big Data”



Date: 17<sup>th</sup> of December, 2013

Time: 09:00 -- 17:55

Venue: Advance Research Building B0110

University of Tsukuba, Japan



筑波大学  
*University of Tsukuba*

University of Tsukuba  
Graduate School of Systems and Information Engineering  
Department of Computer Science  
TCSW2013 HP: <http://www.cs.tsukuba.ac.jp/tcsw2013/>

# Contents

- Welcome message
- Organizers
- Program schedule
- Speaker and talk lists
  - Session 1
    - Mr. Zhiwei Zhang, Chinese University of Hong Kong
    - Mr. Yusuke Kozawa, University of Tsukuba
    - Dr. Chuan Xiao, Nagoya University
    - Mr. Chen-Chen Ding, University of Tsukuba
    - Dr. Kohei Hayashi, National Institute of Informatics
  - Session 2
    - Dr. Volkmar Frinken, Kyushu University
    - Mr. Takahiro Komamizu, University of Tsukuba
    - Dr. Yang Wu, Kyoto University
  - Session 3
    - Mr. Naoya Tsuruta, University of Tsukuba
    - Mr. Ching-Tzun Chang, University of Tsukuba
    - Mr. Mohsin Ali Memon, University of Tsukuba
    - Mr. Rong-Hao Liang, National Taiwan University

# Welcome Message

Welcome to the 1<sup>st</sup> International Tsukuba Computer Science Workshop (TCSW) 2013.

TCSW is an international workshop organized by the graduate students of the Computer Science Department of University of Tsukuba, with the support of the university fund for innovative education. It aims to encourage interaction among the participants: the graduate students of the CS Department, and postdoctoral researchers from various universities and institutions. It is also promising for those who are considering attending doctoral programs to acquire useful information during the interaction.

Theme of TCSW 2013 is "Human Centric Computing and Related Technology of Big Data". Presentations and demonstrations about current researches will be given by 6 researchers invited from various institutions (native or overseas), as well as 6 Ph.D. students from our CS Department.

We thank all the faculty members who have supported this event and all the speakers for their time presenting their research.

We hope that you could enjoy the workshop and have fruitful discussions for supporting your future research activities.

17<sup>th</sup> of December, 2013

Co-General Chair,  
On behalf of the student organizers

# Organizers

## **Co-General Chairs**

- Cheng Haokan (2<sup>nd</sup> year Ph.D. student)
- Chendra Hadi Suryanto (2<sup>nd</sup> year Ph.D student)

## **Session Co-Chairs**

- Savong Bou (2<sup>nd</sup> year Master student)
- Hiroyuki Inoue (2<sup>nd</sup> year Master student)
- Bernardo Bentes Gatto (2<sup>nd</sup> year Master student)

## **Advisors (Faculty members)**

- Prof. Kazuhiro Fukui
- Prof. Toshiyuki Amagasa
- Prof. Shin Takahashi
- Prof. Takeshi Yamada
- Prof. Shinichi Yamagiwa

# Program Schedule

Opening - Co-Chair: Haokan Cheng	
09:00 – 09:05	Opening speech by Co-Chair
Session 1 (Database session) - Session Co-Chair: Hiroyuki Inoue, Savong Bou	
09:05 – 09:40	I/O Efficient: Computing SCCs in Massive Graphs (by Mr. Zhiwei Zhang)
09:40 – 10:15	GPU Acceleration of Probabilistic Frequent Itemset Mining from Uncertain Databases (by Mr. Yusuke Kozawa)
10:15 – 10:50	Processing of Similarity Queries on Unstructured Text Documents (by Dr. Chuan Xiao)
10:50 – 11:25	An Unsupervised Parameter Estimation Algorithm for a Generative Dependency N-gram Language Model (by Mr. Chen-Chen Ding)
11:25 – 12:00	Relational Data Analysis by Tensor Decomposition (by Dr. Kohei Hayashi)
Break	
12:00 – 13:15	** Lunch box is provided for each speaker **
Session 2 (Computer vision session) - Session Co-Chair: Chendra H. Suryanto, Bernardo B. Gatto	
13:15 – 13:50	Long Short-term Memory Neural Networks for Hand Writing Recognition and Language Modelling (by Dr. Volkmar Frinken)
13:50 – 14:25	FUKUWARAI: an Interactive Search System for Face Images (by Mr. Takahiro Komamizu)
14:25 – 15:00	Collaborative Representation for Person Re-identification (by Dr. Yang Wu)
Break	
15:00 – 15:20	Coffee break
Session 3 (Human computer interaction session) - Session Co-Chair: Haokan Cheng, Hiroyuki Inoue	
15:20 – 15:55	Origaminista: An Exploring Tool for Simple Flat Origami with Random Folds (by Mr. Naoya Tsuruta)
15:55 – 16:30	A Remote Communication System to Provide "Out Together Feeling" (by Mr. Ching-Tzun Chang)
16:30 – 17:05	Ensuring Privacy in the Course of Pervasive Logging a Passerby (by Mr. Mohsin Ali Memon)
17:05 – 17:50	GaussSense: Portable and Occlusion-free Near-surface Object Tracking Using Magnetic Sensor Grid (by Mr. Rong-Hao Liang)
Closing - Co-Chair: Chendra Hadi Suryanto	
17:50 – 17:55	Final Question/Answer session for all speakers
17:55 – 18:00	Closing Speech by Co-Chair
Banquet ***Invitation only***	
18:30 – 20:30	Location: 3 <sup>rd</sup> Area Cafeteria, University of Tsukuba

# I/O Efficient: Computing SCCs in Massive Graphs

By Mr. Zhiwei Zhang

A strongly connected component (SCC) is a maximal subgraph of a directed graph  $G$  in which every pair of nodes are reachable from each other in the SCC. With such a property, a general directed graph can be represented by a directed acyclic graph (DAG) by contracting an SCC of  $G$  to a node in DAG. In many real applications that need graph pattern matching, topological sorting, or reachability query processing, the best way to deal with a general directed graph is to deal with its DAG representation. Therefore, finding all SCCs in a directed graph  $G$  is a critical operation. The existing in-memory algorithms based on depth first search (DFS) can find all SCCs in linear time w.r.t. the size of a graph. However, when a graph cannot reside entirely in the main memory, the existing external or semi-external algorithms to find all SCCs have limitation to achieve high I/O efficiency. In this paper, we study new I/O efficient semi-external algorithms to find all SCCs for a massive directed graph  $G$  that cannot reside in main memory entirely. To overcome the deficiency of the existing DFS based semi-external algorithm that heavily relies on a total order, we explore a weak order based on which we investigate new algorithms. We propose a new two phase algorithm, namely, tree construction and tree search. In the tree construction phase, a spanning tree of  $G$  can be constructed in bounded sequential scans of  $G$ . In the tree search phase, it needs to sequentially scan the graph once to find all SCCs. In addition, we propose a new single phase algorithm, which combines the tree construction and tree search phases into a single phase, with three new optimization techniques. They are early acceptance, early rejection, and batch processing. By the single phase algorithm with the new optimization techniques, we can significantly reduce the number of I/Os and CPU cost. We conduct extensive experimental studies using 4 real datasets including a massive real dataset, and several synthetic datasets to confirm the I/O efficiency of our approaches.

## About the speaker:

Zhiwei Zhang received his B.Eng. in computer science and technology from Renmin University of China in 2010. He is currently a Ph.D student in the Department of System Engineering and Engineering Management, The Chinese University of Hong Kong. His major research interests include large-scale graph query processing and I/O efficient graph algorithms.

# GPU Acceleration of Probabilistic Frequent Itemset Mining from Uncertain Databases

By Mr. Yusuke Kozawa

Uncertain databases have been widely developed to deal with the vast amount of data that contain uncertainty. To extract valuable information from the uncertain databases, several methods of frequent itemset mining, one of the major data mining techniques, have been proposed. However, their performance is not satisfactory because handling uncertainty incurs high processing costs. In order to address this problem, we utilize GPGPU (General-Purpose computation on GPU). GPGPU implies using a GPU (Graphics Processing Unit), which is originally designed for processing graphics, to accelerate general purpose computation. In this paper, we propose a method of frequent itemset mining from uncertain databases using GPGPU. The main idea is to speed up probability computations by making the best use of GPU's high parallelism and low-latency memory. We also employ an algorithm to manipulate a bitstring and data-parallel primitives to improve performance in the other parts of the method. Extensive experiments show that our proposed method is up to two orders of magnitude faster than existing methods

## About the speaker:

Yusuke Kozawa is currently a Ph.D. student at University of Tsukuba. He received the Bachelor and M.Eng. degrees from University of Tsukuba in 2011 and 2013, respectively. His research interests include databases, data mining, and parallel computing. More specifically, he is interested in GPU acceleration of database operations and fundamental algorithms in computer science.

# Processing of Similarity Queries on Unstructured Text Documents

By Dr. Chuan Xiao

Recent decades have witnessed a rapid proliferation of unstructured text documents such as Web pages, e-mails, word-processor documents, metadata, etc. Processing such data has attracted extensive attention and many products and tools have emerged. Due to the existence of inevitable inconsistencies in data, such as erroneous data entries, natural noises, and different data representations in heterogeneous data sources, a recent trend is to study the processing of similarity queries. Tackling the inconsistencies with similarities benefits many applications yet imposes considerable technical challenges. This talk targets a series of fundamental problems in processing similarity queries on unstructured text documents, including similarity search and join, approximate entity recognition, and error-tolerant query autocompletion. The progress that has been made on this topic will be explored and a range of open problems will be identified.

## About the speaker:

Dr. Chuan Xiao is a postdoc research associate in the database group at Nagoya University leading by Prof. Yoshiharu Ishikawa. He received PhD degree from The University of New South Wales in 2010, under the supervision of Prof. Xuemin Lin and Prof. Wei Wang. Before that, he received bachelor's degree in Northeastern University, China, in 2005. His main research interests include similarity search, data integration, textual databases, and graph databases.

# An Unsupervised Parameter Estimation Algorithm for a Generative Dependency N-gram Language Model

By Mr. Chen-Chen Ding

We design a language model based on a generative dependency structure for sentences. The parameter of the model is the probability of a dependency N-gram, which is composed of lexical words with four kinds of extra tags used to model the dependency relation and valence. We further propose an unsupervised expectation-maximization algorithm for parameter estimation, in which all possible dependency structures of a sentence are considered. As the algorithm is language-independent, it can be used on a raw corpus from any language, without any part-of-speech annotation, tree-bank or trained parser. We conducted experiments using four languages: English, German, Spanish and Japanese. The results illustrate the applicability and the properties of the proposed approach.

## About the speaker:

2005.09-2009.07: ShanDong University, China. B.S..  
2010.04-2012.03: Department of Computer Science, System and Information Engineering, University of Tsukuba. M.S. of Engineering.  
2012.04-current: PhD candidate.  
2012.09-2013.03: Internship at MSRA (Mentor: Dr. Yuki Arase)  
Research topic: Natural Language Processing. Mainly statistical machine translation, and dependency structure analysis



# Relational Data Analysis by Tensor Decomposition

By Dr. Kohei Hayashi

A collection of relationships among aligned multiple objects, for example, WWW, social networks, and DNA microarray, are called relational data. Generally speaking, the relational data of  $m$ -tuples are represented by an  $m$ -dimensional array or a  $m$ -th order tensor, while its high dimensionality hinders direct analysis. Tensor decomposition is a remedy of this issue, which finds a low-dimensional projection that is useful for data compression, feature extraction, visualization, and data completion. In this talk, we introduce a few basic models of tensor decomposition and several examples of real-world application studied in the fields of machine learning and data mining.

## About the speaker:

Dr. Kohei Hayashi is a project researcher at National Institute of Informatics, Tokyo. He obtained Bachelor degree from Ritsumeikan University, Japan in 2007. His Master and Ph.D. degrees were obtained from Nara Institute of Science and Technology, Japan in 2009 and 2012, respectively. He is currently working on machine learning and data mining, especially on optimization of distributed algorithm for large-scale computation and probabilistic model based data analysis of structured data such as graph and time series.

# Long Short-Term Memory Neural Networks for Handwriting Recognition and Language Modelling

By Dr. Volkmar Frinken

Despite decades of research, the recognition of handwritten text is still a difficult problem. This difficulty arises from the large variety of different writing styles as well as some ambiguity. It turns out that it is difficult to split a text line into words or even characters without knowing the correct transcription, but it is also difficult to transcribe a text line without knowing how to segment it. As a way around this dilemma, sequential methods have been established that treat an entire text line as one element. A novel kind of recurrent neural networks, so-called Long Short-Term Memory (LSTM) neural networks, have shown a very promising performance. In addition, these networks can also be used to model human language and further help in the recognition process.

## About the speaker:

Volkmar Frinken has received his Master degree in 2007 from the University of Dortmund, Germany and his PhD in 2011 from the University of Bern, Switzerland. He did a post-doc at the Computer Vision Center of the Autonomous University of Barcelona, Spain and works currently as a Research Fellow at the Department of Advanced Information Technology of Kyushu University in Fukuoka, Japan. His research interests include Document Analysis, in particular Handwriting Recognition, Language Modelling, and Historical Document processing; recognition methodologies such as neural networks, hidden Markov models, and dynamic time warping; as well as semi-supervised and unsupervised machine learning techniques.



# FUKUWARAI: an interactive search system for face images

By Mr. Takahiro Komamizu

Face image search is one of the most important and difficult tasks in the real world situations, like finding criminals from victims' interviews and finding life partners w.r.t. face impression. This work attempts to make it easy to find people from face image database. To this end, getting features such as sex and colors of skins and eyes can be easy solutions, but more "unexpressive" features, like how his/her hair seemed to be, cannot be included. To this problem, this work proposes FUKUWARAI system which extracts fragments of face images as features for search and employs the fragments as facets in terms of faceted navigation. Although, user study shows FUKUWARAI system works well, feature mismatching problem reduces its performance. For this problem, this work extends FUKUWARAI system to allow more flexible features by probabilistic modeling and proposes probability modification mechanism based on voting. The experimental result shows the modification contributes to improve search performance over FUKUWARAI system.

## About the speaker:

Takahiro Komamizu received B.E and M.E from Department of Computer Science, University of Tsukuba in 2009 and 2011, respectively. Currently, he is a Ph. D candidate at Graduate School of Systems and Information Engineering, University of Tsukuba, Japan. His research interests include database systems, XML data management, data mining, information retrieval and multimedia data management. He is a student member of ACM, IEEE, and DBSJ.

# Collaborative Representation for Person Re-identification

By Dr. Yang Wu

Person re-identification is currently an active research topic in video surveillance, which targets at unobtrusively identifying people at a distance and across cameras in real-world surveillance scenarios. It covers the ubiquitous cases where face recognition is unreliable or infeasible, and technically it is considered to be more challenging than identification using faces. Despite its tremendous challenges which may look hopeless to many researchers, especially for those who have no experiences on this topic, significant progresses have been made in the past five years, and there is strong sign that the increase of publications on it will greatly speed up in the coming future.

In this talk, I will first give a very brief overview of the problem itself, the state-of-the-art progresses and research trends of it, based on my 3 years' research experiences and also the up-to-date books and reviews on this topic. Then, I will introduce my recent contributions, more precisely the collaborative representation based approaches, for solving this problem, which can be very simple, effective, and efficient. I will focus on two representative models which cover both normal person re-identification and transferred re-identification, and have some other extended work briefly mentioned. Finally, I will raise some personal viewpoints, which may be tentative future work, for discussion.

## About the speaker:

Yang is currently a post-doctoral researcher of Academic Center for Computing and Media Studies, Kyoto University. He received a BS degree in information engineering and a Ph.D degree in pattern recognition and intelligent systems from Xi'an Jiaotong University in 2004 and 2010, respectively. From Sep. 2007 to Dec. 2008, he was a visiting student in the General Robotics, Automation, Sensing and Perception (GRASP) lab at University of Pennsylvania. His research is in the fields of computer vision and pattern recognition, with particular interests in the detection, tracking and recognition of humans and also generic objects. He is also interested in image/video search and retrieval, along with machine learning techniques.

# Origaminista: An Exploring Tool for Simple Flat Origami with Random Folds

By Mr. Naoya Tsuruta

Origami is a form of abstract art in which the artists label pieces of folded paper, relating the folded shape with another subject (animals and objects for example). Recent origami design techniques have made possible the design of realistic models. On the other hand, the action of labeling is still important for simple origami models, which are made with a small number of folds. In this presentation, we propose an interactive system for exploring simple origami models by random generation of folded pieces. Our system generates origami pieces using random folds automatically and displays them, so that the user can focus on the labeling process. The system is implemented as a web application. The users can register models they have discovered to a web database and share the diagrams that are automatically generated by the system.

## About the speaker:

Naoya Tsuruta is currently pursuing his Ph.D. in Engineering at the University of Tsukuba, Japan. He received his B.S. in Computer Science and M.S. in Engineering from University of Tsukuba in 2010 and 2012, respectively. His research interests include geometric modeling, optimization, and interactive design systems for the geometry of origami.

# A Remote Communication System to Provide "Out Together Feeling"

By Mr. Ching-Tzun Chang

In this research, we set out to define the out together feeling as the experience when two people at different locations feel as though they are together. In other words, it makes a pair of users, one outdoors and the other indoors, feel as if they are both outdoors together. To determine a set of interaction methods to enable indoor and outdoor users to interact and share the out together feeling, we carried out preliminary experiments to observe the basic elements of communication between people who are really together. We then carried out an experiment in which indoor and outdoor users communicated via a videophone and observed the interaction patterns of each user as they attempted to achieve a given goal. From the analysis of these data, we defined three basic elements that are required to achieve the out together feeling: (1) both users can freely peruse the outdoor user's surroundings, (2) know where each other is looking, (3) and can communicate non-verbally using gestures. Using these basic elements, we designed and implemented a system called WithYou. This consists of two subsystems: a wearable system for the outdoor user and an immersive space for the indoor user. The indoor user wears a head-mounted display (HMD) and watches video from a pan-and-tilt camera mounted on the outdoor user's chest. Thus, the indoor user can look around by simply turning their head. The orientation of the outdoor user's face is also displayed on the HMD screen to indicate where they are looking. We experimentally evaluated the system and, based on an analysis of the subjects' response to questionnaires and video recordings, we were able to assess the level to which the out together feeling was achieved.

## About the speaker:

Ching-Tzun Chang is a PhD candidate in computer science at University of Tsukuba. His research interests include wearable robots, communication support, and tele-presence. He received a BS in computer science at National Taipei University of Technology and a MS in computer science at University of Tsukuba in 2006 and 2011 respectively.

# Ensuring Privacy in the Course of Pervasive Logging by a Passerby

By Mr. Mohsin Ali Memon

Pervasive logging devices are proficient in capturing information passively in the form of images, visited locations, health records, etc. to help in reminiscing past events. The latest available life logging gadgets capture people in the vicinity, producing a great concern in the future if majority of people come to have life log devices and record without any one's consent. Privacy is a state or situation of being free from being observed or captured by a passerby and it can be achieved by acknowledging the privacy concerns of people and protecting them from anonymous logging. We authorize the wearer of life log device to inscribe the restricted places and time durations where and when they would not like others to log them. In addition, the user is empowered to apply restrictions on the individual sensors of neighbor's life log device. We programmed a smart phone to work as a life log device and attached infrared transmitter/receiver pair for identifying human proximity. In addition, an interface is designed for the users to specify their privacy consent in the form of geo-temporal privacy policy. We attempt to instill privacy before capturing rather than using post capture distortion which is incompetent if the algorithm fails due to poor light conditions.

## About the speaker:

Mohsin Ali Memon is a PhD candidate in computer science at University of Tsukuba. His research interests include interaction technologies, life logging and privacy control methods. He received his B.E in Software Engineering and M.E in Information Technology at Mehran University of Engineering & Technology in 2006 and 2009 respectively.

# GaussSense: Portable and Occlusion-free Near-surface Object Tracking Using Magnetic Sensor Grid

By Mr. Rong-Hao Liang

In this talk, I introduce our invention of a portable and occlusion-free object tracking method, GaussSense, which is based on the thin-form analog Hall-sensor grid that allows users to attach to the back of portable displays or arbitrary surfaces easily. For starters, I explain the motivations of developing the new sensing technology with a review of prior mobile object tracking approaches. Then, I showcase the applications by demonstrating several research projects:

- 1) GaussSense (ACM UIST 2012): An add-on magnetic stylus sensing system that extends the stylus input space of portable displays.
- 2) GaussBits (ACM CHI 2013): A system of magnetic tangible design that enriches tangible interactions on and above portable displays.
- 3) FingerPad (ACM UIST 2013): A nail-mounted magnetic sensing device that turns users fingertips into private and rich-haptic 2D touch pads for wearable displays.

In conclusion, I highlight the contributions and the potential future research directions of this sensing technology.

## About the speaker:

Rong-Hao Liang is currently a 3rd-year PhD student of Communications and Multimedia Laboratory in National Taiwan University, Taiwan. He received his M.S. degree in Electrical Engineering from National Taiwan University. His main research interest is on Human-Computer Interaction, with a specific focus on interaction techniques and devices of wearable and tangible user interfaces. He was awarded the 2013 ACM SIGCHI Best Paper Award and the 2012 ACM SIGGRAPH Asia Emerging Technologies Prize.