RESEARCH PRESENTATION ABSTRACTS (as at 10 Mar 2015)

KEYNOTE PRESENTATIONS AND PANEL DISCUSSION

Opening Keynote Address

Date: Saturday, 14 March 2015


Associate Professor Thomas A. Weber
Chair of Operations, Economics and Strategy
Management of Technology and Entrepreneurship Institute
École Polytechnique Fédérale de Lausanne, Switzerland

‘Policy Analytics’ emphasises the use of theory, modelling, simulation and data analysis to identify valuable information and actionable strategies to address business, consumer and social issues of pressing concern to public and private organisations, and society-at-large.

This talk explores the relevance of setting broader interdisciplinary boundaries for policy analytics to create value in managerial decision-making. It will focus on the central importance of economic thinking, game theory modeling, management science and simulation methods as perspectives that have high relevance for the study of business, consumer and social insights research and practice, and produce knowledge and actionable strategies for management. It also will touch on interesting current real-world examples involving digital intermediation and peer-to-peer transaction-making in the sharing economy. They include AirBnB, the collection and dynamic valuation of delinquent credit card debt, and optimal regulation in carbon markets where pollution abatement is affected by technological innovation.

The presentation will be of interest to university-based data analytics researchers, as well as to business strategists and government policy-makers who want to explore new ways to think about the business, consumer and social issues they need to address with effective policy.
Closing Keynote Address and Discussion

Date: Sunday, 15 March 2015

Beyond Data Analytics: Evidence-based Policy-Making and Policy Analytics

Alexis Tsoukiàs
Director, LAMSADE
Université Paris Dauphine, France

Decision analysts often deal with ‘clients’ such as public agencies or business stakeholders who are involved in public and private sector decision processes. These are the people to whom we are expected to provide useful knowledge. But what is useful knowledge in such a context? International agreements, public laws and social norms? Macroeconomic forecasts or the political interests of different stakeholder groups? Or the release of national statistics and business surveys? This talk will focus on ‘evidence-based policymaking’, a contemporary effort to summarise policy-related knowledge as ‘evidence’ that should guide policy formulation and create trust in decision-making. It provides a rational perspective for decision-makers, who may rely on evidence from complex data analytics.

This is an important step forward for decision-makers who need to produce insights that yield acceptable policies and strategic actions. The evidence-based policy-making approach, despite the sophisticated data analytics that underlie it, fails to address a number of issues in the design, implementation and assessment of policies. This failure touches important issue areas that have hard problems to address: designs for sustainable urban living, effective use of natural resources and public spaces, the balance between public and private interests in business logistics, and next generation healthcare and wellness. There is a need to move ahead, beyond the data analytics and the evidence, to the formulation of specialised decision aids that support the provision of relevant and action-oriented information. The wealth of data that we can now collect is a key reason for ensuring that innovations with decision-making practices are developed as rapidly as the data analytics methods that support them.
Panel Discussion

Date: Saturday, 14 March 2015

Innovation from Interdisciplinarity: How Do We Achieve Advances in Business, Consumer and Social Insights Data Analytics Research?

Chair

Pulak Ghosh
Professor of Quantitative Methods and Information Systems
Indian Institute of Management Bangalore, and
Singapore Management University

Panellists

Lau Hoong Chuin
Professor of Information Systems,
Director, Fujitsu-SMU Urban Computing & Engineering Corp Lab, and
Deputy Director, Living Analytics Research Centre
Singapore Management University

Costas Courcoubetis
Professor of Engineering Systems and Design
Singapore University of Technology and Design

Robert Kauffman
Professor of Information Systems
Associate Dean (SIS Interdisciplinary Research), and
Deputy Director, Living Analytics Research Centre
Singapore Management University

The creation of policy-relevant and managerially-useful insights from data analytics is founded on the development and application of new methods from the disciplines of Statistics and Econometrics, Operations Research and Management Science, Economics and Game Theory, Data Mining and Machine Learning, and Psychology, Sociology and the Behavioural Sciences. In addition, Computational Social Science, innovations with the analysis of social media, credit card and mobile phone data, sociophysical and spatiotemporal analytics methods, and closed-loop experiments have had astonishing levels of success in changing how people understand many of the important issues of our time. In this panel discussion, the members will share their thoughts about the production of innovations in data analytics and the new knowledge they have been creating in the suddenly-crowded intersections of interdisciplinary research inquiry, and will respond to questions from the audience.
RESEARCH PRESENTATIONS

Research 1A: Electricity and Urban Analytics

Date: Saturday, 14 March 2015

Electricity Trading and Negative Prices: Storage Versus Disposal

Zhou Yangfang, Helen
Assistant Professor of Operations Management
Lee Kong Chian School of Business
Singapore Management University

Co-authors: Alan Scheller-Wolf, Nicola Secomandi, Stephen Smith (Carnegie Mellon University)

Electricity cannot yet be stored on a large scale, but technological advances leading to cheaper and more efficient industrial batteries make grid-level storage of electricity surpluses a natural choice. Electricity prices can be negative, but it is unclear how the presence of negative prices may affect the storage policy structure known to be optimal when prices are only non-negative, or even how important it is to consider negative prices when managing an industrial battery. For fast storage, when a storage facility can be fully emptied and filled up in one period, we show that negative prices can substantially alter the optimal storage policy structure. As a result, all else equal, it can be optimal to empty an almost empty storage facility and fill up an almost full one. For more typical slow grid-level electricity storage, we find that ignoring negative prices can result in a considerable loss of value when negative prices occur more than five per cent of the time. Negative prices raise another possibility: rather than storing surpluses, a merchant may buy negatively-priced electricity surpluses and dispose of them, for example, by using load banks. We find that the value of such disposal strategy is substantial when negative prices occur 10 per cent of the time, but smaller than that of the storage strategy, when a typical battery is used. Devices for disposal are much cheaper than those for storage though. Our results thus have ramifications for merchants as well as policy-makers.

Smartphone-Enabled Urban Solutions: Mobility-on-Demand and Mobility-as-a-Service

Cheng Shih-Fen
Associate Professor of Information Systems, and
Deputy Director, Fujitsu-SMU Urban Computing and Engineering (UNiCEN) Corp. Lab
Singapore Management University

Co-authors: Pradeep Varakantham, Asrar Ahmed, Duc Thien Nguyen, Lau Hoong Chuin, Chen Cen, Aldy Gunawan, Archan Misra, Koustuv Dasgupta, and Deepthi Chander (Singapore Management University)

One of the prevalent trends in personal communication and computation in the past few years has been the growth of the smart device adoption. With high market penetration and increasingly sophisticated sensors installed in these smart devices, it is becoming possible to provide highly personal and context-aware services. I will discuss two innovative applications that have been introduced to allow urban dwellers to request mobility services and to sell mobility-as-a-service. In research on mobility-on-demand, we have shown that making mobility-on-demand (MoD) services more efficient and integrated with traditional modes of public transport is highly valuable. A decision support system can increase the service level of MoD services, even when all providers are self-interested. Also, ride-sharing-based MoD services are promising for last-mile transfers, when commuters move from public transport hubs to their final destinations. We also have investigated large-scale mobile crowdsourcing, where a large pool of citizen crowd-workers is compensated to voluntarily perform a variety of location-specific urban logistics tasks. We propose TRACCS, a coordinated task assignment approach, in which the crowd-tasking platform assigns a sequence of tasks to each worker, taking into account their expected location trajectory over a wide time horizon.
Through simulations with realistic topologies and commuting patterns, the research demonstrates that our approach increases the fraction of assigned tasks by more than 20 per cent, and reduces the average detour overhead by more than 60 per cent, compared to the current decentralised approach.

Research 1B: Future Perspectives

Date: Saturday, 14 March 2015

Reinventing Service Systems for the Future

Sarah Cheah Lai Yin
Associate Professor (Practice)
School of Business
National University of Singapore

Yang Yinping
Scientist and Capability Group Manager
Institute of High Performance Computing
Agency for Science, Technology and Research (A*STAR), Singapore

Abstract
This presentation provides an overview of the methodological approach and interim findings of a foresighting exercise which aims to reinvent service systems for the manufacturing and service industries through game-changing business models that can be enabled by advanced technologies. We envision that the market in 2025 will be characterized by demand for deeply-personalised, faster and simpler, and more seamless and more reliable service systems. In this research, we have identified commonalities across manufacturing activities and supply chain, public health systems, and social networking service systems. We will report on our assessment of the relevant technology trends and future customer needs. We also will share our findings and views on the capabilities that need to be developed to address the future service systems innovation challenges and create new opportunities for the creation of value.

Discussion
This presentation will be followed by a longer open forum discussion facilitated by session chair, Professor Alexis Tsoukias, Université Paris Dauphine, and discussant, Professor Robert J. Kauffman, Singapore Management University. He served on the A*STAR 2025 Futurescape Panel, and was involved with the formulation of some of the background analysis and ideas that the foresighting exercise delivered. The purpose is to explore other aspects of the technological changes that will affect service systems so that it is possible to bring together the new science of services management and delivery, with the emerging capabilities of a wide range of data and policy analytics approaches.
Research 2A: Technology, Economics and Policy

Managing Seller Heterogeneity in an Online Marketplace

Lin Mei
Assistant Professor of Information Systems
School of Information Systems
Singapore Management University

Co-author: Wu Ruhai (McMaster University)

Online marketplaces, such as eBay and Airbnb, offer products and property rentals through third-party sellers, who have varied capabilities in delivering value to consumers. The different levels of pre- and post-sale services, including responses to inquiries, speed of fulfillment, reliability in payment transaction and hospitality, lead to variation in the sellers’ value offerings to consumers. Marketplace owners have launched programs, such as eBay PowerSellers and Airbnb SuperHost, to offer value-enhancing perks to qualified sellers who are already successful in delivering value to buyers. This research studies how marketplace owners should support sellers in terms of value creation by examining the impact of value variation on the marketplace owner’s profit. We model price competition among a large number of sellers who differ in the type of product offered and the value they create for consumers. We show that the marketplace owner is more profitable when the value of sellers’ offerings is more varied. Thus, by supporting high-value sellers, which increases value variation market-wide, the marketplace owner will achieve higher profit. Furthermore, the value variation among sellers of similar products is important. Such market-niche variation can also raise the marketplace owner’s profit, even when keeping the market-wide variation unchanged.

Patent Competition with Licensing

Ko Chiu Yu
Assistant Professor of Economics
Department of Economics
National University of Singapore

Co-author: Ding Rong (National University of Singapore)

We examine licensing in a patent competition for a cost-reducing innovation among firms competing in a product market. For a licensing auction, licensing is welfare-improving for large innovations but welfare-reducing for small innovations. Moreover, if the licensor has an option to exit the market, a larger magnitude of innovation is needed for the licensing to be welfare-improving. Furthermore, with more firms in the market, the set of innovations leading to a welfare-improving licensing expands. However, for royalty licensing, this approach has no impact on social welfare. This is because consumer surplus and producer surplus remain the same. The surplus of the licensee is fully-extracted by the licensor, but the licensee fee is offset by additional duplication of research effort. The extra gain by the licensee that is due to technology diffusion is fully dissipated by more duplication of research effort.
Research 2B: Social Media and Marketing Analytics

Date: Saturday, 14 March 2015

Social Media Mining and Analysis for Business and Consumer Insights

Zhu Feida
Assistant Professor of Information Systems
School of Information Systems
Singapore Management University

The recent blossoming of social network services has provided an unprecedented level of ease and fun with sharing information of all sorts. Public social data reveal a surprisingly large amount of information about an individual which is otherwise unavailable. The business, consumer and social insights attainable from such big and dynamic social data are critically important and immensely valuable in a wide range of applications for both the private sector and the public sector. In particular, there has been a growing interest in harnessing social media data for financial innovation, among other targeted industry. This presentation will explore some recent advances in this direction, including personal credit scoring, risk management and customer acquisition.

Showrooming Versus Competing: How Does Brand Selection Matter?

Tang Qian
Assistant Professor of Information Systems
School of Information Systems
Singapore Management University

Co-author: Lin Mei (Singapore Management University)

The availability of online and offline markets allows consumers to use the information provided in one market to facilitate purchasing in the other. A well-known consumer behaviour in this setting is called showrooming, for which consumers browse products in traditional stores but make their purchases online. We conduct empirical research to examine the effect of the opening of local shoe stores on the sales of a major competing shoe retailer that operates exclusively online. Both showrooming and competing effects can play a role in the sales volume of the online retailer after the offline store opening; the former leads to a higher online sales volume, whereas the latter substitutes away the online demand. We also are interested in assessing the conditions under which one effect dominates the other. Based on brand selection, offline stores are often single-brand or mixed-brand stores. The results of this research suggest that the showrooming effect is dominant when a single-brand store that opens, but the competing effect seems to dominate when a mixed-brand store opens. Examining the effects of store openings based on the minimum distance to existing shoe stores reveals that, when it is less convenient to purchase from existing offline stores, the competing effect of the new store opening is more pronounced. Examining the opening effects based on the number of existing shoe stores, when there are more stores in an area, the demand for the online retailer’s shoes is less sensitive to new store openings. Our work contributes to the multi-channel literature by studying the channel effects across different retailers, and by identifying both positive and negative effects on online sales.
**Research 3A: Consumer Behavior and Marketing**

**Date:** Saturday, 14 March 2015

**Customer Satisfaction, Costs, and the Moderating Role of Investor Sentiment**

Kapil R. Tuli  
Associate Professor of Marketing  
Lee Kong Chian School of Business  
Singapore Management University

**Co-author:** Rajdeep Grewal (The University of North Carolina at Chapel Hill)

Most of the empirical research that examines financial consequences of marketing actions and assets tends to explore either the top line metrics (e.g., sales) or bottom line metrics (e.g., stock returns, profits, ROI, and cash flows). Surprisingly, few studies directly examine the effects of marketing actions and assets on the operating costs of the firms. This is surprising given that, (i) marketing investments are generally treaded as expense items that are viewed as ‘costs’, (ii) analysts, investors, and senior managers keep a very close eye on operating costs, and (iii) that managing operating costs is viewed increasingly difficult by managers and experts. Against this background, this study presents an empirical examination of the effects of a key marketing metric: customer satisfaction on the operating costs of a firm. Using longitudinal panel data, we find that increases in customer satisfaction have a negative impact on future operating costs of a firm. Importantly, this negative effect of customer satisfaction is strongly moderator by the capital markets performance of the firm. Specifically, the authors find strong empirical support for the argument that negative effects of customer satisfaction on operating expenses are observed only in those cases where the firm experienced either a decline in stock returns or an increase in stock returns volatility. Finally, these results are consistent across different types of costs, such as the cost of selling and the cost of producing, and alternative sample compositions and model specifications.

**Who Should Get Samples? A Test of Consumer Heterogeneity in an Online Physical Product Sampling Campaign**

Lu Xianghua  
Associate Professor of Information Management  
School of Management  
Fudan University, China

**Co-authors:** Phang Chee Wei (Fudan University), Ba Sulin (University of Connecticut)

Consumer insight that can be acquired from traditional sampling campaigns are limited, as it is difficult to keep track of consumers who are trying out the samples that are offered. Online sampling can overcome this limitation with rich ‘pro’ and ‘post’ information on consumers who sampled product. This has the potential to provide important insights so brand marketers can target appropriate consumers in their sampling campaigns. In this product sampling study, we collected data on more than 120,000 participants’ before-campaign and after-campaign campaign information. We evaluated reactions to the sampling of 401 from a leading grocery e-commerce platform in China. We used this as a basis to empirically analyse how online sampling campaigns impacted the sampling study participants’ purchase behaviour. We used a latent class model to test the heterogeneity of sample applicants. We found that information seekers, usually the least experienced of e-commerce customers, were the best candidates for online sampling, and they had the highest purchase probability. They have 27 times higher a purchase probability than ‘opportunists’ did, who are more interested in the monetary value of the free samples. They already had 11 times greater likelihood than ‘party goers.’ They are mature shoppers who just want to make shopping more festive by trying out the samples. In our research, we also found that there were more than 57 per cent of applicants who seemed to be ‘information seekers.’ This study contributes to the sampling literature by investigating individual-level behavioural change. It also provides important implications for sampling marketers.
Research 3B: Urban Analytics

Date: Saturday, 14 March 2015

DIRECT: A Scalable Approach for Route Guidance in Selfish Orienteering Problems

Pradeep Varakantham
Assistant Professor of Information Systems
School of Information Systems
Singapore Management University

Co-authors: Hala Mostafa, Fu Na, Lau Hoong Chuin (Singapore Management University)

We address the problem of crowd congestion at venues such as theme parks, museums and world expos by providing route guidance to multiple selfish users moving through the venue simultaneously. To represent these settings, we introduce the ‘selfish orienteering problem’ (SeOP) that combines two well-studied problems from the literature, namely the ‘orienteering problem’ (OP) and ‘selfish routing’ (SR). OP is a single-agent routing problem where the goal is to minimise latency or maximise reward in traversing a subset of nodes, while respecting budget constraints. SR is a game between selfish agents looking for minimum latency routes from source to destination in a network available to all agents. We employ a Nash Equilibrium solution concept in solving SeOP. A direct mathematical program formulation to find a Nash equilibrium in SeOP cannot scale, however, because the number of constraints is quadratic in the number of paths, which itself is an exponential quantity. To address these issues, we make two key contributions. First, we provide a compact non-pairwise formulation with a linear number of constraints in the number of paths to enforce the equilibrium condition. Second, we introduce DIRECT, an incremental and iterative master-slave decomposition approach to compute an approximate equilibrium solution. We also provide a discussion of our approximation quality and present empirical results on synthetic and real-world graphs to demonstrate the scalability of combining DIRECT with a non-pairwise formulation.

Social and Economic Policy Issues for Net Neutrality

Costas Courcoubetis
Professor of Engineering Systems and Design
Singapore University of Technology and Design

The technical underpinnings of the Internet have caused business leaders, consumers and citizens, and government regulators to scrutinise the kinds of information that are necessary to create meaningful and effective policies. This presentation will illustrate issues that arise through the use of a model. It will set up a discussion of higher-level policy considerations. I will offer observations about Internet access, the limitations of current technologies, and the opportunities for better policy-making in the context of the current debate on net neutrality. My observations support a number of conclusions related to the use of data analytics to support a deeper understanding and consideration of Internet access, service differentiation, and next-generation innovations in digital communications and ubiquitous data content access.
**Research 4A: Innovations for Platforms and Markets**

**Making Recommendations in Crowdsourcing Contests**

**Mo Jiahui**  
Assistant Professor of Information Technology  
Nanyang Business School  
Nanyang Technological University

**Co-authors:** Sumit Sarkar, Syam Menon (University of Texas at Dallas)

Crowdsourcing contests have emerged as an innovative way for firms to acquire ideas to solve business problems from participants external to the firm. As crowdsourcing platforms attract more seekers and solvers, the number of tasks that are open at any time can become quite large. This makes it difficult for solvers to identify tasks to participate in. This research proposes a methodology to develop a system that can recommend tasks to solvers who wish to participate in crowdsourcing contests. A unique aspect of this environment is that it involves competition among solvers. The proposed methodology explicitly models the competition that a solver would face in each open task. It makes recommendations based on the probability of the solver winning an open task. A family of multinomial logit models is found to be appropriate for this environment. We develop these models based on data available on a real crowdsourcing platform, and validate our approach using these data. We further show, using simulated experiments, that deploying such a system should benefit not only the solvers but also the seekers and the platform itself.

**Toward Mechanisms for Faster Settlement on the Payment Platform: Design Choices and Policy Analytics**

**Ma Dan**  
Assistant Professor of Information Systems  
School of Information Systems  
Singapore Management University

**Co-authors:** Guo Zhiling, Robert J. Kauffman, Lin Mei (Singapore Management University)

Recently, accelerated settlement mechanisms for retail payments have been emerging around the world. A specific development is the Society for Worldwide Interbank Telecommunication’s (SWIFT) support for near real-time gross settlement to replace the existing deferred net settlement systems, in which payments are accumulated and settlement occurs only at the end of a day. The choice of a mechanism for payments settlement involves various issues, along with benefits like support for the intra-day financial needs of merchants. But faster payments settlement may cause banks to face higher intraday liquidity risks, since they may experience funds shortages for making frequent outgoing payments to merchants during business hours. Our current work involves comparing different implementations of faster payments systems across a number of developed countries. The basis for this work is founded on value maximisation strategies and the theory of mechanism design in economics. We use this as a lens to understand the current state of adoption in view of the apparent social welfare benefits. But there are inconsistent incentives, which create the need for a central planner to coordinate the settlement system design process. We are in the process of developing a policy-focused managerial decision model for banks on the appropriate frequency of settlement, in light of other key considerations that stakeholders and a central planner need to weigh.
I Want To Be Famous: Multidimensional Sports Star Modeling and Star Power Analytics

Ren Jing
Student of PhD in Information Systems
School of Information Systems
Singapore Management University

One of the most important elements of modern sports is that sports stars gain great fame and attract tremendous public attention in the media. This is not just due to their athletic talent, but also their economic value. Whenever there are interesting updates about top sports celebrities, there is always intense public discussion and scrutiny, and a variety of impacts in the social media and commercial worlds. One of the fundamental research problems in this area is how to characterise a sports star in terms of his or her main qualities. In this presentation, I will introduce a research approach that develops a model to comprehensively characterise different aspects of a sports star, which is based on the heterogeneous and multidimensional features of tennis players. I apply text mining and statistical methods to rich Internet-based data to extract five key measurable dimensions of celebrity, using a learning model that can be adjusted to fit the nature of a specific application, such as tennis, which has unique and recognisable features in its tournaments and competitions. I test and analyse the impact of the top tennis stars related to sports TV viewing for the broadcasts of the play in two Grand Slam tennis tournaments. To leverage my model, I acquired household viewership data on tennis championships from an industry partner. As a result, I was able to identify and estimate the impacts of player characteristics on viewership. In particular, it appears that a tennis player's ranking, the extent of the rivalry between two players, and their position in the tournament (1st round, quarterfinals, semi-finals, finals), tend to affect the extent of TV viewing when they play. The insights from this research can be used to design program recommendations by analysing data on TV viewing behaviour and using the sports star characteristic model.

The Making of Entrepreneurs from the IT Industry: Evidence from a Quasi-Experimental Design

Chen Qing
PhD Student
School of Computing
National University of Singapore

Despite the widely-acknowledged significance of entrepreneurship in driving economic growth and delivering innovations, there is a long-standing debate over whether entrepreneurs are born, or can be taught certain skills to run a business. Recently, the increasing role of information technology in revolutionising existing business has brought about doubts of whether IT-related experiences may drive one's entrepreneurial endeavours. To answer this question, this study investigates how an individual's earlier working experience in the IT industry may influence her to enter entrepreneurship in the later career. We constructed a quasi-experimental design by exploiting the unexpected economic conditions around Internet bubble, which exogenously affects IT graduates' first job placement. We adopted a finite mixture model to explore how this first job's cultivating effect varies across different kinds of IT firms (e.g., leading firms or entrepreneurial firms), with respect to individuals in different latent classes. Further investigation of their career movements (e.g., job hopping within or across employers) identified different career paths to entrepreneurship. Our work is pioneering in empirically documenting the contribution of IT firms for entrepreneurship growth using a large-scale data. We also contribute to the literature of entrepreneur creation, by (1) demonstrating compelling causal evidence on the 'contextual view' and (2) resolving the contrasting 'Xerox view' and 'Fairchild view' by accounting for unobserved individual heterogeneity. Our findings provide critical implications for potential entrepreneurs and policy-makers in career choice and entrepreneurship acceleration.
Product Competition with Network Effects

Chen Cen
Student of PhD in Information Systems
School of Information Systems
Singapore Management University

The increasing connectivity of consumers through their social media interactions has resulted in changing patterns of product evaluation and purchase decisions. Technologies such as cell phones, game consoles, and mobile apps are also responsible for this paradigm shift. When new products are introduced, firms must consider how word-of-mouth effects among consumers influence the efficacy of product pricing and marketing strategy choices. This study examines consumers’ new product adoption behaviour in a social network, by considering two firms that develop competing new products to enter a market that is subject to word-of-mouth influence through consumers’ social interactions. The two firms make decisions about the quality, price, advertising, and market seeding strategy for selling the product. With Facebook social network data to model consumer-to-consumer social interaction, we propose an agent-based simulation model to understand influence propagation, by assessing couponing, awareness building, and other strategies to enhance market share.

The Monetary Value of Twitter Followers: Evidence from NBA Players

Li Zhuolun
PhD Student
School of Computing
National University of Singapore

This research aims to quantify the economic value of the usage and the number of followers of Twitter. The key question is whether NBA players’ participation and popularity on Twitter to help them earn higher salaries. We investigated the impact of a player’s participation in Twitter on his salary using a difference-in-differences model. Our results suggest that participation in Twitter helps a player to increase his salary by US$861,628 on average. Next, to quantify the economic value of popularity, which is measured by the number of Twitter followers, we estimated a Heckman two-stage model. We found that doubling the number of Twitter followers of an NBA player is related to an increase of US$682,122 in his salary on average. Furthermore, this study employs quantile regression to investigate the impact of social media on salary inequality. Our analysis reveals an interesting polarisation pattern: above-average and below-average players benefited more than average players. Our study provides much-needed evidence to demonstrate the business value of Twitter, which is scarce in the IS literature.

Research 5A: Video and Social Media Analytics

Date: Sunday, 15 March 2015

Video Annotation: The New Frontier of Big Video Analytics

Shen Jialie
Assistant Professor of Information Systems
School of Information Systems
Singapore Management University

As the size of video collections grows rapidly, automatic annotation is becoming an increasingly critical technique for a wide range of business intelligence-driven video analytics applications. In this talk, I will focus on three things. First, I will introduce why high quality annotation and annotation schemes are so important for accurate and fast ‘big video’ analytics. Second, I will examine some current systems and research prototypes, focusing on comparing the advantages and disadvantages of the various strategies and schemes they employ for different types of real cases. Also, I will review some of the key challenges for developing systems, and explore how techniques can be used to facilitate different kinds of business intelligence tasks. Finally, I will point out and discuss a few
promising directions and explore potential solutions that may make sense.

**Triggering Individual Attention Based on Common Topic Modelling for a Circle of Friends**

**Dai Bing Tian**  
Research Scientist  
Living Analytics Research Centre  
Singapore Management University

**Co-author:** Lim Ee-Peng (Singapore Management University)

Social networks have become essential tools to attract individuals' attention and to recommend media, products or services that match their individual interests. In a circle of friends, there often are several common topics that they discuss among themselves, indicating that this circle of friends all share an interested in them. If anyone in the circle adopts some new media, recently-released product or a newly-available kind of service, it is likely that other friends will adopt the same things as well. Thus, it is important to first analyse common topics among a circle of friends, and then decide on the appropriate items that match their common topics, attract their attention and will probably be adopted by some of them. This presentation will elaborate on the analysis of common topics in a circle of friends using a specific technique called ‘blockmodelling,’ which will be applied to social interactions on the Twitter network. This will lead to further consideration of how we can make use of the common topics to trigger individuals' attention in their circle of friends, and to do targeted recommendations that are more likely to be effective.

**Research 5B: Data Analytics for Business and Social Value**

**Date:** Sunday, 15 March 2015

**The Knowledge Accumulation and Transfer in Open-Source Software Development**

**Kim Youngsoo**  
Assistant Professor of Information Systems  
School of Information Systems  
Singapore Management University

**Co-author:** Jiang Lingxiao (Singapore Management University)

This research examines the learning curves of individual software developers in open-source software development. We collected a dataset of multi-year code change histories from the repositories for 20 open source software projects involving more than 200 developers. We built and estimated regression models to assess individual developers’ learning progress, especially related to reducing the likelihood that a bug is included in the software they write. The estimation results show that a developer’s coding and indirect bug-fixing experience do not decrease the ratio of bugs to good code, while greater experience with fixing bugs can lead to decreases in the bug ratio in the learning progress. We also find that a developer’s coding and bug-fixing experience in other projects does not decrease the developer’s bug ratio in a focal project. This research empirically confirms the moderating effects of bug types on learning progress. Also, developers exhibit learning effects for some simple bug types (e.g., wrong ‘literals’) or bug types with many instances (e.g., wrong ‘if conditionals’). The results have interesting potential managerial implications and encourage future research on project management about allocating resources on tasks that add new code versus tasks that debug and fix existing code.
Big Data Analytics for Social Value Creation

Pulak Ghosh
Professor of Quantitative Methods and Information Systems
Indian Institute of Management Bangalore and Singapore Management University

It is well known and well documented that big data has great potential to increase the efficiency of solutions that are applied to leading business, consumer and social problems. Some examples include the design of products and services in telecommunications and digital entertainment, the pricing and promotion of product and services settings such as shopping malls and financial services, and the abatement of congestion in commuter rail and municipal rapid transit services. The potential to contribute to the development of society by using data analytics approaches to improve urban planning and transportation, streamline the delivery of government services, enhance the effectiveness of disaster management, and other kinds of activities is enormous. In this presentation, I will describe a number of cases where I have specific knowledge from projects that I have conducted that use big data analytics, and have been important in achieving outcomes involving social value creation.

Research 6A: Analytics Involving Data Mining and Machine Learning

Date: Sunday, 15 March 2015

Machine Learning Application to Urban Analytics

Jamal Atif
Professor of Computer Science
LAMSADE, Université Paris Dauphine, France

Whether in Singapore or London, New York or Beijing, increasing effort on the part of university-based researchers, commercial analytics services providers, and entrepreneurs is being devoted to the development of machine-based data analytics tools to support decisions that will improve the quality of urban living. The economic and social payoffs from such research and commercialisation work are high in value, especially in complex service and social situations, where time-series data need to be analysed to formulate meaningful policy decisions. In this presentation, I will focus on temporal series mining, using data collected with a number of different kinds of acquisition modalities. They include GPS, road networks, energy networks, and other settings. The different contextual data and data streams are large and fast flowing, and current-generation machine learning techniques are beginning to offering significant new capacity to extract meaningful knowledge and support more effective policy-making.

Machine Learning for Mining Big Data to Achieve Informed Policy-Making

Steven Hoi
Associate Professor of Information Systems
School of Information Systems
Singapore Management University

The advent of big data has been presenting a number of challenges and opportunities for R&D for scalable machine learning and data analytics techniques. Conventional batch machine learning techniques suffer from many limitations when they are applied to big data analytics tasks. In this talk, I will introduce a family of classical and recent online learning methods in machine learning, which are promising to tackle the emerging challenges of mining big data, and their applications for timely and informed policy-making.
Sentiment Classification of Web Data and its Application to Public Transportation

Wang Zhaoxia
Social and Cognitive Computing Department
Institute of High Performance Computing (IHPC)
Agency of Science, Technology and Research (A*STAR), Singapore

With the explosion of social media data over the Internet during the past decade, sentiment analysis of Web data is becoming a fast and effective way of evaluating public opinion and sentiment. Machine-learning based classifiers are widely used to perform sentiment classification. However, it is quite difficult to improve the performance of such methods even with large amounts of training data. This research aims to enhance machine-learning methods for sentiment classification of Web data. Building upon prior techniques, including feature selection, negation dealing, and emoticon handling, we propose a new approach to improve the performance of the machine-learning based sentiment analysis methods. We test the model using different text data sets related to urban development in Singapore. The results support the overall effectiveness of the proposed approach. Our method shows good social-sensing capabilities. It helps identify pertinent social issues for prompt public engagement, which has the potential to guide timely and effectively public policy-making. Using MRT data as a case study, we further illustrate the applicability of the proposed method in analysing citizen attitudes and sentiments toward public transportation issues. These analyses provide rich insights into urban transport management.

Web-based GeoVisual Analytics Regionalisation Toolkit for Policy Analytics

Kam Tin Seong
Associate Professor of Information Systems (Practice)
School of Information Systems
Singapore Management University

The uneven distribution of social services and facilities is a well-known fact among public policy analysts and decision-makers. Traditionally, a regionalisation approach is used by planners and policy analysts to delineate the country into smaller regions according to selected urban functions. However, the use of conventional clustering algorithms on geographically-referenced data to delineate planning regions has failed to take into consideration the spatial dependence and heterogeneity of the data. In view of this problem, this presentation will discuss GVAREG, a web-based regionalisation toolkit. It integrates geographically-constrained clustering algorithms and geovisual analytics techniques. I will share the technical approach used to design a related tool, RIGAP, which provides an Internet-based geospatial platform that implements state-of-the-art approaches, and comment on and its software integration as well. Using real-world data, I also will demonstrate the geospatial visual analytics and modelling capabilities of GVAREG, and discuss some directions on the future development direction and enhancements to its features.