2015 IEEE International Congress on Big Data

Pocket Guide

Taipei Satellite Session
May 15-16, 2015
Taipei 101, Taipei, Taiwan
Organization Committee

Honorary Chairs:
- Edward Chang, HTC Corp., Taiwan
- Sy-Yen Kuo, National Taiwan University, Taiwan

Program Committee Chairs:
- Shih-Chia Huang, National Taipei University of Technology, Taiwan
  Email: schuang@ntut.edu.tw
- Patrick C. K. Hung, University of Ontario Institute of Technology, Canada
  Email: patrick.hung@uoit.ca

Publicity Chairs:
- Wenwey Hseush, MacroData Inc., USA
- Shih-Hsuan Yang, National Taipei University of Technology, Taiwan
- Ding-Bing Lin, National Taipei University of Technology, Taiwan

Program Committee Members:
- Wilfred Ng, Hong Kong University of Science and Technology, Hong Kong
- Byron Choi, Baptist University of Hong Kong, Hong Kong
- Alex Huang, Nara Institute of Science and Technology, Japan
- Hong-Va Leong, Hong Kong Polytechnic University, Hong Kong
- Chi-Huang Shih, Hung Kuang University, Taiwan
- Wei-Ho Tsai, National Taipei University of Technology, Taiwan
- Jong Yih Kuo, National Taipei University of Technology, Taiwan
- Ren-Hung Hwang, National Chung Cheng University, Taiwan
- Chang-Hong Lin, National Taiwan University of Science and Technology, Taiwan
- Rua-huan Tsaih, National Cheng Chi University, Taiwan
- Fan-Chieh Cheng, Senao Networks, Inc., Taiwan
- Bo-Hao Chen, National Taipei University of Technology, Taiwan
Friday, May 15
08:30-17:00  Registration (2F, TAIPEI 101)
09:00-17:30 Industry Sessions (36F, TAIPEI 101)

Industry Sessions: Big Data Opportunities and Challenges: Discussions from Investments and Innovation
海量資料的創業與創新
Talk in Mandarin (中文場)

<table>
<thead>
<tr>
<th>Time</th>
<th>Welcome Ceremony</th>
</tr>
</thead>
</table>
| 09:00 - 09:05 | Prof. Sy-Yen Kuo  
Dean, College of Electrical Engineering and Computer Science, National Taiwan University |
| 09:05 - 09:10 | Dr. Edward Chang  
Vice President of Research and Innovation, HTC Corporation |
| 09:10 - 09:15 | Prof. Shih-Chia Huang  
Associate Professor, National Taipei University of Technology |
| 09:15 - 09:20 | Prof. Patrick C. K. Hung  
Associate Professor, University of Ontario Institute of Technology |

Industry Session 1
Session Chair: Dr. Ben Jai, CEO of Hope Bay Technologies, Inc.

<table>
<thead>
<tr>
<th>Time</th>
<th>Welcome Ceremony</th>
</tr>
</thead>
</table>
| 09:20 - 10:20 | Big Data Analytics: Architectures, Algorithms and Applications  
Dr. Edward Chang  
Vice President, HTC Corp. |
| 10:20 - 11:20 | Observation and Challenges for Big Data to Be Applied in the Corporate World  
Mr. Robert Wang  
General Manager, Acer Inc. |
| 11:20 - 12:20 | Architecting Software-Defined Storage with Meta-Data Analytics for the 3rd Platform  
Dr. Eric Chen  
CEO/Founder, ProphetStor Inc. |
| 12:20 - 13:00 | Lunch Buffet |

Industry Session 2
Session Chair: Mr. Chi Chuan Tai, Chairman of EASYCARD Corp.

<table>
<thead>
<tr>
<th>Time</th>
<th>Welcome Ceremony</th>
</tr>
</thead>
</table>
| 13:00 - 14:00 | Big Data Analytics and Manufacturing Intelligence to Empower Digital Manufacturing  
Prof. Chen-Fu Chien  
Chair Professor, National Tsing Hua University |
| 14:00 - 15:00 | BigObject Analytics: In-Place Computing for Interactive Analytics  
Dr. Wenwey Hseush  
CEO/Co-Founder, MacroData Inc. |
| 15:00 - 15:30 | Coffee Break |
| 15:30 - 16:30 | Linked Big Data Challenges and Solutions  
Dr. Ching-Yung Lin  
Chief Scientist, Graph Computing Research, IBM |
| 16:30 - 17:30 | Tech Startups: Hardcore Technologies or Innovative Business Models?  
Dr. Wayne Huang  
CEO/Founder, Armorize Technologies Inc. |
| 18:00 - 20:00 | Conference Banquet |

18:00 - 20:00 Conference Banquet
Saturday, May 16
08:30-17:00  Registration (2F, TAIPEI 101)
09:00-12:20  Academic Session (36F, TAIPEI 101)
13:30-17:20  Research Track Session (36F, TAIPEI 101)

Academic Session: Big Data Research and Development
Talk in English
Session Chair: Prof. Wei-Feng Tung, Fu-Jen Catholic University

09:00 - 9:50  Big Data Applications in Government: Some examples
             Prof. Yen-Nun Huang
             Distinguished Research Fellow, Academia Sinica

9:50 - 10:40  A tale of 2 explorers
              Prof. Shih-Wei Liao
              Associate Professor, National Taiwan University

10:40 - 11:30 Learning for Big Data
             Prof. Hsuan-Tien Lin
             Associate Professor, National Taiwan University

11:30 - 12:20 SSBDS – A High-Performance Big Data Service for Research and Education
              Prof. Yeh-Ching Chung
              Professor, National Tsing Hua University

12:20 - 13:30 Lunch Buffet

Research Track Session
Talk in English
Session Chair: Prof. Patrick C. K. Hung, University of Ontario Institute of Technology

13:30 - 13:50 The Design and Verification of an Optimal Tagging Selection Service, Wei-Feng Tung, Yan-Jun Huang, Hsiao-Yun Chang

13:50 - 14:10 Effects of Word Assignment in LDA for News Topic Discovery, Chuen-Min Huang, Cheng-Yi Wu

14:10 - 14:30 Web Service Recommendations Based on Time-aware Bayesian Networks, Victor Chu, Raymond Wong, Fang Chen, Chi-Hung Chi

14:30 - 14:50 Using Big Data for Profiling Heavy Users in Top Video Apps, Chieh-Hsin Liao, Yu-Heng Lei, Kai-Yu Liou, Jian-Shing Lin, Hsiao-Feng Yeh

14:50 - 15:10 Crowdsourcing Service Design for Social Enterprise Insight Innovation, Wei-Feng Tung, Guillaume Jordann

15:10 - 15:40 Coffee Break

15:40 - 16:00 Ctracer: Uncover C&C in Advanced Persistent Threats based on Scalable Framework for Enterprise Log Data, Kai-Fong Hong, Chien-Chih Chen, Yu-Ting Chiu, Kuo-Sen Chou

16:00 - 16:20 Sequence-Growth: A Scalable and Effective Frequent Itemset Mining Algorithm for Big Data Based on MapReduce Framework, Yen-Hui Liang, Shiow-Yang Wu

16:20 - 16:40 Machine Learning-Based Configuration Parameter Tuning on Hadoop System, Liao Shih-Wei, Chi-Ou Chen, Ye-Qi Zhuo, Chao-Chun Yeh, Chi-Min Lin

16:40 - 17:00 A Criminal Analysis and Visualization System using OLAP and Multidimensional Data Model, Worapot Jakkhupan, Pongsak Klaypaksee

17:00 - 17:20 Analyzing Social Choice and Group Ranking of Online Games for Product Mix Innovation, Wei-Feng Tung, Yu-Jei Lan, Ching-Lung Chang
Industry Sessions

Keynote 1: Big Data Analytics: Architectures, Algorithms and Applications
(05/15 Friday, 09:20-10:20)
Edward Chang
Vice President of Research and Innovation, HTC Corporation, Taiwan

Abstract:
Several studies and real-world applications (e.g., Google Translate and Speech) have demonstrated the effectiveness of the data-driven approach, which uses a large volume of data to significantly improve classification accuracy. This talk first introduces a typical open source stack composed of e.g., Hadoop, Storm, Flume, Impala, Redis, and Scalding for big data analytics (for both query processing and data mining). We then depict key speedup methods for addressing scalability bottlenecks of some key machine learning algorithms, including Frequent Itemset Mining, Spectral Clustering, SVMs, LDA, and Deep Learning. We use a couple of real-world applications to show promising results of directly learning representations from a huge volume of unlabeled data and then using the representations to boost classification accuracy. From architectures, algorithms to applications, this talk addresses both the why and how aspects of big data and its significance and trend.

About the Speaker:
Edward Chang has acted as the Vice President of Research and Innovation at HTC since July 2012, heading software and hardware future technology research and development. Besides having developed and launched competitive camera features (e.g., 360-degree Panoramas, Duo Lens with 3D effects, and Cloud Gallery with content-based image retrieval) for flagship phones, Ed's most notable project is the Tricorder project, which he co-leads (with Prof. CK Peng at Harvard) a team of physicians, scientists, and engineers to design and develop mobile wireless diagnostic instruments that can help consumers make their own reliable health diagnoses anywhere at anytime. The project entered the Qualcomm Tricorder XPRIZE competition in 2013 with 254 other entrants and was selected as one of the ten finalists in August 2014 to advance to the final round (grand prizes to be announced in January 2016). Prior to his HTC post, Ed was a director of Google Research for 6.5 years, leading research and development in several areas including big data mining, indoor localization, social networking and search integration, and Web
search (spam fighting). His contributions in parallel machine learning algorithms and big-data mining are recognized through several keynote invitations (see Stanford MMDS/ACM CIKM/ACM CIVR/ACM MM/AAIM/ADMA keynote deck and tutorial deck for details), and the developed open-source codes (PSVM, PLDA+, Parallel Spectral Clustering, and Parallel Frequent Pattern Mining) have been collectively downloaded over 12,000 times. His work on indoor localization with project X was deployed via Google Maps (see XINX paper and the editor summary of his ASIST/ACM SIGIR/ICADL keynotes). Ed's team also developed the Google Q&A system (codename Confucius), which was launched in 60+ countries including China, Russia, Thailand, Vietnam, and Indonesia, as well as 17 Arab and 40 African nations. Ed's book titled Foundations of Large-Scale Multimedia Information Management and Retrieval provides a good summary of his experience in applying big data techniques to feature extraction, learning, and indexing for organizing multimedia data to support both management and retrieval. Prior to Google, Ed was a full professor of Electrical Engineering at the University of California, Santa Barbara (UCSB). He joined UCSB in 1999 after receiving his PhD from Stanford University, and was tenured in 2003 and promoted to full professor in 2006. Ed has served on ACM (SIGMOD, KDD, MM, CIKM), VLDB, IEEE, WWW, and SIAM conference program committees, and co-chaired several conferences including MMM, ACM MM, ICDE, and WWW. He is a recipient of the NSF Career Award, IBM Faculty Partnership Award, and Google Innovation Award.
Keynote 2: Observation and Challenges for Big Data to Be Applied in the Corporate World (05/15 Friday, 10:20-11:20)
Robert Wang
General Manager, BYOC Business Unit, Acer Inc., Taiwan

Abstract:
The talk will focus on observations and insights from the front line in the battle to push big data technologies and the data-driven approach in the field. We will discuss the elements that are necessary to ensure success, common pitfalls and roadblocks, and practical approaches that work. Focused entirely on practical corporate experience, this talk will be invaluable to those who plan to bring the power of big data analytics to traditional corporations that were not founded on the data-driven approach.

About the Speaker:

Robert Wang was appointed general manager of the BYOC Business Unit at Acer in August 2014, focusing on global cloud computing, Internet of Things and big data integrated services. With two decades of industry experience, he has extensive executive and board-level experience at multi-national IT companies. He currently holds several other roles, such as executive committee member of APICTA (Asia Pacific ICT Alliance) nominated by Taiwan’s Ministry of Economic Affairs, and chief consultant of the Taipei Computer Association, involved in ICT industry advancement activities, policy making and international cooperation. Prior to joining Acer, he has served as director of the board for several publicly listed companies, director of the WiMAX Program Office of Intel Architecture Group, MNC sales director at Intel APAC, and general manager of ADI France.
Keynote 3: Architecting Software-Defined Storage with Meta-Data Analytics for the 3rd Platform (05/15 Friday, 11:20-12:20)

Eric Chen
Founder and CEO of ProphetStor Data Services Inc., USA

Abstract:
Software-defined storage (SDS) is a new generation of storage system. Unlike the traditional embedded-system based storages, the SDS uses a software-stack above commodity hardware to provide more valuable and cost-effective features. This paper argues that the architecture of an SDS platform should have three main characteristics: 1. The separation of the control and data path, 2. Self-configuration of storage resources, and 3. RESTful APIs for new business extension. In this presentation, the architecture of an SDS platform, called Federator, is first presented. A new approach designed within Federator for a self-configured SDS is then described. We then introduce the Meta-Data based analytics to help address the self-configuration issues and achieve much improved resource utilization and planning in storage/computing resources. This approach includes two types of neural network, which provides optimal storage resource configuration for any type of applications. It also includes fuzzy inference, which transforms the resultant optimal configuration into rules for automatic management. SDS has become one of the most important concepts in the storage industry, not only because it is a cost-effective alternative to traditional high-cost storage systems, but also its ease-of-use and the flexibility it can engender in an ever-changing business environment. With the clear separation of the control and the data path, the Meta-Data based analytics, and standard RESTful API, Federator is expected to better meet requirements of new applications in ever-changing computing environments.

About the Speaker:

Dr. Eric Chen (Founder and CEO of ProphetStor), served as co-founder and VP and General Manager, Asia Pacific Operations of FalconStor Software Inc. [NASDAQ: FALC] from 2000 to 2012. He helped establish the company and grew it from zero to a global operation of more than 700 employees with a market capitalization of more than $800M and an average annual growth rate of 35% Y2Y for APAC. Prior to joining FalconStor, he
worked in IBM Networking Systems, Research Triangle Park, North Carolina from 1991 to 1994. As a pioneer in IP SAN technology, Dr. Chen owns more than ten granted/pending patents in the fields of networked storage and data management. Academically, he authored more than 100 papers in networking and storage. He was an Associate Professor at the Department of Computer Science, National Chung-Hsing University, from 1994 to 2002.
Keynote 4: Big Data Analytics and Manufacturing Intelligence to Empower Digital Manufacturing (05/15 Friday, 13:00-14:00)
Chen-Fu Chien
Chair Professor, National Tsing Hua University, Taiwan

Abstract:
High-tech industries are facing challenges in light of the disruptive changes of industry structures due to Internet of Things and the evolution of new technologies such as big data analytics transforming manufacturing in digital era. New ecosystem has emerged, in which value chain positioning of individual firms may be restructured and the firm boundary has become blurred. The introduction of new process technologies and the advances in big data analytics and manufacturing intelligence capabilities are having profound effects on the decisions, management, production and service operations involved in digital manufacturing. There should be a set of systematic methodologies for cross-discipline collaborations to empower smart manufacturing in addition to the underlying technologies. On the basis of our extensive collaborative studies with high-tech industries and the observed trends, this keynote aims to address emerging issues driven by the needs in modeling, big data analytics, and decision analysis that is enabled by the advances of automation, mobile communication, and information technologies in modern manufacturing facilities. Indeed, semiconductor manufacturing is one of the most complicated and fast clock-speed industries driven by Moore’s Law for continuous technology migration and productivity enhancement. In the fully automation facility such as semiconductor fab, various approaches and solutions are developed for data mining and big data analytics to derive effective manufacturing intelligence via mining big data to address new challenges involved in yield enhancement, defect diagnosis, quality control, equipment management, cycle time reduction, cost reduction, human capital and productivity enhancement. In particular, we used several illustrations of innovative manufacturing intelligence approaches via data mining, big data analytics, and decision technologies to empower digital manufacturing. This talk will conclude with discussions of the implications of evolutionary digital manufacturing technologies and applications to foster more discussions of paradigm shift.

About the Speaker:

Dr. Chen-Fu Chien is Tsinghua Chair Professor, in the Department of Industrial Engineering & Engineering Management, National Tsing Hua University (NTHU),
Taiwan. He also teaches EMBA/MBA in the School of Technology Management in NTHU. Professor Chien is the Principal Investigator for the Semiconductor Technologies Empowerment Partners Consortium sponsored by National Science Council (NSC STEP) and Director for the NTHU-TSMC Center for Manufacturing Excellence. He received B.S. with double majors in Industrial Engineering and Electrical Engineering with the Phi Tao Phi Honor from NTHU in 1990. He received M.S. in Industrial Engineering and Ph.D. of Decision Sciences and Operations Research with two minors in Statistics and Business at the University of Wisconsin-Madison, U.S.A., in 1994 and 1996, respectively. He was a Fulbright Scholar in the Department of Industrial Engineering and Operations Research, UC Berkeley, from 2002 to 2003. He also received the Executive Training of PCMPCL from Harvard Business School in 2007. He joined NTHU as Assistant Professor in 1996, was an Associate Professor from 1999 to 2003, and became a Full Professor in 2003. Then, he has been a Tsinghua Distinguished Professor from 2010 to 2013 and become Tsinghua Chair Professor since 2013. He was a Visiting Professor in Institute for Manufacturing, Cambridge University (sponsored by Royal Society, UK), Visiting Professor in Beijing Tsinghua University (sponsored by Chinese Development Foundation), Visiting Professor in Waseda University (sponsored by Japan Interchange Association Young Scholar Fellowship), and Visiting Adjunct Professor in Tianjin University and Zhejiang University. His research mainly concerns the development of decision analytics and optimization solutions for companies confronting with decision problems involved in strategy, manufacturing, and technology that are characterized by uncertain (incomplete information or Big data) and a need for tradeoff among various objectives and the information system for effective implementation. Dr. Chien and his Decision Analysis Lab Associates have been actively involved in many university-industry collaborative research projects in the high-tech industries including semiconductor manufacturing, solar, LED, and computers in which they employ their expertise in solving real problems with domain experts. From 2005 to 2008, he had been on-leave to serve as the Deputy Director of Industrial Engineering Division in Taiwan Semiconductor Manufacturing Company. Dr. Chien has applied several invention patents for semiconductor manufacturing methodologies, seven of which have been granted. Dr. Chien has received many awards including the National Quality Award from the Executive Yuan (2012), twice Distinguished Research Awards (2007, 2011), Tier-One Principal Investigator (2005-2008), and Best Research Awards from the National Science Council, University Industrial Contribution Award from Ministry of Economic Affairs for Individual Contribution (2009) and Group Contribution (2010), Distinguished University-Industry Collaborative Research Award from the Ministry of Education (2001), Distinguished Young Faculty Research Award (2001) and Distinguished University-Industry Collaborative Research Award (2007) by NTHU, Best Paper Award (2001), Distinguished Young Industrial Engineer Award (2001), IE Medal (2010) from Chinese Institute of Industrial Engineers, Best Engineering Paper Award (2002) and Distinguished Engineering Professor (2010) by Chinese Institute of Engineers, TSMC-NTHU Faculty Semiconductor Research Grant
(2004), and the Lu, Feng-Chang Award from Chinese Management Association (2007). Dr. Chien is a member of IEEE, CIIE, SEED, CIDS, and CSMOT and is a Board Member of CIIE, SEED, and CIDS. Dr. Chien has served as the Steering Committee of the Industrial Engineering and Management Program in National Science Council since 2002. He is Associate Editor for IEEE Transactions on Automation Science and Engineering, Area Editor for Flexible Services and Manufacturing Journal and Advisory Board of OR Spectrum. He is also on the editorial board for Computers and Industrial Engineering, International Journal of Operational Research (UK), Journal of the Chinese Institute of Industrial Engineers (EI/TSSCI), Sun Yat-Sen Management Review (TSSCI), Journal of Management and Systems (TSSCI), and Journal of Quality (EI).
Keynote 5: BigObject Analytics: In-Place Computing for Interactive Analytics
(05/15 Friday, 14:00-15:00)
Wenwey Hseush
CEO/Co-Founder of MacroData Inc., USA

Abstract:
In order to address real-time analytic problems for big data, we have introduced a technology, In-Place Computing, a set of principles for storing and computing big data. It defines an abstract model in which data objects live and work on an infinite and persistent memory space. The term in-place indicates data is ready for computing. We have built a system for interactive analytics based on the in-place computing model. The system delivers three different layers of analytic powers, where data analysts can find insightful information from big data by filters, data patterns and associative relations.

About the Speaker:

Dr. Wenwey Hseush, Ph.D. in Computer Science, Columbia University, is the CEO and co-founder of MacroData, a big data technology company. His research falls in the area of experimental research in computer science, including programming language design, in-memory computing, distributed Systems, concurrent Object-Oriented language, database and transaction processing. In 1993-1995, Hseush served as First VP at Lehman Brothers and VP at PaineWebber. In 1995, Hseush founded TimeCruiser Computing, a developer of innovative community and enterprise portals for the education marketplace, for which he served as CEO until 1999. Afterwards, he joined Computer Associates as VP of Advanced Technology, overseeing operations for all Computer Associates invested joint venture companies in Asia. In 2000, Hseush founded eBizprise as a Computer Associates joint venture company in Taiwan. He was dedicated in developing the market of Supply Chain Collaboration, specializing in CPFR® (Collaborative Planning, Forecasting and Replenishment). In 2010, he established Brand Supply Chain Center for eBizprise in Tianjin and served as the Managing Director until the end of 2011. Since 2011, Dr. Hseush has been working on Big Data computing research, developing an in-memory computing technology (BigObject®) based on data transformation-computing programming paradigm. BigObject aims to deliver a high-performance (100x-1000x), powerful data processing model compared against traditional relational database and SQL. The BigObject team was spun off from eBizprise as MacroData Inc. in the beginning of 2014.
Keynote 6: Linked Big Data Challenges and Solutions  
(05/15 Friday, 15:30-16:30)  
Ching-Yung Lin  
Chief Scientist, Graph Computing Research, IBM, USA

Abstract:  
In the Big Data era, data are linked and form large graphs. But, most traditional IT systems were designed for processing independent data, while analyses are mostly done in independent scenarios. Processing connected data has been a big challenge for Big Data Analytics, which requires both the traditional big data platforms for data processes that are easily to be parallelized and the novel graph computing platforms for data that are linked. There are three major aspects of graphs -- graph storage & retrieval, graph topological analysis, and graphical models. Graph database is a tool for efficiently managing large-scale graph data, especially important for contextual and relationship analysis. Graph analytics is important for finding the important vertices or edges that are more central, that are clustered, or that form abnormal patterns. Graphical models are essential to artificial intelligence, information reasoning and predictive analysis, which requires combining many factors to create actionable insights. We have been creating a comprehensive software system for probably all aspects of Graph Computing -- IBM System G. It includes 8 toolkits: Graph Database, Graph Middleware, Graph Analytics, Graph Visualization, Cognitive Networks, Cognitive Analytics, Spatiotemporal Analytics, and Behavioral Analytics. In this talk, I will use our Social Media Solution and Insider Threat Solution as examples to show how these Network Science toolkits work in real-world applications. For instance, the Social Media Solution includes: Live Monitoring, Trend Monitoring, Multimedia Monitoring, Scope Identification, Segment Analytics, Impact Prediction, Person Analytics, Flow Analytics, Target Discovery, and Anomaly Detection.

About the Speaker:

Dr. Ching-Yung Lin is the Manager of the Network Science and Big Data Analytics Department in IBM T. J. Watson Research Center. He is also an Adjunct Assistant/Associate/Full Professor in Columbia University since 2005, in NYU since 2014, and in University of Washington 2003-2009. He received his Ph.D. from Columbia Univ in 2000, M.S. and B.S. from National Taiwan University in 1993 and 1991, respectively, all in Electrical Engineering. His interest is mainly on fundamental research of multimodality signal understanding, network science, and computational social & cognitive science. Since 2011, Lin has been leading a team of more than 40 Ph.D. researchers in worldwide IBM
Research Labs and more than 20 professors and researchers in 10 universities, including Columbia, Northeastern, Northwestern, CMU, U Minnesota, Rutgers, U New Mexico, UC-Berkeley, Stanford Research Institute, and USC. He is an author of 160+ publications and 20+ issued patents. His team recently won the Best Paper Award in IEEE BigData 2013, Best Paper Award in ACM CIKM 2012, and Best Theme Paper Award in ICIS 2011. In 2011, he was the first IEEE Fellow elevated for contributions to Network Science. He is also an IEEE CASS Distinguished Lecturer 2015-2016.
Keynote 7: Tech Startups: Hardcore Technologies or Innovative Business Models? (05/15 Friday, 16:30-17:30)

Wayne Huang
Founder/CEO of Armorize Technologies Inc., Taiwan

Abstract:
Peter Thiel, author of best seller "Zero to One," Founder to Paypal, Palantir and The Founders Fund, is attributed for saying, "We wanted flying cars, instead we got 140 characters" in his Founders' Fund manifesto. In Academia, a fundamental belief is that given time and hard work, humanity will continue to make the impossibles a reality; nothing's too hard for a group of geeks and PhDs, isn't it? And along the way, academia has overcome extremely difficult problems, has had lots of scientific and technology breakthroughs, and has in turn spun off lots of great startups, which in turn have changed our lives for the better. These companies have rightfully grown into what investors call "unicorns," or billion-dollar companies. However, in recent years, many "unicorns" were built not on top of hardcore technology breakthroughs, but instead built on business innovations; for example: Uber, AirBnB, Snapchat, Automattic, Auction.com, Houzz, JustFab, Stripe, SurveyMonkey, Gilt, etc. Even Paypal itself wasn't purely a "technology breakthrough." Have we solved most of the difficult problems? Or are we just at the beginning of humanity's tech boom? In 2015, is it easier to build a startup based on business innovation, or tech breakthrough? What's a suitable path for a hardcore tech guy doing a startup? This talk will overview these topics and anticipate audience participation.

About the Speaker:

Dr. Wayne Huang was Founder & CEO to Armorize Technologies, and is now VP Engineering at Proofpoint. Wayne is a frequent speaker at security conferences including BlackHat (10), DEFCON (10), RSA (07, 10), SyScan (08, 09), OWASP (08, 09), Hacks in Taiwan (06, 07), WWW (03, 04), PHP (07) and DSN (04). A diligent blogger on cyberthreats, Wayne's posts have been covered by the most influential media. Into security since 7th grade, Wayne has lead teams to develop security products ranging from source code analysis, web application firewall, vulnerability assessment, exploit & malware detection, anti-malvertising, email security, and APT defense. Wayne received his PhD in EE from National Taiwan University, and his BS and MS in CS from NCTU. He holds two US patents on source code analysis.
Academic Session

Keynote 1: Big Data Applications in Government: Some examples
(05/16 Saturday, 09:00-9:50)
Yen-Nun Huang
Research Fellow, Academia Sinica, Taiwan

Abstract:
Using big data analytics can improve the quality, reliability and efficiency of government services. In this talk, I will present some examples of big data applications used by Taiwan and other countries to improve the government services. Particularly, I will describe examples of using big data analytics for disaster prevention, crime and traffic predictions, food safety and detection of zero-day security attacks to Taiwan government networks.

About the Speaker:

Dr. Yen-Nun Huang is a Distinguished Research Fellow in Academia Sinica and also an IEEE Fellow. Dr. Huang joined AT&T Bell Labs as a researcher in 1989. He started the AT&T Dependable Computing Research Program and became a Director in 1999. Later on, Dr. Huang became the Executive Director of Dependable Distributed Computing and Communication Research Department. In 2007, Dr. Huang became the Executive VP of Institute for Information Industry. In 2008, Dr. Huang was the President of VeeTIME Co. to build quadruple-play telecom services. Dr. Huang joined Research Center for Information Technology Innovation (CITI) of Academia Sinica in 2011 to lead the security research program. He is also Deputy Executive Secretary in the Office of Science and Technology of Executive Yuan, helping Taiwan government on the Information and Communication Technology (ICT) R&D policy and funding allocation.
Keynote 2: A tale of 2 explorers  
(05/16 Saturday, 9:50-10:40)  
Shih-Wei Liao  
Associate Professor, National Taiwan University, Taiwan

Abstract:  
We examine the effectiveness of different configurations, which can greatly influence the performance of memory system and the overall data center.

About the Speaker:  
Dr. Shih-Wei Liao's lifelong goal is to make impacts through innovation and education, open-source and product-quality systems, so he worked at Stanford, Google headquarter and Intel headquarter for 22 years, culminating at receiving Google Founders’ Award. He received his PhD and MS from Stanford University and his bachelor degree from National Taiwan University.
Keynote 3: Learning for Big Data
(05/16 Saturday, 10:40-11:30)
Hsuan-Tien Lin
Associate Professor, National Taiwan University, Taiwan

Abstract:
Yes, everyone knows that machine learning plays a key role for big data analytics. Professional big data analytics, however, is much more than just applying the best machine learning techniques on data, but requires careful, interactive, and creative human decisions along the analysis procedure. In other words, professional big data analytics rely on both machine and human learning. This talk shares some perspective on what and how humans (i.e., data scientists and engineers) should learn for big data. In particular, the talk highlights the speaker’s subjective choice of four must-learn topics: simple model, feature processing, complexity control, and model selection. The topics represent the speaker’s personal experience on participating in big data competitions, teaching massive online open courses in machine learning, and consulting for big data industry projects.

About the Speaker:

Prof. Hsuan-Tien Lin received a B.S. in Computer Science and Information Engineering from National Taiwan University in 2001, an M.S. and a Ph.D. in Computer Science from California Institute of Technology in 2005 and 2008, respectively. He joined the Department of Computer Science and Information Engineering at National Taiwan University as an assistant professor in 2008, and has been an associate professor since August 2012. Prof. Lin received the Distinguished Teaching Award from the university in 2011, and the Outstanding Mentoring Award from the university in 2013. He co-authored the introductory machine learning textbook "Learning from Data" and offered two popular Mandarin-teaching MOOCs "Machine Learning Foundations" and "Machine Learning Techniques" based on the textbook. His research interests include theoretical foundations of machine learning, studies on new learning problems, and improvements on learning algorithms. He received the 2012 K.-T. Li Young Researcher Award from the ACM Taipei Chapter, and the 2013 D.-Y. Wu Memorial Award from National Science
Council of Taiwan. He co-led the teams that won the third place of KDDCup 2009 slow track, the champion of KDDCup 2010, the double-champion of the two tracks in KDDCup 2011, the champion of track 2 in KDDCup 2012, and the double-champion of the two tracks in KDDCup 2013. He served as the Secretary General of Taiwanese Association for Artificial Intelligence between 2013 and 2014.
Keynote 4: SSBDS – A High-Performance Big Data Service for Research and Education (05/16 Saturday, 11:30-12:20)

Yeh-Ching Chung
Professor, National Tsing Hua University, Taiwan

Abstract:
UniCloud is a distributed cloud system, which consists of a number of cloud platforms at different universities, for research and education in Taiwan. In the UniCloud system, several cloud platforms can form a community cloud. Also, a cloud platform in the UniCloud system can leverage the public cloud resources to form a hybrid cloud. The services provided by the UniCloud system include SSDK – a VM service, SSBox – a cloud storage service, SSDB – a hybrid SQL/NoSQL database service, SSBDS – a big data service, and SSVCS – a virtual cluster service. In this talk, we will focus on the design of SSBDS. The design of SSBDS is to integrate hardware and software to provide a platform, called HPBDA, for big data access, analysis, process, and presentation. The meanings of the term “high-performance” are two-fold. First, the platform equips high-performance hardware devices in storage (SSD), networking (InfiniBand), and computing (GPU). Second, the platform provides an enhanced and optimized Apache Hadoop software stack to achieve satisfactory performance for different big data applications. To reach this goal, we have developed (1) an enhanced HBase with multi-tenancy, fault-tolerance, transaction, and caching mechanisms; (2) an enhanced MapReduce framework with run-time optimization and caching mechanisms; and (3) an enhanced HDFS with heterogeneous storage, multi-tenancy, fault-tolerance, and global address mechanisms. Several big data applications, such as image processing and recognition, social network, and wafer fabrication, etc., are used to verify the design of SSBDS.

About the Speaker:

Dr. Yeh-Ching Chung is a professor in the Department of Computer Science at National Tsing Hua University (NTHU). His research interests are in the areas of parallel and distributed processing, cloud computing, and embedded systems. He is the founder of Taiwan Association of Cloud Computing (http://www.tacc.org.tw), the chief scientist of
UniCloud research group (https://www.unicloud.org.tw), and the deputy director of Computer and Communication Research Center (CCRC) of NTHU. He has delivered an HSA emulator, called HSAemu (hsaemu.org), which was a collaborated work with MediaTek. Dr. Chung received his Ph.D. degree in Computer Science from Syracuse University.
Conference Venue

The IEEE BigData 2015 Taipei Satellite Session will be held in Taipei, in the north of Taiwan. The conference venue is located in the Meeting Room A at 36th floor, TAIPEI 101 (台北101大樓三十六樓A會議室), Exit No.4, MRT Red Line (Xin-Yi Line). TAIPEI 101 Station (台北捷運信義線「台北101站」4號出口). Please note that Conference Registration is located in 2nd floor, TAIPEI 101 International Conference Center.

Transportation

TAIPEI 101 Address: No.7, Sec. 5, Xinyi Rd., Xinyi Dist., Taipei City 110, Taiwan.

By MRT

Take the MRT Red Line (Xin-Yi Line). Then take off at TAIPEI 101 Station (Exit No.4).

Station Location Map: [http://web.metro.taipei/img/ALL/Route2200/100.jpg](http://web.metro.taipei/img/ALL/Route2200/100.jpg)

Station Information Map: [http://web.metro.taipei/img/ALL/INFOPDF/100.pdf](http://web.metro.taipei/img/ALL/INFOPDF/100.pdf)

By Bus

Buses to Taipei 101 can be boarded at Taipei City Hall, the World Trade Center station, Xin-Yi Administrative Center, and the Grand Hyatt Taipei.

By Taxicab

On the side of the FamilyMart in the B1 floor of the Taipei 101.
Banquet Venue

Palace Banquet Hall at DING XIAN 101: The grace and elegance of Baroque art, the grandeur and noble splendor and the exquisite plate presentation. The banquet is about to begin, introducing the magical yet authentic experience that is as extravagant as Versailles Palace.

Location: DING XIAN 101(頂鮮 101), 86 floor of TAIPEI 101(台北 101, 86 樓)
Address: 86 floor, No.7, Sec. 5, Xinyi Rd., Xinyi Dist., Taipei City 110, Taiwan
Date: 15 May 2015, Friday
Time: 18:00 - 20:00

Taipei Hall (台北廳), DING XIAN 101(頂鮮 101), 86 floor of TAIPEI 101(台北 101, 86 樓)

Contact Information
If you have any questions or queries on IEEE BigData Taipei Satellite Session, please send email to schuang@ntut.edu.tw (Prof. Shih-Chia Huang) or patrick.hung@uoit.ca (Prof. Patrick C. K. Hung).
Sponsors

IEEE

IEEE Broadcast Technology Society

IEEE Computer Society

TAIPEI TECH

Academia Sinica

Ministry of Education

Ministry of Science and Technology

HTC

Acer

MacroData

IBM

ProphetStor

armorize

Secure Your Web Applications