SCF2019

2019 Services Conference Federation

ADVANCE PROGRAM

The 2019 International Conference on Web Services (ICWS 2019)
www.icws.org

The 2019 International Conference on Services Computing (SCC 2019)
www.thescc.org

The 2019 International Conference on Cloud Computing (CLOUD 2019)
www.TheCloudComputing.org

The 2019 International Conference on AI and Mobile Services (AIMS 2019)
www.theMobileServices.org or www.ai1000.org

The 2019 International Congress on Big Data (BigData Congress 2019)
www.bigdatacongress.org

The 2019 World Congress on Services (SERVICES 2019)
www.ServicesCongress.org

The 2019 International Conference on Internet of Things (ICIOT 2019)
www.iciot.org

The 2019 International Conference on Edge Computing (EDGE 2019)
www.theEdgeComputing.org

The 2019 International Conference on Cognitive Computing (ICCC 2019)
www.theCognitiveComputing.org

The 2019 International Conference on Blockchain (ICBC 2019)
www.blockchain1000.org

June 25-30, 2019, Sheraton Mission Valley San Diego Hotel, San Diego, CA, 92108, United States

SCF 2019 Leadership Team

General Chairs

ICWS 2019	SCC 2019
Prof. Lakshmish Ramasawamy, University of Georgia,USA	*Prof. Dimitrios Georgakopoulos, Swinburne University of Technology, Australia*
CLOUD 2019	**EDGE 2019**
Prof. Hai Jin, Huazhong University of Science and Technology, China	*Prof. Teruo Higashino, Osaka University, Japan*
BigData 2019	**ICBC 2019**
Prof. Seong-Jong Park, Louisiana State University, USA	*Prof. Bhavani Thuraisingham, The University of Texas at Dallas, USA* *Prof. Jeffrey Tsai, Asia University*
ICIOT 2019	
Dr. Samee U. Khan, National Science Foundation, USA	

Program Chairs

ICWS 2019	SCC 2019
Prof. John Miller, University of Georgia, USA *Prof. Eleni Stroulia, University of Alberta, USA* *Prof. Kisung Lee, Louisiana State University--Baton Rouge, USA*	*Dr. Joao Eduardo Ferreira, Instituto de Matemática e Estatística, Brazil* *Prof. Aibek Musaev, The University of Alabama, USA*
CLOUD 2019	**AIMS 2019**
Prof. Dilma Da Silva, Texas A&M University, USA *Prof. Qingyang Wang, Louisiana State University--Baton Rouge, USA*	*Dr. De Wang, Sunmi US Inc., USA*
BigData 2019	**SERVICES 2019**
Dr. Sangeetha Seshadri, IBM Almaden Research Center, USA *Prof. Keke Chen, Wright State University, USA*	*Prof. Yunni Xia, Chongqing University, China*
ICCC 2019	**EDGE 2019**
Prof. Ruifeng Xu, Harbin Institute of Technology, Shenzhen, China *Dr. Jianzong Wang, Ping An Group, Shenzhen, China*	*Dr. Tao Zhang, Cisco Systems, USA* *Prof. Jinpeng Wei, University of North Carolina at Charlotte, USA*
ICIOT 2019	**ICBC 2019**
Prof. Valerie Issarny, Inria, France *Prof. Balaji Palanisamy, University of Pittsburgh, USA*	*Prof. James Joshi, University of Pittsburgh, USA* *Dr. Surya Nepal, CSIRO's Data61, Australia* *Dr. Qi Zhang, IBM Thomas J. Watson Research Center, USA*

SCF 2019 General Chairs

Prof. Calton Pu, Georgia Tech., USA
Dr. Wu Chou, Vice President-Artificial Intelligence & Software at Essenlix Corporation, USA
Dr. Ali Arsanjani, VP Artificial Intelligence and Machine Learning, 8x8 Cloud Communications, USA

Steering Committee

Calton Pu, Georgia Tech, USA (Co-Chair)
Liang-Jie Zhang, Kingdee International Software Group Co., Ltd, China (Co-Chair)

Program Overview

June 25, 2019 (Tuesday) Services Conference Federation (SCF 2019) (ICBC 2019, CLOUD 2019, ICWS 2019, BigData 2019, SCC 2019, EDGE 2019, ICCC 2019, ICIOT 2019, AIMS 2019, SERVICES 2019)			
8:00-17:00	On-Site Registration (Front Desk)		
8:30-17:00	Services Society Innovation Showcases		
Room	*Navigator*	*Compass Ballroom*	*Destination Ballroom*
9:30-10:30	Summer School Tutorial 1	Social Networking Room	Social Networking Room
10:30-10:50	AM Break (with refreshments)		
10:50-12:00	Summer School Tutorial 1	Social Networking Room	Social Networking Room
12:00-14:00	Lunch (not included)		
14:00-15:00	Summer School Tutorial 2	Social Networking Room	Social Networking Room
15:00-15:20	PM Break (with coffee)		
15:20-16:30	Summer School Tutorial 2	Social Networking Room	Social Networking Room
16:30 --	Enjoy your night in San Diego!		
18:00-20:00	**Services Society Young Scientist Forum (SSYSF, Invitation Only)**		

June 26, 2019 (Wednesday) Services Conference Federation (SCF 2019) (ICBC 2019, CLOUD 2019, ICWS 2019, BigData 2019, SCC 2019, EDGE 2019, ICCC 2019, ICIOT 2019, AIMS 2019, SERVICES 2019)					
8:00-17:00	On-Site Registration (Front Desk)				
8:30-17:00	Services Society Innovation Showcases				
Room	*Destination 1*	*Navigator*	*Compass 1*	*Destination 2*	*Compass 2*
9:00-10:40	**Services Conference Federation (SCF 2019) Opening Steering Committee Chair; General Chairs; PC Chairs; Opening & Keynote Session Chair: Dr. Wu Chou (Essenlix Corporation, USA) Keynote 1: Samee U. Khan (North Dakota State University and National Science Foundation, USA) (Connections Ballroom)**				
10:40-11:00	AM Break (with refreshments)				
11:00-12:00	ICWS Session II	CLOUD Session I	ICIOT Session I	SCC Session I	ICCC Session I
12:00-14:00	Lunch (not included)				
14:00-15:00	AIMS Session II	CLOUD Session III	ICBC Session V	SERVICES Session I	BIGDATA Session III
15:00-15:20	PM Break (with coffee)				
15:20-16:20	**Services Conference Federation (SCF 2019) Keynote Panel 1: Intersection of IoT, Edge, Blockchain, AI, and Big Data (Moderator: Tony Shan, Chief Architect at CTS) (Connections Ballroom)**				
16:20-	Enjoy your night in San Diego!				

June 27, 2019 (Thursday) Services Conference Federation (SCF 2019) (ICBC 2019, CLOUD 2019, ICWS 2019, BigData 2019, SCC 2019, EDGE 2019, ICCC 2019, ICIOT 2019, AIMS 2019, SERVICES 2019)					
8:00-17:00	On-Site Registration (Front Desk)				
8:30-17:00	Services Society Innovation Showcases				
Room	*Destination 1*	*Navigator*	*Compass 1*	*Destination 2*	*Compass 2*
9:00-10:00	ICWS Session IV	CLOUD Session IV	ICBC Session I	SCC Session II	ICCC Session II
10:00-10:20	AM Break (with refreshments)				
10:20-11:40	**Services Conference Federation (SCF 2019) Keynote 2: Yi Pan (Georgia State University, USA) Keynote Session Chair: Dr. Ali Arsanjani (8x8 Cloud Communications, USA) (Connections Ballroom)**				
11:40-13:00	Lunch (not included)				
13:00-14:00	AIMS Session I	CLOUD Session II	ICBC Session II	EDGE Session I	BIGDATA Session II
14:00-14:20	PM Break (with coffee)				
14:20-15:40	AIMS Session III	CLOUD Session V	ICIOT Session II	SERVICES Session II	ICCC Session IV
15:40-16:00	PM Break (with coffee)				
16:00-17:20	**Services Conference Federation (SCF 2019) Keynote 3: Yannis Papakonstantinou (University of California, San Diego, USA) (Connections Ballroom)**				
17:20-18:00	PM Break				
18:00-21:30	**Services Conference Federation (SCF 2019) Banquet (Conference Participants Only, More information On Site)**				

June 28, 2019 (Friday) Services Conference Federation (SCF 2019) (ICBC 2019, CLOUD 2019, ICWS 2019, BigData 2019, SCC 2019, EDGE 2019, ICCC 2019, ICIOT 2019, AIMS 2019, SERVICES 2019)					
8:00-17:00	On-Site Registration (Front Desk)				
8:30-17:00	Services Society Innovation Showcases				
Room	*Destination 1*	*Navigator*	*Compass 1*	*Destination 2*	*Compass 2*
9:00-10:20	**Services Conference Federation (SCF 2019) Keynote 4: Calton Pu (Georgia Institute of Technology, USA) (Connections Ballroom)**				
10:20-10:40	AM Break (with refreshments)				
10:40-11:40	ICWS Session I	CLOUD Session VI	ICBC Session IV	SCC Session III	BIGDATA Session I
11:40-14:00	Lunch (not included)				
14:00-15:00	**Services Conference Federation (SCF 2019) Keynote Panel 2: Global Optimization as a Service (Moderator: Dr. Min Luo, Services Society, USA) (Connections Ballroom)**				
15:00-15:20	PM Break (with coffee)				
15:20-16:40	AIMS Session IV	CLOUD Session VII	ICIOT Session III	SERVICES Session III	ICCC Session III
16:40--	Enjoy your night in San Diego!				
18:00-20:00	**SCF 2019 Organizing Committee Meeting (Invitation Only)**				

June 29, 2019 (Saturday) Services Conference Federation (SCF 2019) (ICBC 2019, CLOUD 2019, ICWS 2019, BigData 2019, SCC 2019, EDGE 2019, ICCC 2019, ICIOT 2019, AIMS 2019, SERVICES 2019)			
8:00-17:00	**On-Site Registration (Front Desk)**		
8:30-17:00	Services Society Innovation Showcases		
Room	***Destination 1***	***Destination 2***	***Destination 3***
9:00-10:00	ICWS Session III	ICCC Session V	EDGE Session II
10:00-10:20	AM Break (with refreshments)		
10:20-11:20	CLOUD Session VIII	ICBC Session III	Cross Topic Session
11:20-13:00	Lunch (not included)		
13:00-14:20	**Services Conference Federation (SCF 2019) Keynote 5: Nalini Venkatasubramanian (University of California, Irvine, USA) Keynote Session Chair: Dr. Liang-Jie Zhang (Kingdee International Software Group Co., Ltd, China) (Connections Ballroom)**		
14:20-14:40	PM Break (with coffee)		
14:40-15:50	**Services Conference Federation (SCF 2019) Award Session (Connections Ballroom)**		
15:50--	Enjoy your night in San Diego!		

June 30, 2019 (Sunday) Services Conference Federation (SCF 2019) (ICBC 2019, CLOUD 2019, ICWS 2019, BigData 2019, SCC 2019, EDGE 2019, ICCC 2019, ICIOT 2019, AIMS 2019, SERVICES 2019)		
8:00-12:00	On-Site Registration (Front Desk)	
Room	*Destination 1*	*Destination 2*
9:00-11:00	Future Forum on new ABC - AI, Blockchain and Cloud (Leading Company Site or Conference Site, To be announced on site)	Future Forum on BigData, IoT and Edge (Leading Company Site or Conference Site, To be announced on site)
11:00-12:00	Services Conference Federation 2020 (SCF 2020) Announcements	
	Have a safe trip home! See you again in 2020!	

** Cross Topic Session (CTS): Per author requests, some papers will be presented in CTS instead of their originally assigned sessions based on topics. Papers to be presented in CTS are marked with ** together with the date to be presented.

About Springer

Springer is a leading global scientific, technical and medical portfolio, providing researchers in academia, scientific institutions and corporate R&D departments with quality content through innovative information, products and services. Springer has one of the strongest STM and HSS eBook collections and archives, as well as a comprehensive range of hybrid and open access journals. Springer is part of **Springer Nature**, a global publisher that serves and supports the research community. Springer Nature aims to advance discovery by publishing robust and insightful science, supporting the development of new areas of research and making ideas and knowledge accessible around the world.

As part of Springer Nature, Springer sits alongside other trusted brands like Nature Research, BMC and Palgrave Macmillan. Visit www.springer.com and follow @SpringerNature

Springer

SCF2019 Proceedings by Springer LNCS

ICWS 2019 – The 2019 International Conference on Web Services
Springer LNCS Volume Number: 11512

CLOUD 2019 – The 2019 International Conference on Cloud Computing
Springer LNCS Volume Number: 11513

BIGDATA 2019 – The 2019 International Congress on Big Data
Springer LNCS Volume Number: 11514

SCC 2019 – The 2019 International Conference on Services Computing
Springer LNCS Volume Number: 11515

AIMS 2019 – The 2019 International Conference on AI & Mobile Services
Springer LNCS Volume Number: 11516

SERVICES 2019 – The 2019 World Congress on Services
Springer LNCS Volume Number: 11517

ICCC 2019 – The 2019 International Conference on Cognitive Computing
Springer LNCS Volume Number: 11518

ICIOT 2019 – The 2019 International Conference on Internet of Things
Springer LNCS Volume Number: 11519

EDGE 2019 – The 2019 International Conference on Edge Computing
Springer LNCS Volume Number: 11520

ICBC 2019 – The 2019 International Conference on Blockchain
Springer LNCS Volume Number: 11521

Table of Contents

2019
Services Conference Federation
Keynote Sessions

Keynote 1: Revised Prespectives on IoT (SCF2019-0001)

(06/26 Wednesday, 9:00-10:40; Connections Ballroom)

Keynote Speaker: Samee U. Khan, North Dakota State University and National Science Foundation

Abstract: Our lives are constantly being driven by decisions supported by data generated by myriad of devices (or things) connected to the Internet. These "things" as they are called, come in various shapes, sizes, and capabilities. In this talk, we will revisit some of the core concepts related to the Internet of Things. In the latter portion of the talk, we will revist some relavant National Science Foundation programs pertaining to the various domain topics discussed.

About the Speaker

Samee U. Khan received a PhD in 2007 from the University of Texas. Currently, he is the Cluster Lead for the Computer Systems Research at the National Science Foundation, and a Full Professor at the North Dakota State University. His research interests include optimization, robustness, and security of computer systems. His work has appeared in over 400 publications. He is on the editorial boards of leading journals, such as Journal of Parallel and Distributed Computing, ACM Computing Surveys, and IEEE IT Pro. He is an ACM Distinguished Speaker and an IEEE Distinguished Lecturer.

Keynote 2: Recent Developments in Deep Learning Research (SCF2019-0002)

(06/27 Thursday, 10:20-11:40; Connections Ballroom)

Keynote Speaker: Yi Pan, Regents' Professor and Chair, Department of Computer Science, Georgia State University, Atlanta, Georgia, USA

Abstract: Neural networks, modeled after the human brain, contain a set of algorithms to recognize patterns via training a data set. Deep learning neural network architectures differ from traditional neural networks because they have more hidden layers and newer training algorithms. Deep learning networks can be trained in an UNSUPERVISED or SUPERVISED manner for both UNSUPERVISED and SUPERVISED learning tasks and hence can be applied in many applications. Deep learning is now producing many remarkable successes in computer vision, automatic speech recognition, natural language processing, audio recognition, bioinformatics and disease prediction and detection. Although various deep learning architectures and novel algorithms have been applied to many big data applications, better explainability, increasing prediction accuracy and speeding up the training process are still challenging tasks among others. In this talk, I will outline recent developments in deep learning research. The topics discussed include proposing more effective architectures, intelligently freezing layers, effectively handling high dimensional data, designing encoding schemes, mathematical proofs, optimization of hyper-parameters, embedding logic and reasoning during training, result explanation and hardware support for deep learning. Some of our solutions and preliminary results in these areas will be presented and future research directions will also be identified in this talk.

About the Speaker

Dr. Yi Pan is currently a Regents' Professor and Chair of Computer Science at Georgia State University, USA and a Member of EU Academy of Sciences. He has served as an Associate Dean and Chair of Biology Department during 2013-2017 and Chair of Computer Science during 2006-2013. Dr. Pan received his B.E. and M.E. degrees in computer engineering from Tsinghua University, China, in 1982 and 1984, respectively, and his Ph.D. degree in computer science from the University of Pittsburgh, USA, in 1991. His profile has been featured as a distinguished alumnus in both Tsinghua Alumni Newsletter and University of Pittsburgh CS Alumni Newsletter. Dr. Pan's research interests include parallel and cloud computing, big data, and bioinformatics. Dr. Pan has published more than 250 journal papers with over 90 papers published in various IEEE journals. In addition, he has published over 150 papers in refereed conferences. He has also co-authored/co-edited 43 books. His work has been cited more than 10,000 times in Google Scholar and his current H-index is 53. Dr. Pan has served as an editor-in-chief or editorial board member for 20 journals including 7 IEEE Transactions. He is the recipient of many awards including IEEE Transactions Best Paper Award, several other conference and journal best paper awards, 4 IBM Faculty Awards, 2 JSPS Senior Invitation Fellowships, IEEE BIBE Outstanding Achievement Award, NSF Research Opportunity Award, and AFOSR Summer Faculty Research Fellowship. He has organized many international conferences and delivered keynote speeches at over 60 international conferences around the world.

Keynote 3: Channlenges and Opportunities of Data Management on the Web (SCF2019-0003)

(06/27 Thursday, 16:00-17:20; Connections Ballroom)

Keynote Speaker: Yannis Papakonstantinou, Professor, Database Research Lab, Computer Science and Engineering Department, University of California, San Diego

About the Speaker

Yannis Papakonstantinou is a Professor of Computer Science and Engineering at the University of California, San Diego. His research is in the intersection of data management technologies and the web, where he has published over ninety research articles that have received more than 12,000 citations, according to Google Scholar. A common theme of his research is the extension of database platforms and query processors beyond centralized relational databases and into semistructured databases, integrated views of distributed databases and web services, textual data and queries involving keyword search, and most recently spatiotemporal sensor data. He has given multiple tutorials and invited talks, has served on journal editorial boards and has chaired and participated in program committees for many international conferences and workshops. He is a co-director and teaches for UCSD's Master of Advanced Studies in Data Science.

Yannis enjoys to commercialize his research and to inform his research accordingly. He was the CEO and Chief Scientist of Enosys Software, which built and commercialized an early Enterprise Information Integration platform for structured and semistructured data. The Enosys Software was OEM'd and sold under the BEA Liquid Data and BEA Aqualogic brand names, eventually acquired in 2003 by BEA Systems. His lab's FORWARD platform was used by many UCSD and commercial applications. He has been involved in data analytics in the pharmaceutical industry, was in the technical advisory board of Brightscope Inc and GraphSQL Inc, and he has been working with Amazon Web Services. He is the inventor of nine patents and has been an expert and consultant in IP litigation cases multiple times.

Yannis holds a Diploma of Electrical Engineering from the National Technical University of Athens, MS and Ph.D. in Computer Science from Stanford University (1997) and an NSF CAREER award for his work on data integration.

Keynote 4: The Millibottleneck Theory of Millisecond-Scale Performance Bugs and Its Experimental Verification (SCF2019-0004)

(06/28 Friday, 9:00-10:20; Connections Ballroom)

Keynote Speaker: Calton Pu, Professor and John P. Imlay, Jr. Chair in Software, School of Computer Science, Georgia Institute of Technology, USA

Abstract: Web-facing applications have complex deployment dependencies and stringent quality of service requirements, e.g., 99.9% of requests with response time within 0.5 seconds. However, despite continued efforts by industry and academic researchers, the latency long tail problem, where a non-trivial fraction of Very Long Response Time (VLRT) requests return after a few seconds, remains a serious research and practical challenge. Latency long tail happens even when the system utilization is still very far from saturation (e.g., 40-60% average CPU utilization). Using automated n-tier application benchmarks, we have reproduced several cases of these VLRT requests (e.g., due to JVM garbage collection and VM-based application consolidation), caused by millibottlenecks (resource saturations that last only tens to hundreds of milliseconds).

The Millbottleneck Theory explains these VLRT requests in a model with two parts. First, a resource millibottleneck is created, and it propagates through a chain of dependencies among system components, accumulating queuing effects that end in VLRT requests. We have released the MilliMonitor toolkit, which is capable of 50ms sampling periods and detailed event monitoring to enable the detection of millibottlenecks and the tracking of chain of dependencies that end in VLRT requests. We will describe a methodical approach to find new millibottlenecks and their chain of dependencies, so we can remove the sources of VLRT requests and improve overall system utilization in data centers while preserving high quality of service.

About the Speaker

Calton Pu was born in Taiwan and grew up in Brazil. He received his PhD from University of Washington and served on the faculty of Columbia University and Oregon Graduate Institute. Currently, he is holding the position of Professor and John P. Imlay, Jr. Chair in Software in the College of Computing, Georgia Institute of Technology. He has worked on several projects in systems and database research. His contributions to systems research include program specialization and software feedback. His contributions to database research include extended transaction models and their implementation. His recent research has focused on automated system management in clouds (Elba project), information quality (e.g., spam processing), and big data in Internet of Things (GRAIT-DM project). He has collaborated extensively with scientists and industry researchers. He has published more than 70 journal papers and book chapters, 280 conferences and refereed workshop papers. He served on more than 120 program committees, including the co-PC chairs of SRDS'95, ICDE'99, COOPIS'02, SRDS'03, DOA'07, DEBS'09, ICWS'10, CollaborateCom'11, ICAC'13, CLOUD'15, BigData Congress'16, CIC'16, and co-general chair of ICDE'97, CIKM'01, ICDE'06, DEPSA'07, CEAS'07, SCC'08, CollaborateCom'08, World Service Congress'11, CollaborateCom'12, IEEE CIC'15, and ICDCS'17. He is a Fellow of AAAS and IEEE.

Keynote 5: Resilient IoT for Community Scale Services (SCF2019-0005)

(06/29 Saturday, 13:00-14:20; Connections Ballroom)

Keynote Speaker: Nalini Venkatasubramanian, Department of Computer Science, University of California, Irvine, USA

Abstract: Advances in technology mobile computing, cyberphysical systems, Internet-of-Things, cloud computing and big data technologies are making available new modalities of information and new channels of communication. It has enabled the interconnection of objects and data to provide novel services that are changing the landscape of cities and communities worldwide. Technologies and services are being rapidly created and repurposed as needed to improve and enrich daily lives of citizens in homes and workplaces. Community lifelines that provide basic utilities such as shelter, food, water and energy are being transformed by technology, but responsible deployment requires an understanding of the vulnerabilities introduced by new

technologies. For example, during large scale disasters and unexpected events such as fires, floods and earthquakes, infrastructure and services are disrupted. The ability to ensure resilient operation under small events and large disasters requires intelligent data collection and data exchange from diverse devices and data sources and interpretation of this information for higher level semantic observations. Drawing on our recent efforts in smartspaces, smart firefighting and smartwater infrastructures, I will discuss the role of IoT and middleware integration technologies to generate situational awareness. The ability to combine novel technologies at multiple layers will open up new possibilities for resilient and scalable communities of the future.

About the Speaker

Nalini Venkatasubramanian is currently a Professor in the School of Information and Computer Science at the University of California Irvine. She has had significant research and industry experience in the areas of distributed systems, adaptive middleware, pervasive and mobile computing, cyberphysical systems, distributed multimedia and formal methods and has over 250 publications in these areas. As a key member of the Center for Emergency Response Technologies at UC Irvine, Nalini's recent research has focused on enabling resilient and sustainable communities using IoT/CPS technologies. In particular, her research addresses scalable observation and analysis of situational information from multimodal input sources; dynamic adaptation of the underlying systems to enable information flow under massive failures and the dissemination of rich notifications to members of the public at large. She is the recipient of the prestigious NSF Career Award, multiple Undergraduate Teaching Excellence Awards and best paper awards. Prof. Venkatasubramanian has served in numerous program and organizing committees of conferences on middleware, distributed systems and multimedia and on the editorial boards of journals. She received and M.S and Ph.D in Computer Science from the University of Illinois in Urbana-Champaign. Her research is supported both by government and industrial sources such as NSF, DHS, ONR, DARPA, Novell, Hewlett-Packard and Nokia. Prior to arriving at UC Irvine, Nalini was a Research Staff Member at the Hewlett-Packard Laboratories in Palo Alto, California.

2019 Services Conference Federation Keynote Panel Sessions

Keynote Panel 1: Intersection of IoT, Edge, Blockchain, AI, and Big Data (SCF2019-0006)

(06/26 Wednesday, 15:20-16:20; Connections Ballroom)

Moderator: Tony Shan, Chief Architect at CTS
Panelists (in alphabetic order of first name):

- Alex Bates, Managing Director of Neocortex Ventures and Technology Chair of The Sandbox
- Arshad Khan, Founder and Chief Scientist at theDevMasters
- Petr Peterka, President and Founder of Relevant Progress
- Raymond Weale, Vice President of Technology and Innovation at Everest Solutions Group
- Tom Caldwell, Group CTO at the Techniche Group

Abstract: It is exciting that technology innovations continue to accelerate at an unprecedented pace. We see tremendous growth and exploration in the hot areas of cloud-native, digital ledger/currency, fog computing, connected devices, deep learning, mixed reality, and advanced analytics. As practitioners and technologists, what are the technical and business opportunities unlocked by these new technologies? How are they related to each other? What are the current barriers and risks that you should be aware of, when you formulate and implement your holistic strategy on emerging technologies? How can we synergize the evolution of IoT, Edge, Blockchain, AI, and Big Data? Why will they converge or diverge?

This panel will delve into the trends of the leading-edge technologies from both a business and a technology perspective. We'll investigate the pragmatic approaches to the successful adoption and interplay of these disciplines. We'll also discuss best practices for integrating these solutions into the smart ecosystem.

About the Moderator:

@tonyshan

Tony Shan is a renowned thought leader and innovative visionary with decades of field experience and guru-level expertise on cutting-edge enterprise computing technologies. He leads incubating and nurturing interdisciplinary practice and enablement on emerging technologies like IoT, big data and AI. He drives award-winning innovation and transformation of most complex enterprise systems. He directs and advises the pragmatic lifecycle design of large-scale distributed solutions on diverse platforms in Fortune 500 companies and public sector organizations. He is a regular speaker and organizer in preeminent conferences. As a book author and an editor/editorial advisory board member of IT research journals, he also founded several user groups/forums.

About the Panelists (in alphabetic order of first name):

@MindAugmented

Alex Bates spent the last decade bringing artificial intelligence and machine learning to the forefront of the industrial market. As an undergraduate he authored 5 peer-reviewed publications (current citation count of 117), performing DARPA-funded research in neural networks, as well as research in memory and computational diagnostics. Next he jumped into the private sector, applying

analytics on some of the world's largest data warehouses at Teradata, a pioneer in big data technology.

Alex co-founded Mtelligence (Mtell) to harness the deluge of sensor data in the industrial IoT, with a mission to create a world that doesn't break down. He is lead inventor on 3 issued patents in the area of sensor networks and machine learning. Mtell was acquired in 2016 by AspenTech (NASDAQ: AZPN), the global leader in process optimization software. Alex received degrees in Mathematics and Computer Science with a concentration in neuroscience. He is a member of the Forbes Technology Council, and the bestselling author of Augmented Mind: AI, Superhumans and the Next Economic Revolution, which explores the use of AI to augment humans to superhuman levels.

@thedevmasters

Arshad Khan is an entrepreneur, with 15-plus years of technology architecture, design, leadership into software development, Artificial intelligence industry. He has Bachelor in Engineering (Computer Science) with concentration into Artificial intelligence. He has MIT professional Education into Data Science & Machine Learning. His background into CSAI concentration has opened doors for him at fortune 500 companies like JAG, Yahoo, Bank of America for more than a decade. Arshad is also part of MIT Sloan Executive Education - Management & Leadership in Artificial Intelligence for business. He has devoted himself to a global cause of improving human life through education using artificial intelligence, Improving and fine-tuning skill sets for superior rapid problem-solving skills for entrepreneurs. He is the founder for sports data science platform eSportsDNA. He has been CTO for various software companies like cumulus data and co-credits. He is now a Founder Chief Scientist at DevMasters & Einstein Assembly.

Petr Peterka has served as CTO of Verimatrix since 2010. In this role, he was responsible for the company's overall security strategy and anti-piracy efforts as well as its advanced technology research and intellectual property development, and participation in standards organizations. He has more than 30 years of experience in the multimedia and entertainment industry and has extensive expertise in content protection techniques for pay-TV and online streaming, as well as forensic watermarking technology and anti-piracy. As the industry migrates to the cloud, Petr's knowledge extends to cloud security, data analytics and machine learning. His research during the last 3-4 years included IoT security, blockchain technology and data analytics.

After a successful M&A transaction, Petr founded an advisory firm Relevant Progress focusing on cyber security, IoT and machine learning as intrusion prevention. Petr earned an MSEE degree from Czech Technical University in Prague and a MSCS degree from University of California in San Diego, and he holds over 30 US patents and numerous patent applications. Petr is also a frequent public speaker at industry events and conferences.

Raymond Weale is a senior technology executive with more than 25 years of experience in innovative solution development, enterprise architecture design and digital transformation leadership. He has led international and multicultural teams throughout his career working with some of the most respected global companies in the airline, consumer packaged goods, banking and financial services industries in the UK, USA, Switzerland and the United Arab Emirates.

Raymond joined the Everest Solutions Group as Vice President of Technology and Innovation as is helping companies go do things better, faster, cheaper, and greener! Leading a team of highly skilled engineers focused on providing technology solutions and services. Engaged with fortune 500 companies as well as small start-ups to build engineering teams responsible for developing world-class products. Prior to

joining Everest Raymond worked in the middle east for Etihad Airways, the world's fastest growing airline, based in the United Arab Emirates serving in a CTO capacity and tasked with driving digital innovation and transforming the customer experience for the 20 million frequent flyer members of Etihad Guest. For 22 years Raymond worked for Nestle and served in progressive technology leadership roles both in the USA as well as at the corporate headquarters in Switzerland.

@cybersdtom

Tom Caldwell, a veteran of Cisco and Microsoft, has deep expertise in delivering Cloud-based software products and large-scale AI/Machine-Learning software systems to enterprise and service providers. Most recently Tom was Co-Founder of CyberFlow Analytics, an AI/ML Behavioral Analytics startup. Webroot acquired CyberFlow in 2016. Tom now leads the Techniche Group as CTO helping to shape a new vision for AI/ML monitoring of light industry IoT. An experienced industry leader in Cyber Security, Cloud infrastructure and AI/ML. With an MS in Computer Science, he has more than 20 years in business and software engineering.

Keynote Panel 2: Global Optimization as a Service (SCF2019-0007)

(06/28 Friday, 14:00-15:00; Connections Ballroom)

Moderator: Dr. Min Luo, Services Society, USA
Panelists (in alphabetic order of first name):

- Ali Arsanjani, 8x8 Cloud Communications, USA
- Yi Pan, Georgia State University, Atlanta, Georgia, USA
- Wu Chou, Essenlix Corpration, USA

Abstract: Global optimization is to find the globally best solution of business or scientific problems/models. Those problems are usually non-linear and constrained, sometimes with competing objectives or constraints. Such problems often rise in many business applications, such as advanced engineering design, biotechnology, data analysis, environmental management, financial planning, process control, risk management, scientific modeling, and others.

Finding the global optimal of such problems is far more difficult than just finding a local optimal solution: analytical methods are frequently not applicable, and the use of numerical solution strategies often leads to many challenges. Recent advances in machine learning and AI based techniques often evolves models with newly gained insights that introduces such difficult but more realistic objectives and constraints that makes existing algorithms inadequate to find solutions for some classes of the problems, significantly impacting the quality and the speed of the solution. The problem-specific nature of the existing algorithms also force business solution providers to fully understand the underlying mathematical nature and limits of those algorithms that itself often leads to issues.

This panel will review optimization techniques applied in the current research and development for machine learning and AI based solutions. The limitation and side-effects of those techniques will be analyzed. Global optimization, on the other hand, can provide a unique, also efficient and effective service as a generalized optimizer that will truly empower adaptive problem solving with better and easy to access solutions.

About the Moderator:

Dr. Min Luo has over 30+ years of successful professional career, including over 25 years of management position at Huawei Central Research Institute, IBM Software Group and Global Services, and two of the

Fortune 500 transportation companies. He is an established expert in the field of next negation software defined networking (SDN), enterprise architecture and information systems, whole life cycle software application and product development, business intelligence, and business process optimization. He is also a pioneer and one of the recognized leading experts and educators in service-oriented architecture (SOA), model/business-driven architecture and development (MDA-D), and component/object-oriented technologies.

About the Panelists (in alphabetic order of first name):

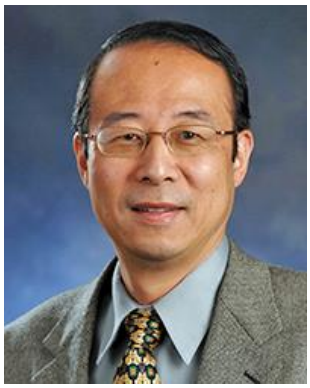
Dr. **Ali Arsanjani** is Vice-president of Artificial Intelligence and Machine Learning at 8x8, the cloud-native, Unified Communications, Collaboration, Contact Center company. He is responsible for a division focused on delivering practical research innovations to infuse AI in the UCaaS and CCaaS space, to create a delightful and proactive customer experience, Enhanced Context-Aware Communications.

Ali is also CTO of Analytics and Machine Learning at Deep Context, a deep learning startup focused on Amalgamation of data for deeper actionable insights. He is an advisor to startups and boards of larger companies.

In his previous role (1998-2018), he was an IBM Distinguished Engineer & CTO for IBM Analytics Hybrid Cloud and Machine Learning responsible for leading services teams in customized machine learning and analytics solutions. He builds teams of teams across multiple geos in large-scale agile solution development His career spans CTO responsibilities for SOA, BPM, Robotic Process Automation, Analytics, Machine Learning and Artificial Intelligence/Cognitive Systems.

He delivering actionable business insights thru' an amalgamation of structured and unstructured data using machine learning and artificial intelligence to augment traditional rule engines, custom/legacy systems. In past jobs he has been CTO for SOA, BPM, Decision Management, Analytics, Content Management.

Ali Arsanjani has chaired standard bodies such as The Open Group and is responsible for co-leading the SOA Reference Architecture, SOA Maturity Model, and Cloud Computing Architecture standards. He & his team specialize in harvesting, developing best-practices for the microservices SOA and Web Services lifecycle on hundreds of projects WW across multiple industries, leading a community of practice of over 6000 people.



Dr. **Yi Pan** is currently a Regents' Professor and Chair of Computer Science at Georgia State University, USA and a Member of EU Academy of Sciences. He has served as an Associate Dean and Chair of Biology Department during 2013-2017 and Chair of Computer Science during 2006-2013. Dr. Pan received his B.E. and M.E. degrees in computer engineering from Tsinghua University, China, in 1982 and 1984, respectively, and his Ph.D. degree in computer science from the University of Pittsburgh, USA, in 1991. His profile has been featured as a distinguished alumnus in both Tsinghua Alumni Newsletter and University of Pittsburgh CS Alumni Newsletter. Dr. Pan's research interests include parallel and cloud computing, big data, and bioinformatics. Dr. Pan has published more than 250 journal papers with over 90 papers published in various IEEE journals. In addition, he has published over 150 papers in refereed conferences. He has also co-authored/co-edited 43 books. His work has been cited more than 10,000 times in Google Scholar and his current H-index is 53. Dr. Pan has served as an editor-in-chief or editorial board member for 20 journals including 7 IEEE Transactions.

He is the recipient of many awards including IEEE Transactions Best Paper Award, several other conference and journal best paper awards, 4 IBM Faculty Awards, 2 JSPS Senior Invitation Fellowships, IEEE BIBE Outstanding Achievement Award, NSF Research Opportunity Award, and AFOSR Summer Faculty Research Fellowship. He has organized many international conferences and delivered keynote speeches at over 60 international conferences around the world.

Dr. **Wu Chou** is an IEEE Fellow and Vice President of AI & Software, Essenlix Corporation. Prior to joining Essenlix, he was VP and CTO of enterprise network at Huawei, with extensive experience and contributions in AI, data networking, services, and IoT (Internet-of-things). He published over 200 conference and journal papers with over 50 granted US and International patents.

2019
Services Conference Federation Tutorials

Tutorial 1: Practical DataOps (SCF2019-0008)

(06/25 Tuesday, 9:30-12:00; Navigator)

Abstract: DataOps has emerged as a systematic method to enhance the speed and quality of analytics through sophisticated data processing and management processes. It is hard to implement DataOps efficiently, due to the complexity and immaturity. This talk presents a methodical framework of strategization and operationalization, comprising the key building blocks of the overarching DataOps practice: Foundation, Anatomy, Discipline, and Enablement (FADE). We will explore the essential elements of DataOps discipline for the pragmatic adoption and brown-field undertakings in complex enterprise environments. Further, we will investigate the techniques and means to expedite the execution in real-world projects. In addition, we will delve into some working examples and real-life cases to illustrate how to apply and accelerate FADE effectively in open source applications. Best practices, guidelines, implementation experience, and lessons learned will be discussed in the session as well.

About the Speaker

@tonyshan

Tony Shan is a renowned thought leader and innovative visionary with decades of field experience and guru-level expertise on cutting-edge enterprise computing technologies. He leads incubating and nurturing interdisciplinary practice and enablement on emerging technologies like IoT, big data and AI. He drives award-winning innovation and transformation of most complex enterprise systems. He directs and advises the pragmatic lifecycle design of large-scale distributed solutions on diverse platforms in Fortune 500 companies and public sector organizations. He is a regular speaker and organizer in preeminent conferences. As a book author and an editor/editorial advisory board member of IT research journals, he also founded several user groups/forums.

Tutorial 2: Cybersecurity Engineering Principles, Process and Standardization (SCF2019-0009)

(06/25 Tuesday, 14:00-16:30; Navigator)

Abstract: In this tutorial, cybersecurity engineering process, principles and standards are being covered. It builds a foundation first on cybersecurity as an engineering process and then discusses general cybersecurity principles and design process. It introduces and navigates various cybersecurity standards defined by various national agencies and institutions such as NIST and IEEE. Participants should have background on system design and information assurance. After the tutorial, participants should have a clear understanding of cybersecurity design principles following published standards and be able to address security issues raised by examples from real cyber world.

About the Speakers

Dr. **Z Chen** has been working on the cybersecurity for over 20 years. His research interests are data security and exploration. He has published over 40 papers. He has worked in IBM research, Shanghai Jiao Tong University and Mercy College. He is the senior members in IEEE and ACM as well. He holds CISSP since 2010. The Cybersecurity Education Center he founded in 2007 is a national center of academic excellence on cyber defense education (CAE-CDE) designated by NSA and DHS. He was awarded several grants from DHA for developing cyber security course ware.

2019 International Conference on Web Services (ICWS 2019)

Session I: Services Modeling (06/28 Friday, 10:40 - 11:40; Destination Ballroom 1)

Chair: Liang-Jie Zhang (Kingdee International Software Group Co., Ltd, China)
20 minutes for each paper.

10:40 Modeling Social Influence in Mobile Messaging Apps (ICWS2019-0001)

Songmei Yu and Sofya Poger (Felician University, USA)

Abstract: Social influence is the behavioral change of a person because of the perceived relationship with other people, organizations and society in general. With the exponential growth of online social network services especially mobile messaging apps, users around the world are logging in to messaging apps to not only chat with friends but also to connect with brands, browse merchandise, and watch content. Mobile chat apps boast a number of distinct characteristics that make their audiences particularly appealing to businesses and marketers, including their size, retention and usage rates, and user demographics. The combined user base of the top four chat apps is larger than the combined user base of the top four social networks. Therefore, it makes great sense to analyze user behavior and social influence in mobile messaging apps. In this paper, we focus on computational aspects of measuring social influence of groups formed in mobile messaging apps. We describe the special features of mobile messaging apps and present challenges. We address the challenges by proposing a temporal weighted data model to measure the group influence in messaging apps by considering their special features, with implementation and evaluation in the end.

11:00 Pricing a digital services marketplace under asymmetric information (ICWS2019-0002)

Pavel Izhutov (Axio Inc & Altos Platform and Stanford University, USA); Haim Mendelson (Stanford University, USA)

Abstract: This paper addresses the pricing problem of an online service marketplace under asymmetric information. An example is an online learning platform such as Coursera that provides courses from suppliers (in this case, universities) to learners. We focus on the matching function of the marketplace whereby it engages in sequential search on behalf of a consumer using partially-observable consumer and supplier attributes. We develop the optimal pricing policies for a general distribution of the unobservable valuations. When these distributions are exponential, it is optimal to charge the {\it same} total fee for each match rather than engage in price discrimination, and this entire fee should be levied on the less elastic side of the marketplace.

11:20 Profit Maximization and Time Minimization Admission Control and Resource Scheduling for Cloud Based Big Data Analytics-as-a-Service Platforms (ICWS2019-0003)

Yali Zhao (University of Melbourne, Australia); Rodrigo N. Calheiros (Western Sydney University, Australia); Athanasios V. Vasilakos (Lulea University of Technology, Sweden); James Bailey and Richard O. Sinnott (University of Melbourne, Australia)

Abstract: Big data analytics typically requires large amounts of resources to process ever-growing data volumes. This can be time consuming and result in considerable expenses. Analytics-as-a-Service (AaaS) platforms provide a way to tackle expensive resource costs and lengthy data processing times by leveraging automatic resource management with a pay-per-use service delivery model. This paper focuses on optimization resource management algorithms for AaaS platforms that automatically and elastically provision cloud resources to execute queries with Service Level Agreement (SLA) guarantees that meet Quality of Service (QoS) requirements of queries. We present admission control and resource scheduling algorithms that serve multiple objectives including profit maximization for AaaS platforms and query time minimization for users. Moreover, to enable queries that require timely responses and/or have constrained budgets, we apply data sampling based admission control and resource scheduling where the accuracy can be traded-off for less costs and quicker responses when necessary. We conduct extensive experimental evaluations for the algorithm performances compared to state-of-art algorithms. The results show that our proposed algorithms perform significantly better in increasing the admission rates, consuming less resources and hence reducing costs, and ultimately provide a more flexible resource scheduling solution for fast, cost-effective, and reliable big data processing.

Session II: Emerging Services Applications (06/26 Wednesday, 11:00 - 12:00; Destination Ballroom 1)

Chair: Kisung Lee (Louisiana State University, USA)
20 minutes for each paper.

11:00 A User Constraint Awareness Approach for QoS-based Service Composition (ICWS2019-0004)

Zhihui Wu, Piyuan Lin, Peijie Huang, Huachong Peng, Yihui He and Junan Chen (South China Agricultural University, China)

Abstract: Web service composition adopts functional features including the inputs and outputs, and non-functional features including Quality of service (QoS), conditional structure constraints, user preferences, and trusts to compose homogeneous or heterogeneous services together in order to create value-added services. However, in some complex practical application scenarios, the web services with the same function can provide the generous differentiated contents, and there is no approach to focus on the user's constraints on the content provided by the web services. In this paper, we focus on handling three composition dimensions simultaneously including functional features, QoS and the user's constraints on the contents provided by the web services. Therefore, an improved genetic algorithm to obtain an optimal solution for this task is applied. In addition, we also take into consideration the over-constrained problem caused by implicit conflicting constraints and improve a constraint correction approach to solve this problem with less cost of consistency checks. Experimental results using the real datasets about travel demonstrate the effectiveness of our approach in creating the fully functional and quality-optimized solutions, on the premise that the user's constraints on the content are satisfied.

11:20 AnomalyDetect：An Online Distance-Based Anomaly Detection Algorithm (ICWS2019-0005)

Wunjun Huo, Wei Wang and Wen Li (Tongji University, China)

Abstract: Anomaly detection is a key challenge in data mining, which refers to finding patterns in data that do not conform to expected behavior. It has a wide range of applications in many fields as diverse as finance, medicine, industry, and the Internet. In particular, intelligent operation has made great progress in recent years and has an urgent need for this technology. In this paper, we study the problem of anomaly detection in the context of intelligent operation and find the practical need for high-accuracy, online and universal anomaly detection algorithms in time series database. Based on the existing algorithms, we propose an innovative online distance-based anomaly detection algorithm. K-means and time-space trade-off mechanism are used to reduce the time complexity. Through the experiments on Yahoo! Web-scope S5 dataset we show that our algorithm can detect anomalies successfully. The comparative study of several anomaly detectors verifies the effectiveness and generality of the proposed algorithm.

11:40 Transitive Pseudonyms Mediated Electronic Health Records Sharing for Very Important Patients (ICWS2019-0006)

Huafei Zhu (Nanyang Technology University, Singapore); Ng Keong (Nanyang Technological University, Singapore)

Abstract: Electronic health record (EHR) greatly enhances the convenience of cross-domain sharing and has been proven effectively to improve the quality of healthcare. On the other hand, the sharing of sensitive medical data is facing critical security and privacy issues, which become an obstacle that prevents EHR being widely adopted. In this paper, we address several challenges in very important patients' (VIPs) data privacy, including how to protect a VIP's identity by using pseudonym, how to enable a doctor to update an encrypted EHR with the VIP's absence, how to help a doctor link up and decrypt historical EHRs of a patient for secondary use under a secure environment, and so on. Then we propose a framework for secure EHR data management. In our framework, we use a transitive pseudonym generation technique to allow a patient to vary his/her identity in each hospital visit. We separate metadata from detailed EHR data in storage, so that the security of EHR data is guaranteed by the security of both the central server and local servers in all involved hospitals. Furthermore, in our framework, a hospital can encrypt and upload a patient's EHR when he/she is absent; a patient can help to download and decrypt his/her previous EHRs from the central server; and a doctor can decrypt a patient's historical EHRs for secondary use under the help and audit by several proxies.

Session III: Services and Cloud (06/29 Saturday, 9:00 - 10:00; Destination Ballroom 1)

Chair: Yishuang Ning (Tsinghua University, China)
20 minutes for each paper.

9:00 A Novel Coalitional Game-theoretic Approach for Energy-Aware Dynamic VM Consolidation in Heterogeneous Cloud Datacenters (ICWS2019-0007)

Xuan Xiao, Yunni Xia (Chongqing Universit, China), Feng Zeng (Discovery Technology (shenzhen) limited, China),Wanbo Zheng (Kunming University of Science and Technology, China), Xiaoning Sun, Qinglan Peng (Chongqing Universit, China), Yu Guo (Sichuan University, China) and Xin Luo (Chongqing Institute of Green and Intelligent Technology, China)

Abstract: Server consolidation technique plays an important role in energy management and load-balancing of cloud computing systems. Dynamic virtual machine (VM) consolidation is a promising consolidation approach in this direction, which aims at using least active physical machines (PMs) through appropriately migrating VMs to reduce resource consumption. The resulting optimization problem is well-acknowledged to be NP-hard optimization problems. In this paper, we propose a novel merge-and-split-based coalitional game-theoretic approach for VM consolidation in heterogeneous clouds. The proposed approach first partitions PMs into different groups based on their load levels, then employs a coalitional-game-based VM consolidation algorithm (CGMS) in choosing members from such groups to form effective coalitions, performs VM migrations among the coalition members to maximize the payoff of every coalition, and close PMs

with low energy-efficiency Experimental results based on multiple cases clearly suggest that our proposed approach outperforms traditional ones in terms of energy-saving and level of load fairness.

9:20 A Web-Service to Monitor a Wireless Sensor Network (ICWS2019-0009)

Rayanne Silveira (Federal University of Maranhão, Brazil); Allyx Fontaine (Université de Guyane & UMR Espace-Dev, French Guiana); Ewaldo Santana (University of State of Maranhao, Brazil); Francisco Alves (Federal University of Maranhão, Brazil)

Abstract: In the last years, the interest in Internet of Things is growing, and WSN is a promising technology that could be applied in different situations. Regardless of the nature of the application, WSNs are often used for data acquisition, to obtain information from an environment of interest, so it is important to consider how this data will be made available to users. Over the last years, an increasing number of web services have been used to deal with databases and final users, providing a familiar interface and a multiplatform access to those data. To address this problem, in this paper is proposed a web application based on MVC architecture to monitor, organize and manage devices and data in a wireless sensor network. A functional evaluation of the proposed system is presented, and the test results are discussed.

Session IV: Services Platform and Architecture (06/27 Thursday, 9:00 - 10:00; Destination Ballroom 1)

Chair: Kisung Lee (Louisiana State University, USA)
20 minutes for each paper.

9:00 Automated Hot Text and Huge Pages: An Easy-to-adopt Solution Towards High Performing Services (ICWS2019-0010)

Zhenyun Zhuang, Mark Santaniello, Shumin Zhao, Bikash Sharma and Rajit Kambo (Facebook, USA)

Abstract: Performance optimizations of large scale services can lead to significant wins on service efficiency and performance. CPU resource is one of the most common performance bottlenecks, hence improving CPU performance has been the focus in many performance optimization efforts. In particular, reducing iTLB miss rates can greatly improve CPU performance and speed up service running. At Facebook, we have achieved CPU reduction by applying a solution that identifies hot-text of the software binary and places the binary on huge pages. The solution is wrapped into an automated framework, enabling service owners to effortlessly adopt it. Our framework has been applied to many services at Facebook, and this paper shares our experiences and findings.

9:20 ThunderML: A Toolkit for Enabling AI/ML Models on Cloud for Industry 4.0 (ICWS2019-0011)

Dhaval Patel (IBM T. J. Watson Research Center, NY & Research Staff Member, USA); Shrey Shrivastava (IBM, USA); Wesley M. Gifford (IBM Research & T. J. Watson Research Center, USA); Stuart Siegel (IBM, USA); Jayant Kalagnanam (IBM Research, USA)

Abstract: AI, machine learning, and deep learning tools have now become easily accessible on the cloud. However, the adoption of these cloud-based services for heavy industries has been limited due to the gap between general purpose AI tools and operational requirements for production industries. There are three fundamentals gaps. The first is the lack of purpose built solution pipelines designed for common industrial problem types, the second is the lack of tools for automating the learning from noisy sensor data and the third is the lack of platforms which help practitioners leverage cloud-based environment for building and deploying custom modeling pipelines. In this paper, we present a system ThunderML that addresses these gaps by providing AI-based solution patterns that allow rapid authoring, training and deployment for Industry 4.0 applications. Importantly, the system also facilitates cloud-based deployments by providing a vendor agnostic pipeline execution and deployment layer.

9:40 Study of Twitter communications on cardiovascular disease by state health departments (ICWS2019-0012)

Aibek Musaev, Rebecca Britt, Jameson Hayes, Brian Britt, Jessica Maddox and Pezhman Sheinidashtegol (The University of Alabama, USA)

Abstract: The present study examines Twitter conversations around cardiovascular health in order to assess the topical foci of these conversations as well as the role of various state departments of health. After scraping tweets containing relevant keywords, Latent Dirichlet Allocation (LDA) was used to identify the most important topics discussed around the issue, while PageRank was used to determine the relative prominence of different users. The results indicate that a small number of state departments of health play an especially significant role in these conversations, and that irregular events like ebola outbreaks also exert a strong influence over the volume of tweets made in general by state departments of health.

10:00 An Efficient Traceable and Anonymous Authentication Scheme for Permissioned Blockchain (ICWS2019-0008)

Qianqian Su, Rui Zhang, Rui Xue and You Sun (The Institute of Information Engineering, Chinese Academy of Sciences, China)

Abstract: Blockchain has become a hot topic in recent years. Many applications apply permissioned blockchain to achieve secure data sharing across organizations such as healthcare blockchain. In the permissioned blockchain, on the one hand, the blockchain system is required to support efficient and dynamic authentication for adding and deleting users in a distributed environment. On the other hand, in some particular applications such as healthcare domain, users prefer to keep anonymity in the process of authentication. Although many solutions for anonymous authentication have been proposed, they often require the participation of a central trusted party in the process of authentication and are not efficient enough. In this paper, we focus on designing an efficient traceable and anonymous authentication scheme, which supports efficient authentication while without revealing user's identity information and does not requires the participation of a central trusted party. While, in case of dispute, the identity of users can be revealed. Moreover, the proposed scheme is able to support dynamic adding and deleting users. Finally, we analyze the security and privacy properties of the proposed scheme and evaluate its performance in terms of computational cost. The experimental results show that the proposed scheme is more efficient than exist schemes and can be easily deployed in the permissioned blockchain.

2019 International Conference on Cloud Computing (CLOUD 2019)

Session I: Cloud Performance (06/26 Wednesday, 11:00-12:00; Navigator)

Chair: Wei Wang (The University of Texas at San Antonio, USA)
20 minutes for each paper.

11:00 Ultra-Low Power Localization System Using Mobile Cloud Computing (CLOUD2019-1001)

Junjian Huang and Yubin Zhao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China); Xiaofan Li (The State Radio Monitoring Center and Testing Center, China); Chengzhong Xu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China)

Abstract: In the existing positioning system based on bluetooth (BT), the interference of the positioning device signal, the slow processing speed of the positioning data and the large energy consumption of the positioning device affect the system positioning accuracy and service quality. In this paper, we propose an Ultra-Low power indoor localization system using mobile cloud computing. The mobile cloud server reduces the signal interference of the positioning device, improves the positioning accuracy and reduces the system energy consumption by controlling the working mode of the positioning device. A simultaneous localization and power adaptation scheme is developed. In the real experiment evaluation, our proposed system can localize the area of a terminal located within 3 m distance with 98% accuracy and average positioning error less then 1.55 m. Compare with other BLE system, 97% average energy consumption of our system is reduced.

11:20 A Method and Tool for Automated Induction of Relations from Quantitative Performance Logs (CLOUD2019-1002)

Joshua Kimball and Calton Pu (Georgia Institute of Technology, USA)

Abstract: Operators use performance logs to manage large-scale web service infrastructures. Detecting, isolating and diagnosing fine-grained performance anomalies require integrating system performance measures across space and time. The diversity of these logs layouts impedes their efficient processing and hinders such analyses. Performance logs possess some unique features, which challenge current log parsing techniques. In addition, most current techniques stop at extraction leaving relational definition as a post-processing activity, which can be a substantial effort at web scale. To achieve scale, we introduce our "perftables" approach, which automatically interprets performance log data and transforms the text into structured relations. We interpret the signals provided by the layout using our template catalog to induce an appropriate relation. We evaluate our method on a large sample obtained from our experimental computer science infrastructure in addition to a sample drawn from the wild. We were able to successfully extract on average over 97% and 85% of the data respectively.

11:40 Systematic Construction, Execution, and Reproduction of Complex Performance Benchmarks (CLOUD2019-1003)

Rodrigo Alves Lima, Joshua Kimball and Calton Pu (Georgia Institute of Technology, USA); J. E. (Joao Eduardo) Ferreira (University of Sao Paulo, Brazil)

Abstract: In this work, the next generation of the Elba toolkit, available under a Beta release, is presented. Over the years, project Elba researchers and collaborators have conducted thousands of large-scale benchmarks to study interesting phenomena in cloud computing environments. However, these benchmarks' hidden and implicit dependencies not only made them difficult to reproduce but also hindered the ability of project Elba researchers to easily increase their configuration spaces. Therefore, a major contribution of this release is WED-Make: a declarative workflow language for the construction, execution, and reproduction of complex benchmarks. A distinguishing characteristic of WED-Make is to facilitate the declaration of benchmark dependencies. Particularly, the execution of benchmarks specified with WED-Make is driven by declared dependencies themselves. The Elba toolkit also comprises fine-grained resource monitors and software specialized with low overhead event loggers, enabling a unique infrastructure for the study of interesting phenomena in cloud computing environments.

Session II: Cloud Infrastructure (06/27 Thursday, 13:00-14:00; Navigator)

Chair: Weipeng Lin (Simon Fraser University, Canada)
20 minutes for each paper.

13:00 Multiple Workflow Scheduling with Offloading Tasks to Edge Cloud (CLOUD2019-1004)

Hidehiro Kanemitsu (Tokyo University of Technology, Japan); Masaki Hanada (Tokyo University of Information

Sciences, Japan); Hidenori Nakazato (Waseda University, Japan)
Abstract: Edge computing can realize a data locality among a cloud and users, and it can be applied to task offloading, i.e., a part of workload on a mobile client is moved to an edge cloud to minimize the response time with reducing energy consumption. However, a policy to determine whether each task in a workflow should be offloaded or not is one of the current challenging issues in task offloading and task scheduling. In this paper, we propose a task scheduling algorithm with task offloading, called priority-based continuous task selection for offloading (PCTSO), to minimize the schedule length with energy consumption at a mobile client is reduced. PCTSO tries to select dependent tasks such that many tasks are offloaded so as to utilize many vCPUs in the edge cloud. Experimental results of the simulation show that PCTSO outperforms other algorithms in the schedule length and satisfies the energy constraint.

13:20 Min-Edge P-cycles: An efficient Approach for Computing P-cycles in Optical Data Center Networks (CLOUD2019-1005)

Amir Mirzaeinia (New Mexico Inst of Mining and Technology, USA); Abdelmounaam Rezgui (Illinois State University, USA); Zaki Malik (Business Analytics, Texas A&M University-Commerce, USA)
Abstract: Effective network protection requires that extra resources be used in failure events. Pre-configured protection cycles (P-cycles) are proposed to protect mesh-based networks using few extra resources. A number of heuristic methods have been developed to overcome the complexity of finding optimum P-cycles in dense optical networks. The processing time of existing approaches depends on the number of working wavelengths. As the number of working wavelengths is increasing in modern networks, the processing time of current P-cycle computing approaches will continue to increase. In this paper, we propose an approach, called Min-Edge P-cycle (MEP), that addresses this problem. The core of the proposed approach is an iterative algorithm that uses the minimum-weight edge in each iteration. Our approach provides the same redundancy requirements as the previously known unity cycle method but it does not depend on the number of working wavelengths. The new approach can significantly reduce the processing time of computing P-cycles in large scale optical, server-centric data center networks, e.g., BCube, FiConn, and DCell networks.

13:40 BlockLite: Toward Accurate and Efficient Emulation of Public Blockchains in the Cloud (CLOUD2019-1006)

Xinying Wang, Abdullah Al-Mamun, Feng Yan and Dongfang Zhao (University of Nevada, Reno, USA)
Abstract: Blockchain is an enabler of many emerging decentralized applications in areas of cryptocurrency, Internet of Things, smart healthcare, among many others. Although various open-source blockchain frameworks are available in the form of virtual machine images or docker images on public clouds, the infrastructure of mainstream blockchains nonetheless exhibits a technical barrier for many users to modify or test out new research ideas in blockchains. To make it worse, many advantages of blockchain systems can be demonstrated only at large scales, e.g., thousands of nodes, which are not always available to researchers. This paper presents an accurate and efficient emulating system, namely BlockLite, to replay the execution of large-scale blockchain systems on tens of thousands of nodes. In contrast to existing work that simulates blockchains with artificial timestamp injection, BlockLite is designed to be executing real proof-of-work workload along with peer-to-peer network communications and hash-based immutability. BlockLite employs a preprocessing approach to avoid the per-node computation overhead at runtime and thus achieves practical scales. We have evaluated BlockLite for emulating up to 20,000 nodes on Amazon Web Services (AWS), showing both high accuracy and high efficiency with millions of transactions.

Session III: Cloud Reliability (06/26 Wednesday, 14:00-15:00; Navigator)

Chair: Abdelmounaam Rezgui (Illinois State University, USA)
20 minutes for each paper.

14:00 Teleportation of VM Disk Images over WAN (CLOUD2019-1007)

Oleg Zaydman (VMware Inc, USA); Roman Zhirin (VMware, Inc, USA)
Abstract: As edge computing and hybrid clouds gain momentum, migrating virtual machines between datacenters is becoming increasingly important. Whether such migration is performed live or not, it starts with a full copy of a virtual disk over the network. This initial copy is consuming the bulk of the transfer time and network use. Improving this copy is the focus of our paper. While compression can somewhat help with this, we propose a novel technique, which we call teleportation. Teleportation assembles disk images directly at the destination from the pieces of other, unrelated disk images already present there. Since the data found at the destination doesn't have to be sent over, our prototype has achieved 3.4x increase in network throughput (comparing to compression).

14:20 Live Migration of Virtual Machines in OpenStack: A Perspective from Reliability Evaluation (CLOUD2019-1008)

Jin Hao (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences & Northeast Normal University, China); Kejiang Ye (Chinese Academy of Sciences, China); Chengzhong Xu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China)

Abstract: Virtualization technology is widely used in cloud data centers and IT infrastructure. A key technology for server virtualization is the live migration of virtual machines (VMs). This technology allows VMs to be moved from one physical host to another while minimizing service downtime. The cloud providers usually use cloud operating system for virtual machine management. Currently the most widely used open source cloud operating system is OpenStack. In this paper, we investigate the reliability of VM live migration in OpenStack by increasing the system pressures and injecting network failures during the migration. We analyze the impact of these pressures and failures on VM live migration performance. The experimental results can be used to guide data center administrators in migration decisions and fault localization. Furthermore, it can help researchers to find bottlenecks and optimization methods for live migration in OpenStack.

14:40 An Approach to Failure Prediction in Cluster by Self-Updating Cause-and-Effect Graph (CLOUD2019-1009)

Yan Yu and Haopeng Chen (Shanghai Jiao Tong University, China)

Abstract: Cluster systems have been widely used in cloud computing, high-performance computing, and other fields, and the usage and scale of cluster systems have shown a sharp upward trend. Unfortunately, the larger cluster systems are more prone to failures, and the difficulty and cost of repairing failures are unusually huge. Therefore, the importance and necessity of failure prediction in cluster systems are obvious. In or- der to solve this severe challenge, we propose an approach to failure prediction in cluster systems by Self-Updating Cause-and-Effect Graph. Different from the previous approaches, the most novel point of our approach is that it can automatically mine the causality among log events from cluster systems, and set up and update Cause-and-Effect Graph for failure prediction throughout their life cycle. In addition, we use the real logs from Blue Gene/L system to verify the effectiveness of our approach and compare our approach to other approaches using the same logs. The result shows that our approach outperforms other approaches with the best precision and recall rate reaching 89% and 85%, respectively.

Session IV: Cloud Security and Privacy (06/27 Thursday, 09:00-10:00; Navigator)

Chair: Krishna Kant (Temple University, USA)
20 minutes for each paper.

9:00 Towards Decentralized Deep Learning with Differential Privacy (CLOUD2019-1010)

Hsin-Pai Cheng (Duke University, USA); Patrick Yu (Monta Vista High School, USA); Haojing Hu (Beihang University of Aeronautics and Astronautics, USA); Syed Zawad and Feng Yan (University of Nevada, Reno, USA); Shiyu Li (Tsinhua University, USA); Hai (Helen) Li and Yiran Chen (Duke University, USA)

Abstract: In distributed machine learning, while a great deal of attention has been paid on centralized systems that include a central parameter server, decentralized systems have not been fully explored. Decentralized systems have great potentials in the future practical use as they have multiple useful attributes such as less vulnerable to privacy and security issues, better scalability, and less prone to single point of bottleneck and failure. In this paper, we focus on decentralized learning systems and aim to achieve differential privacy with good convergence rate and low communication cost. To achieve this goal, we propose a new algorithm, Leader-Follower Elastic Averaging Stochastic Gradient Descent (LEASGD), driven by a novel Leader-Follower topology and differential privacy model. We also provide a theoretical analysis of the convergence rate of LEASGD and the trade-off between the performance and privacy in the private setting. We evaluate LEASGD in real distributed testbed with poplar deep neural network models MNIST-CNN, MNIST-RNN, and CIFAR-10. Extensive experimental results show that LEASGD outperforms state-of-the-art decentralized learning algorithm DPSGD by achieving nearly 40% lower loss function within same iterations and by 30% reduction of communication cost. Moreover, it spends less differential privacy budget and has final higher accuracy result than DPSGD under private setting.

9:20 Dynamic Network Anomaly Detection System by using Deep Learning Techniques (CLOUD2019-1011)

Peng Lin (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences & Shenzhen University, China); Kejiang Ye (Chinese Academy of Sciences, China); Chengzhong Xu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China)

Abstract: The Internet and computer networks are currently suffering from serious security threats. In order to maintain the stability of the network, we designed and implemented a dynamic network anomaly detection system using deep learning methods. We used LSTM (Long Short Term Memory) to build a deep neural network model and added an Attention mechanism to enhance the performance of the model. The SMOTE algorithm and an improved loss function are used to handle the class-imbalance problem in the CSE-CIC-IDS2018 dataset. The experimental results show that the classification accuracy of our model reaches 96.2%, which is higher than other machine learning algorithms used in the experiment. In addition, the class-imbalance problem is alleviated to a certain extent, making our method have great practicality.

9:40 CPR: Client-Side Processing of Range Predicates (CLOUD2019-1012)

Shahram Ghandeharizadeh (USC, USA); Hieu Nguyen (Ebay, USA); Yazeed Alabdulkarim (King Saud University, USA)

Abstract: Range predicates are important to diverse application workloads. A system may process range predicates using either a server-side, a client-side, or a hybrid of these two solutions. This study presents CPR, a client-side solution that caches the result of range predicates and looks up their results. This implementation provides strong consistency and supports alternative write policies. It is embodied in a flexible framework named RangeQP that provides the hybrid solution. We quantify strengths and limitations of CPR when compared with the server-side solution.

Session V: Cloud Storage (06/27 Thursday, 14:20-15:20; Navigator)

Chair: Shahram Ghandeharizadeh (University of Southern California, USA)
20 minutes for each paper.

14:20 Heterogeneity-aware data placement in Hybrid Clouds (CLOUD2019-1013)

Jack Marquez and Juan Gonzalez (Universidad Autonoma de Occidente, Colombia); Oscar H Mondragon (Universidad Autónoma de Occidente, Colombia)

Abstract: In next-generation cloud computing clusters, performance of data-intensive applications will be limited, among other factors, by disks data transfer rates. In order to mitigate those performance impacts, cloud systems offering hierarchical storage architectures are becoming commonplace. The Hadoop File System (HDFS) offers a collection of storage policies that exploit different storage types such as RAM_DISK, SSD, HDD, and ARCHIVE. However, developing algorithms to leverage heterogeneous storage through an efficient data placement has been challenging. This work presents an intelligent algorithm based on genetic programming which allow to find the optimal mapping of input datasets to storage types on a Hadoop file system.

14:40 Towards Automated Configuration of Cloud Storage Gateways: A data driven approach (CLOUD2019-1014)

Sanjeev Sondur and Krishna Kant (Temple University, USA)

Abstract: Cloud storage gateways (CSGs) are an essential part of enterprises to take advantage of the scale and flexibility of cloud object store. A CSG provides clients the impression of a locally configured large size block-based storage device, which needs to be mapped to remote cloud storage which is invariably object based. Proper configuration of the cloud storage gateway is extremely challenging because of numerous parameters involved and interactions among them. In this paper, we study this problem for a commercial CSG product that is typical of offerings in the market. We explore how machine learning techniques can be exploited both for the forward problem (i.e. predicting performance from the configuration parameters) and backward problem (i.e. predicting configuration parameter values from the target performance). Based on extensive testing with real world customer workloads, we show that it is possible to achieve excellent prediction accuracy while ensuring that the model is not overfitted to the data.

15:00 The Case for Physical Memory Pools: A Vision Paper (CLOUD2019-1015)

Heather Craddock, Lakshmi Prasanna Konudula, Kun Cheng and Gökhan Kul (Delaware State University, USA)

The cloud is a rapidly expanding and increasingly prominent component of modern computing. However, monolithic servers limit the flexibility of the cloud-based systems due to the static memory limitations. Developments in OS design, distributed memory systems, and address translation have been crucial in aiding the progress of the cloud. In this paper, we discuss recent developments in virtualization, OS design and distributed memory structures with regards to their current impact and relevance to future work on eliminating memory limits in cloud computing. We argue that creating physical memory pools is essential for cheaper and more efficient cloud computing infrastructures, and we identify research challenges to implement these structures.

Session VI: Cloud and Distributed Systems (06/28 Friday, 10:40-11:40; Navigator)

Chair: Zhilin Zhang (Simon Fraser University, Canada)
20 minutes for each paper.

10:40 A Parallel Algorithm for Bayesian Text Classification Based on Noise Elimination and Dimension Reduction in Spark Computing Environment (CLOUD2019-1016)

Zhuo Tang (Hunan University, China); Wei Xiao, Xin He, Bin Lu, Youfei Zuo (Hunan University, China);Yuan Zhou (Tsinghua University, China) and Keqin Li (Hunan University, China)

Abstract: The Naive Bayesian algorithm is one of the ten classical algorithms in data mining, which is widely used as the basic theory for text classification. With the high-speed development of the Internet and information systems, huge amount of data are being produced all the time. Some problems are certain to arise when the traditional Bayesian classification algorithm addresses massive amount of data, especially without the parallel computing framework. This paper proposes an improved Bayesian algorithm INBCS, for text classification in the Spark computing environment and improves the Naive Bayesian algorithm based on a polynomial model. For the data preprocessing, this paper first proposes a parallel noise

elimination algorithm, and then proposes another parallel dimension reduction algorithm based on Information Gain and TextRank computation in the Spark environment. Based on these preprocessed data, an improved parallel method is proposed for calculating the conditional probability that comprehensively considers the effects of the feature items in each document, class and training set. Finally, through experiments on different widely used corpuses on the Spark computation platform, the results illustrate that INBCS can obtain higher accuracy and efficiency than some current improvements and implementations of the Naive Bayesian algorithms in Spark ML-library.

11:00 On the Optimal Number of Computational Resources in MapReduce (CLOUD2019-1017)

Htway Htway Hlaing (Waseda University, Japan); Hidehiro Kanemitsu (Tokyo University of Technology, Japan); Tatsuo Nakajima and Hidenori Nakazato (Waseda University, Japan)

Abstract: Big data computing in cloud needs faster processing and better resource provisioning. MapReduce is the framework for computing large scale datasets in cloud environments. Optimization of resources requirement for each job to satisfy a specific objective in MapReduce is an open problem. Many factors, e.g., system side information and requirements of each client, must be considered to estimate the appropriate amount of resources. This paper presents a mathematical model for optimal number of map tasks in MapReduce resource provisioning. This mathematical model is to estimate the optimal number of mappers based on resource specification and the size of dataset.

11:20 Class Indistinguishability for Outsourcing Equality Conjunction Search (CLOUD2019-1018)

Weipeng Lin, Ke Wang and Zhilin Zhang (Simon Fraser University, Canada); Ada Waichee Fu (the Chinese University of Hong Kong, Hong Kong); Raymond Chi-Wing Wong (Hong Kong University of Science and Technology, Hong Kong); Cheng Long (Nanyang Technological University, Singapore)

Abstract: Searchable symmetric encryption (SSE) enables a remote cloud server to answer queries directly on encrypted data on a client's behave, therefore, relieves the resource limited client from complicated data management. Two key requirements are a strong security guarantee and a sub-linear search performance. The bucket approach in the literature addresses these requirements at the expense of downloading many false positives or requiring the client to search relevant bucketization ids locally, which limits the applicability of the method. In this paper, we propose a novel approach CLASS to meet these requirements for equality conjunction search while minimizing the client work and communication cost. First, we generalize the standard ciphertext indistinguishability to partitioned data, called class indistinguishability, which provides a level of ciphertext indistinguishability similar to that of bucketization but allows the cloud server to perform search of relevant data and filtering of false positives. We present a construction achieving these goals through a two-phase search algorithm for a query. The first phase finds a candidate set through a sub-linear search. The second phase finds the exact query result using a linear search applied to the candidate set. Both phases are performed by the server and are implemented by plugging in existing search methods. The experiment results on large real-word data sets show that our approach outperforms the state-of-the-art.

Session VII: Cloud Resources Optimization (06/28 Friday, 15:20-16:40; Navigator)

Chair: Simone Raponi (Hamad Bin Khalifa University, Italy)
20 minutes for each paper.

15:20 A Hybrid Approach for Synchronizing Clocks in Distributed Systems (CLOUD2019-1020)

Md Shohel Khan and Ratul Sikder (Bangladesh University of Engineering and Technology, Bangladesh); Muhammad Abdullah Adnan (University of California San Diego, USA)

Abstract: The art of synchronizing clocks across a wide area network has got a new dimension when it comes to the reality of high accuracy synchronization demand, even for the local or small computing systems. Before implementing any clock synchronization system, many aspects must be considered. i.e. communication latency - is it fixed or variable? Is there any reference clock in the system or not? In this paper, we have studied the standard and experimental protocols for synchronizing clocks over a geographically distributed network and implemented the features of the Network Time Protocol (NTP) combined with Global Positioning System (GPS) for synchronizing clients' clocks more accurately. Our proposed system can achieve more accuracy compared to the NTP clock synchronization with the help of decentralized GPS based NTP servers.

15:40 JCallGraph: Tracing Microservices in Very Large Scale Container Cloud Platforms (CLOUD2019-1020)

Haifeng Liu and Jinjun Zhang (JD.com, China); Huasong Shan, Min Li and Yuan Chen (JD Silicon Valley R&D Center, USA); Xiaofeng He and Xiaowei Li (JD.com, China)

Abstract: Microservice architecture splits giant and complex enterprise applications into fine-grained microservices, promoting agile development, integration, delivery and deployment. However, monitoring tens of thousands of microservices is extremely challenging, and debugging problems among massive microservices is like looking for a needle in a haystack.We present JCallGraph, a tracing and analytics tool to capture and visualize the microservice invocation relationship of tens of thousands of microservices with millions of containers at JD.com. JCallGraph achieves three main

goals for distributed tracing and debugging: online microservices invocation construction within milliseconds, minimal overhead without any significant performance impact on real-production applications, and application-agnostic with zero-intrusion to application. Our evaluation shows that JCallGraph can accurately capture the realtime invocation relationship at massive scale and help developers to efficiently understand interactions among microservices, pinpoint root-cause of problems.

16:00 An Overview of Cloud Computing Testing Research (CLOUD2019-1021)

Jia Yao, Babak Maleki Shoja, and Nasseh Tabrizi (East Carolina University, USA)

Abstract: With the rapid growth in information technology, there is a significant increase in research activities in the field of cloud computing. Cloud testing can be interpreted as i) testing of cloud applications, which involves continuous monitoring of cloud application status to verify Service Level Agreements, and ii) testing as a cloud service which involves using the cloud as a testing middleware to execute a large-scale simulation of real-time user interactions. This study aims to examine the methodologies and tools used in cloud testing and the current research trends in cloud computing testing.

Session VIII: Cloud Platform Applications (06/29 Saturday, 10:20-11:20; Destination Ballroom 1)

Chair: Hailu Xu (Florida International University)
20 minutes for each paper.

10:20 A Robust Multi-Terminal Support Method Based on Tele-immersion Multimedia technology (CLOUD2019-1022)

Ronghe Wang, Bo Zhang, Haiyong Xie and Dong Jiao (Academy of Electronics and Information Technology of China Electronics Technology Group Corporation, China); Shilong Ma (Beihang University, China)

Abstract: In this paper, a multi terminal support method based on tele-immersion multimedia technology is proposed. This method provides two functions of synchronous request and asynchronous request, and it can dynamically download and decompress the material by the way of material sub-contracting and on-demand loading, which effectively avoids that the main thread is blocked because of too long request time and that browser crashes because of excessive memory footprint. The system excellently solves the problems of memory, storage and network communication in the development of cross WebGL platform, and optimizes the storage mode of memory, which realizes the high speed calculation and real-time processing of data. The system also integrates the system operation framework, and provides unified framework support for the development of the sub-systems of different platforms. Based on the framework, the secondary development can quickly complete the process development of the sub-system.

10:40 CMonitor: A Monitoring and Alarming Platform for Container-Based Clouds (CLOUD2019-1023)

Shujian Ji (University of Chinese Academy of Sciences & SIAT, China); Kejiang Ye (Chinese Academy of Sciences, China); Chengzhong Xu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China)

Abstract: Container technology has been recently recognized by the industry like Google and Alibaba as an emerging platform for building and deploying applications. Due to the flexibility and low cost, more and more applications are beginning to use container as the underlying resource abstraction platform. In order to maintain the stability of the system and detect any suspected abnormal events or operations, it is necessary to provide a monitoring and alarming mechanism for containers. In this paper, we design and implement a monitoring and alarming platform \textit{CMonitor} for the container-based clouds. It is built upon the interfaces provided by the current Docker containers. Specially, we added some new functions: i) Integrated monitoring service. CMonitor not only monitors the basic resource usages of each container, but also provides hardware-level and application-level monitoring services. ii) Global topology view. CMonitor generates a global topology structure for containers by parsing network traffic among containers. iii) Intelligent alarming mechanism. CMonitor contains several anomaly detection algorithms to identify abnormal behaviors in containers and then notifies an alarm to the users. iv) Rich visualization functions. CMonitor records the runtime log for both system resources and application performance, and generates tables and figures with advanced data visualization techniques. By using CMonitor, users can understand the system runtime status and monitor any abnormal or fault events, making the container-based cloud more stable, efficient and safe.

11:00 Exploiting the Spam Correlations in Scalable Online Social Spam Detection (CLOUD2019-1024)

Hailu Xu, Pinchao Liu, Boyuan Guan and Liting Hu (Florida International University, USA)

Abstract: The huge amount of social spam from large-scale social networks has been a common phenomenon in the contemporary world. The majority of former research focused on improving the efficiency of identifying social spam from a limited size of data in the algorithm side, however, few of them target on the data correlations among large-scale distributed social spam and utilize the benefits from the system side. In this paper, we propose a new scalable system, named SpamHunter, which can utilize the spam correlations from distributed data sources to enhance the performance of large-scale social spam detection. It identifies the correlated social spam from various distributed servers/sources through DHT-based hierarchical functional trees. These functional trees act as bridges among data servers/sources to aggregate, exchange, and communicate the updated and newly emerging social spam with each other. Furthermore, by processing the online

social logs instantly, it allows online streaming data to be processed in a distributed manner, which reduces the online detection latency and avoids the inefficiency of outdated spam posts. Our experimental results with real-world social logs demonstrate that SpamHunter reaches 95% F1 score in the spam detection, achieves high efficiency in scaling to a large number of data servers with low latency.

2019 International Congress on Big Data (BigData 2019)

Session I: Big Data Method (06/28 Friday, 10:40-11:40; Compass Ballroom 2)

Chair: Qingyang Wang (Louisiana State University--Baton Rouge, USA)
20 minutes for each paper.

10:40 Designing and Implementing Data Warehouse for Agricultural Big Data (BIGDATA2019-2001)

Vuong M. Ngo (University College Dublin, Ireland); Nhien-An Le-Khac and M-Tahar Kechadi (University College Dublin, Ireland)

Abstract: In recent years, precision agriculture that uses modern information and communication technologies is becoming very popular. Raw and semi-processed agricultural data are usually collected through various sources, such as: Internet of Thing (IoT), sensors, satellites, weather stations, robots, farm equipment, farmers and agribusinesses, etc. Besides, agricultural datasets are very large, complex, unstructured, heterogeneous, non-standardized, and inconsistent. Hence, the agricultural data mining is considered as Big Data application in terms of volume, variety, velocity and veracity. It is a key foundation to establishing a crop intelligence platform, which will enable resource efficient agronomy decision making and recommendations. In this paper, we designed and implemented a continental level agricultural data warehouse by combining Hive, MongoDB and Cassandra. Our data warehouse capabilities: (1) flexible schema; (2) data integration from real agricultural multi datasets; (3) data science and business intelligent support; (4) high performance; (5) high storage; (6) security; (7) governance and monitoring; (8) replication and recovery; (9) consistency, availability and partition tolerant; (10) distributed and cloud deployment. We also evaluate the performance of our data warehouse.

11:00 FFD: A Federated Learning based method for Credit Card Fraud Detection (BIGDATA2019-2002)

Wensi Yang and Yuhang Zhang (Shengzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China); Kejiang Ye (Shengzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China); Li Li (Shengzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China); Chengzhong Xu (University of Macau, Macau China)

Abstract: Credit card fraud has caused huge loss to both banks and consumers recent years. Thus, an effective Fraud Detection System (FDS) is important to minimize the loss for banks and cardholders. Based on our analysis, the credit card transaction dataset is very skewed, there are much fewer samples of frauds than legitimate transactions. Furthermore, due to the data security and privacy, different banks are usually not allow to share their transaction datasets. These problems make FDS difficult to learn the patterns of frauds and also difficult to detect them. In this paper, we propose a framework to train a fraud detection model using behavior features with federated learning, we term this detection framework FFD (Federated learning for Fraud Detection). Different from the traditional FDS trained with data centralized in the cloud, FFD enable banks to learn fraud detection model with the training data distributed on their own local database. Then, a shared FDS is constructed by aggregating locally-computed updates of fraud detection model. Banks can collectively reap the benefits of shared model without sharing dataset and protect the sensitive information of cardholders. Furthermore, an oversampling approach is combined to balance the skewed dataset. We evaluate the performance of our credit card FDS with FFD framework on a large scale dataset of real-world credit card transactions. Experimental results show that the federated learning based FDS achieves an average of test AUC to 95.5\%, which is about 10\% higher than traditional FDS.

11:20 A Relation Extraction Method Based on Entity Type Embedding and Recurrent Piecewise Residual Networks (BIGDATA2019-2003)

Yuming Wang(Huazhong University of Science and Technology, China; Georgia Institute of Technology, USA), Huiqiang Zhao and Lai Tu (Huazhong University of Science and Technology, China); Jingpei Dan (Chongqing University, China); Ling Liu (Georgia Institute of Technology, USA)

Abstract: Relation extraction is an important while challenging task in information extraction. We found that current solutions can hardly extract correct relation when the sentence is long and complex or the firsthand trigger word does not show. Inspired by the idea of fusing more and deeper information, we present a new relation extraction method that involves the types of entities in the joint embedding. An architecture of Recurrent Piecewise Residual Networks (RPRN) is proposed to cooperate with the joint embedding so that the relation extractor acquires the latent representation underlying the context of a sentence. We validate our method by experiments on public data set of New York Times. Experiment results show that our method outperforms the existing state-of-the-art models.

Session II: Big Data Algorithms (06/27 Thursday, 13:00-14:00; Compass Ballroom 2)

Chair: Seung-Jong Park (Louisiana State University, USA)
20 minutes for each paper.

13:00 Research on Parallel Computing Method of Hydrological Data Cube (BIGDATA2019-2004)

Rong Wang and Dingsheng Wan (Hohai University, China); Zixuan Zhang (Bank of Jiangsu, China); Qun Zhao (Hohai University, China)

Abstract: Data cube technology is a main method for analyzing high-dimensional data, which is of great significance in improving the efficiency of MOLAP. Basic hydrological data are important historical data, including precipitation, water level, discharge, sediment load, evaporation and other hydrological elements. As a kind of high-dimensional data, hydrological data has the characteristics of large amount of information, multiple dimensions, hierarchical and redundant. In order to further compress the time and space consumption in the process of data cube analysis, this paper proposes a segmentation parallel generation method of the dwarf cube based on the Map/Reduce computing model. Experiments show that generating a dwarf cube based on the segmentation parallel method not only reduces the storage space, but also improves the query efficiency.

13:20 A Machine Learning Approach to Prostate Cancer Risk Classification Through Use of RNA Sequencing Data (BIGDATA2019-2005)

Matthew M Casey and Baldwin Chen (Ardsley High School, USA); Jonathan Zhou (Horace Greeley High School, USA); Nianjun Zhou (IBM, USA)

Abstract: Advancements in RNA sequencing technology have made genomic data acquired during sequencing more precise, making models fitted to sequencing data more practical. Previous studies conducted regarding prostate cancer diagnosis have been limited to microarray data, with limited successes. We utilized The Cancer Genome Atlas' (TCGA) prostate cancer sequencing data to test the viability of fitting machine learning models to RNA sequencing data. A major challenge associated with the sequencing data is its high dimensionality. In this research, we addressed two complementary tasks. The first was to identify genes most associated with potential cancer. We started by using the mutual information metric to identify the most significant genes. Furthermore, we applied the Recursive Feature Elimination (RFE) algorithm to reduce the number of genes needed to identify cancer. The second task was to create a classification model to separate potential high-risk patients from the healthy ones. For the second task, we combated the high dimensionality challenge with Principal Component Analysis (PCA). In addition to high dimensionality, another challenge is the imbalanced data set that has a 10:1 class imbalance of cancerous and healthy tissue respectively. To combat this problem, we used the Synthetic Minority Oversampling Technique (SMOTE) to create synthetic observations and equalize the class distribution. We trained and tested our model using 5-fold cross-validation. The results were promising, significantly reducing the false negative rate as compared to current diagnostic techniques. The model showed great improvements over previous machine learning attempts to diagnose prostate cancer. Our model could be applied as part of the patient diagnosis pipeline, helping to improve accuracy.

13:40 User Level Multi-Feed Weighted Topic Embeddings for Studying Network Interaction in Twitter (BIGDATA2019-2006)

Pujan Paudel, Amartya Hatua, Trung T. Nguyen and Andrew H. Sung (The University of Southern Mississippi, USA)

Abstract: Over half a billion tweets on wide range of topics are posted daily by hundreds of millions of Twitter users. A good understanding of Twitter user behaviour and their network interactions can be exploited in practical applications like targeted advertising, viral marketing, and political campaigns, etc. In this paper, we propose a Multi-Feed Weighted Topic Embeddings (MFWTE) model to study user network interaction and topic diffusion patterns on Twitter. Our method extracts topic embeddings from multiple views of a Twitter user feed and weighs them according to their content authoring roles, where the authored tweets, replied tweets, retweeted tweets, and favourited tweets are the views we separate for constructing the embeddings. We then test the proposed method using two different topic modeling algorithms: i) Latent Dirichlet Allocation ii) Twitter-Latent Dirichlet Allocation. The users in our study are divided into multiple hierarchies based on their activity composition regarding individual topics, and the effectiveness of MFWTE is evaluated in the multi-hierarchical setting. The performance of our method on friendship recommendation and retweet behavior prediction task is evaluated using various ranked retrieval measures. We also investigate topic authoring behaviour of different hierarchies of Twitter users using MFWTE and discover information diffusion patterns through topic models. The results indicate that our MFWTE method for topic modeling of Twitter users outperforms various single-feed methods, and topic diffusion patterns can be studied using the method.

Session III: Big Data Application (06/26 Wednesday, 14:00-15:00; Compass Ballroom 2)

Chair: Kisung Lee (Louisiana State University, USA)
20 minutes for each paper.

14:00 Big Data Analytics for Nabbing Fraudulent Transactions in Taxation System (BIGDATA2019-2007)

Priya Mehta, Jithin Mathews (Indian Institute of Technology Hyderabad, India); Sandeep Kumar (Plianto Technologies, India); Suryamukhi K , Ch.Sobhan Babu (Indian Institute of Technology Hyderabad, India); S.V. Kasi Visweswara Rao (Government of Telangana, India)

Abstract: This paper explains an application of big data analytics to detect illegitimate transactions performed by fraudulent communities of people who are engaged in a notorious tax evasion practice called circular trading. We designed and implemented this technique for the commercial taxes department, government of Telangana, India. This problem is solved in two steps. In step one, the problem is formulated as detecting fraudulent communities in a social network, where the vertices correspond to dealers and edges correspond to sales transactions. In step two, specific type of cycles are removed from each fraudulent community, which were identified in step one, to detect the illegitimate transactions. We used RHadoop framework for implementing the same.

14:20 Multi-Step Short Term Traffic Flow Forecasting using Temporal and Spatial Data (BIGDATA2019-2008)

Hao Peng and John A. Miller (University of Georgia, USA)

Abstract: Short term traffic flow forecasting is valuable to both governments for designing intelligent transportation systems and everyday commuters or travelers who are interested in the best routes to their destinations. This work focuses on forecasting traffic flow in major freeways in southern California using large amounts of data collected from the Caltrans Performance Measurement System. Both statistical models and machine learning models are considered. The statistical models include seasonal ARIMA, Exponential Smoothing, and seasonal VARMA. The machine learning models include Support Vector Regression, feed forward Neural Networks, and Long Short-Term Memory Neural Networks. Forecasting is performed in both a univariate manner by relying on the historical temporal data of a particular sensor as well as in a multivariate manner by considering a neighborhood of three closely located sensors. Multivariate forecasters generally improve upon their univariate counterparts and Long Short-Term Memory Neural Networks lead in overall performance.

14:40 Towards Detection of Abnormal Vehicle Behavior Using Traffic Cameras (BIGDATA2019-2009)

Chen Wang, Aibek Musaev, Pezhman Sheinidashtegol, and Travis Atkison (The University of Alabama, USA)

Abstract: Throughout the world, many surveillance cameras are being installed every month. For example, there are over 18,000 publicly accessible traffic cameras in 200 cities and metropolitan areas in the United States alone. Live video streams provide real-time big data about behavior happening in the present, such as traffic information. However, until now, extracting intelligence from video content has been mostly manual, i.e. through human observation. The development of smart real-time tools that can detect abnormal vehicle behaviors may alert law enforcement and transportation agencies of possible violators and can potentially avoid traffic accidents. In this study, we address this problem by developing an application for detection of abnormal driving behavior using traffic video streams. Evaluation is performed using real videos from traffic cameras to detect stalled vehicles and possible abnormal vehicle behavior.

2019 International Conference on Services Computing (SCC 2019)

Session I: Services Linkage (06/26 Wednesday, 11:00-12:00; Destination Ballroom 2)

Chair: Aibek Musaev (The University of Alabama, USA)
20 minutes for each paper.

11:00 A Quality-Aware Web API Recommender System for Mashup Development (SCC2019-3001)

Kenneth K Fletcher (University of Massachusetts Boston, USA)

Abstract: The rapid increase in the number and diversity of web APIs with similar functionality, makes it challenging to find suitable ones for mashup development. In order to reduce the number of similarly functional web APIs, recommender systems are used. However, various existing web API recommendation methods attempt to improve recommendation accuracy by mainly using some discovered relationships between web APIs and mashups. Such methods are basically incapable of recommending quality web APIs because they fail to incorporate web API quality in their recommender systems. In this work, we propose a method that considers the quality features of web APIs, to make quality web API recommendations. Our proposed method uses web API quality to estimate their relevance for recommendation. Specifically, we propose a matrix factorization method, with quality feature regularization, to make quality web API recommendations and also enhance recommendation diversity. We demonstrate the effectiveness of our method by conducting extensive experiments on a real-world dataset from www.programmableweb.com. Our results not only show quality web API recommendations, but also, improved recommendation accuracy. In addition, our proposed method improves recommendation diversity by mitigating the negative Matthew effect of accumulated advantage issue intrinsic to most existing web API recommender systems. We also compare our method with some baseline recommendation methods for validation.

11:20 Practical Verification of Data Encryption for Cloud Storage Services (SCC2019-3002)

Jinxia Fang (University of Chinese Academy of Sciences, China); Limin Liu and Jingqiang Lin (Institute of Information Engineering, Chinese Academy of Sciences, China)

Abstract: Sensitive data is usually encrypted to protect against unauthorized access and leakage for cloud storage services. Generally, remote user has no knowledge of the actual data format stored in the cloud, even though a cloud server promises to store the data with encryption. Although a few works utilize data encapsulation and remote data checking to detect whether the sensitive data is protected securely in the cloud, they still suffer from a number of limitations, such as heavy computational cost at the user side and poor practicality, that would hinder their adoptions. In this paper, we propose a practical verification scheme to allow users to remotely evaluate the actually deployed data encryption protection in the cloud. We employ pseudo-random number generator and present a data encapsulation solution, which can benefit users with significant cost savings. By imposing a monetary rewards or penalties, our proposed scheme can help ensure that economically rational cloud servers store data encrypted at rest honestly. Extensive experiments are conducted to further demonstrate the efficiency and practicality of the proposed scheme.

11:40 Generating Personalized and Certifiable Workflow Designs: a Prototype (SCC2019-3003)

Manon Froger (Ecole Nationale Supérieure des Mines d'Albi & Interopsys, France); Fréderick Bénaben and Sébastien Truptil (Ecole Nationale Supérieure des Mines Albi, France); Nicolas Boissel-Dallier (Interopsys, France)

Abstract: As the first level of a BPM strategy, being able to design event-oriented models of processes is a must-have competence for every modern business. Unfortunately, industrial procedures have reached a certain complexity making the designing task complex enough to discourage businesses facing the blank page. Moreover, the 21st century witnesses the emergence of myriads of norms and external regulations that businesses want to abide by. Although domain experts have a limited process modelling and norm interpretation knowledge, they know how to describe their activities and their sequencing. With progresses made in the artificial intelligence, particularly in the natural language processing domain, it becomes possible to automatize the task of creating a process in compliance with norms. This paper presents a business-oriented prototype assisting users in getting certifiable specific business processes. We detail the metamodel used to separately model norms and business' existing procedures and then, the algorithm envisaged to deduce a corresponding cartography of processes.

Session II: Analytics and Recommendation (06/27 Thursday, 09:00-10:00; Destination Ballroom 2)

Chair: Yishuang Ning (Tsinghua University, China)

20 minutes for each paper.

9:00 An Empirical Investigation to Real-world QoS of Web Services (SCC2019-3004)

Yang Syu and Chien-Min Wang (Academia Sinica, Taiwan)

Abstract: Quality of Service (QoS) is a critical non-functional property and criterion of Web Services (WSs), prompting the development of many QoS- aware or QoS-based approaches. However, the current literature lacks a systematic investigation and analysis of real-world QoS data. This paper investigates the properties of WS QoS and presents results of interest to WS/QoS researchers. Experiments are performed on two real-world, large-scale QoS datasets are used as the basis for experiments. The main contribution of this paper is to empirically confirm some useful and interesting properties of real-world WS QoS. For example, it is found that the distance between a service consumer and its invoked WS does not influence the invocation failure rate, but that shorter distances lead to shorter response times and higher throughputs (i.e., better performance) of WSs.

9:20 Towards the Readiness of Learning Analytics Data for Micro Learning (SCC2019-3005)

Jiayin Lin, Geng Sun, Jun Shen, Tingru Cui and Ping Yu (University of Wollongong, Australia); Dongming Xu (, unknown); Li Li (Southwest University, China); Ghassan Beydoun (University of Technology Sydney, Australia)

Abstract: With the development of data mining and machine learning techniques, data-driven based technology-enhanced learning (TEL) has drawn wider attention. Researchers aim to use established or novel computational methods to solve educational problems in the 'big data' era. However, the readiness of data appears to be the bottleneck of the TEL development and very little research focuses on investigating the data scarcity and inappropriateness in the TEL research. This paper is investigating an emerging research topic in the TEL domain, namely micro learning. Micro learning consists of various technical themes that have been widely studied in the TEL research field. In this paper, we firstly propose a micro learning system, which includes recommendation, segmentation, annotation, and several learning-related prediction and analysis modules. For each module of the system, this paper reviews representative literature and discusses the data sources used in these studies to pinpoint their current problems and shortcomings, which might be debacles for more effective research outcomes. Accordingly, the data requirements and challenges for learning analytics in micro learning are also investigated. From a research contribution perspective, this paper serves as a basis to depict and understand the current status of the readiness of data sources for the research of micro learning.

9:40 Personalized Service Recommendation based on User Dynamic Preferences (SCC2019-3006)

Benjamin A Kwapong, Richard Anarfi and Kenneth K Fletcher (University of Massachusetts Boston, USA)

Abstract: In order to personalize users' recommendations, it is essential to consider their personalized preferences on non-functional attributes during service recommendation. However, to increase recommendation accuracy, it is essential that recommendation systems include users' evolving preferences. It is not sufficient to only consider users' preferences at a point in time. Existing time-based recommendation systems either disregard rich and useful historical user invocation information, or rely on information from similar users, and thus, fail to thoroughly capture users' dynamic preferences for personalized recommendation. This work proposes a method to personalize users' recommendations based on their dynamic preferences on non-functional attributes. First, we compose a user's preference profile as a time series of his/her invocation preference and pre-invocation dependencies (i.e. the different services that were viewed prior to invoking the preferred service). We model a user's invocation preference as a combination of non-functional attribute values observed during service invocation, and topic distribution from WSDL of the invoked service using hierarchical dirichlet process (HDP). Next, we employ long short-term memory recurrent neural networks (LSTM-RNN) to predict the user's future invocation preference. Finally, based on the predicted future invocation preference, we recommend service(s) to that user. To evaluate our proposed method, we perform experiments using real-world service dataset, WS-Dream.

Session III: Business Solutions (06/28 Friday, 10:40-11:40; Destination Ballroom 2)

Chair: Aibek Musaev (The University of Alabama, USA)
20 minutes for each paper.

10:40 Better Service Performance Management via Workload Prediction (SCC2019-3007)

Hachem Moussa, I-ling Yen and Farokh Bastani (University of Texas at Dallas, USA); Yulin Dong and Wei He (Shandong University, China)

Abstract: In this paper, we consider managing service performance starting from the composition time, aiming to reducing the risk of execution failures during service composition. We use ARIMA to predict workloads of the services at the time when they are likely to be invoked and subsequently predict the response time and chances that the requests for accessing the services may be declined due to admission control. The in-depth analysis can help avoid timing failures during service execution. However, these analyses may incur overhead and we introduced a two-phase composition algorithm to reduce the potential overhead. Our system also consider continuous monitoring and service recomposition to greatly increase the

probability of completing the service execution within the deadline. Experimental results show that our service management approach can greatly improve the success rate for meeting the deadline as well as reduce the overall composition overhead.

11:00 Chatbot Assisted Marketing in Financial Service Industry (SCC2019-3008)

Jon Quah and Yanwei Chua (Singapore Institute of Management, Singapore)

Abstract: The rise of chatbots in the finance sector is the latest disruptive force that has change the way customers interact. The adoption of Artificial Intelligence powered chatbots particularly in the banking industry has changed the face of communication interface between bank and customers. This paper explores the effectiveness of the current use of chatbot in Singapore's banking industry. The banking sector in Singapore plays a significant role in Singapore economy. It also investigates the current chatbot functionality to determine if it can meet the ever-changing expectation of customers.

11:20 Application of Deep Learning in Surface Defect Inspection of Ring Magnets (SCC2019-3009)

Xu Wang and Pan Cheng (Sankyo Precision Co., Ltd, China)

Abstract: We present a method of inspecting surface defects of ring magnets by using deep learning technology, and the inspection system developed utilizing this method has achieved much better accuracy and speed than human inspectors in actual production environment, while such accuracy and speed are essential for such systems. The proposed method can also be used for the surface defect inspection of many other industrial products and systems.

2019 International Conference on AI & Mobile Services (AIMS 2019)

Session I: AI Modeling (06/27 Thursday, 13:00-14:00; Destination Ballroom 1)

Chair: De Wang (Sunmi US Inc., USA)
20 minutes for each paper.

13:00 Population-Based Variable Neighborhood Descent for Discrete Optimization (AIMS2019-4001)

Petar Afric, Adrian S Kurdija and Lucija Šikić (University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia); Marin Šilić (University of Zagreb, Croatia); Goran Delac, Klemo Vladimir and Sinisa Srbljic (University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia)

Abstract: Many problems in smart solution development make use of discrete optimization techniques. It is expected that smart cities will have a constant need for parcel delivery and vehicle routing which is heavily reliant on discrete optimization. In this paper we present an improvement to the Variable neighborhood descent (VND) algorithm for discrete optimization. Our method makes the search procedure more exhaustive at the expense of time performance. Instead of keeping track of a single solution which is being improved, we allow branching of the solution into at most M promising solutions and keep track of them. Our experiments show that the proposed method produces results superior to VND. We analyze the impact on time complexity and give general usage guidelines for our method.

13:20 The Constrained GAN with Hybrid Encoding in Predicting Financial Behavior (AIMS2019-4002)

Yuhang Zhang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences & University of Chinese Academy of Sciences, China); Wensi Yang (Shenzhen Institutes of Advanced Technology，Chinese Academy of Sciences & University of Chinese Academy of Sciences, China); Wanlin Sun (Shenzhen Institutes of Advanced Technology，Chinese Academy of Sciences, China); Kejiang Ye (Chinese Academy of Sciences, China); Ming Chen (Shenzhen Institutes of Advanced Technology，Chinese Academy of Sciences, China); Chengzhong Xu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China)

Abstract: Financial data are often used in predicting users' behaviors in business fields. The previous work usually focuses on the positive samples which means those specific persons can bring the profit to companies. However, in most cases, the proportion of positive samples is very small. The traditional algorithms perform not well when the positive and negative samples are extremely unbalanced. To solve this problem, we propose an integrated network to build the classification network. Considering the original dataset including both objective and index data, we integrate the one-hot encoding and float encoding together to uniform the data, that we call it hybrid encoding. Then, we use GAN framework to enrich the positive samples. Considering the attributes of financial data are not the same, we implement Mahalanobis Distance (MD) as a constraint in GAN framework to generate high quality positive samples. Finally, we use voting rules to put both data sensitive and data insensitive classifiers together to make a strong classifier. We evaluate the performance of our framework on a real world dataset. The experimental results show our method are effective.

13:40 ECGAN: Image Translation with Multi-Scale Relativistic average Discriminator (AIMS2019-4003)

Weihao Xia (Tsinghua University, China); Yujiu Yang (Tsinghua University, Shenzhen, China); Xianyu Bao (Hefei University of Technology, China)

Abstract: Image-to-image translation aims to transform an image of one source domain to another target domain, which can be applied to many specific tasks, such as image restoration and style transfer. As one of the most popular frame- works, CycleGAN with cycle consistency loss can transform between different domains with unpaired training data, which greatly expands its application effi- ciency. However, due to its under-constrained mapping and dissatisfactory net- work design, the results of CycleGAN are sometimes unnatural and unreal. In some cases, there are visible artifacts in the synthetic images. In this paper, we propose an Enhanced CycleGAN (ECGAN) framework with multi-scale rela- tivistic average discriminator, which integrates the loss function design and network structure optimization to make the generated images have more natural textures and fewer unwanted artifacts. In the evaluation, besides using quantitative full reference image quality assessment metrics (such as PSNR, SSIM), we also conduct an evaluation on the Frechet Inception distance (FID), which are more consistent with human subjective assessment according to natural, realistic and diverse. Experiments on two benchmark datasets, including CMP facades and CUHK face datasets, show that the proposed ECGAN framework outperforms the state-of-the-art methods in both quantitative and qualitative evaluation.

Session II: AI as a Service (06/26 Wednesday, 14:00-15:00; Destination Ballroom 1)

Chair: De Wang (Sunmi US Inc., USA)
20 minutes for each paper.

14:00 Pork Registration Using Skin Image with Deep Neural Network Features (AIMS2019-4004)

Daohang Song, Cheng Cai and Zeng Peng (Northwest A&F University, China)

Abstract: Pork food safety is not optimistic in China. Some pork carrying viruses can damage the liver and kidneys of consumers. At present, the pork traceability system in China mainly relies on the information of the stamp on the pork skin. However, there is no unified stamp standard in China to regulate pork market, and information on the stamp is likely to be destroyed during the circulation process. In this study, deep convolutional neural network (DCNN) was used to extract the features of pork skin, and dynamic inlier selection was used to register pork skin images to achieve accurate traceability of pork. The data set consists of 810 images, which containing images captured from three angles and nine position. The results show that features extracted using DCNN are better than SIFT features, and the average matching rate is 92.59%. Compared with KNN (K-Nearest Neighbor), CPD (Coherent Point Drift), ICP (Iterative Closest Point), our dynamic inlier selection has better registration effect. In our pork skin dataset, the success rate of registration reached 86.67%, which provided a reference for subsequent pork traceability research.

14:20 ORB-based Multiple Fixed Resolution Approach for On-Board Visual Recognition (AIMS2019-4005)

Daniel Phillips and Akash Pooransingh (The University of the West Indies, Trinidad and Tobago); Sinem Guven (IBM T. J. Watson Research Center, USA)

Abstract: Maintenance and troubleshooting of hardware at large scale pose a challenge in deploying expert technicians at multiple sites. Augmented Reality-based technology support equips the technicians with the skills they need to solve hardware problems even without expert level training, thereby reducing training time and cost to the vendor. Enabling Augmented Reality for technology support requires the ability to visually recognize the hardware in real time using mobile devices, and train the underlying algorithms at scale. This paper proposes a novel approach to address these issues. Our ORB-based fixed multi-resolution recognition algorithm achieves over 95% accuracy at a resolution scale of 0.2, and an approximately 60% faster recognition time than the next best comparable method. We also demonstrate the real-world applicability of our algorithm through an implementation of an Augmented Reality application.

14:40 GRASP Method for Vehicle Routing with Delivery Place Selection (AIMS2019-4006)

Petar Afric, Adrian S Kurdija and Lucija Šikić (University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia); Marin Šilić (University of Zagreb, Croatia); Goran Delac, Klemo Vladimir and Sinisa Srbljic (University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia)

Abstract: In this paper we present a greedy randomized adaptive search procedure (GRASP) for solving a vehicle routing problem (VRP) for package delivery with delivery place selection. The problem can be solved by stepwise optimization, i.e., first selecting delivery sites and then defining routes based on that selection. Alternatively, it can be solved by jointly optimizing delivery site selection and routing. We investigate the effects of stepwise optimization in comparison to joint optimization. The evaluation results show that our proposed stepwise approach, while expectedly producing longer routes than joint approach (by 4% on aver- age), can provide a solution 1000× faster than the previous benchmark approach. The proposed procedure is therefore well suited for the dynamic environment of package delivery which is widespread in modern cities as a consequence of e-commerce.

Session III: AI Algorithms and Mobile Services (06/27 Thursday, 14:20-15:20; Destination Ballroom 1)

Chair: De Wang (Sunmi US Inc., USA)
20 minutes for each paper.

14:20 Indexed Search (AIMS2019-4007)

Erdal Kose (Fairleigh Dickinson University, USA)

Abstract: The domain of Artificial Intelligence (AI) Search algorithms is a well-studied area. A variety of AI search algorithms have been developed in the past decades. Some of them are considered to be among the main AI search algorithms such as Best First Search, A*, Iterative Deepening Search, etc. The others use these main algorithms as a basis and include more heuristics to generate newer algorithms such as Iterative Deepening A* (IDA*), Branch and Bound Search, etc. the innovation of this research is a new approach to the AI search algorithms, we called it indexed search. The methods assigns integer indices to the states of the problem and instead of searching for a goal state, it searches for the index of the goal state. Once the index of a goal state found, it converts the index from decimal base to the base of the branching factor of the problem, and this new number is the solution path from goal to the start node. The new approach eliminates the Closed List, which employed for most AI search algorithms to store the explored nodes. The new approach also generates a solution path faster than the respective versions of it and it also use less memory space.

14:40 Cognitively-inspired Agent-based Service Composition for Mobile & Pervasive Computing (AIMS2019-4008)

Oscar J Romero (Carnegie Mellon University, USA)

Abstract: Automatic service composition in mobile and pervasive computing faces many challenges due to the complex and highly dynamic nature of the environment. Common approaches consider service composition as a decision problem whose solution is usually addressed from optimization perspectives which are not feasible in practice due to the intractability of the problem, limited computational resources of smart devices, service host's mobility, and time constraints to tailor composition plans. Thus, our main contribution is the development of a cognitively-inspired agent-based service composition model focused on bounded rationality rather than optimality, which allows the system to compensate for limited resources by selectively filtering out continuous streams of data. Our approach exhibits features such as distributedness, modularity, emergent global functionality, and robustness, which endow it with capabilities to perform decentralized service composition by orchestrating manifold service providers and conflicting goals from multiple users. The evaluation of our approach shows promising results when compared against state-of-the-art service composition models.

15:00 Accelerating Deep Learning inference on mobile systems (AIMS2019-4009)

Darian Frajberg, Carlo Bernaschina, Christian Marone and Piero Fraternali (Politecnico di Milano, Italy)

Abstract: Artificial Intelligence on the edge is a matter of great importance towards the enhancement of smart devices that rely on operations with real-time constraints. We present PolimiDL, a framework for the acceleration of Deep Learning on mobile and embedded systems with limited resources and heterogeneous architectures. Experimental results show competitive results with respect to TensorFlow Lite for the execution of small models.

Session IV: AI & Mobile Applications (06/28 Friday, 15:20-16:40; Destination Ballroom 1)

Chair: De Wang (Sunmi US Inc., USA)

20 minutes for each paper.

15:20 Design of Mobile Service of Intelligent Large-Scale Cyber Argumentation for Analysis and Prediction of Collective Opinions (AIMS2019-4010)

Najla Althuniyan, Joseph Sirrianni, Md Mahfuzer Rahman and Xiaoqing "Frank" Liu (University of Arkansas, USA)

Abstract: Issues of national and international importance attract the attention of millions of people who want to share their opinions online. These discussions among a large number of people contain rich information, from which we want to extract the crowd wisdom and collective intelligence. Most of these discussions happen in social media platforms such as Facebook and Twitter. While these platforms are accessible, they not structured in a way to effectively facilitate these large-scale discussions, leading many discussions to be fragmented, difficult to follow, and nearly impossible to analyze collective opinions. Alternatively, some people use debate-centric platforms for online discussions. While these platforms offer structured discussions, they usually represent issues with binary solutions and modest analytics. In the cyber-argumentation, a sub-field of AI, argumentations platforms have been developed to facilitate online discussion effectively. These platforms provide structured argumentation frameworks, which allows for meaningful analytics models to mine the argumentation. How-ever, few platforms have mobile service applications and those that do provide only basic statistical analytics. In this paper, we present the design of a mobile application service to support intelligent cyber argumentation. This service is de-signed to facilitate large-scale discussion and report complex analytics on a handheld screen size. The platform has several integrated analytical models, which use AI techniques, to capture collective opinions, detect opinion polarization, and predict missing user opinions. An example of a large-scale discussion is used to demonstrate the effectiveness of bringing the intelligent cyber-argumentation into the mobile space.

15:40 ResumeVis: Interactive Correlation Explorer of Resumes (AIMS2019-4011)

Xiaohui Wang and Jingyan Qin (University of Science and Technology Beijing, China)

Abstract: Resumes are critical for individuals to find jobs and for HR to select staffs. Almost all key events about career and demographic information are recorded on the resumes, which are commercial secrets. Too many valuable patterns can be explored from resumes. In this paper, we collect 372,829 Chinese resumes with complete attributes, such as income and Chinese ID number. Then a system called ResumeVis is developed to explore the correlations among the attributes. In the visualization system, we propose a new correlation representation -- parallel coordinates with multi-valued attributes to adapt to the characteristics of resumes. Besides, user-friendly interactions, such as filter elements, reorder attributes, brushing and linking, are integrated to provide an easy-use interface. Lots of patterns and inspiring results can be found by using the system. Three case studies are illustrated as examples to validate the system usability. The system can be used as a recommendation system for job seekers and HRs.

16:00 Named Entity Recognition in Clinical Text based on Capsule-LSTM for Privacy Protection (AIMS2019-4012)

Changjian Liu *(Harbin Institute of Technology (Shenzhen), China);* Jiaming Li *(Harbin Institute of Technology,*

Shenzhen & South China University of Technology, China); Yuhan Liu *(Harbin Institute of Technology, Shenzhen & Chinese Nankai University, China); Jiachen Du (Harbin Institute of Technology, Shenzhen, China);* Buzhou Tang *(Harbin Institute of Technology (Shenzhen), China);* Ruifeng Xu *(Harbin Institute of Technology, Shenzhen, China)*

Abstract: Clinical Named Entity Recognition for identifying sensitive information in clinical text, also known as Clinical De-identificaion, has long been critical task in medical intelligence. It aims at identifying various types of protected health information (PHI) from clinical texts and then replace them with special tokens. Along with the development of deep learning technology, lots of neural-network-based methods have been proposed to deal with Named Entity Recognition. As one of the state-of-the-art methods to address this problem, Bi-LSTM-CRF has become the mainstream due to its simplicity and efficiency. In order to better represent the entity-related information expressed in the context of clinical texts, we design a novel Capsule-LSTM network that is able to combine the great expressivity of capsule network with the sequential modeling capability of LSTM network. Experiments on 2014 i2b2 dataset show that the proposed method significantly outperform the baseline and thus reveal the superiority of the newly proposed Capsule-LSTM network.

16:20 Domain Knowledge Enhanced Error Correction Service for Intelligent Speech Interaction (AIMS2019-4013)

Yishuang Ning and Chunxiao Xing (Tsinghua University, China); Liang-Jie Zhang (Kingdee International Software Group Co., Ltd, China)

Abstract: Intelligent speech interaction systems have gained great popularity in recent years. For these systems, the accuracy of automatic speech recognition (ASR) has become a key factor of determining user experience. Due to the influence of environmental noise and the diversity and complexity of natural language, the performance of ASR still cannot meet the requirements of real-world application scenarios. To improve the accuracy of ASR, in this paper, we propose a domain knowledge enhanced error correction method which first the improved phonetic editing distance to select the candidates which have the same or similar phonetics with the error segment, and then adopts language model the find the most appropriate one from the domain knowledge set as the final result. We also encapsulate the method as a service with the Flask + Gunicorn + Nginx framework to improve the high concurrency performance. Experimental results demonstrate that our proposed method outperforms the comparison methods over 48.4% in terms of accuracy and almost 20-40 times concurrency performance.

2019 World Congress on Services (SERVICES 2019)

Session I: Services Research (06/26 Wednesday, 14:00-15:00; Destination Ballroom 2)

Chair: Yunni Xia (Chongqing University, China)
20 minutes for each paper.

14:00 SMT-based Modeling and Verification of Cloud Applications (SERVICES2019-5001)

Xiyue Zhang and Meng Sun (Peking University, China)

Abstract: Cloud applications have been rapidly evolving and gained more and more attention in the past decade. Formal modeling and verification of cloud services are necessarily needed to guarantee their correctness and reliability of complex cloud applications. In this paper, we present a formal framework for modeling and verification of cloud applications based on the SMT solver Z3. Simple cloud services are specified as the basis for the modeling of composition and more complex cloud services. Three different classes Service, Composition and Cloud indicating simple cloud services, composition patterns and composed cloud services are defined, which facilitates the further development of attributes and methods. We also propose an approach to check the refinement and equivalence relations between cloud services, in which counter examples can be automatically generated when the relation is not valid.

14:20 On Efficiency of Scrambled Image Forensics Service using Support Vector Machine (SERVICES2019-5002)

Sahibzada Muhammad Shuja, Raja Farhat Makhdoom Khan, Munam Ali Shah, Hasan Ali Khattak and Assad Abbass (COMSATS University Islamabad, Pakistan); Samee U. Khan (North Dakota State University, USA)

Abstract: Images can be a very good evidence during investigation of a crime scene. At the same time they can also contain very personal information which should not be exposed without the consent of the involved people. In this paper, We have presented here a practical approach to protect privacy of under investigation images with the use of Arnold's Transform (AT) scrambling and Support Vector Machine, We also pro-vide a new approach towards the whole forensics service provided by the designated agencies with the help of implementation of our approach. We enhanced the security of AT and provided privacy preserving mechanism to ensure protection of privacy. In literature only policies are defined to protect the privacy and lack of a solid approach which we have tried to resolve with a proof of concept implementation. In short, we have provided a full image forensics framework for illegal image detection while preserving the privacy.

14:40 Investigating Personally Identifiable Information Posted on Twitter Before and After Disasters (SERVICES2019-5003)

Pezhman Sheinidashtegol, Aibek Musaev and Travis Atkison (The University of Alabama, USA)

Abstract: Through social media, users may consciously reveal information about their personality; however, they could also unintentionally disclose private information related to the locations of where they live/work, car license plates, signatures or even their identification documents. Any disclosed information can endanger an individual's privacy, possibly resulting in burglaries or identity theft. To the best of our knowledge, this paper is the first to claim and demonstrate that people may reveal information when they are vulnerable and actively seek help; evidently, in the event of a natural disaster, people change their behavior and become more inclined to share their personal information. To examine this phenomenon, we investigate two hurricane events (Harvey and Maria) and one earthquake (Mexico City) using datasets obtained from Twitter. Our findings show a significant change in people's behavioral pattern in the disaster areas, regarding tweeting images that contain Personally Identifiable Information (PII), before and after a disaster event.

15:00 Utilizing Blockchain Technology to Enhance Halal Integrity: The Perspectives of Halal Certification Bodies (SERVICES2019-5009)

Marco Tieman, Mohd Ridzuan Darun, Yudi Fernando and Abu Bakar Ngah (Universiti Malaysia Pahang, Malaysia)

Abstract: Brand owners face multiple challenges in establishing end-to-end halal supply chains and in halal issue management. To this end, leading Halal Certification Bodies (HCBs) set up a discussion group to achieve consensus on a way forward and to examine the potential role of the halal blockchain in resolving these issues, as well as they key parameters, and segregation and communication requirements of the blockchain. Halal issues can be divided broadly into three areas: contamination, non-compliance and perception. Only cases involving contamination and non-compliance need to be reported to the HCB. Consensus has been achieved in the segregation of halal supply chains in terms of designated halal transport, storage and halal compliant terminals, for Muslim (majority) countries, whereas in non-Muslim (majority) countries greater leniency is possible. Effective segregation is only possible with effective communication, whereby the term 'halal supply chain' is encoded in freight documents, on freight labels and within the ICT system.

Session II: Services and Data Intelligence (06/27 Thursday, 14:20-15:40; Destination Ballroom 2)

Chair: Yunni Xia (Chongqing University, China)
20 minutes for each paper.

14:20 Current Trends in Collaborative Filtering Recommendation Systems (SERVICES2019-5004)

Sana Abida Amin (Florida International University, USA); James Philips and Nasseh Tabrizi (East Carolina University, USA)

Abstract: Many different approaches for designing recommendation systems exist, including collaborative filtering, content-based, and hybrid approaches. Following an overview of different collaborative filtering recommendation system design methodologies, this paper reviews 71 journals articles and conference papers to provide a detailed literature review of model-based collaborative filtering. The articles selected for this review were published within the last decade between 2008-2018. They are classified by database, application field, methodology, and publication year. Papers using Clustering, Bayesian, Association Rule, Neural Networks, Regression, and Ensemble methodologies are surveyed. Application areas include books, music, movies, social networks, and business. This survey also analyzes the type of the data that was used for application field. This literature review identifies trends for model-based collaborative filtering and through empirical results gives insight into future research trajectories in this field.

14:40 Big Data Quality: A Data Quality Profiling Model (SERVICES2019-5005)

Ikbal Taleb (Concordia University, Canada); Mohamed Adel Serhani (UAE University, United Arab Emirates); Rachida Dssouli (Concordia University, Canada)

Abstract: Big Data is becoming the standard data model and it is gaining wide adoption in the digital world. This is due to the volume of generated and processed Data from diverse application domains, and data sources. Estimating the Quality of Big Data is identified to be essential for data management and governance. To ensure a fast and efficient data quality assessment represented by its dimensions we need to extend the data profiling model to incorporate also quality profiling. The latter encompasses more value-added quality processes that go beyond data and its corresponding metadata. In this paper, we propose a Data Quality Profiling Model (DQPM) for Big Data that involves modules from sampling, profiling, exploratory quality profiling, data quality profile, and a quality profile repository (QPRepo). The repository plays an important role in managing quality dimensions and their related metrics, predefined quality actions, preprocessing activities (PPA) and their functions (PPAF), and the data quality profile. Our exploratory quality profiling discovers PPAF from systematic quality actions to leverage the quality trends of any dataset and show the cause and effects of such a process on the data. We conducted a series of experiments to evaluate different features of the DQPM including sampling and profiling, quality evaluation, and exploratory quality profiling for Big Data quality improvement. The results prove that quality profiling traces quality at the earlier stage of Big data lifecycle leading to quality improvement and enforcement.

15:00 On Development of Data Science and Machine Learning Applications in Databricks (SERVICES2019-5006)

Wenhao Ruan, Yifan Chen and Babak Forouraghi (Saint Joseph's University, USA)

Abstract: Databricks is a unified analytics engine that allows rapid development of data science applications using machine learning techniques such as classification, linear and nonlinear regression, clustering, etc. Existence of myriad sophisti-cated computational options, however, can become overwhelming for design-ers as it may not always be clear what choices can produce the best predictive model given a specific data set. Further, the mere high dimensionality of big data sets is a challenge for data scientists to gain a deep understanding of the results obtained by a utilized model. This paper provides general guidelines for utilizing a variety of machine learning algorithms on the cloud computing platform, Databricks. Visualiza-tion is an important means for users to understand the significance of the un-derlying data. Therefore, it is also demonstrated how graphical tools such as Tableau can be used to efficiently examine results of classification or cluster-ing. The dimensionality reduction techniques such as Principal Component Analysis (PCA), which help reduce the number of features in a learning exper-iment, are also discussed. To demonstrate the utility of Databricks tools, two big data sets are used for performing clustering and classification. A variety of machine learning algo-rithms are applied to both data sets, and it is shown how to obtain the most accurate learning models employing appropriate evaluation methods.

15:20 Concept Drift Adaptive Physical Event Detection for Social Media Streams (SERVICES2019-5007)

Abhijit Suprem and Calton Pu (Georgia Institute of Technology, USA); Aibek Musaev (University of Alabama, USA)

Abstract: Event detection has long been the domain of physical sensors operating in a static dataset assumption. The prevalence of social media and web access has led to the emergence of social, or human sensors who report on events globally. This warrants development of event detectors that can take ad-vantage of the truly dense and high spatial and temporal resolution data provided by more than 3 billion social users. The phenomenon of concept drift, which causes terms and signals associated with a topic to change over time, renders static machine learning ineffective. Towards this end, we present an application for physical event detection on social sensors that improves traditional physical event detection with concept drift adaptation. Our approach continuously updates its machine learning classifiers automatically, without the need

for human intervention. It integrates data from heterogeneous sources and is designed to handle weak-signal events (landslides, wildfires) with around ten posts per event in addition to large-signal events (hurricanes, earthquakes) with hundreds of thousands of posts per event. We demonstrate a landslide detector on our application that detects almost 350% more land-slides compared to static approaches. Our application has high performance: using classifiers trained in 2014, achieving event detection accuracy of 0.988, compared to 0.762 for static approaches.

Session III: Application Services (06/28 Friday, 15:20-16:40; Destination Ballroom 2)

Chair: Yunni Xia (Chongqing University, China)
20 minutes for each paper.

15:20 ClientNet Cluster an Alternative of Transferring Big Data Files by use of Mobile Code (SERVICES2019-5008)

Waseem Akhtar Mufti (Aalborg University, Denmark)

Abstract: Big Data has become a nontrivial problem in the field of business as well as in scientific applications. It becomes more complex with the growth of data and scaling of data entry points. These points refer to the remote and local sources where huge data is generated within tiny slots of time. This may also refer to the end user devices including computers, sensors and wireless gadgets. As far as scientific applications are concerned, for example, Geo Physics applications or real time weather forecast requires heavy data and complex mathematical computations. Such applications generate large chunks of data that needs to transfer it through conventional computer networks. Problem with Big Data applications emerges when heavy amount of data is transferred or downloaded (files or objects) from remote locations. The results drawn in real-time from large data files/sets become obsolete due to the fact data keeps on adding new data into the files and the downloading by remote machines remains slower as compared to file growth. This paper addresses this problem and provides possible solution through ClientNet Cluster of remote computers, Specialized Cluster of Computers, as one of the alternative to deal with real-time data analytics under the hard constraints of network. The idea is moving code, for analytic processing, to the remotely available big size files and returning the results to distributed remote locations. The Big Data file does not need to move around network for uploading or downloading whenever the processing is required from distributed locations.

15:40 Maintaining Fog Trust through Continuous Assessment (SERVICES2019-5010)

Hasan Ali Khattak, Muhammad Imran and Assad Abbas (COMSATS University, Islamabad, Pakistan); Samee U. Khan (North Dakota State University, USA)

Abstract: Cloud computing continues to provide flexible and efficient way for delivery of services, meeting user requirements and challenges of the time. Software, Infrastructures, and Platforms are provided as services in cloud and fog computing in a cost-effective manner. Migration towards fog instigate new aspects of research for security & privacy. Trust is dependent on measures taken for availability, security, and privacy of users' services as well as data in fog as well as sharing of these statistics with stakeholders. Any type of lapses in measures for security & privacy shatter user's trust. In order to provide a trust worthy security and privacy system, we have conducted a thorough survey of existing techniques. A generic model for trustworthiness is proposed in this paper. This model yields a comprehensive component-based architecture of a trust management system to aid fog service providers to preserve users' Trust in a fog computing environment.

16:00 Study on the Coordination Degree between FDI and Modern Service Industry Development in Shenzhen (SERVICES2019-5011)

Zhu Xiangbo(Shenzhen Polytechnic, China; China Three Gorges University, China); Ou Guoliang and Wu Gang (Shenzhen Polytechnic, China)

Abstract: Foreign direct investment (FDI) is an important driving force for economic development. In the period of rapid development and growth of Modern Service Industry (MSI), whether FDI had any impact on it was not clear. This paper selects evaluation indicators based on scale, structure and performance, and constructs a "FDI-MSI" coordination degree evaluation model. Based on the foreign direct investment (FDI) data of Shenzhen City in 2010-2018 and the development data of modern service industry, the development coordination degree between Shenzhen Foreign direct investment (FDI) and modern service industry (MSI) is measured.Research results show that the foreign direct investment (FDI) and the modern service industry are generally coordinated in Shenzhen, but the overall level is not high. The structure of foreign capital is not reasonable, and the current inefficiency in the use of foreign capital by the service industry is the main reason.In the process of utilizing foreign capital in the future, Shenzhen should strengthen the guidance of foreign investment with the goal of modern service industry development, create a favorable environment for foreign capital to flow into modern service industry, and promote the coordinated development of foreign capital utilization and modern service industry.

2019 International Conference on Cognitive Computing (ICCC 2019)

Session I: Cognitive Modeling (06/26 Wednesday, 11:00-12:00; Compass Ballroom 2)

Chair: Liang-Jie Zhang (Kingdee International Software Group Co., Ltd, China)
20 minutes for each paper.

11:00 Efficient gene assembly and identification for many genome samples (ICCC2019-6001)

Zhichun Zheng (University of Chinese Academy of Sciences & Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China); Ning Guo and Konda Mani Saravanan (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China); Yanjie Wei (Shenzhen Institutes of Advanced Technology, CAS, China)

Abstract: The development of the next generation sequencing technology (NGS) has advanced the genomics research in many application domains. Meta-genomics is one such powerful approach to study large community of microbial species. For the unknown species in the metagenomic samples, gene assembly and identification without a reference genome is a very challenging problem. To overcome this issue, distributed gene assembly software handling multiple metagenome samples can be used. In this paper, based on our highly scalable gene assembly software SWAP, we present a work flow called WFswap to assemble large genomic data based on many samples and to identify more genes. Our computational analysis reveals that the proposed workflow WFswap showed better performance than WFsoap, a similar work-flow that used SOAPdenovo2 for gene assembly. Our results suggested that WFswap is able to identify 94.2% of the benchmark genes when tested on the 19 metagenomic samples that contain Bifidobacterium animalis subsp. lactis CNCM I-2494.

11:20 A Neural Framework for Joint Prediction on Intent Identification and Slot Filling (ICCC2019-6002)

Jiawei Shan, Huayun Xu, Zeyang Gong and Hanchen Su (University of International Relations, China); Xu Han (Capital Normal University, Beijing, China); Binyang Li (University of International Relations, China)

Abstract: In task-oriented dialog systems, understanding of users' queries (expressed in natural language) is a process of parsing users' queries and converting them into some structure that machine can handle. The understanding usually consists of two parts, namely intent identification and slot filling. To address this problem, we propose a neural framework, named SI-LSTM, that combines two tasks and integrates CRF into LSTM network, where the slot information is extracted by using CRF, and the intent will be identified by using LSTM. In our approach, the slot information is used for determining the intent, while the intent type is used to rectify the slot filling deviation. Based on the dataset provided by NLPCC 2018, SI-LSTM achieved 90.71% on intent identification, slot filling and error correction in terms of accuracy.

11:40 Conditional Joint Model for Spoken Dialogue System (ICCC2019-6003)

Changliang Li (Kingsoft AI Lab, China); Yan Zhao and Dong Yu (Beijing Language and Culture University, China)

Abstract: Spoken Language Understanding (SLU) and Dialogue Management (DM) are two core components of a spoken dialogue system. Traditional methods model SLU and DM separately. Recently, joint learning has made much progress in dialogue system research via taking full advantage of all supervised signals. In this paper, we propose an extension of joint model to a conditional setting. Our model does not only share knowledge between intent and slot, but also efficiently make use of intent as a condition to predict system action. We conduct experiments on popular benchmark DSTC4, which includes rich dialogues derived from real world. The results show that our model gives excellent performance and outperforms other popular methods significantly, including independent learning methods and joint models. This paper gives a new way for spoken dialogue system research.

Session II: Sensing Intelligence (06/27 Thursday, 09:00-10:00; Compass Ballroom 2)

Chair: Huan Chen (Kingdee International Software Group Co., Ltd, China)
20 minutes for each paper.

9:00 Small Traffic Sign Detection and Recognition in High-Resolution Images (ICCC2019-6004)

Lei You, Hongpeng Wang and Wenhu You (Harbin Institute of Technology, China); Bo Wu (Shenzhen Polytechnic, China); Xinghao Song (Harbin Institute of Technology, China)

Abstract: Traffic sign detection and recognition is a research hotspot in the computer vision and intelligent transportation systems fields. It plays an important role in driver-assistance systems and driverless operation. Detecting signs, especially small ones, remains challenging under a variety of road traffic conditions. In this manu-script, we propose an end-to-end deep learning model for detecting and recogniz-ing traffic signs in high-resolution images. The model consists of basic feature extraction and multi-task learning. In the first part, a network with fewer parame-ters is proposed, and an effective

feature fusion strategy is adopted to gain a more distinct representation. In the second part, multi-task learning is conducted on dif-ferent hierarchical layers by considering the difference between the detection and classification tasks. The detection results on two newly published traffic sign benchmarks (Tsinghua-Tencent 100K and CTSD) demonstrate the robustness and superiority of our model.

9:20 A Feature Extraction Algorithm for the Detection of "Five-Distances" in Hazardous Chemical Storage and Stacking in Night Vision Environment (ICCC2019-6005)

Xuejun Liu, Yuchen Wei, Jiandong Zhang, Bo Li, Jin Wang, Bo Dai and Cuiqing Li (Beijing Institute of Petrochemical Technology, China);

Abstract: In recent years, safety distance monitoring of hazardous chemical stacking has become an increasingly important problem to be solved. It is a promising solution to build "Five-Distances" of binocular camera detection system based on cloud platform for intelligent safety distance warning. Aiming at the problem of feature extraction in "Five-Distances" detection of stacking in night vision environment, this paper proposes the construction of two-dimensional OSTU average gray function window and the average algorithm of threshold data results. The threshold value of feature extraction is calculated by combining the probability density distribution of foreground and background. Experiments on standard data sets and simulated experimental scene images show that this algorithm significantly reduces the number of noise points and improves the accuracy of subsequent edge detection and corner point detection compared with the gray straight square homogenization algorithm.

9:40 Structural-based Graph Publishing under Differential Privacy (ICCC2019-6006)

Yiping Yin, Qing Liao, Yang Liu and Ruifeng Xu (Harbin Institute of Technology (Shenzhen), China)

Abstract: Network data (e.g. online social networks, collaboration networks and so on) ,which is large and structural, is the core for many high impact applications. Their use requires techniques that meet the demand of applications while offering rigorous privacy guarantee. Differential Privacy has achieved remarkable results of protecting privacy on data publishing. However, current differential privacy methods are often difficult to adapt a very large network because of its high computation expense and do not learn the structure of networks which involves too much noise. Therefore, we develop new algorithmic method for publishing network data by combining state-of-the-art community detection method under the framework of differential privacy based on hierarchical representation, sanitizing a network by adding different degree of noises based on the structural information.

Session III: Cognitive Computing Technologies and Frameworks **(06/28 Friday, 15:20-16:40; Compass Ballroom 2)**

Chair: Huan Chen (Kingdee International Software Group Co., Ltd, China)

20 minutes for each paper.

15:20 Face detection using bionic cascaded Framework (ICCC2019-6007)

Jin Li, Ziyue Chen, Shunxin Ouyang, Jingyu Xie, Yue Hu and Hui Lv (Hubei University of Technology, China)

Abstract: Face interaction plays an irreplaceable role in the service robots human-robot in-teraction, while face detection in this kind of scenario are challenging due to re-strictions on computing capabilities and power, the character of real-time and re-quirements of the interaction pattern. Recent studies show that deep learning ap-proaches can achieve impressive performance on these kinds of tasks. In this pa-per, we propose a bionic cascaded framework which exploits saccade and atten-tion mechanism of human eyes to balance the performance. In particular, our framework adopts a cascaded structure with two stages of carefully designed face detectors. In addition, in the working process of the service robots, we propose a new online sampling strategy that can improve the performance of interaction pat-terns. Our method achieves real time face detection in service robot systems, and more natural human-robot interaction pattern.

15:40 A Neural Rumor Detection Framework by Incorporating Uncertainty Attention on Social Media Texts (ICCC2019-6008)

Yan Gao (University of International Relations, China); Xu Han (Capital Normal University, China); Binyang Li (University of International Relations, China)

Abstract: Automatic rumor detection technology has become a very urgent need, as rumors can arise and spread dauntingly fast in social media, which bring unforeseeable and devastating impacts. However, current approaches mainly capture the event semantics or us-er-based features for rumor detection, but neglect the uncertainty expressions that strongly indicate the unverified nature of a rumor. As a result, these methods perform suboptimal when the topics of being verified rumors are changing wildly. In this paper, we present a neural rumor detection framework, namely NERUD. In NERUD, both uncertainty semantics and the event semantics of a word are represented by the attention mechanisms to generate a rumor representation for rumor detection. Experiments were conducted on the benchmark dataset and the Chinese Rumor Corpus (CRC), and the results showed that our NERUD out-performed state-of-the-art approaches on CRC dataset, and the uncertainty semantics was proven effective on rumor detection task.

16:00 Using GAN to Generate Sport News from Live Game Stats (ICCC2019-6009)

Changliang Li (Kingsoft AI Lab, China); Yixin Su (University of Melbourne, Australia); Ji Qi (Kingsoft AI, China); Min Xiao (Beijing University of Posts and Telecommunications, China)

Abstract: One goal in artificial intelligence field is to create well-formed and human-like natural language text given data input and a specific goal. Some data-to-text solutions have been proposed and successfully used in real applied domains. Our work focuses on a new domain, Automatic Sport News Generating, which aims to produce sport news immediately after each match is over so that both time and labor can be saved on writing the news articles. We propose to use Generative Adversarial Networks (GAN) architecture for generating sport news based on game stats. Our model can automatically determine what is worth reporting and generate various appropriate descriptions about the game. We apply our approach to generate NBA (National Basketball Association) game news. Especially, This paper focuses on reporting the summary of game result and performance of players. Our model achieves good results on both tasks. To our best knowledge, this is the first work based on GAN to generate sports news using game statistics.

16:20 Speech Emotion Recognition Using Multi-Granularity Feature Fusion through Auditory Cognitive Mechanism (ICCC2019-6010)

Cong Xu and Haifeng Li (Harbin Institute of Technology, China); Hongjian Bo (Shen Zhen Academy of Aerospace Technology, China); Lin Ma (Harbin Institute of Technology, China)

Abstract: In this paper, we focus on the problems of single granularity in feature extraction, loss of temporal information and inefficient use of frame features in discrete speech emotion recognition. Firstly, preliminary cognitive mechanism of auditory emotion is explored through cognitive experiments, and then a multi-granularity fusion feature extraction method inspired by the mechanism for discrete emotional speech signals is proposed. The method can extract 3 different granularity features, including short-term dynamic features of frame granularity, dynamic features of segment granularity and long-term static features of global granularity. Finally, we use the LSTM network model to classify emotions according to the long-term and short-term characteristics of the fusion features. We implement experiment on the discrete emotion datasets of CHEAVD(CASIA Chinese Emotional Audio-Visual Database) released by the Institute of automation, China Research Academy of Sciences, and achieved improvement in recognition rate, increasing the MAP by 6.48%.

Session IV: Sensing Intelligence Application (06/27 Thursday, 14:20-15:40; Compass Ballroom 2)

Chair: Sheng He (Tsinghua University, China)

20 minutes for each paper.

14:20 Learning Contextual Features with Multi-head Self-attention for Fake News Detection (ICCC2019-6011)

Yangqian Wang and Hao Han (Harbin Institute of Technology（Shenzhen）, China); Ye Ding (Dongguan University of Technology, China); Xuan Wang (Harbin Institute of Technology Shenzhen Graduate School, China); Qing Liao (Harbin Institute of Technology（Shenzhen）, China)

Abstract: Automatic fake news detection has attracted great concern in recent years due to it's tremendous negative impacts on public. Since fake news is usually written to mislead readers, lexical features based methods have great limitations. Previous work has proven the effective-ness of contextual information for fake news detection. However, they ignore the in uence of sequence order when extract features from contextual information. Inspired by transformer technique, we propose Contextual Features with Multi-head Self-attention model(CMS) to extract features from contextual information for fake news detection. CMS can automatic capture the dependencies between contextual information and learning a global representation from contextual information for fake news detection. Experimental results on the real-world data demonstrate the effectiveness of the proposed model.

14:40 Reduce the Warfighters' Cognitive Burden (ICCC2019-6012)

Matthew Maher and Rob Orlando (Processus Group & Corvus Inc, USA)

Abstract: As the DoD and Military acquisitions shift the paradigm from antiquated software capabilities to artificial intelligence (AI) and machine learning (ML) capabilities, they need to ensure key infrastructure (big data and knowledge artifacts) is also mature enough to handle the transition. Before any organizations or domains develop or adopt a strategy based on AI, ML, or any other form of intelligent system, they need to do a self -assessment that answers three key questions. How well does your organization/domain share knowledge? What is the state of your data? Do you understand the cognitive requirements? Answering and understanding these questions is the key to knowing where you are now (point A) so you can take actionable and progressive steps towards where you want to go (point B) to make your organization or domain AI-capable. To provide the answers and/or solutions to the questions above we execute a Mission Engineering (ME) process that analyzes the organization's or domain's state of readiness. The ME process is a detailed analysis of an organization/domain providing an assessment of how the people, systems, knowledge, data, and workflows are aligned to the operational processes. ME adds a layer of operational viability to existing engineering processes with the goal of ensuring the program enhances knowledge sharing, reduces cognitive burdens, and increases semantic understanding. The ME process has roots in tactical operations and arose out of an operational need. A need to combine data science with our

operational experience. The requirement for this pairing first became obvious when far too many of our after-action reviews in Bosnia, Iraq and Afghanistan focused on the lack of information and proper analysis of the information to support decision making. On one specific operation, while in 3rd Special Forces Group (A), we were told that the objective was supported by only light contingent of forces. However, as we approached the target area, we saw heavy equipment at the target site. During the debriefing we realized we had the information that a heavy force was possible, but because individuals responsible for informing us had too much data to analyze and not enough time it was never included in the planning. They had a heavy cognitive burden based on the amount of data, information and intelligence. Also, they had to apply their knowledge and understanding without any system that could lessen the burden. From then on, we started developing an ME process that would lessen the cognitive burden on the warfighter while still applying knowledge and understanding that comes with experience. Over the years, the ME process has evolved. The output of the process is three key components: Semantic Data Model (SDM), Cognitive Architecture (CA), and Knowledge Graph (KG). The three components work together to enhance the cognitive development in the following ways: • SDM informs the AI/ML solution on the critical structured and unstructured data that is required to access for intelligent systems. The SDM defines the semantic (machine readable) definition, provenance and relationship of the data to core operational processes to reduce the cognitive burden of a systems and warfighter. • CAs to provide a blue print for AI/ML developers to understand the how they develop systems to a Mission Engineering standard. The CA defines the key Data, Information, Knowledge, Understanding, and Intelligence (DIKUI) knowledge resources that are required for cognitive problem solving for domain and multidomain operations. • KGs to provide the level of detail for how the operations is executed and maps how each function, task, systems and personnel depend, create, change and access the distributed data required to make an operational decision. Together these three deliverables will provide a proven and operationally accurate cognitive problem-solving technique, a predictive analytic capability that provides the semantics and patterns for confident decision making, and a cognitive model tool and will reduce the cognitive workload.

15:00 An ensemble method: Case-Based Reasoning and the Inverse Problems in Investigating Financial Bubbles (ICCC2019-6013)

Francis Ekpenyong and Stelios Kapetanakis (University of Brighton, United Kingdom (Great Britain)); Georgios Samakovitis (University of Greenwich, United Kingdom (Great Britain)); Miltos Petridis (University of Brighton, United Kingdom (Great Britain))

Abstract: This paper presents an ensemble approach that uses Case Base Reasoning and Inverse Problem formulation to describe and model abnormal stock market fluctuations in time series datasets from historical stock market prices (often associated with asset bubbles). The framework proposes to use a rich set of past observations and a geometric pattern description and apply CBR to formulate the forward problem; Inverse Problem formulation will then be used to identify a set of parameters that can statistically be associated with the occurrence of the observed patterns. The proposed use of the CBR-IP ensemble is a novel methodology that has promised superior performance over existing methodologies.

15:20 The Development Trend of Intelligent Speech Interaction (ICCC2019-6014)

Yishuang Ning, Sheng He and Chunxiao Xing (Tsinghua University, China); Liang-Jie Zhang (Kingdee International Software Group Co., Ltd, China)

Abstract: To make the computers have capabilities of listening, speaking, understanding and even thinking is the lastest development direction of human-computer interaction. As one of the most convenient and natural ways for communication, speech has become the most promising way of human-computer interaction in the future, which has more advantages than other interaction ways. As one of the most popular artificial intelligence (AI) technologies, intelligent speech interaction technology has been widely applied in many industries such as electronic commerce, smart home and intelligent industry as well as manufacturing. It will change the user behavior habits and become the new mode of human input and output. In this paper, we state the current situation of intelligent speech interaction at home and abroad, take many examples to illustrate the application scenarios of speech interaction technology and finally introduce its development trend in the future.

Session V: Cognitive Intelligence Application (06/29 Saturday, 09:00-10:00; Destination Ballroom 2)

Chair: Sheng He (Tsinghua University, China)
20 minutes for each paper.

9:00 Imagery Signal-based Deep Learning Method for Prescreening Major Depressive Disorder (ICCC2019-6015)

Hyunjin Kwon, Jinhyeok Park, Seokhwan Kang and Youngho Lee (Gachon University, Korea)

Abstract: Depression is a high-risk mental illness that can lead to suicide. However, due to a variety of reasons, such as a negative perception of mental illness, most patients with depressive symptoms are reluctant to go to the hospital and miss appropriate treatment. Therefore, a simple prescreening method to identify depression that an individual can use is needed. Most EEG measurement devices that individuals use have fewer channels. However, most studies using EEG to diagnose depression have been conducted in a professional multi-channel EEG environment. Therefore, it is difficult for individuals to prescreen depression based on the results of the studies. Therefore, in this study, we proposed a model that predicts

depression by using EEG data measured by a few channels so that it can measure depression using the EEG data measured by an individual. In this study, brain waves measured in two channels were imaged using STFT transform and spectrogram. Then the EEG image data was used in a deep learning model. As a result of the performance evaluation, 75% accuracy was shown for the classification of image depression EEG and normal image type EEG. As a result, the use of low channel EEG data for deep learning can be used as an auxiliary tool to pro-actively diagnose depressed patients.

9:20 Visual Cognition Mechanism Guided Video Shot Segmentation (ICCC2019-6016)

Chenzhi Shao, Haifeng Li and Lin Ma (Harbin Institute of Technology, China)

Abstract: Shot segmentation of video sequences is one of the key technologies in video information processing, especially video retrieval. Traditional shot segmentation methods have low detection rate for the gradient shot and the abrupt shot, especially in a single scene. To deal with this problem, this paper proposes a video segmentation method based on visual cognition mechanism. This method proposes a block granularity color histogram to strengthen the visual salient area, and a highlight measure to describe the difference between the front and back frames. This brings great improvements to the accuracy of detecting shot switching in a single scene. In addition, based on the brightness visual perception in video, the difference between adjacent multi-frames in the sliding window is used to capture the brightness change for the gradient shots. Comparing with traditional methods, the proposed algorithm achieves better segmentation effect and has higher precision and recall rate.

9:40 A best detecting synchrony method in audio STROOP EEG based on wavelet coherence (ICCC2019-6017)

Kang Liu and Chunying Fang (Heilongjiang University of Science and Technology, China); Haifeng Li (Harbin Institute of Technology, China); Tingpeng Li (Heilongjiang University of Science and Technology, China)

Abstract: Advanced brain function requires different levels of integration and coordination between multi-regional nervous systems, the underlying mechanism is the simul-taneous oscillation of various neural networks. EEG is an increasingly method to detect brain function with high temporal resolution and low cost. How to analyze the synchronization phenomenon is the focus of cognitive neuroscience research based on EEG signals. Wavelet coherence is a classical method to evaluate EEG synchronization, but it is uncertain how to use. In this paper, this requires knowledge of the true relationship between signals, hence we compare different measures of functional connectivity on simulated data (unidirectional coupled Hénon maps, and the auditory Stroop EEG), including wavelet cross-spectrum, wavelet correlation, wavelet coherence and FFT coherence. To determine whether synchrony is detected, surrogate data were generated and analyzed, and FFT co-herence measures performed best on simulated data. Above all, the parameter op-timization method of the wavelet cross-spectrum is proposed with many samples. It is found that the optimized wavelet coherence performed most reliably than FFT coherence.

2019 International Conference on Internet of Things (ICIOT 2019)

Session I: IoT Security (06/26 Wednesday, 11:00-12:00; Compass Ballroom 1)

Chair: Samee U. Khan (North Dakota State University, USA)
20 minutes for each paper.

11:00 Underground Environment Aware MIMO Design Using Transmit and Receive Beamforming in Internet of Underground Things (ICIOT2019-7001)

Abdul Salam (Purdue University, USA)

Abstract: In underground (UG) multiple-input and multiple-output (MIMO), the transmit beamforming is used to focus energy in the desired direction. There are three different paths in the underground soil medium through which the waves propagates to reach at the receiver. When the UG receiver receives a desired data stream only from the desired path, then the UG MIMO channel becomes three path (lateral, direct, and reflected) interference channel. Accordingly, the capacity region of the UG MIMO three path interference channel and degrees of freedom (multiplexing gain of this MIMO channel requires careful modeling). Therefore, expressions are required derived the degrees of freedom of the UG MIMO interference channel. The underground receiver needs to perfectly cancel the interference from the three different components of the EM-waves propagating in the soil medium. This concept is based upon reducing the interference the undesired components to minimum at UG receiver using the receive beamforming. In this paper, underground environment aware MIMO using transmit and receive beamforming has been developed. The optimal transmit beamforming and receive combining vectors under minimal inter-component interference constraint are derived. It is shown that UG MIMO performs best when all three component of the wireless UG channel are leveraged for beamforming. The environment aware UG MIMO technique leads to three-fold performance improvements and paves the wave for design and development of next generation sensor-guided irrigation systems in the field of digital agriculture.

11:20 Evaluation of Heterogeneous Scheduling Algorithms for Wavefront and Tile Parallelism in Video Coding (ICIOT2019-7002)

Natalia Panagou, Maria Koziri, Panos Papadopoulos and Panagiotis Oikonomou (University of Thessaly, Greece); Nikos Tziritas (University of Thessaly, Greece); Kostas Kolomvatsos and Thanasis Loukopoulos (University of Thessaly, Greece); Samee U. Khan (North Dakota State University, USA)

Abstract: Video is by far the "biggest" Big Data, stretching network and storage capacity to their limits. To handle the situation, video compression has been an active field of study for many years, producing output of huge commercial interest, e.g., MPEG-2 and DVD. However, video coding is a computationally expensive process and for this reason, parallelization was proposed at various granularity levels. Of particular interest, are block level methods implemented in HEVC (High Efficiency Video Coding) which was designed to be the successor of H.264/AVC for the 4K era. Parallelization in HEVC is supported by the following three modes: slices, tiles and wavefront. While considerable research was conducted on the parallelization options of HEVC, it was focused on the case of homogeneous processors. In this paper we consider video coding parallelization when the processing elements are heterogeneous. In particular, we focus on wavefront and tile parallelism and measure the performance of scheduling schemes for the induced subtasks. Through simulation experiments with dataset values obtained from common benchmark sequences, we conclude on the relevant merits of the evaluated scheduling algorithms.

11:40 A Method to Secure IoT devices Against Botnet Attacks (ICIOT2019-7003)

Trusit Shah (The University of Texas at Dallas, USA); Subbarayan Venkatesan (University of Texas at Dallas, USA)

Abstract: An unsecured or weak authentication system between an IoT device and a user provides an opportunity to an attacker to manipulate and use the IoT device as a botnet. The proliferation of IoT devices with an unsecured/weak authentication mechanism has increased the threat of using a huge number of botnets for large-scale DDoS attacks. Default credential pairs (like 'root-root' or 'admin-admin') for the Telnet or SSH connections are still part of a large group of IoT products, and many malwares have exploited this vulnerability to capture the IoT devices and use them as botnets. In the recent past, Mirai malware acquired more than 600k IoT devices by brute-forcing 62 pairs of default credentials. In this paper, we present a concept called 'digital captcha' to mitigate this problem. Just like normal captcha, 'digital captcha' is a problem presented during the login process to the remote device to reduce the number of log-in attempts. Digital captcha is a computer science problem with a varying complexity, which the login entity needs to solve before logging into any IoT device. Every unsuccessful log-in attempt increases the complexity of solving the digital captcha for the next attempt. If each IoT device had used digital captcha, Mirai attack would have required almost two months to acquire devices, while it acquired them in 20 hours.

Session II: IoT Services (06/27 Thursday, 14:20-15:40; Compass Ballroom 1)

Chair: Balaji Palanisamy (University of Pittsburgh, USA)
20 minutes for each paper.

14:20 Clearer than Mud: Extending Manufacturer Usage Description (MUD) for Securing IoT Systems (ICIOT2019-7004)

Simran Singh, Ashlesha Atrey, Mihail Sichitiu and Yannis Viniotis (North Carolina State University, USA)

Abstract: Internet of Things (IoT) devices, expected to increase exponentially over the next several years, are easy targets for attackers. To make these devices more secure, the IETF's draft of Manufacturer Usage Description (MUD) provides a means for the manufacturer of an IoT device to specify its intended purpose and communication patterns in terms of access control lists (ACLs), thereby defining the device's normal behaviour. However, MUD may not be sufficient to comprehensively capture the normal behaviour specification, as it cannot incorporate variable operational settings that depend on the environment. Further, MUD only supports limited features. Our approach overcomes these limitations by allowing the administrator to define the normal behaviour by choosing combinations from a wider set of features that includes physical layer parameters, values of packet headers, and flow statistics. We developed and implemented a learning-based system that captures and demodulates wireless packets from IoT devices over a period of time, extracts the features specified in the normal behaviour specification, and uses a learning algorithm to create a normal model of each device. Our implementation also enforces these normal models by detecting violations and taking appropriate actions, in terms of ACLs on an Internet Gateway, against the misbehaving devices. Hence, our framework makes the specification tighter and clearer than what is possible with MUD alone, thereby making IoT systems more secure.

14:40 A WS-Agreement based SLA Ontology for IoT Services (ICIOT2019-7005)

Fan Li, Christian Cabrera and Siobhán Clarke (Trinity College Dublin, Ireland)

Abstract: In the Internet of Things (IoT), billions of physical devices, distributed over a large geographic area, provide a near real-time state of the world. These devices' capabilities can be abstracted as IoT services and delivered to users in a demand-driven way. In such a dynamic large-scale environment, a service provider who supports a service level agreement (SLA) can have a comprehensive competitive edge in terms of service quality management, optimized resource allocation, service customization, and trustworthiness. However, there is no consistent way of drafting an SLA with respect to describing heterogeneous IoT services, which obstructs automatic service selection, SLA negotiation, and SLA monitoring. In this paper, we propose an ontology, WIoT-SLA, to achieve semantic interoperability. We combine IoT service properties with two prominent web service SLA specifications: WS-Agreement and WSLA, to take advantage of their complementary features. This ontology is used to formalize the SLAs and SLA negotiation offers, which further facilitates the candidate service selection and automatic SLA negotiation. It can also be used by a monitoring engine to detect SLA violations by providing the semantics of service level objectives (SLOs) and quality metrics. To evaluate our work, a prototype is implemented to demonstrate its feasibility and efficiency.

15:00 IoT for fault detection in Thailand (ICIOT2019-7006)

Anurak Choeichum, Yutthana Krutgard1 and Wichan Inyoo (Provincial Electricity Authority, Thailand)

Abstract: Fault location and notification are one of the most challenges in Thailand, which Provincial Electricity Authority (PEA) must develop, and in particular fault detection in a power distribution system that long line and so many locations. The power outage events after the drop out fuse (DF) device in the distribution line fault often occur many times, and take a long time to find the fault location, these things making the power system unreliable. There is no notification system and long time to fault section finding, if there is a sudden drop in DF detectors, it will greatly reduce the power outage duration (SAIDI). The objective of this innovation is to provide with a new way to detect, send and monitor the fault signal and show the exact section of occurred fault immediately and accurately. This innovation employs the current signal to locate the fault sector based on the internet of things (IoT), wireless sensor network (WSN), and send the information to monitoring centers (SCADA center) or service team with newly implemented LoRa technology for massive and widespread installation. This is breaking the limits in signal sending, and using a magnetic field, solar energy, and data sending via wireless networks. This design concept used many devices, such as hall-effect current sensor, microcontroller for IoT, wireless networks, and LoRa & GSM module, etc. There are creating the web server and web application for online fault monitoring. When the fault occur after DF in distribution line, this system will detect the fault signal with WSN client node, send the data to LoRa base station with WSN master node, then send the data to GSM base station, and send information & display the power measurements, device status, and fault location to the dispatching center (SCADA center) or service team. In this implementation, we proposed concept using internet of things that could sense the event in the electrical line via internet network.

15:20 IoT-Based Monitoring and Control Systems for Window Energy Management: Design and Implementation (ICIOT2019-7007)

Yoon G Kim and Shenghui Chen (Calvin College, USA); Timothy J Dykhuis (Calvin College & Mackinac Technology

Company, USA); John Slagter and Matthew Nauta (Mackinac Technology Company, USA)
Abstract: This paper presents the design and implementation of an IoT-based system for window energy management. The system computes thermal transmittances of window units and solar heat gain coefficients through window units from measured sensor readings. It is critical to have energy efficient windows in homes and offices so as to conserve energy. In order to test the performance of the window units in realistic environments we created a system which can be used outdoors at any time. It is convenient for a user to control the system and access the measured data from a remote or even mo-bile site. The goal of this work was to design and implement a system that measures sensor data, transmits the data over a wireless communication link, distributes the data through the Internet, and stores the data in a database for analysis at any time and anywhere. The IoT is a computer network, in which anyone and anything can connect together anytime and anywhere [1]. It is realizable through sensing and communication technologies [2]. The IoT approach was adopted to achieve our goal due to the availability of sensing technology, wireless communication technologies, and standard computer networking protocols. We designed and implemented a system for calculating the thermal energy related parameters of the window units. The measured data from sensors were transmitted to a cloud server over cellular networks and the Internet. The data in the server can be accessed, stored, and displayed remotely. As long as electricity and cellular phone networks are available, the system can connect sensors to users. Extensive testing was conducted to verify the operation of the system. The testing and measurement results show that the system successfully performs the necessary operations to achieve the goal.

Session III: Smart IoT (06/28 Friday, 15:20-16:40; Compass Ballroom 1)

Chair: Balaji Palanisamy (University of Pittsburgh, USA)
20 minutes for each paper.

15:20 Online Conditions Monitoring of End-Mill Based on Sensor Integrated Smart Holder (ICIOT2019-7008)

Zhaowu Zhan, Kai Xie, Letian Rong and Wei Luo (Fuhuake Precision Industry (Shenzhen) Co., Ltd & Foxconn Technology Group, China)
Abstract: A intelligent IoT (Internet of Things) system dedicated to online conditions monitoring of end-mill has been developed in this research work. The smart holder integrated with triaxial acceleration sensor, MCU (Micro Controller Unit), wireless radio chip and battery is demonstrated. The system collects high resolution end-milling data during cutting without interfering with the machining process. In order to extract the data sequence containing the information highly related to the conditons of the end-mill, a simple but effective data extraction algorithm based on sinusoidal correlation is proposed. By realtime exploring the extracted data sequence via the developed data mining algorithm on the CNC (Computerized Numerical Control) edge server, the conditions of the end-mill can be monitored online.

15:40 Context-Aware Continuous Authentication and Dynamic Device Pairing for Enterprise IoT (ICIOT2019-7009)

Na Yu, Jia Ma, Xudong Jin, Jian Wang and Ken Chen (Samsung Research America, USA)
Abstract: Enterprise IoT is a special type of Internet of Things (IoT). It focuses on connecting the enterprise assets and their devices with backend application services and frontend user interactions. While it is more complicated and more impactful than smart home, it faces several challenges such as a large number of IoT devices deployed at various locations and also multiple users to be granted with different permissions. To solve these challenges, we propose a multi-domain enterprise IoT system in consideration of both user movement and IoT device relocation. We propose a context-aware continuous authentication method to authenticate mobile users to the IoT domains based on the context inferred from various sensors on the smartphones and in the IoT domains. We also propose a dynamic device pairing method to support the mobility of users and IoT devices based on data (e.g., contexts, access history, etc.) sharing among the IoT domain, the enterprise server, and the mobile users.

16:00 Smart IoT In-Car Life Detector System to Prevent Car Deaths (ICIOT2019-7010)

Nesreen Alsbou and Ka Hei Samuel Chen (University of Central Oklahoma, USA); Mohamed Afify (UCO, USA)
Abstract: Unfortunately, many deaths are reported every year during the summer for infants, kids and pets who were left in cars and suffered from heatstroke. This paper is proposing a simple smart IoT system for in car detection of a living subject and taking the necessary actions to prevent their life loss. The system has a wide range of sensors including microwave sensors to collect a large set of data inside the car. The system will analyze the data and make the appropriate actions to prevent any death. These actions include sending an alert message to the parents and the emergency department, rolling down the windows, or starting the engine and turning on the air conditioner.

16:20 Cloud-Based IoT Smart Parking System for Minimum Parking Delays on Campus (ICIOT2019-7011)

Nesreen Alsbou (University of Central Oklahoma, USA); Mohamed Afify (UCO, USA); Imad Ali (University of Oklahoma Health Sciences Center, USA)
Abstract: Growing cities always have parking challenges and they are in need for creative ideas to solve this issue and avoid the time wasted in searching for empty parking spots. To overcome the problem, this paper proposes a simple solution using

a low-cost cloud-based system design. The design will be initially implemented on campus in one parking lot at the University of Central Oklahoma. The goal is to make the faculties and students life easier by guiding them to empty parking spots. The design of the proposed system is discussed in this paper and preliminary data are presented including the cost function. The system will guide the users through the web based application.

2019 International Conference on Edge Computing (EDGE 2019)

Session I: Edge Computing in IoT (06/27 Thursday, 13:00-14:00; Destination Ballroom 2)

Chair: Huan Chen (Kingdee International Software Group Co., Ltd, China)
20 minutes for each paper.

13:00 Characterization of IoT Workloads (EDGE2019-8001)

Daniel A Menasce and Uma Tadakamalla (George Mason University, USA)

Abstract: Workload characterization is a fundamental and necessary step in carrying out any performance and Quality of Service engineering study. The workload of a system is defined as the set of all inputs received by the system from its environment during one or more time windows. The characterization of the workload entails determining the nature of its basic components as well as a quantitative and probabilistic description of the workload components in terms of both the arrival process, event counts, and service demands. Several workload characterization studies were presented for a variety of domains, except for IoT workloads. This is precisely the main contribution of this paper, which also presents a capacity planning study based on one of the workload characterizations presented here.

13:20 Latency Control for Distributed Machine Vision at the Edge through Approximate Computing (EDGE2019-8002)

Arun Ravindran and Anjus George (University of North Carolina at Charlotte, USA)

Abstract: Multicamera based Deep Learning vision applications sub- scribe to the Edge computing paradigm due to stringent latency require- ments. However, guaranteeing latency in the wireless communication links between the cameras nodes and the Edge server is challenging, especially in the cheap and easily available unlicensed bands due to the interference from other camera nodes in the system, and from exter- nal sources. In this paper, we show how approximate computation tech- niques can be used to design a latency controller that uses multiple video frame image quality control knobs to simultaneously satisfy latency and accuracy requirements for machine vision applications involving object detection, and human pose estimation. Our experimental results on an Edge test bed indicate that the controller is able to correct for up to 164% degradation in latency due to interference within a settling time of under 1.85 seconds.

Session II: Edge Comminications and Security (06/29 Saturday 09:00-10:20; Destination Ballroom 3)

Chair: Huan Chen (Kingdee International Software Group Co., Ltd, China)
20 minutes for each paper.

9:00 Stackelberg Game-Theoretic Spectrum Allocation For QoE-Centric Wireless Multimedia Communications (EDGE2019-8004)

Krishna Murthy Kattiyan Ramamoorthy and Wei Wang (San Diego State University, USA); Kazem Sohraby (South Dakota School of Mines and Technology, USA)

Abstract: Multimedia Quality of Experience (QoE) is a predominant factor that drives customer satisfaction and user experience in the future wireless networks. For example, Dynamic Adaptive Streaming over HTTP (DASH) has been widely adapted for multimedia content delivery over internet and allows the client devices to choose the bitrates of the subsequent frames based on channel condition. This paper proposes a Stackelberg game theoretic spectrum allocation approach for QoE-centric wireless multimedia communication rather than the traditional data traffic. Here, we introduce the cost of utilizing the spectrum as a factor in the utility of the service provider and the client device. Both service provider and client devices are assumed rational and selfishly look to maximize their utility in a non-cooperative manner. Stackelberg game is used to formulate the interaction between the service provider and the client device, and to derive the Nash Equilibrium for the utility maximization problem. The paper proves existence of a Stackelberg game solution such that the utility of both client device and the service provider is maximized. The simulation results demonstrate that QoE and fairness can be achieved by the proposed game-theoretic spectrum allocation scheme.

9:20 Intrusion Detection at the Network Edge: Solutions, Limitations, and Future Directions (EDGE2019-8005)

Roberto Di Pietro (Hamad Bin Khalifa University, Qatar); Maurantonio Caprolu (HBKU-CSE, Qatar); Simone Raponi (Hamad Bin Khalifa University, Italy)

Abstract: The low-latency, high bandwidth capabilities promised by 5G, together with the diffusion of applications that require high computing power and, again, low latency (such as videogames), are probably the main reasons-though not the

only one-that have led to the introduction of a new network architecture: Fog Computing, that consists in moving the computation services geographically close to where computing is needed. This architectural shift moves security and privacy issues from the Cloud to the different layers of the Fog architecture. In this scenario, IDS are still necessary, but they need to be contextualized in the new architecture. Indeed, while on the one hand Fog computing provides intrinsic benefits (e.g., low latency), on the other hand it introduces new design challenges. In this paper, we provide the following contributions: we analyze the possible IDS solutions that can be adopted within the different Fog computing tiers, together with their related deployment and design challenges; and, we propose some promising future directions, by taking into account the challenges left uncovered by the considered solutions.

9:40 Volunteer Cloud as an Edge Computing Enabler (EDGE2019-8006)

Tessema M Mengistu and Dunren Che (Southern Illinois University at Carbondale, USA)

Abstract: The rapid increase in the number of devices connected to the Internet, due to the Internet of Things, demands new ways of processing data produced by the devices. Edge Computing is one of the solutions that tries to process data close to the origin, which is the edge of networks. Emerging cloud systems, such as volunteer clouds can also be used towards the pro- cessing of data produced by IoT devices. This paper proposes a Volunteer Computing as a Service (VCaaS) based Edge Computing infrastructure. The paper addresses the architectural design of the proposed system together with its research and technical challenges.

10:00 Energy-Aware Capacity Provisioning and Resource Allocation in Edge Computing Systems (EDGE2019-8003)

Tayebeh Bahreini, Hossein Badri and Daniel Grosu (Wayne State University, USA)

Abstract: Energy consumption plays a key role in determining the cost of services in edge computing systems and has a significant environmental impact. Therefore, minimizing the energy consumption in such systems is of critical importance. In this paper, we address the problem of energy-aware optimization of capacity provisioning and resource allocation in edge computing systems. The main goal is to provision and allocate resources such that the net profit of the service provider is minimized. We formulate the problem as a mixed integer linear program and prove that the problem is NP-hard. We develop a heuristic algorithm to solve the problem efficiently. We evaluate the performance of the proposed algorithm by conducting an extensive experimental analysis on problem instances of various sizes. The results show that the proposed algorithm has a very low execution time and is scalable with respect to the number of users in the system.

2019 International Conference on Blockchain (ICBC 2019)

Session I: Blockchain and Consensus (06/27 Thursday, 09:00-10:00; Compass Ballroom 1)

Chair: Surya Nepal (CSIRO, Australia)
20 minutes for each paper.

9:00 Blockchain: an empirical investigation of its scope for improvement (ICBC2019-9001)

Evgeniia Filippova (WU Vienna University of Economics and Business Administration, Austria); Arno Scharl (MODUL University Vienna, Austria); Pavel Filippov (Technical University Vienna, Austria)

Abstract: General Purpose Technologies, or GPTs are defined in the economic literature as the key technologies that shape the economy. Despite the large conceptual literature base on Blockchain potential to revolutionize the current economic system, there is a lack of empirical research on its economic nature and the course of technological development. The paper at hand covers this research gap by providing the quantitative approach aimed at understanding the evolutionary path of Blockchain and its scope for improvement - an acknowledged feature of a GPT - in line with the industrial dynamics and GPT literature. The longitudinal analysis of Blockchain-related patents from PATSTAT and their rule-based classification both from technological and application perspectives is complemented by the study of Blockchain media landscape to provide insights into the social context in which it emerges. The increasing amount of patents related to essential technical issues, such as security, scalability, and usability contribute to wider adoption of the technology, whereas the positive sentiment in the media associated with Blockchain creates beneficial social context for its development. The empirical results advance the claim that Blockchain does show a positive scope for improvement peculiar to the GPTs in the making and, therefore, deserves attention as a technology that will define macroeconomic dynamics in a long term.

9:20 Establishing Standards for Consensus on Blockchains (ICBC2019-9002)

Derek Sorensen (Pyrofex Corporation, USA)

Abstract: We survey seven popular blockchain consensus algorithms, showing that the blockchain community does not have well-established theoretical foundations. In this paper, we consolidate and unify these foundational notions, establishing high quality standards and exploring the comparative relationship between foundations used on these different algorithms. The framework established here is meant to be used by academic and industrial blockchain researchers as a foundation for the theory of consensus.

9:40 Performance Benchmarking and Optimization for Blockchain: A Survey (ICBC2019-9003)

Rui Wang (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China); Kejiang Ye (Chinese Academy of Sciences, China); Chengzhong Xu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China)

Abstract: Blockchain is a decentralized infrastructure widely used in emerging digital cryptocurrencies. With the continuous development of blockchain technology, there are many achievements in different fields of practice and research. As a decentralized, shared and encrypted distributed ledger technology, blockchain has three distinctive features: decentralization, traceability, and non-tampering. Therefore, blockchain technology has transaction autonomy and saves regulatory costs. The advantages of excellent confidentiality and security, its birth was even considered as the fourth industrial revolution. However, the blockchain has only been a low-level technology for a long time, and the progress of the application scenario has not been so smooth. Blockchain has also been questioned because of low performance issues. In order to better understand the development status of the blockchain, we introduced the architecture and consensus protocol of the current mainstream blockchain system, analyzed some open source blockchain benchmarking tools, and summarized some blockchain systems optimization plan. Finally, we have put forward some new ideas on the development of blockchain technology.

10:00 Performance Evaluation of Hyperledger Fabric with Malicious Behavior (ICBC2019-9015)

Shuo Wang (Tsinghua University, China)

Abstract: Hyperledger Fabric is a widely-used permissioned blockchain platform for enterprise consortium applications. It adopts Practical Byzantine Fault Tolerance (PBFT) algorithm as the consensus protocol in its version 0.6. Faulty replicas can intentionally delay messages, be not responsive and send inconsistent messages to different replicas, which significantly undermines the system. However, the existing performance evaluation for Fabric is accomplished in a fault-free environment without malicious behaviors. In this paper, we analyzed the impact of malicious behavior on consensus process, design malicious behavior patterns and test the blockchain performance on Hyperledger Fabric.

Session II: Innovative Services (06/27 Thursday, 13:00-14:00; Compass Ballroom 1)

Chair: Surya Nepal (CSIRO, Australia)
20 minutes for each paper.

13:00 Layered Consensus Mechanism in Consortium Blockchain for Enterprise Services (ICBC2019-9004)

Sheng He and Yishuang Ning (Tsinghua University, China); Huan Chen (Kingdee International Software Group Co., Ltd, China); Chunxiao Xing (Tsinghua University, China); Liang-Jie Zhang (Kingdee International Software Group Co., Ltd, China)

Abstract: The concept of blockchain born out of Bitcoin in just a decade ago has greatly attracted the attentions of industry and academia. The third generation of blockchain is believed to be able to support a large number of commercial and social applications with security, scalability, speed and developer friendly. To maintain a distributed ledger in many independent peer nodes without central authority, the consensus mechanism is the key protocol to construct a blockchain system with multiple technologies. However compared to the well-studied consensus protocols in public blockchain, the customized consensus mechanism is still lack of research for the consortium blockchain in both industry and academia, especially in the enterprise scenarios of applications and services. The layered consensus mechanism in the consortium blockchain is discussed after considering and combining the common characteristics of enterprise applications and services. Based on the famous Service-Oriented Architecture (SOA), the new layered and service-oriented consensus mechanism should be more practical to apply the blockchain technology to much diversified enterprise services with decentralization and extensibility in the principle of enterprise-level security.

13:20 Digital-Physical Parity for Food Fraud Detection (ICBC2019-9005)

Sin Kuang Lo (Data61, CSIRO & University of New South Wales, Australia); Xiwei Xu (Data61-CSIRO & UNSW, Australia); Chen Wang (Data61, CSIRO, Australia); Ingo Weber (Data61/CSIRO, Australia); Paul Rimba (Data61, CSIRO, Australia); Qinghua Lu (CSIRO, Australia); Mark Staples (Data61, CSIRO, Australia)

Abstract: Food fraud has an adverse impact on all stakeholders in the food production and distribution process. Lack of transparency in food supply chains is a strong factor contributing to food fraud. With limited transparency, the insights on food supply chains are fragmented, and every participant has to rely on trusted third parties to assess food quality. Blockchain has been introduced to the food industry to enable transparency and visibility, but it can only protect the integrity of a digital representation of physical food, not the physical food directly. Tagging techniques, like barcodes and QR codes that are used to connect the physical food to its digital representation, are vulnerable to attacks. In this paper, we propose a blockchain-based solution to link physical items, like food, to their digital representations using physical attributes of the item. This solution is generic in its support for different methods to perform the physical checks; as a concrete example, we use machine learning models on visual features of food products, through regular and thermal photos. Furthermore, we use blockchain to introduce a reward system for supply chain participants, which incentivizes honesty and supplying data. We evaluate the technical feasibility of components of this architecture for food fraud detection using a real-world scenario, including machine-learning models for distinguishing between grain-fed and grass-fed beef.

13:40 Blockchain Interoperable Digital Objects (ICBC2019-9006)

Babu Pillai (Griffith University, Australia); Kamanashis Biswas (Australian Catholic University, Australia); Vallipuram Muthukkumarasamy (Griffith University, Australia)

Abstract: The future of distributed ledger technology such as blockchain is dependent on its ability to interact and integrate with other systems. Therefore, interoperability has become a fundamental issue that needs to be addressed. The emerging category of crypto-assets are managed and understood using different frameworks. There is, therefore, a need for a unified classification of crypto-assets. This work aims to bring some clarity to and understanding on interoperable crypto-assets and their characteristics. This paper categorizes digital crypto-assets for the purpose of implementing interoperability. The categorization of crypto-assets are based on their functionalities and their purpose. An interoperability scenario has been given for the defined crypto-asset classes.

Session III: Blockchain Applications (06/29 Saturday, 10:20-11:20; Destination Ballroom 3)

Chair: Liang-Jie Zhang (Kingdee International Software Group Co., Ltd, China)
20 minutes for each paper.

10:20 Patient Privacy and Ownership of Electronic Health Records on a Blockchain (ICBC2019-9007)

Debasish Ray Chawdhuri (Talentica Software (India) Pvt. Ltd., India)

Abstract: Blockchain technology has found application outside of cryptocurrency in recent times. The development of patient-centric storage of medical records on a blockchain has recently gained momentum. However, there have been few developments in providing a solution towards giving up full control to the public so that the concentration of power by only a small group of validators can be avoided. In this paper, we discuss a solution that enables patient-driven interoperability of

medical records on a public blockchain while maintaining privacy using new cryptographic constructs and truly giving up control to the general public.

10:40 Blockchain Federation for Complex Distributed Applications (ICBC2019-9008)

Zhitao Wan, Minqiang Cai, Xianghua Lin and Jinqing Yang (Institute of Advanced Technology Research, Ge Lian Corporation, China)

Abstract: Blockchains are immutable distributed ledger systems usually without a central authority. Blockchains enables people to establish trusted application among untrusted parties. But, the performance of blockchain is a challenge for massive applications. There are many researches to improve the performance of blockchain including side blockchain, interconnection of blockchains. In fact, a distributed application usually needs resource of computing, storage and transportation, even with possible permission access. It means that any current blockchain cannot satisfy all the demands simultaneously. This paper proposes a new blockchain federation that consolidates several blockchains to support complex distributed applications. Two typical application scenarios are implemented following the proposed concept of blockchain federation. The new emerging blockchain technology are combined together to meet the demands of complex peer-to-peer applications. And, the conclusions are drawn and future direction of blockchian federation evolution, federation blockchain, is discussed as well.

11:00 Enriching Smart Contracts with Temporal Aspects (ICBC2019-9009)

Fabiana Fournier and Inna Skrabovsky (IBM Research - Haifa, University of Haifa, Israel)

Abstract: Blockchain technology provides a platform for the decentralized execution of smart contracts. A smart contract is an agreement that is automatically executed when certain conditions are met. Current smart contracts use business rules to express conditions in which transactions occur but lack the capability to reason over time. We present a model-driven approach for creating blockchain applications that can run temporal logic. We present the new concepts required at the model level and explain how these concepts can be translated to a blockchain solution applying Hyperledger Fabric, Hyperledger Composer, and PROTON complex event processing. We illustrate the approach using a pharma cold chain scenario. Cold chain scenarios are typical examples of blockchain cases in which temporal reasoning is required. Here, the blockchain provides transparency, a non-repudiation process, full traceability and trackability of the transactions (provenance), shorter lead times, and authentication and verification of the parties in the supply chain network at each asset transfer. The temporal reasoning is required since decisions about pharmaceutical quality is based on time windows in which the goods have been exposed to certain conditions. Our solution includes the model, the blockchain backend, and a mobile user interface that allows each certified and authorized party in the network to initiate actions, finish their trans-actions, and track progress. It also includes temperature monitoring to ensure compliance with the proper conditions for transportation and asset transfer.

Session IV: Linkable Blockchain (06/28 Friday, 10:40-11:40; Compass Ballroom 1)

Chair: James Joshi (University of Pittsburgh, USA)
20 minutes for each paper.

10:40 CloudAgora: Democratizing the Cloud (ICBC2019-9010)

Katerina Doka, Tasos Bakogiannis, Ioannis Mytilinis and Georgios Goumas (National Technical University of Athens, Greece)

Abstract: In this paper we present CloudAgora, a platform that enables the realization of a democratic and fully decentralized cloud computing market where participating parties enjoy significant advantages: On one hand, cloud consumers have access to low-cost storage and computation without having to blindly trust any central authority. On the other hand, any individual or company, big or small, can potentially serve as cloud provider. Idle resources, be it CPU or disk space, are monetized and offered in competitive fees, regulated by the law of supply and demand. In the heart of the platform lies the blockchain technology, which is used to record commitment policies, publicly verify off-chain services and trigger automatic micropayments. Our prototype is built on top of the Ethereum blockchain and is provided as an open source project.

11:00 Dual Token Blockchain Economy Framework, The Garment Use case (ICBC2019-9011)

Magdi ElMessiry (Alexandria University, USA); Adel ElMessiry (USA & Garment Chain, USA); Malak ElMessiry (Vanderbilt, USA)

Abstract: Over the last few years, the introduction of blockchain technology has brought forth a new set of challenges to understand the acute nature of blockchain token economics and how to best model the underlying business needs. Many Initial Coin Offerings (ICOs) have focused on a single token that facilitated the initial project funding but complicated the actual implementation. Business arrangements can be complex and result in disputes between even the most well-intentioned parties. Supply chain transparency requires careful modeling of a specific implementation of the blockchain technology to correctly capture it. In this work, we propose a dual token model that takes into consideration both the financial aspect and the non-fungible nature of complex, real-world industries. Furthermore, we demonstrate the efficacy of

our model to accommodate the stages hidden under the umbrella term supply chain transparency as it comes to the costing involved in producing the firm's products internal supply chain transparency. We show how this approach not only increases the transparency of the manufacturing process but also the profit and the transparency for external stakeholders.

11:20 An Experimental Evaluation of BFT Protocols for Blockchains (ICBC2019-9012)

Mohammad Jalalzai, Golden G. Richard and Costas Busch (Louisiana State University, USA)

Abstract: Byzantine Fault Tolerant (BFT) protocols have been used in blockchains due to their high performance and fast block acceptance. However, their weakness is on lack of scalability to support large number of nodes in the network due to message demanding broadcasts. There have been recent improvements to the classic Practical Byzantine Fault Tolerant (PBFT) protocol. Evaluating the performance and reliability of the different BFT based protocols in the context of blockchains will give the users a better picture on the behaviour and scalability of these protocols under different circumstances. For this purpose, we implemented and evaluated the performance of different BFT based protocols for blockchain under normal conditions as well as when byzantine failures are encountered in the network. Furthermore, we also calculated the reliability of each protocol under the desired throughput.

Session V: Decentralized Economy (06/26 Wednesday, 14:00-15:00; Compass Ballroom 1)

Chair: James Joshi (University of Pittsburgh, USA)
20 minutes for each paper.

14:00 Developing a Vehicle Networking Platform based on Blockchain Technology (ICBC2019-9013)

Siu-Yeung Cho, Ningyuan Chen and Xiuping Hua (University of Nottingham Ningbo China, China)

Abstract: Blockchain technology has been developing very rapidly through the entire world and is widely seen as one of the most disruptive forces in modern businesses. The applications of blockchain technology to the finance, insurance and automobile industries create extra values in innovation and motivation to the business. This paper is to describe the research and development work called iCarChain which is based on vehicle networking and blockchain technology, through distributed ledger, consensus mechanism, smart contract and other underlying architectures, using open and transparent value acquisition and an artificial intelligence based allocation algorithm to achieve value realization of data on vehicle and driving behavior. Each user (vehicle owner or driver) is referred to install an intelligent ODB device. The device captures the real-time data from vehicles to produce value identification. With such technology, the values of data acquisition and transparent can be transferred through the token carriers called iCarToken which is an incentive unit launched for users to contribute their behavior value, which is generated by blockchain and smart contract, and is gained through the value of contribution behavior.

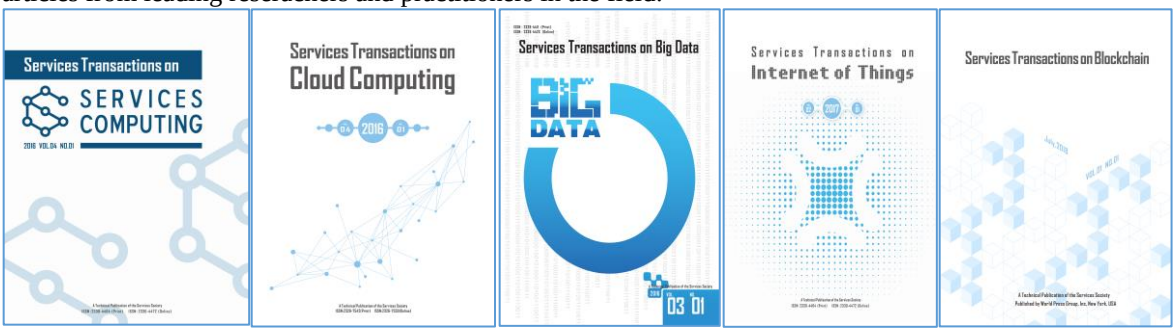
14:20 Decentralized Identity Authentication with Trust Distributed in Blockchain Backbone (ICBC2019-9014)

Jiahe Wang, Songjie Wei and Haozhe Liu (Nanjing University of Science & Technology, China)

Abstract: To overcome the difficulties in traditional identity authentication procedure based on password which leads to high security risk and authentication complexity, and the problems arising from a series of new technologies such as using biometrics as key to cause privacy disclosure, this paper proposes a universal solution to decentralized identity authentication by blockchain combined with trusted execution environment, through a separate hardware to establish a secure area, enabling identity anonymity and avoiding feature disclosure. Meanwhile, smart contract and consensus mechanism are adopted for building trusted identity nodes be-tween the decentralized identity nodes by constructing a traceable identity data structure of blockchain. The evaluation results show that the solution proposed can realize user identity registration, cross-platform authentication, and guarantee security in the whole procedure.

Journals on Services Innovations

The *Services Transactions* series journals only consider extended papers of reputable conferences and invited articles from leading reserachers and practitioners in the field.



Services Transactions on Services Computing (STSC, ISSN 2330-4472)
http://www.hipore.com/stsc
The *Services Transactions on Services Computing (STSC)* covers state-of-the-art technologies and best practices of Services Computing, as well as emerging standards and research topics which would define the future of Services Computing.

Services Transactions on Cloud Computing (STCC, ISSN 2326-7550)
http://www.hipore.com/stcc
The *Services Transactions on Cloud Computing (STCC)* is a refereed, international publication featuring the latest research findings and industry solutions involving all aspects of cloud computing. STCC's major topics include but not limited to architecture, modeling, development, deployment, analytics, optimization, outsourcing, delivery model, consulting, and solutions of Cloud Computing.

Services Transactions on Big Data (STBD, ISSN 2326-442X)
http://www.hipore.com/stbd
The *Services Transactions on Big Data (STBD)* aims to provide the first Public Access publication channel for all authors working in the field of all aspects of Big Data. STBD's major topics include but not limited to architecture, modeling, development, deployment, analytics, optimization, outsourcing, delivery model, consulting, and solutions of big data.

Services Transactions on Internet of Things (STIOT, ISSN 2577-1310)
http://www.hipore.com/stiot
The Services Transactions on Internet of Things (STIOT) aims to be a core resource providing leading technologies, development applications, empirical studies, and future trends in the field of Internet of Things (IOT).

Services Transactions on Blockchain (STBC, ISSN 2577-7238 Print, 2577-7246 Online)
http://www.hipore.com/stbc
The *Services Transactions on Blockchain (STBC)* aims to provide the first Public Access publication channel for all authors working in the field of all aspects of Blockchain. STBC's major topics include but not limited to enabling technologies, applications, and standards of blockchain.

International Journal of Web Services Research (JWSR, EI & SCI indexed)
http://www.igi-global.com/journal/international-journal-web-services-research/1079

International Journal of Business Process Integration and Management (IJBPIM)
http://www.inderscience.com/jhome.php?jcode=ijbpim

Services Society

The Services Society (S2) is a non-profit professional organization that has been created to promote worldwide research and technical collaboration in services innovations among academia and industrial professionals. Its members are volunteers from industry and academia with common interests. S2 is registered in the USA as a "501(c) organization", which means that it is an American tax-exempt nonprofit organization. S2 collaborates with other professional organizations to sponsor or co-sponsor conferences and to promote an effective services curriculum in colleges and universities. The S2 initiates and promotes a "Services University" program worldwide to bridge the gap between industrial needs and university instruction. The Services Society has formed 10 Special Interest Groups (SIGs) to support technology and domain specific professional activities.

The Services Society has been sponsoring various international conferences such as the International Conference on Web Services (ICWS), International Conference on Services Computing (SCC), World Congress on Services (SERVICES), International Conference on Cloud Computing (CLOUD), International Conference on AI and Mobile Services (AIMS), International Congress on Big Data (BigData Congress), International Conference on Internet of Things (ICIOT), International Conference on Cognitive Computiong (ICCC), International Conference on Edge Computing (EDGE) and International Conference on Blockchain (ICBC).

The Services Society is a technical sponsor of a series of international journals such as Services Transactions on Big Data (STBD, http://www.hipore.com/stbd), Services Transactions on Cloud Computing (STCC, http://www.hipore.com/stcc), Services Transactions on Services Computing (STSC, http://www.hipore.com/stsc), and Services Transactions on Blockchain (STBC, http://www.hipore.com/stbc).

S2 also offers volunteer services to support conferences and builds professional communities like special interest groups, local chapters, Services Society Young Scientist Forum (SSYSF). The SSYSFs have been proactive in helping organizing Services Conference Federation (SCF) satellite sessions and other professional activities.

In 2013, the *YouthLancer Foundation* was founded to proof read papers for professors or researchers whose native language is not English. As an organization composed entirely of volunteers, there is no charge for correcting papers.

The Services Society opens to everyone, 7/24!

Introduction of Big Data Certification

http://www.servicessociety.org/

Big Data includes but not limited to Data as A Service (DaaS), massive data, related technologies, and applications. Nowadays, professionals from academia and industry are exploring potential values of big data in various scenarios. New computing platforms such as Cloud Computing, Mobile Internet, and Social Networking are leveraged to create innovations around big data.

Services Society (S2) has been leading the platform of exchanging information and knowledge on the latest service-centric technologies and solutions. In order to address the shortage of Big Data professionals, Services Society led the creation of Body of Knowledge on Big Data (BoK-BD) in 2014.

Based on the first version of BoK-BD, Services Society launched International Certification Program for Big Data professionals (ICP4BD), which includes the following two certification categories.

>>Category 1: Big Data for Job Positions

1.1 Big Data Business Analyst (BDBA)

Key Skills: data warehouse, data minning, machine learning, data analysis algorithm, requirement-driven data modeling, data-driven business decision making

1.2 Big Data Engineer (BDE)

Key Skills: distributed system (e.g. Hadoop, Hive, Storm, Spark), scalable database programming, Web programming, mobile programming

1.3 Big Data Architect (BDA)

Key Skills: data warehouse design, cloud storage, computing architecture (e.g. SOA, Cloud), ETL architecture, distributed system architecture (e.g. Hadoop, Hive, Storm, Spark), system redesign, and independent product design.

Key Requirements for Category 1:

1). 5 years working experience in related areas;
2). Completed related scores (e.g. scores from selected MOCC courses, scores derived from published research journals);

>>Category 2: Big Data for Professional Grade

2.1 Certified Big Data Professional (CBDP)

With this certificate, you can serve as a chief architect, or data scientist in an organization.

Key Requirements for Category 2:

1). 5 years working experience;
2). Completed at least two certificates defined in Category 1;
3). Completed related scores (e.g. scores from selected MOCC courses, scores derived from published research journals);
4). 3 references from CBDP holders or at least 2 references from the program committee of ICWS, SCC, CLOUD, BigData Congress, SERVICES, S2 ICIOT, or editorial board members from the Services Transactions on Big Data, Services Transactions on Cloud Computing, Services Transactions on Services Computing, Services Transactions on Internet of Things, and Services Transactions on Blockchain.

Services Society
Special Interests Groups (SIGs)

The Services Sector has account for 79.5% of the GDP of United States in 2016. The world's most services-oriented economy, with services sectors accounting for more than 90% of GDP. To rapidly respond to the changing economy, the Technical Activities Board at the Services Society (http://ServicesSociety.org/) has created the following 10 Special Interest Groups (SIGs) for our worldwide Services Innovations Community members.

Special Interest Group on Web Services (SIG-WS)

Special Interest Group on Services Computing (SIG-SC)

Special Interest Group on Services Industry (SIG-SI)

Special Interest Group on Big Data (SIG-BD)

Special Interest Group on Cloud Computing (SIG-CLOUD)

Special Interest Group on Artificial Intelligence (SIG-AI)

Special Interest Group on Edge Computing (SIG-EC)

Special Interest Group on Cognitive Computing (SIG-CC)

Special Interest Group on BlockChain (SIG-BC)

Special Interest Group on Internet of Things (SIG-IOT)

Services Confererence Federation

As the founding member of the Services Conference Federation (SCF), the first International Conference on Web Services (ICWS) was held in June 2003 in Las Vegas, USA. Meanwhile, the First International Conference on Web Services - Europe 2003 (ICWS-Europe'03) was held in Germany in Oct, 2003. ICWS-Europe'03 is an extended event of the 2003 International Conference on Web Services (ICWS 2003) in Europe. In 2004, ICWS-Europe was changed to the European Conference on Web Services (ECOWS), which was held at Erfurt, Germany. SCF 2018 was held successfully in Seattle, USA. To celebrate its 17-year-old birthday, SCF 2019 will be held on June 25 - June 30, 2019, San Diego, USA.

The past 16 years, ICWS community has been expanded from Web engineering innovations to scientific research for the whole services industry. The service delivery platforms have been expanded to mobile platforms, Internet of Things, cloud computing, and edge computing. The services ecosystem is gradually enabled, value added, and intelligence embedded through enabling technologies such as big data, artificial intelligence, and cognitive computing. In the coming years, all the transactions with multiple parties involved will be transformed to blockchain.

Based on the technology trends and best practices in the field, the Services Conference Federation (SCF) will continue serving as the conference umbrella's code name for all services-related conferences. SCF 2019 defines the future of New ABCDE (AI, Blockchain, Cloud, BigData & IOT). We are very proud to announce that SCF 2019's 10 co-located theme topic conferences will all center around "services", while each focusing on exploring different themes (web-based services, cloud-based services, Big Data-based services, services innovation lifecycle, AI-driven ubiquitous services, blockchain driven trust service-ecosystems, industry-specific services and applications, and emerging service-oriented technologies).

Some highlights of SCF 2019 are shown below:

- Bigger Platform: The 10 collocated conferences (SCF 2019) get sponsorship from the Services Society which is the world-leading not-for-profits organization (501 c(3)) dedicated for serving more than 30,000 worldwide Services Computing researchers and practitioners. Bigger platform means bigger opportunities to all volunteers, authors and participants. Meanwhile, Springer provides sponsorship to best paper awards and other professional activities. All the 10 conference proceedings of SCF 2019 will be published by Springer and indexed in ISI Conference Proceedings Citation Index (included in Web of Science), Engineering Index EI (Compendex and Inspec databases), DBLP, Google Scholar, IO-Port, MathSciNet, Scopus, and ZBlMath.

- Brighter Future: While celebrating 2019 version of ICWS, SCF 2019 highlights the Second International Conference on Blockchain (ICBC 2019) to build the fundamental infrastructure for enabling secure and trusted services ecosystems. It will also lead our community members to create their own brighter future.

- Better Model: SCF 2019 will continue to leverage the invented Conference Blockchain Model (CBM) to innovate the organizing practices for all the 10 theme conferences. Senior researchers in the field are welcome to submit proposals to serve as CBM Ambassador for an individual conference to start better interactions during your leadership practices for organizing SCF 2019.

====== Member of SCF 2019 ======

Services Conference Federation (SCF) includes 10 service-oriented conferences: ICWS, CLOUD, SCC, BigData Congress, AIMS, SERVICES, ICIOT, EDGE, ICCC and ICBC.

[1] 2019 International Conference on Web Services (ICWS 2019, http://icws.org/2019) is the flagship theme-topic conference for Web-centric services, enabling technologies and applications.

[2] 2019 International Conference on Cloud Computing (CLOUD 2019, http://thecloudcomputing.org/2019) is the flagship theme-topic conference for resource sharing, utility-like usage models, IaaS, PaaS, and SaaS.

[3] 2019 International Congress on Big Data (BigData Congress 2019, http://bigdatacongress.org/2019) is the theme-topic conference for data souring, data processing, data analysis, data-driven decision making, and data-centric applications.

[4] 2019 International Conference on Services Computing (SCC 2019, http://thescc.org/2019) is the flagship theme-topic conference for leveraging the latest computing technologies to design, develop, deploy, operate, manage, modernize, and redesign business services.

[5] 2019 International Conference on AI & Mobile Services (AIMS 2019, http://ai1000.org/2019) is the theme-topic conference for artificial intelligence, neural networks, machine learning, training data sets, AI scenarios, AI delivery channels, and AI supporting infrastructure as well as mobile internet services. The goal of AIMS is to bring AI to any mobile devices and other channels.

[6] 2019 World Congress on Services (SERVICES 2019, http://servicescongress.org/2019) puts its focus on all innovations of services industry that includes financial services, education services, transportation services, energy services, government services, manufacturing services, consulying services, and other industry services.

[7] 2019 International Conference on Cognitive Computing (ICCC 2019, http://thecognitivecomputing.org/2019) puts its focus on leveraging the latest computing technologies to simulate, model, implement, and realizate of cognitive sensing and brain operating systems.

[8] 2019 International Conference on Internet of Things (ICIOT 2019, http://iciot.org/2019) puts its focus on the science, technology, and applications of IOT device innovations as well as IOT services in various solution scenarios.

[9] 2019 International Conference on Edge Computing (EDGE 2019, http://theedgecomputing.org/2019/) is a theme-topic conference for leveraging the latest computing technologies to enable localized device connections, edge gateways, edge applications, edge-cloud interactions, edge-user experiences, and edge business models.

[10] 2019 International Conference on Blockchain (ICBC 2019, http://blockchain1000.org/2019/) concentrates on all aspects of blockchain that includes digital currency, distributed application development, industry-specific blockchains, public blockchains, community blockchains, private blockchains, blockchain-based services, and enabling technologies.

SCF 2019 Satellite Session
(May 30, 2019, Shenzhen, China)

The China Big Data Industry Summit organized by the Services Society is the earliest big data industry summit in China. Since 2013, it had been successfully held in Shenzhen China for six times, becoming one of the top technology and business events in the Internet big data field. Previous conferences attracted a large number of industry elites, authoritative experts in the field of Internet big data, and outstanding entrepreneurs to discuss and exchange Internet big data and related fields, to share the industry's latest trends, trends, and experiences, and had aroused strong reactions in the industry, becoming a platform for practitioners and entrepreneurs in big data related industries to interact and connect with each other.

Conference Room Floor Plan