

Guest Editorial: Special Issue on Data Analytics and Machine Learning for Network and Service Management - Part II

Nur Zincir-Heywood, *Member, IEEE*, Giuliano Casale, *Member, IEEE*, David Carrera, *Member, IEEE*, Lydia Y. Chen, *Senior Member, IEEE*, Amogh Dhamdhere, *Member, IEEE*, Takeru Inoue, *Member, IEEE*, Hanan Lutfiyya, *Senior Member, IEEE*, Taghrid Samak, *Member, IEEE*

I. INTRODUCTION

Network and Service analytics can harness the immense stream of operational data from clouds, to services, to social and communication networks. In the era of big data and connected devices of all varieties, analytics and machine learning have found ways to improve reliability, configuration, performance, fault and security management. In particular, we see a growing trend towards using machine learning, artificial intelligence and data analytics to improve operations and management of information technology services, systems and networks.

Research is therefore needed to understand and improve the potential and suitability of data analytics and machine learning in the context of services, systems and network management. This will provide deeper understanding and better decision making based on largely collected and available operational and service data. It will also present opportunities for improving machine learning and data analytics algorithms and methods on aspects such as reliability, dependability and scalability, as well as demonstrate the benefits of these methods in management and control systems. Moreover, there is an opportunity to define novel platforms that can harness the vast operational data and advanced data analysis algorithms to drive management decisions in networks, data centers, and clouds.

This special issue of IEEE Transactions on Network and Service Management presents novel research tackling the above challenges. It is the fourth special issue in this area to appear in this series, after issues published in [1], [2], [3]. The collection of works we present illustrates recent trends, novel solutions and approaches to leverage Data analytics and Machine Learning in Network and Service management, as well as to extract insights from data that can guide system operators and network managers in their daily activities.

The special issues consists of two parts. In Part II, presented here, we have accepted 23 papers out of 84 papers submitted to the open call for novel contributions addressing the underlying challenges of *Data Analytics and Machine Learning for*

Network and Service management. Part I was published in the previous issue (December 2020 issue [4]).

II. SPECIAL ISSUE OVERVIEW

The special issue papers span three central areas of *Data Analytics and Machine Learning for Management*: (i) Data Analytics and Machine Learning for Network Management, (ii) Data Analytics and Machine Learning for Service Management, and (iii) Advanced Security Management based on Data Analytics and Machine Learning.

A. Data Analytics and Machine Learning for Network Management

Ten papers in this special issue focus on data analytics and machine learning for management of networks.

In "Tensor-Based Recurrent Neural Network and Multi-modal Prediction With Its Applications in Traffic Network Management", Wu *et al.* [item 1] in the Appendix] presents a tensor-based recurrent neural network approach to predict traffic flows. Empirical evaluations on the metro traffic flow dataset demonstrate that the proposed approach can improve the traffic flow prediction accuracy compared to the traditional approaches on the same dataset.

In "Detection and Characterization of Network Anomalies in Large-Scale RTT Time Series", Hou *et al.* [item 2] in the Appendix] propose an unsupervised learning based approach for the detection and characterization of general network anomalies. Then, they analyze the relations between links with state changes and localize the entities that most likely cause the corresponding event.

In "Adaptive Network Latency Prediction from Noisy Measurements", Tripathi and Rajawat [item 3] in the Appendix] explore the estimation and prediction of network latencies from a sequence of noisy and incomplete latency matrices collected over time. Theoretical and empirical evaluations demonstrate the viability of the proposed approach as a network monitoring tool.

In "Host Behavior in Computer Network: One-Year Study", Jirsik and Velan [item 4] in the Appendix] present their study of the host behaviors on a one-year-long real-world network dataset. They inspect the availability of the data for host profiling, identify the temporal patterns in host behavior,

N. Zincir-Heywood is with Dalhousie University, Canada; G. Casale is with Imperial College London, UK; D. Carrera is with Barcelona Supercomputing Center, Spain; Lydia Y. Chen is with Delft University of Technology, Netherlands; A. Dhamdhere is with Amazon Web Services, USA; T. Inoue is with NTT Network Innovation Laboratories, Japan; H. Lutfiyya is with Western University, Canada; T. Samak is with Google, USA.

introduce a method for stable labeling of the hosts, and assess the variability of the host characteristics.

In "Online anomaly detection leveraging stream-based clustering and real-time telemetry", Putina and Rossi [item 5] in the Appendix] benchmark seven unsupervised learning algorithms to the streaming flow of control and data-plane telemetry data with the purpose of real-time anomaly detection. Five of these algorithms are designed for evolving data streams whereas two are not. Their results show that DenStream, one of the five unsupervised learning algorithms designed for evolving data, outperforms the other six.

In "An LSTM Framework for Software-Defined Measurement", Lazaris and Prasanna [item 6] in the Appendix] present a deep learning based framework for scalable software-defined measurement for several network management tasks from traffic engineering to load balancing. Empirical evaluations using real network traces show that the proposed approach outperforms the baselines employed.

In "Intelligent Routing based on Reinforcement Learning for Knowledge-Defined Networking", Casas-Velasco *et al.* [item 7] in the Appendix] introduce a machine learning approach for routing in Software-Defined Networking (SDN), namely Reinforcement Learning and Software-Defined Networking Intelligent Routing. They capitalize on the interaction with the environment using reinforcement learning, and the global view and control of the network provided by Software Defined Networking, to compute and install optimal routes in the forwarding devices, in advance.

In "Burst Traffic Scheduling for Hybrid E/O Switching DCN: An Error Feedback Spiking Neural Network Approach", Yu *et al.* [item 8] in the Appendix] presents a feedback-based spiking neural network approach for high accuracy burst traffic prediction. They then design a prediction-assisted scheduling algorithm to manage the worst-case burst traffic. The simulation results show that the approach can efficiently integrate a spiking neural network into the traffic scheduling scheme.

In "STAD: Spatio-Temporal Anomaly Detection Mechanism for Mobile Network Management", Dridi *et al.* [item 9] in the Appendix] explore a dynamic on-line data mining technique to detect network anomalies allowing operators to pro-actively monitor and control a variety of real-world phenomena. Based on real cellular communication traces, they propose STAD, an automated framework that aims to ensure spatio-temporal detection of outliers using a combination of machine learning techniques.

In "Characterization and Prediction of Mobile-App Traffic using Markov Modeling", Aceto *et al.* [item 10] in the Appendix] seek to analyze publicly available mobile-app traffic, namely MIRAGE-2019, using Markov Chains and Hidden Markov Model learning algorithms. They discuss and empirically evaluate the suitability of the learning algorithms for different network management tasks.

B. Data Analytics and Machine Learning for Service Management

Six papers in this special issue focus on data analytics and machine learning for management of services.

In "QoS Time Series Modeling and Forecasting for Web Services: A Comprehensive Survey", Syu and Wang [item 12] in the Appendix] review and investigate the current web services quality of service time series modeling and forecasting research in the literature. They classify and discuss the current studies in terms of the four components that are identified and provide overall guidelines for the researchers in this area.

In "CEDULE+: Resource Management for Burstable Cloud Instances Using Predictive Analytics", Pincirolu *et al.* [item 12] in the Appendix] present CEDULE+ predictive data analytics to optimize the management of burstable instances in cloud workload variations. They evaluate the proposed system on Amazon EC2 and assess its efficiency and high accuracy through real-case scenarios.

In "Machine Learning-based Scaling Management for Kubernetes Edge Clusters", Toka *et al.* [item 13] in the Appendix] introduce a Kubernetes scaling engine that enables the automatic scaling decision parameters to be set dynamically for managing the variability of incoming requests. This engine uses various machine learning forecast methods that compete with each other via a short-term evaluation loop to suit to the request dynamics.

In "Profit Maximization of Online Service Function Chain Orchestration in an Inter-Datcenter Elastic Optical Network", Yu *et al.* [item 14] in the Appendix] explore online service function chain provisioning in inter-datcenter elastic optical networks. They design and evaluate time-efficient orchestration algorithms for online service function chain requests.

In "Deep-FDA: Mixing Functional Data Analysis and Neural Networks to Characterize Network Services", Perdices *et al.* [item 15] in the Appendix] introduce a deep learning based approach for network service modeling using functional data analysis. They evaluate and demonstrate the applicability of the proposed approach on synthetic and real-world data, and compare to other state-of-the-art alternatives.

In "Mosaic: Advancing User Quality of Experience in 360-degree Video Streaming with Machine Learning", Park *et al.* [item 16] in the Appendix] present a comprehensive approach called Mosaic that combines a neural network based viewport prediction with a rate control mechanism for streaming 360-degree panoramic videos. They provide a comprehensive performance evaluation of Mosaic along with five other streaming techniques.

C. Advanced Security Management based on Data Analytics and Machine Learning

Seven papers in this special issue focus on Advanced Security Management based on Data Analytics and Machine Learning.

In "Detecting Anomalies at a TLD Name Server Based on DNS Traffic Predictions", Madariaga *et al.* [item 17] in the Appendix] propose a near real-time anomaly detection based on prediction approach to detect anomalies in DNS traffic. They show that the proposed approach improves upon the current state-of-the-art anomaly detection in authoritative TLD name servers.

In "Hierarchical Anomaly-Based Detection of Distributed DNS Attacks on Enterprise Networks", Lyu *et al.* [item 18] in

the Appendix] seek to detect distributed DNS attacks using a hierarchical graph structure combined with machine learning. The evaluations are performed on a month worth of DNS data from the two enterprises and the results are compared against blacklists and firewall logs. Results show the ability of the system in detecting distributed attacks while maintaining a reasonable real-time performance.

In "Uncovering Lateral Movement using Authentication Logs", Bian *et al.* [item 19] in the Appendix] employ a machine learning based approach to detect hosts in a network that are targets of an advanced persistent threat attack. They evaluate several machine learning classifiers to detect susceptible hosts in the Los Alamos National Lab dataset.

In "Adaptive Protection of Scientific Backbone Networks using Machine Learning", Mogyorósi *et al.* [item 20] in the Appendix] utilize a machine learning scheme to achieve a backbone protection scheme that periodically re-allocates the unused capacity to meet the service availability requirements. They demonstrate and evaluate their scheme on the real traffic from Energy Sciences Network (ESnet), which is a high-speed, international scientific backbone network.

In "WIDS: An Anomaly Based Intrusion Detection System for Wi-Fi (IEEE 802.11) Protocol", Satam and Hariri [item 21] in the Appendix] introduce a wireless intrusion detection system using an anomaly behavior analysis approach. They represent the normal behavior using n-grams and use machine learning models to classify Wi-Fi traffic flows. The proposed system is evaluated on the University of Arizona and AWID datasets.

In "Comparative Assessment of Process Mining for Supporting IoT Predictive Security", Hemmer *et al.* [item 22] in the Appendix] present the exploitability and performance of a process mining approach for detecting misbehaviors in Internet-of-Things systems. They describe a proof-of-concept prototype security management system, and evaluate it on different industrial datasets.

In "Gradient Boosting Feature Selection with Machine Learning Classifiers for Intrusion Detection on Power Grids", Upadhyay *et al.* [item 23] in the Appendix] present an integrated framework for an intrusion detection system for smart grids which combines feature engineering with machine learning classifiers. They implement and evaluate various decision tree based machine learning techniques after obtaining the most promising features of the power grid dataset.

ACKNOWLEDGMENT

We sincerely thank the authors for contributing their papers and the reviewers for their thorough assessment and their work to improve the quality and presentation of each paper. We are very grateful to the Editor-in-Chief Filip De Turck for his continuous support throughout the process and to Janine Bruttin and Catherine Van Sciver for their help with the administrative tasks associated to this special issue.

APPENDIX

- 1) Q. Wu, Z. Jiang, K. Hong, H. Liu, L. T. Yang and J. Ding, "Tensor-Based Recurrent Neural Network and Multi-modal Prediction With

- Its Applications in Traffic Network Management," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 2) B. Hou, C. Hou, T. Zhou, Z. Cai and F. Liu, "Detection and Characterization of Network Anomalies in Large-Scale RTT Time Series," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 3) R. Tripathi and K. Rajawat, "Adaptive Network Latency Prediction from Noisy Measurements", IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 4) T. Jirsik and P. Velan, "Host Behavior in Computer Network: One-Year Study", IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 5) A. Putina and D. Rossi, "Online anomaly detection leveraging stream-based clustering and real-time telemetry," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 6) A. Lazaris and V. K. Prasanna, "An LSTM Framework for Software-Defined Measurement," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 7) D. M. Casas-Velasco, O. M. C. Rendon and N. L. S. da Fonseca, "Intelligent Routing based on Reinforcement Learning for Software-Defined Networking," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 8) A. Yu, H. Yang, K. K. Nguyen, J. Zhang and M. Cheriet, "Burst Traffic Scheduling for Hybrid E/O Switching DCN: An Error Feedback Spiking Neural Network Approach," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 9) A. Dridi, C. Boucetta, S. E. Hammami, H. Afifi and H. Mounghla, "STAD: Spatio-Temporal Anomaly Detection Mechanism for Mobile Network Management," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 10) G. Aceto, G. Bovenzi, D. Ciuonzo, A. Montieri, V. Persico and A. Pescapé, "Characterization and Prediction of Mobile-App Traffic using Markov Modeling", IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 11) Y. Syu and C. -M. Wang, "QoS Time Series Modeling and Forecasting for Web Services: A Comprehensive Survey," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 12) R. Pincirolì, A. Ali, F. Yan and E. Smirni, "CEDULE+: Resource Management for Burstable Cloud Instances Using Predictive Analytics," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 13) L. Toka, G. Dobreff, B. Fodor and B. Sonkoly, "Machine Learning-based Scaling Management in Kubernetes Edge Clusters," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 14) H. Yu, Z. Chen, G. Sun, X. Du and M. Guizani, "Profit Maximization of Online Service Function Chain Orchestration in an Inter-Datacenter Elastic Optical Network," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 15) D. Perdices, J. E. L. de Vergara and J. Ramos, "Deep-FDA: Using Functional Data Analysis and Neural Networks to Characterize Network Services Time Series," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 16) S. Park, A. Bhattacharya, Z. Yang, S. R. Das and D. Samaras, "Mosaic: Advancing User Quality of Experience in 360-Degree Video Streaming with Machine Learning," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 17) D. Madariaga, J. Madariaga, M. Panza, J. Bustos-Jiménez and B. Bustos, "Detecting Anomalies at a TLD Name Server Based on DNS Traffic Predictions," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 18) M. Lyu, H. H. Gharakheili, C. Russell and V. Sivaraman, "Hierarchical Anomaly-Based Detection of Distributed DNS Attacks on Enterprise Networks," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 19) H. Bian, T. Bai, M. A. Salahuddin, N. Limam, A. A. Daya and R. Boutaba, "Uncovering Lateral Movement using Authentication Logs," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 20) F. Mogyorósi, A. Pašić, R. Cziva, P. Revisnyei, Z. Kenesi and J. Topolcai, "Adaptive Protection of Scientific Backbone Networks using Machine Learning," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 21) P. Satam and S. Hariri, "WIDS: An Anomaly Based Intrusion Detection System for Wi-Fi (IEEE 802.11) Protocol," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 22) A. Hemmer, M. Abderrahim, R. Badonnel, J. François and I. Christment, "Comparative Assessment of Process Mining for Supporting IoT Predictive Security," IEEE Trans. Netw. Service Manag., vol. 18, no. 1, March 2021.
- 23) D. Upadhyay, J. Manero, M. Zaman and S. Sampalli, "Gradient Boosting Feature Selection with Machine Learning Classifiers for Intrusion

REFERENCES

- [1] G. Casale, Y. Diao, H. Lutfiyya, P. Owezarski, and D. Raz, “Guest Editors’ Introduction: Special Issue on Big Data Analytics for Management,” *IEEE Trans. Netw. Service Manag.*, vol. 13, no. 3, Sep. 2016.
- [2] G. Casale, Y. Diao, M. Mellia, R. Ranjan, and N. Zincir-Heywood, “Guest editorial: Special section on advances in big data analytics for management,” *IEEE Trans. Netw. Service Manag.*, vol. 15, no. 1, pp. 10–12, 2018.
- [3] D. Carrera, G. Casale, T. Inoue, H. Lutfiyya, J. Wang, and N. Zincir-Heywood, “Guest editorial: Special section on novel techniques in big data analytics for management,” *IEEE Trans. Netw. Service Manag.*, vol. 16, no. 3, pp. 797–799, 2019.
- [4] N. Zincir-Heywood, G. Casale, D. Carrera, L. Y. Chen, A. Dhamdhere, T. Inoue, H. Lutfiyya, and T. Samak, “Guest editorial: Special section on data analytics and machine learning for network and service management—part i,” *IEEE Transactions on Network and Service Management*, vol. 17, no. 4, pp. 1971–1974, 2020.

BIOGRAPHIES



Nur Zincir-Heywood is a Full Professor of Computer Science with Dalhousie University, Canada. She received the Ph.D. degree in computer science and engineering from Ege University, Turkey, in 1998. Her research interests include machine learning and data mining techniques for network management and cybersecurity, topics on which she has published over 200 fully reviewed papers. She is a recipient of several best paper awards as well as the supervisor for the recipient of the IFIP/IEEE

IM 2013 Best Ph.D. Dissertation Award in Network Management. She received the 2017 Women Leaders in the Digital Economy Award. She is a co-editor of the book “Recent Advances in Computational Intelligence in Defense and Security” by Springer. She is an Associate Editor of the *IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT* and is the General Co-Chair of the 16th International Conference on Network and Service Management 2020. She has been a Co-Organizer for the IEEE/IFIP International Workshop on Analytics for Network and Service Management since 2016.



Giuliano Casale joined the Department of Computing at Imperial College London in 2010, where he is currently a Senior Lecturer in modeling and simulation. Previously, he worked as a scientist at SAP Research UK and as a consultant in the capacity planning industry. He teaches and does research in performance engineering and cloud computing, topics on which he has published more than 130 refereed papers. He has served as program co-chair for several conferences in the area of performance engineering, such as ACM SIGMETRICS/Performance

and IEEE MASCOTS. His research is recipient of multiple awards, recently the best paper award at ACM SIGMETRICS 2017. He serves on the editorial boards of *IEEE TNSM* and *ACM TOMPECS* and as current chair of *ACM SIGMETRICS*.



David Carrera received the MS degree at the Technical University of Catalonia (UPC) in 2002 and his PhD from the same university in 2008. He is an associate professor at the Computer Architecture Department of the UPC. He is also the Head of the “DataCentric Computing” research group at the Barcelona Supercomputing Center (BSC). His research interests are focused on the performance management of data center workloads. In 2015, he was awarded an ERC Starting Grant for the project HiEST and ICREA Academia award, and an ERC

Proof of Concept grant (‘Hi-OMICS’) in 2017 to explore the commercialization of an SDI orchestrator for genomics workloads. He has participated in several EU-funded projects and has led the team at BSC that has developed the Aloja project (aloja.bsc.es) and the servIoTicy platform (servioticy.com). He is the PI for several industrial projects and collaborations with IBM, Microsoft and Cisco among others. He was a summer intern at IBM Watson (Hawthorne, NY) in 2006, and a Visiting Research Scholar at IBM Watson (Yorktown, NY) in 2012. He received an IBM Faculty Award in 2010. He is an IEEE and ACM member.



Lydia Y. Chen Lydia Y. Chen is an Associate Professor in the Department of Computer Science at the Technology University Delft. Prior to joining TU Delft, she was a research staff member at the IBM Zurich Research Lab from 2007 to 2018. She received Ph.D. from the Pennsylvania State University and B.A. from National Taiwan University in 2006, and 2002, respectively. Her research interests center around dependability management, resource allocation and privacy enhancement for large scale data processing systems and services. More specifically, her work focuses on developing stochastic and machine learning models

and applying these techniques to application domains, such as datacenters and AI systems. She has published more than 80 papers in journals, e.g., *IEEE Transactions on Distributed Systems*, *IEEE Transactions on Service Computing*, and conference proceedings, e.g., *INFOCOM*, *Sigmetrics*, *DSN*, and *Eurosys*. She was a co-recipient of the best paper awards at *CCGrid’15* and *eEnergy’15*. She received TU Delft Professor fellowship in 2018. She was program co-chair for Middleware Industry Track 2017 and *IEEE ICAC 2019* and track vice-chair for *ICDCS 2018*. She has served on the editorial boards of *IEEE Transactions on Service Computing* and *IEEE Transactions on Network and Service Management*. She is a senior IEEE member.



Amogh Dhamdhere received the B.E. degree from Mumbai University, Mumbai, India, in 2002, and the Ph.D. degree from the College of Computing, Georgia Institute of Technology, Atlanta, in 2009, both in computer science. Currently, he is a Principal Research Scientist at Amazon Web Services, USA. Before joining to Amazon Web Services, he was a researcher with the Cooperative Association for Internet Data Analysis (CAIDA), USA. His current research focuses on the structure and dynamics of the Internet topology, interdomain traffic characteristics, Internet economics, and the management/troubleshooting of IP networks.



and Communication Engineers (IEICE).

Takeru Inoue (M'04) received the B.E. and M.E. degrees in engineering science and the Ph.D. degree in information science from Kyoto University, Japan, in 1998, 2000, and 2006, respectively. He joined Nippon Telegraph and Telephone Corporation (NTT) Laboratories in 2000 and is now a Senior Researcher. He was an ERATO Researcher with the Japan Science and Technology Agency, from 2011 to 2013. His research interests widely cover algorithmic approaches in computer networks. He is a member of the Institute of Electronics, Information,



Strategic Grants Committee. She was a member of the Computer Science Accreditation Council. She is currently an Associate Editor of the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, and has recently served as the Program Co-Chair for the IEEE/IFIP Network Operations and Management Symposium and the IEEE International Conference on Network and Service Management. She is currently on the steering committee for the Ontario Celebration of Women in Computing Conference.

Hanan Lutfiyya is a Professor and the Chair of the Department of Computer Science, Western University, Canada. Her research interests include Internet of Things, software engineering, selfadaptive and self-managing systems, autonomic computing, monitoring and diagnostics, mobile systems, policies, and clouds. She was a recipient of the UWO Faculty Scholar Award in 2006. She is a past member of the Natural Science and Engineering Research Council of Canada (NSERC) Discovery Grant Committee, and a past member and the Chair of an NSERC



in Computing organization. Dr. Samak holds a doctorate degree in computer science from DePaul University in Chicago, where she worked as a teaching assistant, research assistant and then lecturer. She holds BSc and MSc in computer science from Alexandria University, Egypt and is currently pursuing her Juris Doctorate degree at University of San Francisco school of law. In her free time, Dr. Samak volunteers as a mentor for various women in computing organizations.

Taghrid Samak is a senior data analyst at Google. She applies statistical modeling for a diversity of network applications from capacity planning to wireless networks. Prior to Google, she worked as research scientist at Lawrence Berkeley National Laboratory where her research focused on applying data analysis and machine learning to enable cross-discipline scientific discovery, from modeling application behavior in large-scale systems, to enabling statistical analysis for genomics. She is co-founder and steering committee member of the Arab Women