

# JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

## *Preface*

For the last decades, we witness an explosive growth in the volume, velocity, and variety of the data available on the Internet and other global networks. Terabytes of data originated from multitude of sources including mobile devices, sensors, individual archives, social networks, Internet of Things, enterprises, cameras, software logs, is created on a daily basis. Modern intelligent networks cannot be considered in today's data-aware computing as the systems for such data storage and transmission. Network end-users and devices can also generate the data, such as in the smart city environment, where most of the data is generated by the sensors. Thus, the recent challenging research issues in today's intelligent networking include the effective and optimal management techniques of such a large amount of data and information processing methodologies, as well as new ways to analyze large amounts of data for unlocking information.

One of the most important problem in global and highly distributed networks is a huge traffic generated by billions of connected devices. This devices amount is expected to triple or even quadruple over the next several years as connectivity rapidly expands beyond phones and tablets. The impact on the existing infrastructure will create profound challenges for equipment providers and network operators alike. In such a case, scalability and reliability of the whole infrastructure in massive data processing are the major issues.

Recently, Cloud Computing gives application developers the ability to marshal virtually infinite resources with an option to pay-per-use and as needed, instead of requiring upfront investments in resources that may never be optimally used. Once applications are hosted on cloud resources, users are able to access them from anywhere at any time, using devices ranging from wide class of mobile devices (smartphones, tablets) to desktop computers. The data centre cloud provides virtual centralization of application, computing, and data. While cloud computing optimizes the use of resources, it does not (yet) provide an effective solution hosting the massive data applications.

Despite the above technological advances in intelligent networking and data processing paradigms, large-scale reliable system-software for massive data applications are yet to become commonplace. There are multiple reasons for this state of affairs including: (i) lack of system-software frameworks that allow portability of such applications across multiple distributed data hosts and data centers, (ii) inefficient fault-tolerance mechanism for the improvement of the network reliability and security, (iii) manual approaches leading to

non-optimal hardware and software resource provisioning, data and application management and network monitoring; and (iv) lack of a right set of programming abstractions, which can extend the capability of existing information and data processing paradigms to large-scale dynamic network infrastructures with mobile and remotely accessed services and devices and mobile users.

This issue contains thirteen research papers reporting the recent results on models, solutions, and techniques from a wide area of data-intensive computing and intelligent networking, ranging from conceptual and theoretical developments to advanced technologies and innovative applications and tools.

In the first five papers, the authors present the novel solutions for the improvement of the fault detection and fault-tolerant mechanism in various types of distributed networks and infrastructures. Iacono and Marone in the first paper addressed the problem of Web service selection and integration in Internet. Nita *et al.* developed a Fault Injector Module for cloud systems. Both models are responsible for supporting the cloud developers in the validation of their infrastructure. Nita *et al.* work on CloudSim simulator, where the system faults are generated by using various statistical distributions. In the experimental section, the impact of the statistical distributions on the cloud failures has been analyzed. Careful Web data scanning can improve the efficiency of the malware detection in large-scale networks. Malware can disrupt or even damage the whole infrastructure and is especially dangerous in massive data processing. Kruczkowski and Niewiadomska-Szynkiewicz investigated a novel malware detector based on three supervised learning methods for data mining, namely Support Vector Machine, Naive Bayes and  $k$ -Nearest Neighbors techniques. They evaluated their model on realistic data originated from the devices located in several units, organizations and monitoring systems serviced by CERT Poland. Similar problem has been studied by Conti *et al.* defined the new biometric authentication technology for mobile devices with Android system. They showed that their technology can rapidly increase the security of the data and messages exchanged between the mobile telecommunication network nodes. Finally, Lokesh and Nalini developed the novel fault detection and recovery to achieve fault tolerance mechanisms in Distributed Sensor Network (DSN) by using Bayesian Networks (BNs). It is assumed that sensors in DSN can be located randomly in the dynamic environment. Bayesian Network is used to aid reasoning and decision making under uncertainty of the generated data and resources.

The decision making is a very complex process in massive data processing in large-scale networks. There are several abstract models which can be used for supporting the users and system managers' decisions at different network levels and clusters. Multiagent systems (MAS) for instance is commonly used for simulation the users' behavior and choices, users' interactions and network management components. Mariano and Correia developed game-based model supporting the agents' decisions on the selection of the partners for collaboration based on the reputation parameters. They defined  $n$ -players game for illustration the agent's selection results in the dynamic populations (of the variable sizes), where the agents can interact with each other, reproduce and die. Such MAS model and collaboration behavior can be applied in many data-intensive processes and scenarios, which is confirmed in the paper of Tasquier and Aversa. The authors present the agent-based interaction model in smart house environment. The agents' decisions are made on the optimal (in the sense of energy consumption) selection of the electrical devices based on the current users needs and preferences, and state of the environment.

Scheduling the complex applications in highly distributed computing environments has been usually studied separately from the required data scheduling. Szmajduch developed a novel generic scheduling model in Infrastructure-as-a-Service cloud layer, where both applications and data are scheduled simultaneously in one joint process. The provided simulation results show the importance of considering various scenarios of data transfer to the computational network nodes. In the next paper, Abaev *et al.* studied the delays in message and data processing Session Initiation Protocol (SIP) proxy servers. They modeled such processing in SIP servers by using a specialized G|PH|c queuing system. A comprehensive theoretical analysis showed that measured waiting time and minimum transit time through SIP server can be approximated by acyclic phase-type distributions.

The last part of the issue is composed of four papers on miscellaneous topics related to the wide area of intelligent networking. Le, in the first paper in this part, focused on the optimization of the design of FTTH (Fiber To The Home) access points for popular telecom-

munication networks. El-Bendary and El-Tokhy address the problem of the optimization of retransmission times of Enhanced Data Rate (EDR) packets in Wireless Personal Area Networks. They performed a simple probabilistic analysis for Packet Error Probabilities over the Additive White Gaussian Noise (AWGN) and channels (with Hamming encoding method) for Bluetooth (EDR) packets. The provided experiments show high efficiency of the proposed method in the reduction of the energy utilization in the Bluetooth data transmission. Sudhir and Manvi developed for the improvement of the GPS-based global satellite navigation systems. They designed and implemented a new model of dual antenna GPS-GLONASS navigation receiver. Their receiver improved significantly the performance of the navigation system in critical realistic scenarios (lost of the link, lost of the signal). Finally, Kazakova *et al.* present a short survey of the information recovery system models based on the data generated by the monitoring systems of sensor networks. The paper contains a simple theoretical specification and analysis of the addressed problem.

I strongly believe that this issue will bring all the readers (students, researchers, and industry practitioners interested or currently working in the evolving and interdisciplinary areas of data intensive intelligent networking, and many others) new inspirations for their further research and developments.

I am very grateful to the authors for their hard work and sharing so many interesting ideas and results. My sincere thanks go to all reviewers, who have helped us to ensure the quality of the papers. I thank very much the Journal editorial team for a great support throughout the entire publication process.

Joanna Kołodziej  
Guest Editor  
Institute of Computer Science  
Cracow University of Technology, Cracow, Poland  
E-mail: [jokolodziej@pk.edu.pl](mailto:jokolodziej@pk.edu.pl)  
URL: <http://www.joannakolodziej.org>

