Tutorial T-13: Big Sensed Data: Evolution, Challenges and Practical Directions

Presenters: Hossam Hassanein & Sharief Oteafy (School of Computing, Queen's University, Canada)

Tutorial Overview

The growing realizations of Wireless Sensor Networks (WSNs), sensing over smart devices (tablets, smartphones) and wired sensors, are all generating an exponentially increasing amount of data. The ensuing advent of Big Sensed Data (BSD) is generating critical challenges. First, collected data is mainly insightful to each deployed network, any "sense-making" processes to be built upon heterogeneously collected data faces significant interoperability problems, exposing challenges with varying quality, data-labelling inconsistencies, inaccuracies, time-sensitivities and different reporting granularities. Second, sensing systems inherently adopt a collect-and-report model, whereby collected data is indiscriminately pushed onto the networking infrastructure, regardless of the Quality of Information (QoI) or its value (VoI). Not only do we face scalability issues, but establishing reliable Information Services on top of BSD is not attainable over inconsistently collected, validated and reported data. Thus, the future of Big Data is hampered by the sheer volume of reported data, its uncalibrated discrepancies, and worse by the flood of redundant and lower quality data. Real-time decision making is inherently built on the efficacy of ubiquitous sensing systems, not on the aggregation of devices that are isolated in operation and management. In a time when important applications such as health Informatics and emergency services require rapid and scalable access to contextual information about patients, mobile crowds and the general public, the status quo falls significantly short.

In **Part I** we will present a chronological evolution of WSN advancements, and the core challenges that yielded the primitive views on *de facto* protocols. This will entail discussions on evolution of MAC, routing and cooperative sensing protocols. This brief introduction will facilitate an in-depth discussion of why we ended up with current classical trade-offs in WSN design, and the disparity in the current status quo. This part will delve into the WSN primitives of the IoT, and how different projects and research directions have hindered its realization.

Part II will cover two components. First, we will detail challenges that are causing current bottlenecks in WSNs, especially in terms of scalability, lack of interoperability, and the rise of Big Data and communication-infrastructure ailments. This entails election schemes to reduce and/or aggregate sensor readings, and current efforts to establish Quality of Information (QoI) and Quality of Resource (QoR) metrics to govern the viability of data collected from heterogeneous sensors. Second, we will cover recent directions in literature (both from academia and industry) in addressing the phases of data accessibility, polling, classification, calibration and pruning. This includes incentive schemes to solicit crowd-based data, and engaging the larger public in contributing real-time data instead of pre-deploying WSNs to cover all potential regions of interest. We advocate for severing sense-making processes, especially critical ones with time latency constraints, from pre-deployed architectures. This part of the tutorial will detail the impact of realizing sensor networks on-the-fly and the potential proliferation of information services based on this paradigm.

In **Part III**, we will introduce novel paradigms that promise synergistic operation across WSN architectures, and the premise of building scalable Information Services. Specifically, we will detail recent efforts in standardizing access to data, provisioning of sensor-oriented services, and the

hierarchical naming conventions that will enable access to crowd-solicited/heterogeneous data sources. This encompasses details on standardized formats, formal naming conventions, and the adoption of attribute-based naming systems. In this part we will delve into efforts in realizing Sensing in the Cloud, especially in diverting selective operations and processing on the Edge over Fog Computing architectures. This direction will cover recent directions in multi-tiered Edge offloading, and the integration premise of Cloudlets and end systems with sensing systems. This direction details the potential ubiquity and variable accessibility to sensor networks over the Cloud infrastructure, as a potential parallel to direct accessibility (Via 6LowPAN, BLE and ZigBee). We will also delve into novel models addressing the Smog of Communication caused by the multiplicity of sensors and low-power devices in the era of IoT, and how we can approach an evolutionary view of IoT based on the aggregated capabilities of its encompassed Dynamic WSNs, while reducing the overhead on ensuing latency-aware MAC protocols.

The tutorial plan is as follows:

Part I – Introduction and Evolution of BSD	35 mins
Part II – Challenges in BSD proliferation in the IoT era	40 mins
Break	20 mins
Part II – Practical directions for Surviving the Big Data tide	30 mins
Part III – Building Information Services over Fog-Sensing Architectures	35 mins
Questions and Discussion	20 mins

Presenter Biographies

Dr. Hossam Hassanein is a leading authority in the areas of broadband, wireless and mobile networks architecture, protocols, control and performance evaluation. His record spans more than 500 publications in journals, conferences and book chapters, in addition to numerous keynotes and plenary talks in flagship venues. Dr. Hassanein has received several recognitions and best papers awards at top international conferences. He is the founder and director of the Telecommunications Research Lab (TRL) at Queen's University School of Computing, with extensive international academic and industrial collaborations. Dr. Hassanein is a senior member of the IEEE, and is a former chair of the IEEE Communication Society Technical Committee on Ad hoc and Sensor Networks (TC AHSN). He has delivered numerous tutorials and keynote talks on Internet of Things and its technologies. Dr. Hassanein is an IEEE Communications Society Distinguished Speaker (Distinguished Lecturer 2008-2010).

Dr. Sharief Oteafy is an Adjunct Assistant Professor at Queen's School of Computing, where he obtained his PhD in 2013. His research interests span novel paradigms in sensor networks, advancing core technologies in the Internet of Things and synergy efforts in next generation context-aware sensing systems and content-centric networks. He co-authored a book on "Dynamic Wireless Sensor Networks", and over 35 referred Journal and Conference papers. During his PhD studies at Queen's, he spearheaded the establishment of the Graduate Computing Society (GCS), and chaired the first GCS-conference in 2010. Dr. Oteafy has chaired a number of workshops at IEEE ICC and IEEE LCN, and is actively engaged in the IEEE Communications society, and an IEEE, ACM and SIAM member since 2008. Dr. Oteafy is actively engaged in both the AHSN and WC Technical committees at ComSoc conferences.