Big Sensed Data: Evolution, Challenges, and a Progressive Framework

Sharief M.A. Oteafy and Hossam S. Hassanein

The rise of sensor deployments, the uptake of the IoT, and new manifestations of sensing systems have resulted in a tide of sensed data that is potentially drowning our communication resources and hindering big data analytics with superfluous data. The authors argue that efficient management of IoT systems in smart communities and cities lies not in sensing systems alone, but in the expedited funneling and processing of data as we attempt to prune the unnecessary and build on the valuable.

Abstract

The rise of sensor deployments, uptake of the Internet of Things (IoT), and new manifestations of sensing systems (e.g., crowd sensing, M2M-driven sensing, cloud sensing) has resulted in a tide of sensed data that is potentially drowning our communication resources and hindering big data analytics with superfluous data. We argue that efficient management of IoT systems in smart communities and cities lies not in sensing systems alone, but in the expedited funneling and processing of data as we attempt to prune the unnecessary and build on the valuable. The quest for energy efficiency that dominated sensor networks for so long is now matched with a more pressing demand for access ubiquity and real-time operation. We highlight how big data became a challenge in sensing systems, then elaborate on the status quo in managing this challenge under different research umbrellas. We draw upon three planes that encompass current and future developments for the management of big sensed data (BSD), namely resources, data, and information planes, detailing their pertinent challenges and how evolving solutions can streamline their contributions in light of others. We conclude by highlighting core challenges rising across these three planes, and potential solutions to addressing synergy in coping and scaling with BSD.

Understanding Big Sensed Data

The generation of sensed data is growing at unprecedented levels, raising multi-faceted challenges in how underlying communication infrastructures can cope with this growth, and how services could build on top of this heterogeneous wave of data. Specifically, data generated over diverse sensing systems, Internet of Things/machine-to-machine (IoT/M2M) systems, and wireless sensor networks (WSNs) are increasingly disparate in quality and value.

We highlight the growing challenge of handling sensed data flows, and potential directions in enabling real-time sense making services over such data after it is fused and quantified. This is not to undermine the challenge of growing data size under traditional big data research, but to elicit the elastic mandates of collecting, aggregating, reporting, pruning, and communicating *sensed data* in a scalable and ubiquitous framework.

Sensing systems extend far beyond the traditional view of WSNs. With the rapid increase in IoT devices that are capable of sensing and reporting, many sensing architectures are built on devices with abundant resources (e.g., wearable devices, vehicles). IoT and M2M systems alone are projected to dominate 46 percent of global connections and generate 6.3 exabytes by 2020 [1].

Today, the proliferation of sensing systems are spawning many sensing paradigms, which include public sensing, participatory sensing, and cloud-based sensing, to name a few. They mainly differ in how we engage users in sensing, that is, by active participation to collect specific data, or by mandating a group of devices to passively sense and report data. An interesting overview of these systems is presented in [2]. Cloud-based sensing systems leverage cloud access to report data for (mostly) offline processing and queries, whereas public/participatory sensing systems typically build on proprietary networks with varying connectivity mandates.

A major hindrance in most of these sensing systems lies in the inherent framework: crowd-solicited devices seldom enable real-time access to data, as they rely on participation from users. More critically, the fidelity, trust, and accuracy of data is always marred by the fact that it is "publicly" solicited, except for scenarios when the application mandates the use of specific (pre-calibrated) sensors [3]. Moreover, the utility and uptake of these sensing systems is challenged by privacy preservation for contributing devices, as well as security concerns.

Emerging research in establishing crowd-attributed trust levels and verifying data contributions across trusted vs new users is pushing the envelope in these sensing systems. However, the main challenge in these systems lies in interoperability, in addition to competing for the sensing resources/devices that would "adopt" their service architecture.

In light of all these islandic systems, we define big sensed data (BSD) as the exponential growth of data collected from heterogeneous data sources, and the ensuing challenges in sense-making applications. That is, BSD is manifested in the evolution of sensing systems, in both sensing resources and data produced, that are presenting critical challenges in heterogeneous resource management and adaptive frameworks for operation across sensing systems. BSD is a growing

Digital Object Identifier:
10.1109/MCOM.2018.1700557

Sharief M. A. Oteafy is with DePaul University; Hossam S. Hassanein is with Queen's University.

 0163-6804/18/$25.00 © 2018 IEEE

phenomenon that transcends networking and storage primitives, to more pressing challenges in building applications and services on top of heterogeneously produced data.

Open Challenges in BSD

As data trickles through the network from sensors all the way to services, many intermediate architectures are involved, raising issues of redundancy, pruning discrepancies, data loss, intermediate storage and accessibility, lossy fusion, as well as access rights. Ultimately, we need to carefully calibrate the quality of information (QoI) that results from these sensors to better serve applications and services built on top of this information.

In coping with the BSD in smart cities, there are multiple challenges to be addressed. We hereby present the core challenges, and then overview the status quo in managing BSD under different research domains. This brief survey leads to a discussion on the three planes that encompass BSD research. This tri-plane view presents a layered approach to understanding current voids in BSD management, and guides potential directions under each plane in light of cross-plane and intra-plane interactions.

Where Shall We Process?

As we evolve into BSD, the locality of data processing (aggregation, fusion, sense-making processes, and service composition) are highly application-dependent, and directly impact our access and transport of data. The rise of cloud/fog computing infrastructures [4] now enable data processing much closer to the region of interest, thereby retaining more context for pruning and fusion, thus reducing traffic load and the ensuing big data impact. The rise of cloudlet access to facilitate near-field offloading to cloud services, along with recent research on edge analytics and edge computing to support IoT services and architectures [5], present unique frameworks for leveraging BSD proliferation in decentralized operation.

Funneling Effect (Hierarchical Sensor Fusion)

Big data has to be tackled closer to the source, with proactive mechanisms for fusion and pruning [6] aimed at reducing irrelevant/inferior content to relieve the network backbone. This directly impacts BSD scalability and access time for information by ubiquitous services, and aids rapid distribution of content for real-time sense making processes, especially for emergency response systems. However, heterogeneous architectures inherently introduce hierarchical frameworks of operation, whereby not all data sources are directly accessible, and much of the context is only visible on partial layers in these systems. Therefore, a core challenge lies in addressing real-time hierarchical sensor fusion in BSD, with meta-tagging of information and ensuring loss-less context fusion.

Regulating Flow across Sensing Systems

A critical challenge lies in naming and identifying data flows, based on content, source, destination service/application, or a combination of these factors. Ultimately, the quest for data calibration at each stage — across sensing systems — raises compatibility and interoperability challenges, especially as the end user is only interested in quality data under clear service level agreements (SLAs), despite the heterogeneity of underlying infrastructure(s).

Control flow is another challenge in realizing BSD, especially as we envision real-time feedback cycles to control the flow of data and interplay of resources across BSD planes. A major research challenge lies in regulating the policies and access management frameworks to probe, enlist, calibrate, fuse, valuate, and compensate resources in BSD, while managing their operational mandates in decentralized schemes.

A critical challenge lies in naming and identifying data flows, based on content, source, destination service/application, or a combination of these factors. Ultimately, the quest for data calibration at each stage — across sensing systems — raises compatibility and interoperability challenges.

Coping with Heterogeneous Data Sources

As we diversify the sensing systems that provide data in any given environment, we are faced with the challenge of heterogeneous data sources that often conflict, and mandate rigorous measures for weeding out false positives and anomalies. That is, establishing adaptive measures for the quality of data (QoD) raises multiple challenges. This entails gauging QoD thresholds, especially as some applications mandate lower QoD than others. Also, a challenge lies in defining the metrics that affect QoD. If QoD is confined to "expected" reports based on a given neighborhood, how does that affect detecting true-positive anomalies? This yields multiple challenges in defining context-aware QoD, which reacts to variations in the environment, and is not simply mandated by the services.

Trust, Privacy, and Security Barriers

A foundational challenge in soliciting crowd data is establishing trust in the data we collect, ensuring the privacy of those who provide it (even via seamless/passive sensing), and preventing security breaches of the collected data and its providers. Recent advances in cloud-based security and IoT security are picking up pace [7], such as the adoption of the Jasper system by Cisco for their IoT platform, where connection security and data encryption drive IoT communication.

However, there are core challenges with scaling these systems over BSD resources. First, as more heterogeneous and multi-proprietary devices report data, there is a question of resourcefulness to carry out data encryption and ensure secure communication in relaying that data. More critically, a growing need for distributed policy management should address access rights and soliciting (potentially third-party) components to anonymize data and still maintain links to the producing resources to enable quality of resource (QoR)-based calibration and trust management [2]. We argue for incorporating degrees of trust, encryption, secure communication, and anonymization in the QoR calibration meta-data to empower receiving services in selecting the data/resources that meet its required SLAs.

Enabling BSD Information Services

Most information services are limited by the resources deployed at the design stage. A new service typically mandates deploying its own new sensing resources. To provide truly scalable information services, the underlying sens-

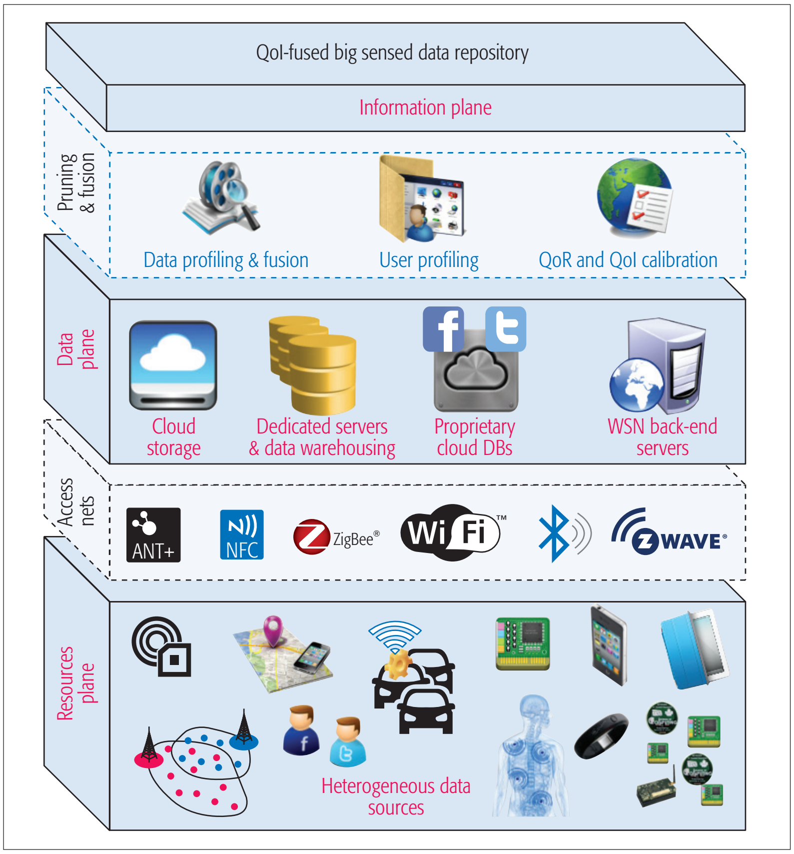

Figure 1. Big sensed data planes and intermediate infrastructures.

ing architecture should inherently feed data into a scalable information repository on which services can be built. Thus, a new service might only require augmenting current resources with those that suffice to meet the new service. This opens the door for simple services (requiring only specific data feeds) and even allows for the orchestration of services, which are built on data feeds from raw sensors as well as other services. For example, a simple heat monitoring service would mandate feeds from distributed temperature sensors, but an advanced weather prediction system would require a combination of underlying services to yield a comprehensive weather information service.

The Status Quo

Many of the aforementioned challenges have been disparately addressed under different research umbrellas. At the heart of data processing lies a significant body of research on big data analytics. While we are focused on BSD from the data generation point of view, it is important to understand the machine learning tools (e.g., support vector machines and *K*-nearest neighbors) that are heavily utilized in big data analytics, and how they are used to prune and analyze big data streams. Moreover, significant efforts are being invested in big data analytics solutions, such as Hadoop, Spark, and MapReduce, to handle the increasing influx of data streams that require cleaning and sense-making. To summarize, Table 1 overviews major research areas that share common challenges with BSD management, and which of the three planes may yield insights on their development. This primarily contrasts existing research directions against a holistic view of BSD management.

A Tri-Plane Approach to BSD

Our quest for scaling with BSD lies not in managing and pruning individual flows of data from all sensing systems, but inherently in synergizing them to pool data in manageable and non-redundant flows. All of our requirements for granularity, time latency, accuracy, and reliability should form a rigorous scale for *tagging* reported data and pruning data flows closer to sources.

The spectrum of challenges lend themselves to three distinct planes, as they mandate minimal control across the planes, and more autonomous operation and management in every plane. That involves viewing BSD challenges in order of soliciting data from heterogeneous resources (*resources plane*), identifying the quality and usability of data on a scalable *data plane*, and quantifying the quality and value of information polled from all these resources toward a scalable *information plane*. Such a tri-plane framework is depicted in Fig. 1.

Resources in the bottom plane are probed via a number of access schemes (e.g., NFC for very short-range communication, Bluetooth and ZigBee for short-range, and WiFi for long-range), which then provide crude data to the data plane.

BSD challenge	Addressed in research under:	Contrasting solution contexts	Potential synergy with BSD planes
Processing tiers	Cloud/edge computing	Attempts to balance load, capacity, and communication cost	Insights into information and data planes from offloading thresholds
Data processing from source to service	Stream computing	Focused on allocating computing resources to carry out hierarchical data analytics	Data and information planes
Heterogeneous data sources	Multi-modal fusion	Focused on data alignment and event classification from multiple feeds about a given event	Resource and data planes when fusing single-stream data
Trust, privacy, and security	Access control in distributed systems	Focused on policy management and updates in distributed environments	Resource plane as data is produced and pruned by access rights
Information services	Ubiquitous service discovery	Mostly focused on matching of service requests with providers under cost and quality constraints	Data and information planes, as they establish real-time SLA negotiation

Table 1. Overview of research areas addressing similar or complementary challenges to BSD.

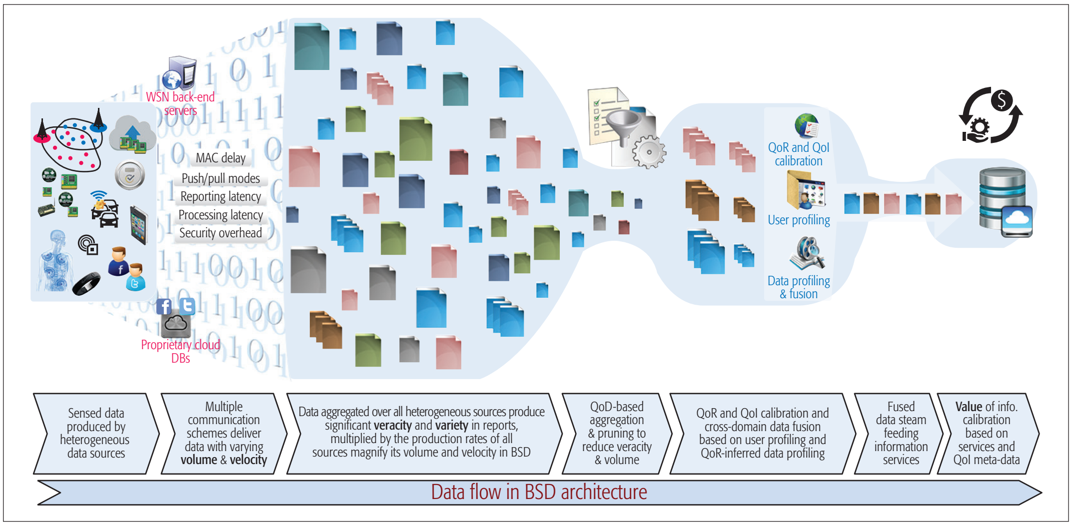

Figure 2. Data flow from resources to information services in BSD, highlighting the unique impact of coping with sensed data from heterogeneous resources in smart sensing systems, and the pruning, calibration, profiling, and fusion stages that introduce bottlenecks for selective data propagation in BSD toward the final valuation of data for information services in smart cities.

At this stage, hierarchical fusion techniques take place in light of reported resource attributes, and across different data repositories that collect data from individual sensing systems and heterogeneous data sources. Finally, a more broad context-based analysis of QoD and QoI aid in fine-tuned data fusion to feed into the information plane. We envision that services will run on top of the information plane having mandated specific thresholds for requested QoI indicators.

In Fig. 1 we highlight the intermediate infrastructures required to communicate the data (access networks), and carry out data fusion and pruning, before we can realize ubiquitous QoI-BSD repositories. This flow of data is presented in Fig. 2. The remainder of this article details these three planes, and then concludes with major challenges in the quest to address and build upon BSD. Under each plane, we delve into the core challenges that face ubiquitous BSD scalability, and address the interplay of these challenges in light of current research directions and future outlook.

Resources Plane

To appreciate the scope of BSD, we must encompass all sensed data sources. We broadly refer to all data producers as resources, which could be physical (e.g., sensors) or virtual (e.g., smart meter status reports or Twitter-based feeds). Understanding the capacity and attributes of a data source is pivotal to determining the QoD it produces. This spans a spectrum of attributes that dictate the QoR and its usability. Generally, each resource must be identified by its location, functional capacity, operation levels, temporal availability, access rights, energy consumption for every operation level, and ultimately its region of fidelity.

Capitalizing on Heterogeneous Resources

The heterogeneity of resources feeding into BSD presents a double-edged reality. On one hand, we can capitalize on diverse resources to provide sensed data, with varying qualities and in different contexts (e.g. angle of view, type of hardware used, trust level in resource). However, their diversity presents a challenge in calibrating the QoR to aid resilient sense-making processes. For example, a road monitoring application may only wish to rely on resources "verified" by the Department of Transportation, or a surveillance system may wish to exclude cameras that produce inferior images. This results in most sensing systems depending on proprietary resources that are calibrated to the desired application.

A scalable BSD ecosystem must capitalize on the heterogeneity of resources deployed, rather than mitigate variability in data. That is, instead of aiming to deploy new resources to yield the sensed data required, we should investigate the resources already in place and attempt to calibrate their QoR in light of our sensing requirements. This view of how sensing systems should thrive and grow in collaboration and synergy is detailed in our earlier work [8].

A scalable BSD ecosystem must capitalize on the heterogeneity of resources deployed, rather than mitigate variability in data. That is, instead of aiming to deploy new resources to yield the sensed data required, we should investigate the resources already in place and attempt to calibrate their QoR in light of our sensing requirements.

Variability vs. Redundancy

As we deploy more resources, we face the challenge of resource redundancy across sensing systems. We address the abundance of resources deployed under one paradigm (overlapping WSNs) and across sensing systems (e.g., WSNs co-located with crowd-sensing systems). The notion of redundancy is only truly captured when we abstract the definition of a sensing resource and detail its attributes.

A core challenge lies in rigorous attribute models for establishing the QoR deployed/probed in a given region and their yielded QoD. Then we can uniformly assess the QoR across their variability, and decide if we indeed need to introduce redundant resources to achieve new functional gains. Variability refers to the new set of attributes a resource demonstrates in variance to existing resources. For example, a new resource with the same functional transceiver (e.g., ZigBee) but with extended range or less energy consumption introduces variability in the functional spectrum of communication [9].

We dissect data representation into two distinct phases: its form as reported directly from the data sources and its representation on back-end systems. It is important to distinguish between both phases as they typically reside on different resources, and affect ensuing calibration and fusion mechanisms.

Quest for Uniform QoR Measurement

Establishing a uniform scale for expressing the attributes of current resources, and gauging the QoRs and QoD in a given region of interest, are critical research directions. Existing quality of service (QoS) definitions in networks have focused on perceived measures of performance, including metrics such as packet loss rates, jitter, transfer bit rates, and delays in communication. Newer models, such as the QoXphere [10], are more inclusive of user-centric views of quality.

In the context of BSD, it is important to build on intrinsic performance metrics of network components, in addition to hardware profiles of sensing resources, to provide a uniform and robust attribute-based representation of QoR. This representation must aid in rapid calibration of resources to judge both the QoR and resulting QoD as it feeds into the BSD cycle. In Fig. 1, this is not only confined to the resources plane, but must also be incorporated in fusion frameworks prior to pushing the data to BSD data repositories (the top level).

Data Plane: Sensing vs. Datafication

We broadly classify data sources into two categories: explicit sensing and "datafication" of the environment. A phenomenon, an anomaly, and sheer sampling of physical properties in a given environment are all triggers for sensors. In the former category, we aim to improve our capacity to detect such triggers, and report them under mandates of accuracy, timeliness, localization, and often the pertinent energy footprint.

Datafication, in its broader sense, spans capturing data on everyday processes that are not necessarily part of a physical phenomenon [11]. This includes M2M control and maintenance information, data flows from social network feeds (e.g., trends and popularity scores of content), and reports from sensors in wearable technologies. In fact, recent market surveys are citing wearables as the IoT domain of highest market penetration, with the spectrum of e-health growing to include many promising directions [12]. Such user-centric data offers great insights in applications that span epidemiology, crowd behavior, commute patterns, and general societal well being.

Data Representation

A core challenge in BSD is the inherent disparity in data representation. As sensing systems evolved, each design entailed its own representation for data. We dissect data representation into two distinct phases: its form as reported directly from the data sources (e.g., sensor node, smart device) and its representation on back-end systems (e.g., sink, cloud/server). It is important to distinguish between both phases as they typically reside on different resources, and affect ensuing calibration and fusion mechanisms.

The first phase is crucial to data fusion schemes; for example, simple WSNs aimed at minimizing packet sizes to reduce transmission time. Thus, much of the context of each sensed report had to be inferred by the receiving node(s), or ultimately at the sink. Any analytics mandate more information to be passed on by the originating device, hence incurring energy and processing load on the path to the source device. The gain was demonstrated early on in [13], where the need for aggregation as a means of reducing total communication in a network was important. Today, the sheer volume of redundant data is witnessed across deployments, even if individual networks implement fusion and/or aggregation.

The second phase, arguably more critical in BSD, is how we represent the data in each repository. Any effort to fuse data and enable sense-making services over multiple data sources mandates consistency in data representation. In this direction we could build on recent efforts in information-centric networks, which addressed naming consistency [14], especially as we scale with multiple producers in a paradigm that is content-centric.

QoD Calibration

Most efforts in measuring the quality of sensed data are confined to homogenous networks. This includes data filtering, detecting anomalies, faulty nodes, and cross-validation across nodes in a given region; that is, those expected to report similar results. As data is aggregated toward collection points (e.g., sink/base station/cloudlet), fusion techniques mostly entail direct averaging of data after excluding outliers, or in more recent research adopting a fuzzy logic approach [6] to assign different weights to "more trusted" reports (e.g., by anchor nodes, or in correlation with a fitness function pertinent to remaining energy or reporting history).

A measure of quality is thus attributed to the resources themselves and the data they produce in contrast to neighboring nodes. However, this method is heavily dependent on the assumption of homogeneity. More importantly, it often offloads most of the data calibration task to the sink, which is both a waste of resources (to communicate low-quality data) and risks losing the context of original data points. In addition, as we head into heterogeneous systems that tap directly (or even via sinks) to sensor nodes, this decision should be made closer to the sensing region to ensure that only high-quality data is allowed to traverse the network; low-quality nodes can be duty cycled to conserve power and reduce medium contention, and heterogeneous cross-validation is enabled to improve QoD calibration closer to the source.

Data Fusion Granularities

Data fusion involves combining data from multiple sources to reach a decision. Many definitions arise to specify the multiplicity of sensor inputs, their heterogeneity, and whether or not it depends on a priori knowledge about the context of each data source [15]. In the context of sensing systems, sensor fusion has almost always been a quest in solidifying partial information from multiple sources, in contrast to aggregation schemes that assume full data but aim to reduce its size.

As BSD scales with homogeneous and heterogeneous networks, encompassing low-end nodes that have very limited resources, hierarchical fusion rises as another challenge. It is necessary to explore fusion as a factor of the underlying resources, and in correlation with the vitality of

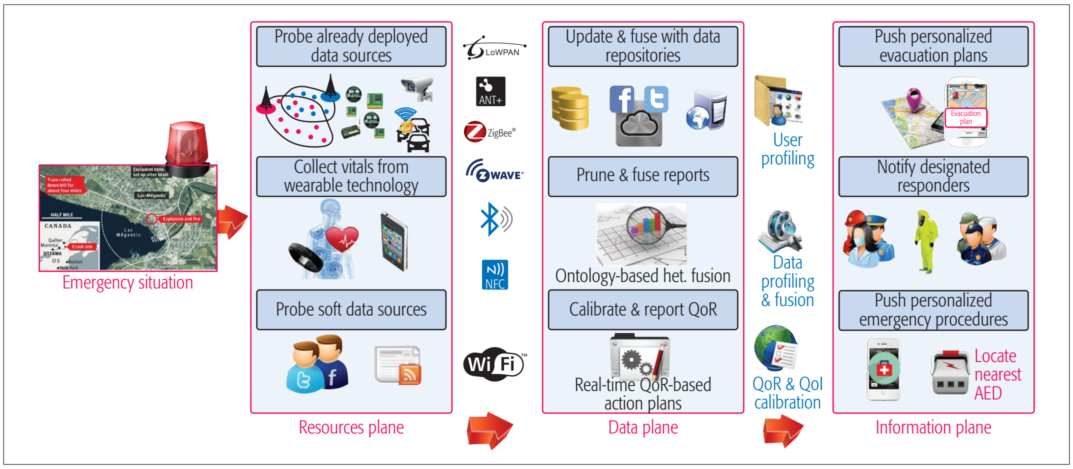

Figure 3. An instance of the BSD framework reacting to an emergency situation in smart cities.

data. Simply put, often simple averaging would suffice, especially in resource-constrained sensors. The potential for QoD-based fusion with a feedback cycle in BSD is a core challenge, and underlines multiple directions of research, most prominently addressing the viability of introducing an active pruning mechanism to silence nodes and prune data along forwarding paths, and introducing reactive systems to calibrate and update QoD measures in real time based on data context and ensuing ontologies.

Information Plane

At the information plane, it is vital to balance QoI based on QoD metrics, and on SLA mandates for services built on such data. This is challenging, especially as the demand grows for ubiquitous services. Figure 2 presents a thorough view of data flow toward BSD information services.

Information services are coupled with data; thus, mandates for QoI and monetary exchange often come into play. The more sources, and the better the QoD and ensuing QoI, the better the service. Hence, a foundational block to establishing ubiquitous information services is active calibration of the QoI feeds.

The challenge in establishing QoI metrics with BSD lies in heterogeneity as well as context variation. Even if rigorous measures of QoI are in place to account for QoR and QoD, at given times (e.g., in emergency situations) we are interested in all data, rather than only ones that meet a preset threshold of QoI. A scale has to be established that uniformly gauges QoI across heterogeneous resources, feeds, and ultimately providers. However, different SLAs might couple QoI calibration with data feeds, which compromises the goal of BSD in uniformly quantifying QoD and QoI across all resources and data flows. Although this has been partially addressed in ontology-based data integration, the challenge remains open in BSD frameworks.

Information processing is highly dependent on spatial and temporal properties of data, and must be captured in the evaluation of QoI. For example, a redundant resource becomes more viable at times of high variation in reported data, at locations where fewer data sources are currently viable, or in emergency situations. Thus, establishing QoI measures must adapt to the context.

BSD Proof of Concept

To demonstrate the interactions in a BSD framework, Fig. 3 depicts a proof-of-concept emergency situation and the ensuing BSD reaction phases across the aforementioned planes. An emergency situation triggers the BSD framework to engage all viable resources in the region of concern, and via different access networks (e.g., 6LowPan and ZigBee for short-range sensors, NFC and BLE for wearable sensors, and WiFi and LTE for web-based access and long-range reporting). All collected data will be pushed onto data warehousing mechanisms that will store and fuse data with current repositories, which will all feed into pruning and fusion mechanisms that adopt ontology-based heterogeneous fusion. This forms the basis of all real-time QoR calibration and action plans (e.g., silencing faulty/redundant resources or waking up/soliciting dormant resource as needed).

The ensuing QoI and profiling mechanisms will be triggered to evaluate the viability of information before triggering and feeding into emergency services, which span notifying only the needed first responders, creating personalized evacuation plans, notifying those in need of nearby devices to use (e.g., automatic defibrillators), and pushing tailored emergency procedures (e.g., tending to a personal chronic condition triggered by the event).

BSD Outlook and Conclusions

We argue for a shift in focus from application-specific and tailored sensing applications, to a broader view of synergistic sensing systems. Specifically, this means targeting the growing demand for establishing quality indicators for data that potentiate its usability for real-time services which require scrutiny of their data sources. This is most prevalent in sense-making systems, especially in light of IoT developments and proliferation. The islandic development of current sensing systems is presenting critical challenges in handling the spur of BSD.

The future lies in coping with (and exploiting) existing resources to establish and benchmark QoRs, anchor validation parameters, establish trust in resources, and reduce the footprint of new applications by capitalizing on resources already there instead of deploying new ones.

Information services are coupled with data; thus, mandates for QoI and monetary exchange often come into play. The more sources, and the better the QoD and ensuing QoI, the better the service. Hence, a foundational block to establishing ubiquitous information services is active calibration of the QoI feeds.

The future lies in coping with (and exploiting) existing resources to establish and benchmark QoRs, anchor validation parameters, establish trust in resources, and reduce the footprint of new applications by capitalizing on resources already there instead of deploying new ones.

True ubiquity in sensing infrastructures is inherently coupled with crowd support and poses game changing designs in BSD. The case for incentivizing smart devices and crowds [2] to feed data into the BSD architecture is both critical and difficult. For one, there must be a monetary/bartered value associated with reported data, correlated with its quality, timeliness, accuracy, and even reputation, in light of the available hardware.

Ultimately, the end user might decide on the acceptable QoI thresholds for input data, whether or not this is a design function of the service in question. The user could actively include/exclude certain data sources/flows that fail a given threshold, or the service/application can pursue calibration without user involvement, depending on current context, user profile, and other settings. At some level in BSD, user-centric QoI calibration must take place, and this remains an open challenge, not necessarily because it is new, but mostly because it has thus far been application-dependent.

In adopting a scalable BSD framework, the anticipated leverage in economic, technological, and societal impacts are immense. To name a few, we will be able to reduce the cost of deployments, maintenance, and communication of devices to only those required to augment existing resources, which will inevitably improve the feasibility of many information services as they would require a lower starting budget and upkeep. More importantly, other crowd-based services will be better incentivized to take part in more important sensing applications (e.g., reporting hazardous situations) as they enlist their resources (e.g., smartphones) in the resource pool, and provide rapid dissemination of information before municipal services can even reach the area of concern.

References

[1] Cisco Visual Networking Index, "Global Mobile Data Traffic Forecast Update 2015–2020 White Paper," Feb. 2016; http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html (2016), accessed June 2016.

[2] H. Ma, D. Zhao, and P. Yuan, "Opportunities in Mobile Crowd Sensing," *IEEE Commun. Mag.*, vol. 52, no. 8, Aug. 2014, pp. 29–35.

[3] Z. Sanaei *et al.*, "Heterogeneity in Mobile Cloud Computing: Taxonomy and Open Challenges," *IEEE Commun. Surveys & Tutorials*, vol. 16, no. 1, 2014, pp. 369–92.

[4] S. Oteafy and H. Hassanein, "IoT in the Fog: A Roadmap for Data-Centric IoT Development," in IEEE Communications Magazine, vol. 56, no. 3, Mar. 2018, pp. 157–63.

[5] M. Satyanarayanan *et al.*, "Edge Analytics in the Internet of Things," *IEEE Pervasive Computing*, vol.14, no. 2, June 2015, pp. 24–31.

[6] I. Goodman, R. Mahler, and H. Nguyen, *Mathematics of Data Fusion*, vol. 37, Springer Science & Business Media, Harvard, 2010.

[7] M. Nanavati *et al.*, "Cloud Security: A Gathering Storm," *Commun. ACM*, vol. 57, no. 5, 2014, pp. 70–79.

[8] S. Oteafy and H. Hassanein, "Resource Re-Use in Wireless Sensor Networks: Realizing a Synergetic Internet of Things," *J. Commun.*, vol. 7, no. 7, July 2012, pp. 484–93.

[9] S. Oteafy and H. Hassanein, "Resilient IoT Architectures over Dynamic Sensor Networks with Adaptive Components," *IEEE Internet of Things J.*, vol. 4, no. 2, Apr. 2017, pp. 474–83.

[10] E. Ibarrola *et al.*, "A New Global Quality of Service Model: QoXphere," *IEEE Commun. Mag.*, vol.52, no.1, Jan. 2014, pp. 193–99.

[11] J. Dijck, "Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology," *Surveillance & Society*, vol. 12, issue 2, 2014, pp. 197–208.

[12] M. Domingues and A. Radwan, *Optical Fiber Sensors for IoT and Smart Devices*, Springer, Feb. 2017. ISBN: 978-3-319-47348-2.

[13] J. Heidemann *et al.*, "Building Efficient Wireless Sensor Networks with Low-Level Naming," *ACM Symp. Operating Systems Principles*, vol. 35, issue 5, Dec. 2012, pp. 146–59.

[14] F. Bari *et al.*, "A Survey of Naming and Routing in Information Centric Networks," *IEEE Commun. Mag.*, vol. 50, no. 12, Dec. 2012, pp. 44–53.

[15] A. Al-Fuqaha *et al.*, "Toward Better Horizontal Integration among IoT Services," *IEEE Commun. Mag.*, vol. 53, no. 9, Sept. 2015, pp. 72–79.

Biographies

Sharief M. A. Oteafy [M] is an assistant professor at the School of Computing, DePaul University. He co-authored a book, *Dynamic Wireless Sensor Networks*, and has presented over 40 publications and delivered multiple ComSoc tutorials on sensing systems and IoT. He has co-chaired a number of IEEE symposia and workshops in conjunction with IEEE ICC and IEEE LCN. He serves as the IEEE ComSoc AHSN Standards Liaison, on the ComSoc Tactile Internet Standards WG, and serves as an Associate Editor of both *IEEE Access* and Wiley's *Internet Technology Letters*.

Hossam S. Hassanein [F] is a professor and director of the School of Computing at Queen's University, Canada. His record spans more than 500 publications, in addition to numerous keynotes and plenary talks. He is the founder and director of the Telecommunications Research Lab at Queen's University. He is an IEEE Communications Society Distinguished Speaker (Distinguished Lecturer 2008–2010). He has received several recognitions and best paper awards at top international conferences, and led a number of symposia at flagship ComSoc conferences.