Internet-of-Everything Oriented Implementation of Secure Digital Health (D-Health) Systems

Grayson Honan, Alex Page, Ovunc Kocabas
University of Rochester, ECE
Rochester, NY 14627
{ghonan2,apage4,okocaba}@ece.rochester.edu

Tolga Soyata
SUNY Albany, ECE
Albany, NY 12222
tsoyata@albany.edu

Burak Kantarci
Clarkson University, ECE
Potsdam, NY 13699
bkantarc@clarkson.edu

***Abstract*—The past few decades have witnessed incredible advances in human health care, owing to the invention of devices such as MRI scanners, which allow physicians to monitor personal health in more detail than was ever previously possible. Such advances have drastically improved diagnostic quality and patient health care. Central to this incredible progress was the uncanny ability of technologists and academics to invent ever more useful tools to help physicians, be it the X-ray machine, CT, or MRI scanner. Whereas the aforementioned past-decades' tools aimed at acquiring personal data, the advent of the Internet-of-Things, vast computational power available in the cloud, and new data analytics algorithms will completely change the way we acquire and process medical data to improve health care going forward. In this paper, we conduct a quantitative feasibility study of a Digital Health (D-Health) system that is aimed at acquiring and processing health data using the emerging Internet-of-Everything paradigm. We specifically investigate the technological feasibility of communication, software, and data privacy aspects.**

***Index Terms*—decision support; Internet of Everything (IoE); visualization; analytics; remote health monitoring.**

I. Introduction

It is hard to believe that in 2016, cardiac diagnoses are mostly based on physical examinations and visual inspection of electrocardiograms [1], [2]; such methods could almost be considered "vintage" when compared to the non-medical world's technology. Real-time means to assess and predict the risk of cardiac diseases that can lead to chronic heart failures, and methods to permit therapeutic intervention are no more than research topics [3]. The pathological progression of many diseases requires long term observations of a patient to gather sufficient data to make accurate statistical inferences related to the onset of the disease at hand [4]. A D-Health system that provides automated remote health monitoring of clinically-relevant bio-markers could provide invaluable diagnostic information [5]–[7] and translate to health care cost savings of up to $300 B [8].

Adoption of D-Health systems and their use in improving health care will advance simultaneously with trends that are in motion already; commercially available *personal health and fitness monitoring* devices such as Applewatch [9], Fitbit [10], and Jawbone [11] are becoming the "next iPhone," and are even considered fashionable [12]. Off-the-shelf *advanced personal health monitoring* devices also exist for Glucose monitoring [13] or ECG monitoring [14], [15]. Significantly more sophisticated bio-patches are also becoming commercialized [16]–[18] that provide *clinical grade remote health monitoring* of advanced bio-markers such as gait, posture, body temperature, and surface EMG. Acquisition of the data that can be used to improve health care is not limited to personal data; the *crowdsensing* phenomenon promises to acquire data from the environment that can be used to determine environmental factors that are affecting our health [19], such as air or water quality.

The final destination of this acquired data is the cloud, where machine learning algorithms [20] make statistical inferences on the data to provide decision support to the health care professionals [21]–[24]. Using an expanding set of medical databases will open the door to discoveries of new treatments for diseases [25] and a better understanding the way the human body works [26]. In this paper, we investigate how these individual trends can be incorporated into a holistic system designed to develop effective, commercially-accepted D-Health systems that can improve our health care.

Our contributions in this paper are: In Section II-A, we draw a conceptual diagram of a D-Health system that can be applied to a generalized set of applications aimed at improving our health care. We highlight technical challenges in realizing a D-Health system in Section II-B and provide technical feasibility studies for each technical challenge: Issues related to providing a reliable communication infrastructure are elaborated on in Section III. In Section IV, issues related to data privacy, system-level security and the correctness of acquired data are studied quantitatively. In Section V, quantitative case studies are provided for visualization and analytics algorithms. Our concluding remarks in Section VI outline our position regarding the future of D-Health systems.

II. D-Health System Structure and Challenges

A conceptualized D-Health system is depicted in Fig. 1, which consists of two sections: *Front End* section is responsible for acquiring, aggregating, and pre-processing the data. *Back End* section is responsible for processing the data to extract useful information for use in health care. We will now detail the components and challenges of a D-Health system.

A. Conceptualized Structure of a D-Health System

Data Acquisition: Personal health monitoring is achieved using Wireless Body Area Networks (WBANs). WBANs consist of lightweight wearable sensors [5], [27] or more ad-

978-1-5090-0679-3/16/$31.00 ©2016 IEEE

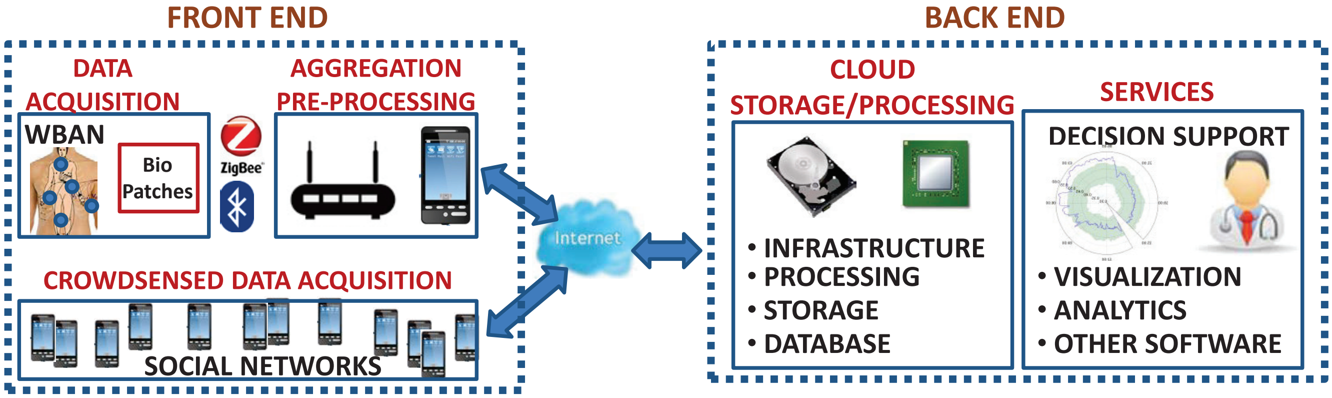

Fig. 1. A conceptual structure of a D-Health system, consisting of two parts: *Front End* is where the data is acquired, aggregated and pre-processed. *Back End* is where the data is stored/processed to provide services to health care organizations.

vanced clinical grade sensors (bio-patches) that are capable of measuring bio-markers such as EMG, gait, and blood pressure [17], [28]. BAN sensors utilize low-power Bluetooth or ZigBee protocols, based on standards such as the IEEE 802.15.6 [29], which prescribes radio frequency (RF)-based ultrawideband (UWB) and narrowband communication standards and RF-based human body communication standards [30].

Data Aggregation/Pre-processing: A WBAN consists of severely battery-power-restricted sensor devices and passive RF devices. Aggregation and pre-processing of the acquired data is necessary to reduce the data volume being handled and transmitted. This is achieved by concentrators [31], [32], and cloudlets [33]–[35], or smartphones acting as cloudlets, because sensors aren't computationally capable enough for D-health data processing and have more limited battery life.

Crowdsensed Data Acquisition is an emerging phenomenon [36], [37] that promises to enable the concurrent acquisition and aggregation of data — such as temperature, air quality, or humidity — from a wealth of capable, sensor-rich "crowd" resources [38], such as smartphones and tablets.

Cloud Infrastructure functions provided by a cloud service include multiple servers on one or more racks, storage space, virtualization, and other components to enable the *back end functionality* of a D-Health system. This infrastructure must be compliant with government health data regulations.

Database-Oriented Storage structures medical data in a standardized format to be rapidly queried for analytics purposes. Structuring the data in a standard database format also enables the *fusion* of similar data from multiple sources to enrich the data quality.

Services such as decision support for health-care professionals can be provided once the data is stored in a structured way and potentially fused with the crowdsensed data and applied as input to statistical inference algorithms.

B. Technical Challenges in Building D-Health Systems

Wireless Standards and Interoperability are discussed in Section III-A. Multiple wireless services can operate on the Industrial Scientific and Medical (ISM) 2.4 GHz band leading to co-channel interference for on-body networks.

Protocol Design Challenges and solutions are addressed in Section III-B. Physical layer challenges aim at addressing path loss and low power gain in on-body sensor networks whereas MAC layer challenges deal with urgency-based resource allocation for message frames. Network layer challenges include thermal-aware routing algorithms to avoid tissue damage.

Data Privacy Challenges involve protecting data from adversaries attempting to obtain it without authorization and are primarily *crypto-level* challenges. We detail the encryption schemes that are used to ensure data privacy in Section IV-B.

System Level Security is examined in IV-C in terms of side channel attacks because they attempt to obtain the secret keys by using system-level operational run-time information, such as server power consumption during crypto operations.

Data Trustworthiness Challenges include distinguishing between sensor malfunction and intentional sensor tampering. These challenges are exacerbated in crowdsensing settings. A detailed quantitative study is provided in Section IV-D.

Database Challenges: Health records in many different formats must be parsed and aggregated into a database system that is well-suited for tasks such as statistical analysis and machine learning.

Visualization of medical data reduces the *data burden* for the doctor and allows fast data handling for multiple patients. A quantitative case study is provided in Section V-A.

Decision Support by using machine learning algorithms reduce the *statistical inference burden* by taking advantage of the vast processing capability of computers. A quantitative case study is provided in Section V-B.

III. Communication Challenges

We break communication challenges in D-Health systems into two main categories, namely i) the wireless standards for D-Health systems and ii) the protocol design challenges. Wireless communications are widely used for sensor readings and actuation signals in D-Health systems [39], [40].

A. Wireless Standards and Interoperability

The scope of IEEE 802.15.6 includes radio frequency (RF)-based ultrawideband (UWB) and narrowband communication standards, as well as RF-based human body communication standards. RF-based human body communications utilize the 21 MHz centered frequency band with data rates of 164-1312.5 Kbps. UWB and narrowband-based human body communications utilize frequency bands between 402 MHz and 10 Ghz. UWB operates at data rates between 395 Kbps and 12.636 Mbps, and narrowband-based communications operate at 100 Kbps and 1000 Kbps [29].

RF-based communications have been reported to communicate through the air with high attenuation due to body shadowing at data rates up to 13 Mbps, whereas on-body communications solutions communicate through the body with low signal attenuation at data rates below 2 Mbps [40]. The protocols should be built on the communication standards for on-body networks, and address energy efficiency, security, privacy and low electromagnetic interference.

Cognitive and opportunistic solutions have become popular to address the interoperability challenges. Several system architectures have been developed on IEEE 802.22, which specifies the standard for Wireless Cognitive Radio Network Medium Access Control [41].

B. Protocol Design Challenges

PHY layer Challenges: The study in [42] uses a biofeedback control loop through sensor and actuator nodes. The proposed on-body network operates in the low bit rate medical implant communication service (MICS) 402–405 MHz frequency band with maximum bandwidth of 300 KHz [43]. The sensor nodes perform continuous health monitoring while the actuator nodes are responsible for medical drug delivery for patients who are in critical condition.

As an alternative to the MICS-based wearable and implantable systems such as [42], UWB-based implantable body area networks are also popular [44], [45]. A grand challenge in an RF-based body area network is the path loss and signal attenuation due to the physical characteristics of the medium such as blood circulation, respiration, and temperature variation throughout the body. Floor et al. derived a path loss model for UWB-based in-vivo communication systems [46] and showed that low frequencies such as 1–3 GHz reduced the transmission power in implanted networks as the path loss was remarkably low at these frequency levels. Furthermore, the study shows that the higher the number of on-body receiver antennas, the better the power gains (i.e., $\geq$ 3dB) as long as they are placed close enough to each other. Propagation paths are highly correlated; therefore, this phenomenon has to be taken into account in designing communication protocols and algorithms.

MAC layer Challenges: In [47], MAC layer protocols have been surveyed within the context of Machine-To-Machine (M2M) communications; hybrid protocols have also been studied as a solution to cope with the performance issues experienced under either contention-based or scheduling-based MAC protocols. As an example hybrid protocol, Hybrid MAC (HyMAC) consolidates the advantages of CDMA with TDMA and FDMA; each node is assigned a time slot and a frequency in response to its bandwidth request by using contention-based access [48]. Applicability of hybrid MAC protocols to D-Health systems has been extensively discussed in [47] in the context of M2M communications, and the authors have concluded that hybrid MAC protocols would have scalability issues due to the dense deployment of M2M networks. Random access-based reservation of slots, codes, and frequencies lead to bottlenecks under hybrid MAC protocols. Furthermore, overheads due to system reconfiguration lead to a large number of wasted time slots when compared to conventional wireless sensor networks. Hence, the authors advocate that TDMA-based MAC protocols are more suitable to M2M communication systems.

Network-layer Challenges: Thermal-aware routing algorithm (TARA) has been proposed to avoid hotspots in a body area network [49]. TARA defines a temperature threshold to identify a region in the network as a hot-spot (i.e., above the threshold); the packets are routed around the hotspot regions. If a region is identified as a hotspot in the network by a node, the protocol withdraws all the packets destined to that region, and sends them back to the source node. In [50], the hotspot preventing routing algorithm (HPR) selects the shortest path to the destination node if the destination is not in a hotspot region. While selecting the next hop, if the temperature of the next hop is not above the threshold, the packet is sent to that node, otherwise, the packet is sent to the coolest neighbor unless the next hop which is a hot spot is not the destination. Scalability and longer network lifetime is guaranteed by HPR (compared to TARA) at the expense of an overhead due to carrying the temperature information forward as a packet propagates towards its destination [51]. As another alternative to TARA, Adaptive Least Temperature Routing (ALTR) sends the packets to the coolest neighbor [52]; as soon as the number of hops exceeds a pre-defined threshold, the algorithm switches to shortest hop routing.

IV. Security Challenges

In this section, we study the security mechanisms that allow a D-Health system to guarantee privacy, prevent side channel attacks, and ensure the correctness of crowdsensed data.

A. Cryptographic Challenges

We study medical data privacy from three different aspects:

Data Storage Privacy refers to the assurance that encrypted data cannot be accessed unless an adversary obtains access to the private key, which is necessary for decryption. Conventional encryption schemes such as Advanced Encryption Standard (AES) and Elliptic Curve Cryptography (ECC) can provide data privacy and their details are given below.

Data Sharing Privacy refers to the guarantee that when multiple users must share data — within a list of authorized users — no additional user can access the data. This is feasible by Attribute Based Encryption (ABE) schemes that allow access to data based on user credentials, i.e., *attributes*. Two existing ABE family encryption schemes, namely KP-ABE and CP-ABE are detailed below.

Data Computation Privacy refers to the protection of data privacy during *computation*. Homomorphic encryption schemes (e.g. Paillier and Fully Homomorphic Encryption) can achieve data computation privacy and allow computations on medical data to be performed on *encrypted data* [53], [54]. These schemes are detailed below.

B. Encryption Schemes to Protect Data Privacy

Encryption schemes can be categorized as *conventional* and *emerging*. Conventional encryption schemes — AES and ECC — find widespread acceptance due to their resource-friendliness; however, they only provide data storage privacy. Emerging schemes — ABE and homomorphic — provide data sharing and data computation privacy but they are significantly more resource-intensive. We detail these schemes below.

Advanced Encryption Standard (AES): AES [55] is one of the most commonly-used symmetric key conventional encryption schemes for industry and government security needs. AES uses lightweight functions including XOR, data shuffling, and replacement-by-lookup, so the algorithm is both fast and power efficient.

Elliptic Curve Cryptography (ECC): ECC is a public key conventional encryption scheme that can achieve the level of security provided by 1024-bit RSA using only a 160-bit prime p. This vast improvement on RSA's key sizes allows significant savings in bandwidth and storage when using public key cryptosystems. One of ECC's most common implementations is the Elliptic Curve Integrated Encryption Scheme (ECIES) [56], which makes use of Diffie-Hellman key exchange to generate a shared secret. ECIES is much more computationally expensive than plain AES; in a generic C implementation [57], ECIES takes 3 orders-of-magnitude longer for encryption and decryption than plain AES. Additionally, ciphertext in ECIES requires approximately $6\times$ more space than a generic C implementation of AES.

Attribute-based Encryption (ABE): ABE improves on the data sharing capabilities of conventional encryption schemes (e.g., AES and ECIES) through the use of access policies. ABE exists in two variants, based on the placement of the access policy: Ciphertext-Policy ABE (CP-ABE) and Key-Policy ABE (KP-ABE).

In CP-ABE, users' private keys are associated with their credentials [58]. The ciphertext specifies an access policy, and only the users whose credentials satisfy the requirements of the access policy can decrypt it. Encryption and decryption in CP-ABE take 6 orders-of-magnitude longer than plain AES and ciphertexts require two orders-of-magnitude more storage.

In KP-ABE, access policies are placed on users' private keys and attributes are associated with the ciphertexts. Encryption and decryption in KP-ABE take four orders-of-magnitude longer than plain AES, and ciphertexts occupy approximately $40\times$ more space than plain AES.

Homomorphic Encryption: Absent from the encryption schemes we've examined thus far is the ability to operate on encrypted data; homomorphic encryption (HE) enables computation without observing decrypted data. At the least, an HE scheme implements either homomorphic addition or homomorphic multiplication, which translate to addition and multiplication on plaintext, respectively. A homomorphic scheme is defined as Fully Homomorphic Encryption (FHE) when it implements both homomorphic addition and homomorphic multiplication, and is thus able to evaluate arbitrary functions.

Paillier HE [59] is a lightweight, additively-homomorphic encryption scheme used for many practical applications. Its performance is similar to CP-ABE; encryption and decryption take 6 orders-of-magnitude longer than plain AES and ciphertexts require 2 orders-of-magnitude more storage.

FHE schemes are fairly resource-intensive for current generation D-Health systems [60]–[62], even when using the state-of-art Brakerski-Gentry-Vaikuntanathan (BGV) scheme [63]. Using the HElib implementation as a benchmark [64], BGV takes nearly 6 orders-of-magnitude longer for encryption and 6 orders-of-magnitude longer for decryption when compared to a generic C implementation of AES. Additionally, BGV ciphertexts require 6 orders-of-magnitude more storage than AES, and BGV homomorphic computation is $3100\times$ slower than Paillier computation [65].

C. System-Level Security Challenges

Chief among the security concerns associated with D-Health system design are various side channel attacks that exploit systemic information leaks. Vulnerabilities in the system's software and hardware implementations can enable these attacks as we detail below.

Cache Attacks: Cache attacks work by observing the cache access latency of the cryptographic instructions to recover the cache lines that store the secret key [66], [67]. Some hardware offers built-in defenses against this attack. The Intel AES-NI CPU instruction set [68], for example, makes cache access latency independent of data and calculates substitution results in hardware, rather than using a lookup table.

Timing Attacks: Timing attacks attempt to discover the secret key of a cryptosystem by observing the execution time of operations performed during encryption or decryption. If the execution time of operations varies based on the bits of the secret key [69], a timing attack will be effective. In ECC timing attacks, the execution time of scalar multiplication operations can leak information. This leak can be prevented by using a multiplication method that performs the operation independent of the bits in the secret key, such as the Montgomery multiplication method [70].

Power Analysis Attacks: If power consumption of a cryptosystem varies based on the bit values of a secret key, adversaries can discover the key by observing the power usage of the device (simple power analysis) or by using statistical methods of differential power analysis for more noise-tolerant measurements. When using AES, such attacks can be prevented by using randomized masks on AES operations [71] to remove the correlation between power consumption, the AES secret key, and the data being acted upon. In ECC power analysis attacks, randomizing the intermediate computations has been shown to remove the correlation between power consumption and sensitive key information.

Fault-Based Attacks: Through the application of a power glitch, magnetic field, or other stimulus to a cryptosystem, errors may be generated that reveal the secret key to an adversary. To prevent such attacks from being effective against AES-based cryptosystems, [72] proposes checking the correctness of results at various stages. An alternative presented in [73] is based on error detecting codes (EDC). Fault-based attacks for ECC-based schemes attempt to produce a point that is not on the elliptic curve during decryption [74]; these attacks can be thwarted by checking if the result is a point on the elliptic curve, and if not, discarding the result.

Data Rate Attacks: If the amount of data being transferred from a remote medical sensor depends on any physiological parameters, then an attacker may be able to learn some health information simply from the data transfer rate. For example, if a packet is sent after every heart beat, the heart rate can be easily inferred from the number of packets being sent per minute. Defense against this type of attack involves using techniques such as padding to maintain the same data rate independent of physiological events.

D. Data Correctness (Trustworthiness) Challenges

In crowdsensing-assisted data acquisition via social communities, trustworthiness of crowdsensed data should focus on reputation of sensing devices and their corresponding sensing accuracy [75], [76]. In trustworthy crowdsensing, instruments (nodes) are recommended to be recruited based on their reputation [77]. Percentage of "positive" readings — excluding outliers via an outlier detection algorithm [78] — denotes the reputation of a node [19]. Although tracking positive/negative readings may improve trustworthiness of crowdsensed data, the system is still prone to Sybil-like attacks [79]; a newly joining mobile device (i.e., sensing node) builds its reputation based on the votes of other devices.

Vote-based trustworthiness of node i is defined as $\Re_i(t)$ and is calculated as the total vote from the neighbors averaged by their total voting capacities. However, this may still lead to biased calculation of trustworthiness of crowdsensed data. Two solutions can be considered against this challenge. First, some trustworthy nodes, called *anchor nodes*, can be initially recruited with 100% trustworthiness and 100% vote capacity no matter what they report as sensing data [37]. Alternatively, collaborative trustworthiness can also be considered as a hybrid of the vote-based and statistical trustworthiness assessment [80]. Assessing trustworthiness of crowdsensed data requires assessing sensing node reputations according to

$$\begin{aligned}\Re_i(t) &= \sigma\cdot\Re_i(t^-)+(1-\sigma)\cdot\Re_i(t)\\ &=\sigma\cdot\Re_i(t^-)+(1-\sigma)\big[(1-\delta)\Re_i^{stat}(t)+\delta\cdot\Re_i^{voted}(t)\big]\end{aligned}\quad(1)$$

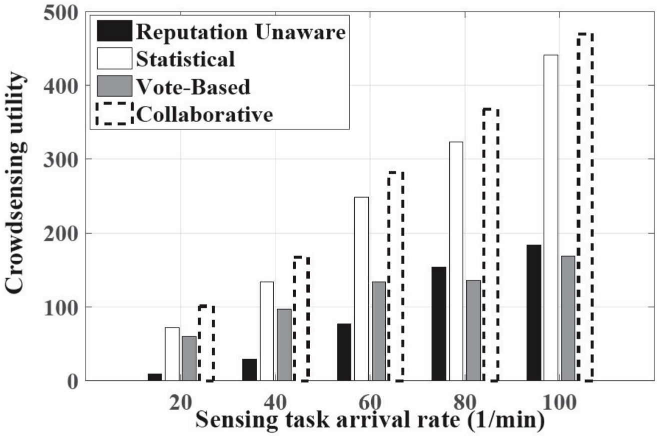

Fig. 2. A case study of crowdsensing utility under various trustworthiness assessment approaches.

$$\Re_i(t)=\sigma\cdot\Re_i(t^-)+(1-\sigma)\cdot\Bigg[(1-\delta)\cdot\Big(\frac{p_i(t)+\epsilon}{p_i(t)+n_i(t)+\epsilon}\Big)+\delta\cdot\frac{\displaystyle\sum_{j|T_{\{i\}}\cap T_{\{j\}}\neq\varnothing}(\omega_j\cdot\chi_j^i\cdot\Re_j)}{\displaystyle\sum_{j|T_{\{i\}}\cap T_{\{j\}}\neq\varnothing}(\omega_j\cdot\Re_j)}\Bigg]\quad(2)$$

where $\Re_i(t)$ is the reputation of node i and is a compound function of the statistical reputation ($\Re_i^{stat}(t)$ i.e., ratio of positive readings ($p_i(t)$) to the total readings ($p_i(t)+n_i(t)$)) and social reputation ($\Re_i^{voted}(t)$). σ and δ are weight factors that are used to quantify the transition speeds of node reputations [37]. (1) can be expanded as (2), where $T_{\{i\}}$ denotes the set of data sensed by node i such that $T_{\{i\}}\cap T_{\{j\}}$ represents the intersection of the sets of data sensed by node i and node j. In the vote-based component of the reputation assessment, ω_i denotes the current vote capacity of user i, and χ_j^i denotes the vote of node j for node i.

When health data is acquired through crowdsensing, the sensing nodes that are recruited for data acquisition need to be rewarded based on their sensing costs and the usefulness of the data they have provided. Figure 2 illustrates a comparison between the statistical, vote-based and collaborative trustworthiness assessment approaches in terms of crowdsensing utility where reputation-unaware data acquisition is used as a benchmark. Utility is defined as the difference between the total usefulness of the acquired data and the rewards/-compensation made to the sensing nodes. Figure 2 is based upon a simulation study in a $1000\,\mathrm{m}\times1000\,\mathrm{m}$ terrain with 1000 nodes and a sensing range of 30m. Amongst the 1000 nodes, 5% report wrong sensing data intentionally whereas the rest of the nodes report accurate sensor readings 97-98% of the time. The upper bound for the usefulness of acquired data and the sensing costs are set at 5 and 10, respectively. The 30-minute monitoring period under various sensing task arrival rates shows the viability of collaborative trustworthiness assessment for the sensing nodes. Furthermore, solely vote-based assessment leads to biased votes under heavy data acquisition rates leading to lower crowdsensing utility.

V. Software Implementation Challenges

In this section, we quantitatively study two key software-side technical challenges of designing a D-Health system. First, the large volume of sensor data produced from long-term, persistent patient monitoring would easily overwhelm a physician caring for 20-30 patients; therefore, new data visualization methods must be introduced to present medical data in an intuitive, summarized format. Second, decision support based on statistical trends in a patient cohort has the potential to increase diagnostic accuracy and clinical predictive capabilities, but significant challenges exist (including the assurance of data privacy).

A. Data Visualization

A novel visualization mechanism is introduced in [81] that is capable of presenting multi-modal medical data on a scale of ≥24 hours. The authors generate these visualizations through several stages of preprocessing, which transform the raw sensor data into filtered clinical markers for plotting [82]. This preprocessing step also addresses the issue of data volume by simplifying raw data into a summarized and practical format for clinical use. A quantitative example of data visualization for the QTc clinical market is given in Fig. 3 using the open source code provided in [81]. In this example, the polar plot shows intervals measured beat-to-beat during a 24 hour ECG of a patient. While the patient's daytime QTc values are only somewhat alarming, the nighttime values are distinctly life-threatening.

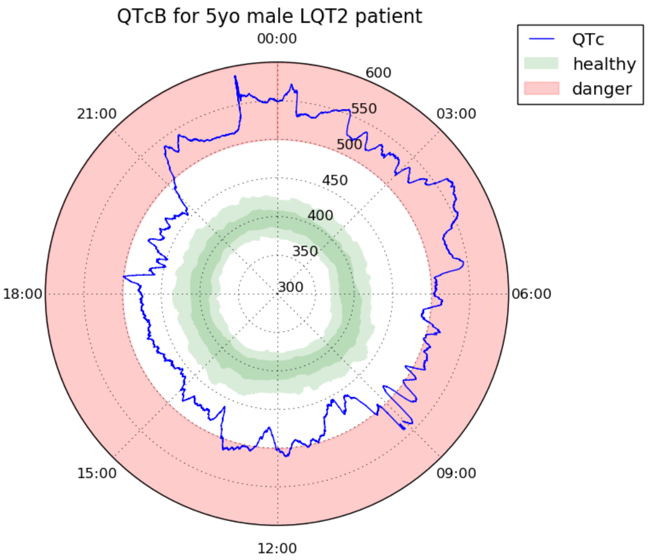

Fig. 3. Example 24-hour QTc plot in the "ECG Clock" format [81]. QTc for a healthy patient should normally fall into the green range for the entire day. While this patient shows a borderline high QTc range during daytime, QTc becomes clearly abnormal at night, indicating a potential cardiac hazard. Increased risk during sleep is consistent with this patient's LQT2 diagnosis.

B. Decision Support

We tested several machine learning (ML) algorithms on 639 24-hour Holter ECG recordings. 145 of the recordings came from LQTS type 2 patients, 294 came from LQTS type 1 patients (both are genetic disorders affecting cardiac function), and 200 recordings came from healthy patients. We used the scikit-learn Python library [83] to provide decision support. 70% of the data set is used for training the ML algorithms, and the resulting model is tested on the remaining 30% of the dataset. Classifier performance was characterized based on 20 trials with Holter recordings randomly split between "training" and "testing" during each trial. On average, classification of "healthy" vs. "long QT" was relatively accurate (around 90%). Additionally, differentiation between type 1 and type 2 LQTS is found to be 70–75% accurate with Support Vector Machine (SVM) and Random Forest ML algorithms.

All classifiers were generally effective, especially when optimized attributes were used; for example, setting the `coef0` attribute for Polynomial SVM to 1.0 and `dual` for Linear SVM to `False` improved the scores by ∼4%. Random Forest and SVM generally proved superior to other algorithms. The ability to change the SVM classification method by simply changing the kernel attribute offers great versatility; in our case, polynomial SVM performed slightly better than the linear or radial basis function (RBF) SVM.

VI. Conclusion and Ongoing Work

The digital health (D-Health) revolution is propelled forward by the Internet-of-Everything (IoE) paradigm, leading to the creation of advanced D-health systems capable of remote monitoring, analytics, visualization, and decision support. In this paper, we have studied the feasibility of a holistic framework for D-Health systems where data acquisition is based on IoE and assisted by mobile crowdsensing, and processing and storage are handled at a cloud platform to provide services such as visualization, analytics and decision support. The proposed D-Health framework consists of *Front End* and *Back End* sections. The front end is responsible for data acquisition via IoE sensors (i.e., on-body sensors and crowdsensing smartphones) and incorporates a cloudlet which performs aggregation and pre-processing. The back end consists of the cloud platform, which primarily provides storage and processing for services that include visualization, analytics, and so on. We have thoroughly investigated the challenges faced in the implementation of this framework, and we have discussed the possibility of integration of existing solutions to those challenges. To this end, we have studied the feasibility of crowdsensed data acquisition under various correctness assessment techniques, and we have concluded that the collaborative approaches perform better when data acquisition is assisted by crowdsensing nodes. As for the back end processing, we have used measurements from 24-hour Holter ECG recordings to test the performance of various data classifiers, and found that SVM and Random Forest based classifiers were superior to other approaches in this case study. Finally, as a service component in the back end, we have presented an intuitive way to visualize the continuously monitored data to the end user.

We are planning to integrate these pieces on a real testbed, where on-body sensors are interfaced by front end circuitry and communicate with the cloudlet through low-power Bluetooth. We will also attempt to address trustworthiness/tamper-resistance of sensor data, as conceptualized in [4]. Furthermore, social networks currently serve as the data publishing layer in the data acquisition block of the front end plane. Therefore, we are planning to extend the role of social networks to an unstructured knowledge-base which will be analyzed by the cloudlet in order to retrieve useful data.

Acknowledgments

This work is supported in part by the National Science Foundation grants CNS-1239423 and CNS-1464273. Tolga Soyata was an Assistant Professor - Research during the preparation of this manuscript. He was the PhD adviser of Alex Page and Ovunc Kocabas, as well as the undergraduate adviser of Grayson Honan. Before the publication of this manuscript he joined SUNY Albany, ECE as an Associate Professor.

References

[1] E. J. Petr, C. R. Ayers, A. Pandey, J. A. Lemos, T. Powell-Wiley, A. Khera, D. M. Lloyd-Jones, and J. D. Berry, "Perceived lifetime risk for cardiovascular disease (from the dallas heart study)," *The American Journal of Cardiology*, vol. 114, no. 1, pp. 53 – 58, 2014.

[2] J. Saul, P. J. Schwartz, M. J. Ackerman, and J. K. Triedman, "Rationale and objectives for ECG screening in infancy," *Heart Rhythm*, vol. 11, no. 12, pp. 2316 – 2321, 2014.

[3] W.-H. Lin, H. Zhang, and Y.-T. Zhang, "Investigation on cardiovascular risk prediction using physiological parameters," *Computational and Mathematical Methods in Medicine*, vol. 2013, no. 1, pp. 1–21, 2013.

[4] A. Page, M. Hassanalieragh, T. Soyata, M. K. Aktas, B. Kantarci, and S. Andreescu, "Conceptualizing a Real-Time Remote Cardiac Health Monitoring System," in *Enabling Real-Time Mobile Cloud Computing through Emerging Technologies*. IGI Global, 2015, ch. 1, pp. 1–34.

[5] A. Pantelopoulos and N. Bourbakis, "A survey on wearable sensor-based systems for health monitoring and prognosis," *IEEE Trans. Sys., Man, and Cybernetics, Part C: Applic. and Reviews*, vol. 40, no. 1, pp. 1–12, Jan 2010.

[6] R. Paradiso, G. Loriga, and N. Taccini, "A wearable health care system based on knitted integrated sensors," *IEEE Trans. Info. Tech. in Biomedicine*, vol. 9, no. 3, pp. 337–344, Sept 2005.

[7] A. Milenkovi, C. Otto, and E. Jovanov, "Wireless sensor networks for personal health monitoring: Issues and an implementation," *Comput. Commun.*, vol. 29, no. 1314, pp. 2521 – 2533, 2006, wirelsess Senson Networks and Wired/Wireless Internet Communications. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140366406000508

[8] Goldman Sachs, Inc. Digital Healthcare could save America $300 Billion. http://www.businessinsider.com/goldman-digital-healthcare-is-coming-2015-6.

[9] Apple Inc., "Apple watch," accessed April 2015. [Online]. Available: https://www.apple.com/watch/

[10] FitBit Inc., "flex: Wireless activity + sleep wristband," accessed April 2015. [Online]. Available: https://www.fitbit.com/flex

[11] Jawbone Inc., "Jawbone fitness trackers," accessed April 2015. [Online]. Available: https://jawbone.com/up/trackers

[12] A. Schneider, "Tech makeover: The days of tech being a mere practical application of science are over. fashionistas, take note : Sartorial has turned cyber," *In New York*, pp. 26–31, June 2015.

[13] Sensys Medical, Inc. Near-Infrared Spectroscopy. http://www.diabetesnet.com/diabetes-technology/meters-monitors/future-meters-monitors/sensys-medical.

[14] Alivecor. (2013) ECG screening made easy. http://www.alivecor.com/.

[15] C. Leaf, "World's Thinnest 3-Lead ECG Patch," http://www.clearbridgevitalsigns.com/brochures/CardioLeaf_ULTRA_Brochure.pdf.

[16] S. Xu, Y. Zhang, L. Jia, K. E. Mathewson, K.-I. Jang, J. Kim, H. Fu, X. Huang, P. Chava, R. Wang, S. Bhole, L. Wang, Y. J. Na, Y. Guan, M. Flavin, Z. Han, Y. Huang, and J. A. Rogers, "Soft microfluidic assemblies of sensors, circuits, and radios for the skin," *Science*, vol. 344, pp. 70–74, 2014.

[17] D. Son, J. Lee, S. Qiao, R. Ghaffari, J. Kim, J. E. Lee, C. Song, S. J. Kim, D. J. Lee, S. W. Jun, S. Yang, M. Park, J. Shin, K. Do, M. Lee, K. Kang, C. S. Hwang, N. Lu, T. Hyeon, , and D.-H. Kim, "Multifunctional wearable devices for diagnosis and therapy of movement disorders," *Nature Nanotechnology*, pp. 1–8, 2014.

[18] D.-H. Kim, R. Ghaffari, N. Lu, and J. A. Rogers, "Flexible and stretchable electronics for biointegrated devices," *Annual Review of Biomedical Engineering*, pp. 113–128, 2012.

[19] B. Kantarci and H. Mouftah, "Trustworthy Sensing for Public Safety in Cloud-Centric Internet of Things," *IEEE Internet of Things Journal*, vol. 1, no. 4, pp. 360–368, Aug 2014.

[20] Committee on the Analysis of Massive Data, *Frontiers in Massive Data Analysis*. National Academies Press, 2013.

[21] M. Hassanalieragh, A. Page, T. Soyata, G. Sharma, M. K. Aktas, G. Mateos, B. Kantarci, and S. Andreescu, "Health Monitoring and Management Using Internet-of-Things (IoT) Sensing with Cloud-based Processing: Opportunities and Challenges," in *IEEE International Conference on Services Computing (SCC)*, Jun 2015, pp. 285–292.

[22] S. Earley, "The Promise of Healthcare Analytics," *IEEE Computing Edge*, pp. 27–29, June 2015.

[23] L. Wang and R. Ranjan, "Processing distributed internet of things data in clouds," *IEEE Computing Edge*, pp. 12–16, Jun 2015.

[24] A. Page, S. Hijazi, D. Askan, B. Kantarci, and T. Soyata, "Research Directions in Cloud Based Decision Support Systems for Health Monitoring Using Internet-of-Things Driven Data Acquisition," *International Journal of Services Computing (IJSC)*, vol. 4, no. 4, pp. 18–34, 2016.

[25] E. Eskin, "Discovering Genes Involved in Disease and the Mystery of Missing Heritability," *Communications of the ACM*, vol. 58, no. 10, pp. 80–87, 2015.

[26] Multiple, "Hacking the Human OS," *IEEE Spectrum*, pp. 31–48, June 2015.

[27] A. Benharref and M. Serhani, "Novel cloud and SOA-based framework for E-Health monitoring using wireless biosensors," *IEEE Journal of Biomed. and Health Inf.*, vol. 18, no. 1, pp. 46–55, Jan 2014.

[28] S. Babu, M. Chandini, P. Lavanya, K. Ganapathy, and V. Vaidehi, "Cloud-enabled remote health monitoring system," in *Int. Conf. on Recent Trends in Inform. Tech. (ICRTIT)*, July 2013, pp. 702–707.

[29] "IEEE Standard for Local and metropolitan area networks - Part 15.6: Wireless Body Area Networks," *IEEE Std 802.15.6-2012*, pp. 1–271, Feb 2012.

[30] T. Soyata, L. Copeland, and W. Heinzelman, "RF Energy Harvesting for Embedded Systems: A Survey of Tradeoffs and Methodology," *IEEE Circuits and Systems Magazine*, vol. 16, no. 1, pp. 22–57, Feb 2016.

[31] W. Zhao, C. Wang, and Y. Nakahira, "Medical application on internet of things," in *IET Int. Conf. on Com. Tech. and Application (ICCTA 2011)*, Oct 2011, pp. 660–665.

[32] F. Hu, D. Xie, and S. Shen, "On the application of the internet of things in the field of medical and health care," in *IEEE Int. Conf. on and IEEE Cyber, Physical and Social Computing Green Computing and Communications (GreenCom),(iThings/CPSCom)*, Aug 2013, pp. 2053–2058.

[33] N. Powers, A. Alling, K. Osolinsky, T. Soyata, M. Zhu, H. Wang, H. Ba, W. Heinzelman, J. Shi, and M. Kwon, "The Cloudlet Accelerator: Bringing Mobile-Cloud Face Recognition into Real-Time," in *Globecom Workshops (GC Wkshps)*, San Diego, CA, Dec 2015.

[34] T. Soyata, R. Muraleedharan, C. Funai, M. Kwon, and W. Heinzelman, "Cloud-Vision: Real-Time Face Recognition Using a Mobile-Cloudlet-Cloud Acceleration Architecture," in *IEEE Symposium on Computers and Communications (ISCC)*, Cappadocia, Turkey, Jul 2012, pp. 59–66.

[35] T. Soyata, H. Ba, W. Heinzelman, M. Kwon, and J. Shi, "Accelerating mobile cloud computing: A survey," in *Communication Infrastructures for Cloud Computing*, H. T. Mouftah and B. Kantarci, Eds. IGI Global, Sep 2013, ch. 8, pp. 175–197.

[36] X. Sheng, J. Tang, X. Xiao, and G. Xue, "Sensing as a Service: Challenges, Solutions and Future Directions," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3733–3741, Oct 2013.

[37] M. Pouryazdan, B. Kantarci, T. Soyata, and H. Song, "Anchor-Assisted and Vote-based Trustworthiness Assurance in Smart City Crowdsensing," *IEEE Access*, vol. 4, pp. 529–541, Mar 2016.

[38] T. Soyata, *Enabling Real-Time Mobile Cloud Computing through Emerging Technologies*. IGI Global, Aug 2015.

[39] Y. Liu, T. Ketterl, G. Arrobo, and R. Gitlin, "Modeling the wireless in vivo path loss," in *IMWS-Bio)*, Dec 2014, pp. 1–3.

[40] A. Karargyris and A. Koulaouzidis, "OdoCapsule: Next-Generation Wireless Capsule Endoscopy With Accurate Lesion Localization and Video Stabilization Capabilities," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 1, pp. 352–360, Jan 2015.

[41] "Standard for Wireless Regional Area Networks Part 22 : Cognitive Wireless RAN Medium Access Control(MAC) and Physical Layer (PHY) specifications: Policies and procedures for operation in the TV Bands," *IEEE Std 802.22-2011*, Jul 2011.

[42] Q. Fang, S.-Y. Lee, H. Permana, K. Ghorbani, and I. Cosic, "Developing a Wireless Implantable Body Sensor Network in MICS Band," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 4, pp. 567–576, Jul 2011.

[43] US Federal Communications Commission (FCC), "Medical Device Radiocommunications Service ," https://www.fcc.gov/encyclopedia/medical-device-radiocommunications-service-medradio.

[44] Y. Gao, Y. Zheng, S. Diao, W.-D. Toh, C.-W. Ang, M. Je, and C.-H. Heng, "Low-Power Ultrawideband Wireless Telemetry Transceiver for Medical Sensor Applications," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 768–772, Mar 2011.

[45] M. Yuce, "Recent wireless body sensors: Design and implementation," in *IEEE MTT-S International Microwave Workshop Series on RF and Wireless Technologies for Biomedical and Healthcare Applications (IMWS-BIO)*, Dec 2013, pp. 1–3.

[46] P. Floor, R. Chavez-Santiago, S. Brovoll, O. Aardal, J. Bergsland, O.-J. Grymyr, P. Halvorsen, R. Palomar, D. Plettemeier, S.-E. Hamran, T. Ramstad, and I. Balasingham, "In-Body to On-Body Ultrawideband Propagation Model Derived From Measurements in Living Animals," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 938–948, May 2015.

[47] A. Rajandekar and B. Sikdar, "A Survey of MAC Layer Issues and Protocols for Machine-to-Machine Communications," *IEEE Internet of Things Journal*, vol. 2, no. 2, pp. 175–186, Apr 2015.

[48] M. Salajegheh, H. Soroush, and A. Kalis, "HYMAC: Hybrid TDMA/FDMA Medium Access Control Protocol for Wireless Sensor Networks," in *IEEE Intl. Sympp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep 2007.

[49] Q. Tang, N. Tummala, S. Gupta, and L. Schwiebert, "TARA: Thermal-Aware Routing Algorithm for Implanted Sensor Networks," in *Distributed Computing in Sensor Systems*, V. Prasanna, S. Iyengar, P. Spirakis, and M. Welsh, Eds., 2005, vol. 3560, pp. 206–217.

[50] A. Bag and M. A. Bassiouni, "Hotspot Preventing Routing algorithm for delay-sensitive applications of in vivo biomedical sensor networks," *Information Fusion*, vol. 9, no. 3, pp. 389–398, 2008.

[51] R. Kamal, M. Rahman, and C. S. Hong, "A lightweight temperature scheduling routing algorithm for an implanted sensor network," in *International Conference on ICT Convergence*, Sep 2011, pp. 396–400.

[52] A. Bag and M. Bassiouni, "Energy Efficient Thermal Aware Routing Algorithms for Embedded Biomedical Sensor Networks," in *IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS)*, Oct 2006, pp. 604–609.

[53] O. Kocabas and T. Soyata, "Medical Data Analytics in the cloud using Homomorphic Encryption," in *Handbook of Research on Cloud Infrastructures for Big Data Analytics*, P. R. Chelliah and G. Deka, Eds. IGI Global, Mar 2014, ch. 19, pp. 471–488.

[54] A. Page, O. Kocabas, T. Soyata, M. K. Aktas, and J. Couderc, "Cloud-Based Privacy-Preserving Remote ECG Monitoring and Surveillance," *Annals of Noninvasive Electrocardiology (ANEC)*, vol. 20, no. 4, pp. 328–337, 2014.

[55] National Institute of Standards and Technology, "Advanced encryption standard (AES)," Nov 2001, FIPS-197.

[56] D. Hankerson, A. J. Menezes, and S. Vanstone, *Guide to elliptic curve cryptography*. Springer Science & Business Media, 2006.

[57] J. A. Akinyele, C. Garman, I. Miers, M. W. Pagano, M. Rushanan, M. Green, and A. D. Rubin, "Charm: a framework for rapidly prototyping cryptosystems," *Journal of Cryptographic Engineering*, vol. 3, no. 2, pp. 111–128, 2013.

[58] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *IEEE Symposium on Security and Privacy*, 2007, pp. 321–334.

[59] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *EUROCRYPT*, 1999, pp. 223–238.

[60] O. Kocabas and T. Soyata, "Utilizing Homomorphic Encryption to Implement Secure and Private Medical Cloud Computing," in *IEEE 8th International Conference on Cloud Computing (CLOUD)*, New York, NY, Jun 2015, pp. 540–547.

[61] O. Kocabas, T. Soyata, J. Couderc, M. K. Aktas, J. Xia, and M. Huang, "Assessment of Cloud-based Health Monitoring using Homomorphic Encryption," in *IEEE International Conference on Computer Design (ICCD)*, Ashville, VA, Oct 2013, pp. 443–446.

[62] A. Page, O. Kocabas, S. Ames, M. Venkitasubramaniam, and T. Soyata, "Cloud-based Secure Health Monitoring: Optimizing Fully-Homomorphic Encryption for Streaming Algorithms," in *Globecom Workshops (GC Wkshps)*, Austin, TX, Dec 2014, pp. 48–52.

[63] Z. Brakerski, C. Gentry, and V. Vaikuntanathan, "(Leveled) fully homomorphic encryption without bootstrapping," in *ITCS*, 2012, pp. 309–325.

[64] S. Halevi and V. Shoup, "HElib," https://github.com/shaih/HElib.

[65] O. Kocabas, T. Soyata, and M. Aktas, "Emerging Security Mechanisms for Medical Cyber Physical Systems," *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 2016.

[66] D. J. Bernstein, "Cache-timing attacks on AES," 2005.

[67] D. A. Osvik, A. Shamir, and E. Tromer, "Cache attacks and countermeasures: the case of AES," in *Topics in Cryptology–CT-RSA*, 2006, pp. 1–20.

[68] S. Gueron, "Intels new AES instructions for enhanced performance and security," in *Fast Software Encryption*, 2009, pp. 51–66.

[69] P. C. Kocher, "Timing attacks on implementations of diffie-hellman, rsa, dss, and other systems," in *CRYPTO*, 1996, pp. 104–113.

[70] P. L. Montgomery, "Speeding the pollard and elliptic curve methods of factorization," *Mathematics of computation*, vol. 48, no. 177, pp. 243–264, 1987.

[71] T. S. Messerges, "Securing the AES finalists against power analysis attacks," in *Fast Software Encryption*, 2001, pp. 150–164.

[72] R. Karri, K. Wu, P. Mishra, and Y. Kim, "Fault-based side-channel cryptanalysis tolerant rijndael symmetric block cipher architecture," in *Defect and Fault Tolerance in VLSI Systems, 2001. Proceedings. 2001 IEEE International Symposium on*, 2001, pp. 427–435.

[73] G. Bertoni, L. Breveglieri, I. Koren, P. Maistri, and V. Piuri, "Error analysis and detection procedures for a hardware implementation of the advanced encryption standard," *Computers, IEEE Transactions on*, vol. 52, no. 4, pp. 492–505, 2003.

[74] I. Biehl, B. Meyer, and V. Müller, "Differential fault attacks on elliptic curve cryptosystems," in *CRYPTO*, 2000, pp. 131–146.

[75] L. Kazemi, C. Shahabi, and L. Chen, "GeoTruCrowd: Trustworthy Query Answering with Spatial Crowdsourcing," in *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2013, pp. 314–323.

[76] C. Shahabi, "Towards a Generic Framework for Trustworthy Spatial Crowdsourcing," in *MobiDE*, Jun 2013, pp. 1–4.

[77] X. Sheng, J. Tang, X. Xiao, and G. Xue, "Sensing as a Service: Challenges, Solutions and Future Directions," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3733–3741, Oct 2013.

[78] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier Detection Techniques for Wireless Sensor Networks: A Survey," *IEEE Communications Surveys Tutorials*, vol. 12, no. 2, pp. 159–170, 2010.

[79] B. Kantarci, K. G. Carr, and C. D. Pearsall, "SONATA: Social Network Assisted Trustworthiness Assurance in Smart City Crowdsensing," *Intl. Journal of Distributed Systems and Technologies*, vol. 7, no. 1, pp. 64–84, Jan-Mar 2016.

[80] B. Kantarci, P. M. Glasser, and L. Foschini, "Crowdsensing with social network-aided collaborative trust scores," in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2015.

[81] A. Page, T. Soyata, J. Couderc, and M. K. Aktas, "An Open Source ECG Clock Generator for Visualization of Long-Term Cardiac Monitoring Data," *IEEE Access*, vol. 3, pp. 2704–2714, Dec 2015.

[82] A. Page, T. Soyata, J. Couderc, M. Aktas, B. Kantarci, and S. Andreescu, "Visualization of Health Monitoring Data acquired from Distributed Sensors for Multiple Patients," in *IEEE Global Telecommunications Conference (GLOBECOM)*, San Diego, CA, Dec 2015.

[83] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.