Challenges in Data Quality Assurance in Pervasive Health Monitoring Systems

Janani Sriram[1] · Minho Shin[1] · David Kotz[1] · Anand Rajan[2]
Manoj Sastry[2] · Mark Yarvis[2]

[1]Institute for SecurityTechnology Studies
Dartmouth College, Hanover, NH, USA
{Janani.C.Sriram | mhshin | dfk}@cs.dartmouth.edu

[2]Intel Corporation Hillsboro, OR, USA
{anand.rajan | manoj.r.sastry | mark.d.yarvis}@intel.com

Abstract

Wearable, portable, and implantable medical sensors have ushered in a new paradigm for healthcare in which patients can take greater responsibility and caregivers can make well-informed, timely decisions. Health-monitoring systems built on such sensors have huge potential benefit to the quality of healthcare and quality of life for many people, such as patients with chronic medical conditions (such as blood-sugar sensors for diabetics), people seeking to change unhealthy behavior (such as losing weight or quitting smoking), or athletes wishing to monitor their condition and performance. To be effective, however, these systems must provide assurances about the quality of the sensor data. The sensors must be applied to the patient by a human, and the sensor data may be transported across multiple networks and devices before it is presented to the medical team. While no system can guarantee data quality, we anticipate that it will help for the system to annotate data with some measure of confidence. In this paper, we take a deeper look at potential health-monitoring usage scenarios and highlight research challenges required to ensure and assess quality of sensor data in health-monitoring systems.

1 Introduction

The advent of portable computing devices and miniature sensing devices presents many new opportunities for personal healthcare. Formerly, most medical sensing devices were used in a hospital setting under the care of trained medical and technical personnel; soon, many devices will be worn throughout a patient's daily life or installed at home and in assisted-living settings. These devices will collect health related data for many purposes, by patients with chronic medical conditions (such as blood-sugar sensors for diabetics), people seeking to change behavior (e.g., losing weight or quitting smoking), or athletes wishing to monitor their condition and performance. The resulting data may be used directly by the person, or shared with others: with a physician for treatment, with an insurance company for coverage, with the adult children of elderly parents, or by a coach.

Clearly, to be useful these systems must assure that high-quality information reaches the data user – or, at least, the system must be able to express some level of confidence in the data being presented. Failure to consider the confidence level can directly lead to incorrect medical and financial decisions. For exam-

D. Gawrock, H. Reimer, A.-R. Sadeghi, C. Vishik (Editors): Future of Trust in Computing, Vieweg+Teubner (2009), 129-142

ple, if a physician bases a medical decision on information that is inaccurate, stale, or irrelevant to the patient at hand, he could end up prescribing incorrect medication to a patient, or even worse, perform surgery on the wrong patient. Similarly, an insurance company could collect fees or render payments that do not accurately reflect the patient's health status and treatments. Confidence in data in the area of emerging pervasive medical applications can return to medical providers some of the same assurances available in a personal interaction with the patient.

Outside the hospital setting, in particular, the sensors may be applied by the patient or family members; the data may be gathered through a personal mobile device (such as a mobile phone), over a personal network (such as a wireless network at home), or over the public Internet. Clearly, the accuracy and availability of sensor data is difficult to ensure. To evaluate the trustworthiness of medical data gathered in this manner, we need a holistic system view that inspects contributions to risk and error in medical data as it flows from the patient to the caregiver.

We motivate this topic area with three use cases in Section 2. We list several factors involved in data quality (Section3) and describe research challenges related to each factor (Section 4). We survey related work in Section 5. Finally, in Section6 we summarize.

2 Use cases

We set the stage for our discussion by presenting three scenarios that motivate many of the research challenges.

2.1 Scenario 1

Jack's doctor suspects that Jack has episodes of low blood-glucose levels not immediately visible from the manual measurement method and recommends a five-day glucose sensor test. The test uses an implanted glucose sensor and an electronic recorder to continuously monitor glucose levels, allowing physicians to study patterns over an extended period using more frequent samples.[1]

The doctor inserts the sensor under the skin of the patient's abdomen and connects the wire from the sensor to the recorder. A nurse calibrates the sensor by entering the blood-glucose level obtained from the manual method into the recorder. Jack is advised to always carry the recorder, and to keep it from getting wet.

During the five-day test, the sensor takes glucose readings every five minutes and sends it across the wire to the recorder. Jack also records glucose levels using the manual method four times a day, along with insulin injection times, meals, and exercise.

After the five-day test is complete, the doctor removes the implanted sensor and uploads data from the recorder to the doctor's computer. The doctor compares results from both types of tests to diagnose the condition or prescribe changes to insulin type or dosage and diet.

2.2 Scenario 2

Following minor surgery Jane awoke in a hospital room. A nurse applied a suite of sensors to her, including electro-cardiogram, blood oxygenation, and blood pressure. The nurse waved each device in

1 This scenario is based on a real "Five-Day Glucose Sensor Test" program at the Dartmouth-Hitchcock Medical Center (http://tinyurl.com/6y2tk6).

front of a barcode reader on the bedside monitor. This step helped ensure that the monitor was receiving data wirelessly from the right set of sensors; the barcode also encoded the date of the last calibration of each sensor. Any sensor due for periodic maintenance would be automatically rejected. Soon her vitals were continuously being displayed on the bedside monitor.

Back at the nurse's station, telemetry from the room monitor was flowing into a display that monitored all patients on the floor. Validating against the vitals she had just seen in the room, the nurse uses the nurse's station computer to enter Jane's identity, associating the data stream with Jane. Data entering the database could now be routed to Jane's medical record. From the nurse's station, the nurse is able to view sensor data from all of the patients on the hospital ward. Each display includes patient identifying information as well as confidence information to reduce false alarms (false positives) and missed alarms (false negatives).

Meanwhile, on the floor of the ward, a physician is making rounds. When he approaches Jane's bed, his mobile tablet associates with Jane's bedside monitor and sensor and displays Jane's medical record as well as her current vitals. A built-in confidence indicator assures the physician that this is indeed Jane's data and that the sensors are relevant, accurate, functional, and correctly applied.

Concerned about her recovery, Jane's doctor refers her to a cardiologist for an expert opinion. Rather than traveling to see Jane, the cardiologist first reviews her medical record. Via a telepresence capability built into her room, the cardiologist is able to speak with Jane, viewing her condition first-hand and discussing her symptoms, while simultaneously viewing her continuous vitals stream. Based on this information, and a built-in trust assessment of the data, the doctor prescribes medication for Jane with confidence.

2.3 Scenario 3

John's insurance company promises to reduce his insurance rates if he would quit smoking. The insurance company provides a wrist-mounted device that contains sensors to detect heart beats, blood oxygenation, accelerometers, and a smoke sensor to ensure that John is true to his word. The device includes several tamper-evident features that would allow the insurance company to determine at periodic intervals that the device (sensors, processor, and software) have not been modified.

A nurse visits John to program the sensing device by connecting it via a temporary wire to her own trusted device. During this training phase, the device measures small changes in John's heart rate, establishing a signature for John's heart-rate variability. Then John was asked to perform several tasks, including mimicking his smoking behavior, to calibrate an activity inference. Before the nurse left, she reminded John to wear the device at all times, to avoid smoking or being around others who smoked, and to periodically connect it to his PC for data upload. John was told that if he did not wear his device, or if he gave the device to a friend to wear, he would not receive his discount.

The device constantly monitored its sensors, ensuring that John was wearing it and there were no telltale signs of smoking, such as detection of smoke in the air, drops in blood oxygenation without corresponding exercise activities, or detection of smoking gestures. Every week, John would connect the device to his PC to upload data to his insurance company.

John's insurance company input the periodic data from John into a validation algorithm to assess the probability that he was smoking. The algorithm includes correlation across multi-modal sensors, such as detecting smoke in the air or changes in physiological data, and recognizing smoking gestures, to determine the probability that John is smoking. In addition, the results indicate the probability that John

has applied the sensors correctly, based on both a qualitative assessment of the data as well as self-test features in the sensor. Finally, the algorithm assesses the probability that John is trying to actively cheat, testing the authenticity of the devices that originated the data, validating the attestable state of the devices, and matching patterns in the sensor data (e.g., signatures of activity and heart rate variability) against John's previously recorded patterns.

3 Data Quality

Health-monitoring scenarios like those above require high-quality data from medical sensing devices or sensors. *Data quality* refers to the accuracy, authenticity, and appropriateness of a set of data for a given purpose. Ideally, a pervasive health-monitoring system assures high-quality data through a design and implementation that *ensures* the authenticity and integrity of the sensor data. Since it is impossible for a system to ensure perfect data quality, it is important to be able to *assess* the degree of confidence in the data, and to express that confidence in a way that allows the user to interpret the data in context. *Confidence* in sensor data represents the belief that the sensor reading accurately reflects the desired value; by *assessing quality* we mean quantifying confidence in sensor data and effectively presenting the results of the assessment to the user of the sensor data.

In this context, *facts* are data that are presumed correct a priori, not obtained from any of the sensors in question. For example, facts include information from the patient's medical history.

Finally, we define *trust* as the system's belief that a person or component will behave as expected.

3.1 Factors

An assessment of data quality entails an understanding of the various factors that might affect data quality; we consider six sensor factors, two human factors, and three system architecture factors. Ultimately, in future research, we hope to find a quantifiable metric or rubric for each factor.

3.1.1 Sensor factors

(Factor S1) Sensor design: Each sensor is characterized by a baseline measurement error inherent in the engineering, design, type, and purpose of the sensor. Confidence depends on both the *precision* (granularity of its reading) and *accuracy* (potential deviation from the true value).For example, a household bathroom scale may report weight in tenths of pounds (precision) and may be expected to be within two pounds of the correct weight (accuracy).To ensure high-quality sensor data, well-designed sensors with fine precision and high accuracy must be selected with due consideration to the medical needs of the situation, cost, and convenience. In Scenario 1, for example, the implanted sensor provides higher temporal precision (many readings per day) and possibly higher accuracy than the manual method.

(Factor S2) Sensor manufacture: Quality of manufacture reflects the trust in the sensor manufacturer and its manufacturing process; confidence may be based on past interaction or reputation.

(Factor S3) Sensor calibration: Sensor accuracy degrades with time, requiring periodic recalibration. Confidence in the sensor's calibration depends on the time since the last calibration, rate of drift away from calibration, and the reliability of the calibration authority.

(Factor S4) Sensor application: Many medical sensors must be applied correctly to provide meaningful results. For example, a thermometer may need to be applied directly to bare skin; a pulse oximeter may need to be applied to a specific part of the body. Confidence derives from trust in the patient or car-

egiver's ability and reliability in applying the sensor (see below) or from corroborating evidence (e.g., by secondary sensors that confirm the proper application of the primary sensor).

(Factor S5) Sensor integrity: In some settings, such as Scenario 3, there may be concern that the patient or another party may tamper with the sensor or the sensing system. Confidence in the sensor's integrity may derive from tamper-resistant or tamper-evident hardware, including trusted-platform modules that can attest to the integrity of the sensor system software.

(Factor S6) Sensor data correlation: One mechanism to assess confidence in sensor readings is to compare those readings against other data. *Redundant sensing* correlates a sensor reading against other sensors of the same type, with sensors from different parts of the body (spatial correlation), or with historical values from the same sensor (temporal correlation). *Multi-modal sensing* correlates a sensor reading against other sensors of different types, building on known correlations between sensor types. In Scenario 3, for example, a smoking event creates smoke in the air, a drop in blood oxygenation, and smoking gestures by the patient. *Fact checking* correlates a sensor reading with facts, such as information (e.g., age) from medical records.

3.1.2 Human factors

Any health-monitoring system involves human participants (patients and caregivers) and must necessarily trust these participants to carry out specific roles in each usage scenario [8], [24]. Some participants may be more trustworthy than others, for certain roles. Confidence in the sensor data derives from the level of trust in the participant(s), specifically, the system's ability to believe in the participant's identity (authenticity), responsibility (performing the role when expected), competence (performing the role correctly), and motivation (willingness to perform the role).

(Factor H1) Trust in patient: Consider the role of applying the sensor, which raises the following fundamental trust issues. Identity: are we sensing the right patient? Responsibility: does the patient regularly apply the sensor? Competence: does the patient tend to apply the sensor correctly? Motivation: does the patient have incentives to cheat? In some usage scenarios the patient may be the only participant involved, monitoring his own health. These trust issues ultimately affect the quality of data from the patient's sensor.

(Factor H2)Trust in caregiver: One or more caregivers are responsible for the initial configuration of a sensor, and (in some cases) for the periodic application or adjustment of the sensor. For instance, in the smoking cessation scenario data quality is affected by trust in the nurse who calibrates and provides the sensor to the insured. In other settings, a caregiver may be a physician, a technician, or a lay person such as a family member. The trust issues mirror those with patients.

3.1.3 System architecture factors

Some envision a three-tiered architecture for pervasive health monitoring: sensing, storing, and delivering health data [16]. The specific architectural choices will depend substantially on the needs of the situation. An architecture suitable for use in an emergency room is likely to be different from that used in an assisted living environment or a personal health monitoring system for an athlete. Regardless, the architecture must be robust and available to ensure timely delivery of data and secure to ensure data quality; we highlight three common factors here.

(Factor A1) Networking: From patient to caregiver, sensor data may travel on many networks: the patient's home network, public networks such as the Internet, or private networks such as coffee-shop

Wi-Fi networks. Despite the threats to sensor/system communications we desire data to arrive intact and without delay.

(Factor A2) Device platform: We anticipate that devices other than health sensors, such as the patient's mobile phone, will be involved in a typical deployment to provide computation and storage for the sensors. Data quality may be affected by the choice of device hardware and software platforms; for example, a platform with higher computation power can afford more sophisticated data-protection mechanisms. Confidence depends on the robustness and integrity of the device.

(Factor A3) Data pre-processing: System components pre-process sensor data for various purposes. *Data aggregation* combines multiple sensor values into a new statistical value (such as an average over time), or into an informative metric (such as an activity level from accelerometer data). *Data fusion* combines sensor data from multiple noisy sensors to derive information that is more concise and less noisy. Confidence in sensor data depends on the choice of aggregation or fusion methods and the location of the data processing along the data path.

4 Challenges

Were visit each of the factors described above, identifying the key challenges (denoted by[C]) involved in ensuring and assessing data quality and recognizing some of the technical wrinkles (denoted by [W]) that may have to be ironed out. The common research challenge across factors is assigning a metric to each. And, given some metric for each factor, how do we derive concise confidence metrics from multiple factors— and how does the passage of time modify our earlier confidence estimates? The resulting confidence level could expressed by a single metric (number between 0 and 1) or multiple metrics. How would the system present data confidence alongside data values, in a display meaningful to the data user?

4.1.1 Sensor challenges

(Challenge S1) Sensor design: We assume that a sensor's designed-in quality metrics (precision and accuracy) can be measured by an accredited lab that publishes the results as facts. [C1] When receiving the data, then, assessment reduces to a question of identification: how can the system authenticate the source of the data as being from a specific sensor model? [W1] If sensor data are collected in a storage device for later retrieval, what confidence do we have in the chain of custody for that device, and its integrity against tampering? [W2] If sensor data are cryptographically signed by a remote sensor, then what confidence do we have in the validity and integrity of the signing key? See also Challenge S5.

(Challenge S2) Sensor manufacture: The best solutions presumably rest on professional engineering standards and quality-metrics organizations. Similarly, the methods to assess confidence in a given manufacturer are non-technical, based on reputation or on a history of high-quality products. [C2] The technical challenge, however, is to find a way to *quantify* confidence in manufacturers, at least as far as needed to assess this factor alongside the others. [W3] How do we evolve our confidence measures, particularly on historical data, when there is new information about the manufacturer? [W4] How does quality of manufacture affect factory calibration of the sensor and how can this be factored into our belief in the calibration state of the sensor?

(Challenge S3) Sensor calibration: Calibration is necessary to configure a sensor to achieve its design specifications for precision and accuracy, and is initially performed by the manufacturer. Most sensors require periodic re-calibration, however, to accommodate natural drift in the sensor's capability or the effects of temperature, air pressure, or other environmental factors. Hospitals have trained technical

staff and a careful inventory system to ensure that all medical devices are tested and calibrated often. At home, the responsibility of getting the sensor recalibrated may rest on a patient or caregiver (Challenges H1, H2). Confidence in the calibration state of a sensor, then, reduces to the authentication of a calibration authority and the time since the most recent calibration.

If active mechanisms are possible, the system can trigger an instantaneous self-test in which a test signal is given to the sensor; the system can verify that the sensor detects the test signal or use the results of the self test to dynamically adjust future readings. How can the system accomplish this test in an environment with potential for adversaries to interfere? [C3] The key challenge is: how do we assess the confidence in the calibration state of the sensor? How do we know that the sensor been calibrated correctly and sufficiently recently? [W5] How do we represent the calibration results? If a sensor attaches a digital certificate with its sensor data, how should calibration results be encoded? How should the system validate this certificate and assess confidence from it? [W6] How do we assess our confidence in the calibration process? How do we model the accuracy of calibration, trust in the calibration authority and rate of calibration drift? How do we model environmental effects, and use input from auxiliary sensors (see Challenge S6)?For sensors that include a self-calibration mechanism, such as a scale that "zeroes" itself before use, how do we assess our confidence in the sensor's self-calibration, and the risks of tampering in that process (Challenges S4, S5)?

(Challenge S4) Sensor application: The sensor must be applied correctly and its position stabilized when the patient is mobile. To ensure proper application requires training of (and trust in) the medical personnel, caregiver, or patient who applies the sensor (Challenges H1, H2). [C4] The key challenge is: How can we validate that the sensor is applied correctly and remains stabilized in position? One approach is to use auxiliary sensors to validate correct application of the primary sensor. They might be packaged with the primary sensor or worn separately, or be embedded in the room (to measure ambient temperature or light). For example, the wrist device in Scenario 3 can make use of tilt sensors that ensure that the device is oriented correctly; pulse oximeters can be coupled with contact pressure sensors to ensure that optimum contact is maintained for reliable estimation of blood oxygenation. Another approach is to identify causal relationships between incorrect sensor application and lack of sensor data correlation[23].For instance, in Scenario 2 if one of the vital sign sensors exhibits a physiologically impossible change from past values the system can prompt the nurse to reapply the sensor. [W7] How can the assessment of sensor application be represented internally? Correct application of the sensor may depend on several parameters such as skin contact, orientation, lack of motion, pressure, or even environmental factors (for example, ambient light levels for pulse oximeters). If so, how can we combine individual assessments? [W8] Is the assessment static or dynamic? If dynamic, how does the time since last assessment factor into the metric? How do we trigger reassessment balancing the overhead involved with the need to ensure that the sensor is in place since it was applied? Do we reassess periodically or reassess when an event (e.g., motion) occurs? [W9] How do we know who applied the sensor and how does trust in the person who applied the sensor factor into the assessment of sensor application?

(Challenge S5) Sensor integrity: There is always the risk that a sensor, or its associated computing and communications capabilities, may be damaged or even manipulated to produce incorrect results. [C5] How can we ensure sensor integrity, using tamper-resistant hardware and secure embedded software? [W10] How can the sensor attest to its integrity, for example, through cryptographic statements made under the protection of a trusted hardware platform [32]? [W11] How can the system assess and quantify confidence in the sensor values, say, based on the attestations of the sensor's integrity? For devices that store data for later retrieval, how can we assess the integrity of the data while stored, and assess evidence of tampering with the sensor or device?

(Challenge S6) Sensor data correlation: Although assessment of the above factors may provide some degree of confidence in the sensor data, ultimately it is important to determine whether a given sensor value is somehow corroborated by other sensor values: either redundant sensors of the same type, complementary sensors of a different type, or known facts about the patient or the environment. [C6] How do we identify and model the correlations, and quantify confidence from the correlations observed? This requires a thorough understanding of the usage scenario, careful physiological study with human subjects, and a collaboration between bio-engineers and medical practitioners. Context awareness introduces possibilities for correlation. Context information such as patient motion or environmental changes can be used to correlate or correctly interpret the sensor data [33], [7], [20]. For example, if the motion sensor detects activity this information can be used to rationalize a sudden increase in sensed heart rate. [W12] How can we assess the degree of correlation among redundant sensors? Complementary multimodal sensors? How do we assess confidence in the sensor data across time? Measurement error or patient activity may disrupt one set of readings, but the prior and subsequent readings may fit the general trend. [W13] How can we combine these assessments into an assessment of overall correlation? How does the assessment weigh the relative importance of each correlation method? [W14] How do we represent the results of correlation to the system? Do we report the data with low confidence or reject non-correlating data? Correlation may also be used to address the challenges involved in some of the other factors such as the use of multi-modal sensors for ChallengeS4.Or,a non-correlating sensor may indicate a need for recalibration. How can we assess the reliability of these mechanisms?

4.1.2 Human challenges

A key question, involving both technical and non-technical considerations, is how the system should balance *trust* (expecting good behavior from the actors), *enforcement* (ensuring good behavior through sensor design or cryptographic protocols), and *assessment* (expecting good behavior, but assessing the results carefully). The right balance depends on the nature of the scenario, the motivations of actors, and the risk of incorrect decisions based on invalid data. The following discussion highlights some of the trust issues that impact data quality.

(Challenge H1) Trust in patient: We cannot *ensure* that the patient will be trustworthy, that is, that the patient will fulfill her role properly. We must therefore trust the patient to fulfill her role, whatever it might be, and then *assess* our confidence in the patient based on a priori information (such as the patient's prior history of compliance or capability with sensor devices) and based on dynamic information (such as data from contextual sensors that corroborate the primary sensor). [C7] How do we quantify confidence in the data based on fuzzy notions of trust in the patient [38], [30]?

[W15] What we may be able to ensure, to some degree, is the identity of the patient. How can we ensure that the sensor is applied to the correct patient, or that the sensed data is labeled with the correct identity when stored or transmitted? Regarding assessment, how can we determine whether the sensed data indeed comes from the desired patient? These challenges may be easier in some settings than others. In Scenario 2, for example, the identity of patient is established with bar-coded sensors and a one-time validation of observed vital signs of the patient. In Scenario 3, however, John may try to cheat by having someone else wear the sensor. In such settings, we may be able to use health-sensor data as a biometric identifier [6], [15], [14]. [W16] The patient's role may be minimal (e.g., an in-patient laying in a hospital bed may simply be expected not to remove the sensor device), or extensive (e.g., the glucose-monitor patient who must conduct the finger-prick test four times daily, and keep the embedded sensor dry). How can we assess whether the patient has fulfilled his or her roles responsibly and competently? How can we leverage information about past experience with this patient to do so? [W17] In settings where the patient may be motivated to provide incorrect data, how can we model these risks and use other evidence to validate that she has fulfilled her roles? [W18] Do we evolve trust in the patient based on new

information about the patient? If so, how does this affect data quality? In the case of long-term monitoring, for instance, if we know nothing about the patient *a priori* we may decide to make only minimal trust assumptions about her behavior and later evolve the trust as we learn from later interactions. For short-term monitoring, however, we may have to make an active effort to get useful information about the patient to establish a basis of trust.

(Challenge H2) Trust in caregiver: [C8] Ultimately, how do we model these different roles, and how do we assess and quantify the effects on our confidence in the data quality? How would the system know which caregiver is involved, and in what way? It may be hard to identify, let alone authenticate, a caregiver assisting the patient. [W19] As with the patient, above, how do we quantify confidence in the data based on fuzzy notions of trust in the caregiver? How do we relate this confidence to other factors, such as our confidence in the sensor's calibration state?

4.1.3 System architecture challenges

(Challenge A1) Networking: Since health information is sensitive, health-care providers are required to comply with HIPAA privacy policies [12]. Thus, the system should have no weak links that leak health information or that are susceptible to side-channel analysis, for example, discovering the number and nature of health sensors or medical servers in use by traffic analysis. [C9] How can we ensure the confidentiality and integrity of health-sensor communications in low-resource devices? [W20] How do we ensure availability of network links for timely arrival of data? How do we ensure robustness of the network in the face of faulty links, network latency or malicious denial of service attacks? [W21] Sensor devices may use a wireless network to communicate, such as Bluetooth or Wi-Fi. These network protocols (and their implementations) have known vulnerabilities; can we prevent (or at least detect) an adversary who cracks into a sensor device through one of these vulnerabilities? How can we provide high availability and low latency in the face of adversaries who jam wireless networks?

(Challenge A2) Device platform: [C10] The challenge is to develop mechanisms to protect data quality on the mobile platform and to assess their state when the device is used, and protocols for communicating that state to the system. [W22] How can trusted hardware(such as a Trusted Platform Module[32]) be used to secure the mobile platform?

[W23] How can the platform attest its state to the system? Remote attestation mechanisms on the device can significantly improve the reliability of the device and hence the system's confidence in the data quality. The precise form of attestation depends on the system's methods for assessing data quality.

(Challenge A3) Data pre-processing: Data pre-processing techniques can reduce false alarms that may be caused by outliers. For example, the inaccuracies introduced in certain physiological signals due to bodily motion, known as motion artifacts, can be modeled in different ways [35], [37]. These models help to recover the original physiological signal from the motion-distorted sensor data. [C11] What pre-processing techniques are useful, and how do we assess confidence in the result? [W24] Where do we perform pre-processing? On the sensor device itself, on a personal device that collects data from the body-area network, or in back-end servers? Or a combination? The choice impacts our confidence in data quality. [W25] How does the system recognize where and how data pre-processing has occurred? Do we trust the components that perform data pre-processing? [W26] How much trust does the system place in these data pre-processing services? How can the system assess confidence in derived data, depending on its trust in the pre-processing servers, without knowledge of all of the raw sensor readings? How can confidence assessments in the raw sensor readings be factored into confidence in the pre-processed data? How do we deal with the potential data loss?

5 Related Work

Living, in-patient monitoring, sleep apnea monitoring and continuous blood glucose monitoring. An analysis of the risks to data quality should begin with a deeper understanding of the needs of a specific usage scenario and their implications on potential deployments. In this section we introduce existing literature helpful in understanding the design space of pervasive health monitoring systems followed by a discussion of other frameworks that have been proposed to analyze threats to data quality in pervasive health monitoring.

5.1 The design space of pervasive healthcare systems

Muras et al. [21] present a novel taxonomy of pervasive health monitoring that helps understand the breadth of the problem space. The taxonomy is based on the international classification of functioning, disability, and health, and provides a framework for describing different categories of user requirements within the healthcare domain. The taxonomy identifies a set of properties to describe various types of pervasive healthcare systems and serves as a useful guide in understanding where the system fits in within the broad spectrum of healthcare applications and characteristics of its operating environment.

The US Department of Health and Human services has released a detailed use case for remote patient monitoring [24] that describes the requirements of the problem space, issues and stakeholders involved and identifies typical information flows. The scope of the use case includes remote collection and communication of physiological, diagnostic, device tracking information and "activities of daily living" information. In an effort to standardize care co-ordination among different organizations, the document identifies specific roles of different stakeholders. The document also outlines the issues and obstacles that have to be overcome for effective adaptation of the new healthcare paradigm by all stakeholders. Particularly valuable to our work are the descriptions of candidate information flows from monitoring device to patient's electronic health record. The discussion details the primary and contextual flows and identifies system capabilities that support the flow at each step. The document serves as an important first step in identifying fundamental vulnerabilities in remote monitoring infrastructure and addressing them suitably.

Varshney [34] identifies different flavours of health monitoring and classifies existing projects in that space. Geer [4] discusses non-invasive pervasive medical devices and opportunities for cost-effective improved healthcare. Kulkarni et al. [18] discuss the design space of pervasive healthcare in the context of body sensor networks of non-invasive, portable sensors. Halperin et al. [11] discuss the unique challenges presented by wireless implantable medical devices. Baker et al. [1] describe five different prototypes that converge to an effective healthcare paradigm design, including infant and firefighter vital-sign monitoring.

5.2 Data assurance in pervasive health monitoring

Several recent studies have analyzed different categories of risks in pervasive healthcare systems. These threats can be viewed from system security, patient privacy and data integrity standpoints. Our view of data quality is similar to the data-centric trust approach proposed for a vehicular sensor network [29], that is, factoring assessment of different categories of risks into confidence in the reported sensor data. Our data assurance framework provides a holistic view of the associated risks in a pervasive healthcare scenario so that suitable countermeasures can be employed. While there are useful overlaps with existing literature, to the best of our knowledge, our framework covers a broader spectrum of factors and relationships between the various factors can be explored.

As part of the on going work in the Trusted Software Systems and Services project, Presti et al. [26], [27], [28] have developed a framework for analyzing trust issues in a pervasive computing. Their view of trust is a human-centric, composite and evolving belief; hence, trust issues are considered from perspectives of the different stakeholders, including patients and caregivers. For them, "trust" comprises trust in system components, data components and subjective components. Their approach involves scenario analysis to highlight trust issues and categorize them into a proposed trust-analysis grid. The range of issues presented in the framework show significant overlap with our data assurance factors. These issues are, however, factored into human-centric trust rather than confidence in reported data. The proposed trust-analysis grid may be useful in the design of trustworthy systems, but it is not clear how trust could be quantified in an ongoing manner based on observations from an existing system.

Maglogiannis et al. [19] describe a Bayesian network modelling approach to performing a risk analysis of health information systems. The model concisely presents the causes of and interactions between undesirable events within the system to identify and prioritize risks based on probability of occurrence. They present a prototype patient monitoring system, namely the VITAL-Home System, developed and maintained for a private medical center (Medical Diagnosis and Treatment S.A.), and apply the proposed framework to identify and prioritize associated risks. The proposed model considers threats to the system assets and other vulnerabilities from a system architecture standpoint.

The Warfighter Physiological Status Monitoring (WPSM) [2] is part of the US Army's research effort towards reliable physiological monitoring for warfighters. A Bayesian network is used to assess the status of the soldier and report confidence in the diagnosis based on clinical uncertainty and system reliability diagnostics such as sensor failure.

The Advanced Instrumentation group at the University of Oxford is investigating the design of self validating sensors using online uncertainty metrics and developing prototype applications [13]. A self validating sensor performs a set of assessments regarding its internal state and consistency checks on measured values to report quality metrics, such as online uncertainty, along with its measurements. Although traditionally applied to sensors in mechanical control systems, Peter et al. [25] recently demonstrated the application of their sensor validation approach to a wearable system that measures physiological parameters for emotion sensing. Sensor data is validated against previously received data and stored information about the measured variable. Sensor device status is also validated using a self test. Each self-validating sensor reports sensor data together with uncertainty based on the two kinds of validation results.

Other related studies use quality-driven sensor data acquisition by exploiting relationships among sensor data to perform validation checks. Tatbul et al.[31]have proposed a data-confidence model-driven method for physiological sensor data acquisition, which reports data only if the confidence level is acceptable. The confidence is derived from other observations such as data from multiple sensors.

Several data-validation schemes to obtain high confidence data have also been proposed. The data fusion architecture proposed by Carvalho et al. [3] uses evidence from redundant and multimodal sensors to obtain high-confidence data. The proposed data-fusion architecture is applied to a prototype health monitoring application to obtain high-confidence heart rate measurements using pulse oximeter and ECG sensors. Donoghue et al. [22] propose a real-time sensor-data validation framework for a home health monitoring system by correlating data using known boundary values, values from other sensors and patient information. The data-validation reports are used to estimate sensor reliability and presented to the caregiver.

Depending on the nature of the physiological signal being sensed, knowledge about the dynamics of the sensed signal can be leveraged for validation. Several recent papers correlate ECG or heart-rate sensor

data with accelerometer data to obtain reliable readings in the presence of interference due to activity [5], [35], [36], [10]. Another recent paper applies clinical assessment techniques to mathematically model the accuracy of continuous glucose sensor data [17]. The factors considered in modeling the accuracy of the data are quality of calibration, physiology of glucose dynamics and sensor engineering. The C-BICC and MIMIC projects [23], [9] are investigating probabilistic models and machine-learning techniques for representing and reasoning about physiological data from critical-care sensors using knowledge of human physiology and sensor dynamics. Other related papers employ additional information, such as context of sensing to obtain reliable data from the health-monitoring sensors [33], [7], [20].

6 Summary

High-quality data is critical for many pervasive health-monitoring applications. We recognize that no system can ensure perfect data quality, and we highlight the need to assess *confidence* in the sensor data. In this paper we outline the key challenges related to ensuring or assessing the quality of sensor data in such applications. We identify six factors related to confidence in the sensors (Sensor Design, Sensor Manufacture, Sensor Calibration, Sensor Application, Sensor Integrity, Sensor Data Correlation), two types of factors related to human interactions(Trust in patient, Trust in caregiver),and three system architecture factors (Networking, Device Platform, Data Pre-processing). In the context of each factor, we identify and discuss research challenges in ensuring and assessing data quality. The actual impact of each factor (and the associated challenges) on data-quality assurance depends on the needs of the situation. Hence only a subset of these factors may be relevant in any given scenario. We recognize that to walk the fine line between enforcement and assessment it is important to understand the sources of risks and threats to data quality in each specific situation. The model of risks and threats for an actual deployment must account for the special needs of each usage scenario, from the point of sensor-data capture to presentation to the data user. As researchers, then, we must seek general-purpose frameworks that can capture and evaluate potential solutions. Ultimately, by resolving such challenges, we can help to provide quality healthcare in an effective and timely manner.

References

[1] Chris R. Baker, Kenneth Armijo, Simon Belka, Merwan Benhabib, Vikas Bhargava, Nathan Burkhart, Artin Der Minassians, Gunes Dervisoglu, Lilia Gutnik, M. Brent Haick, Christine Ho, Mike Koplow, Jennifer Mangold, Stefanie Robinson, Matt Rosa, Miclas Schwartz, Christo Sims, Hanns Stoffregen, Andrew Waterbury, Eli S. Leland, Trevor Pering, and Paul K. Wright. Wireless sensor networks for home health care. In *AINAW'07: Proceedings of the 21st International Conference on Advanced Information Networking and Applications Workshops*, pages 832–837. IEEE Computer Society, 2007.

[2] Maurizio Borsotto, C.T. Savell, Jaques Reifman, Reed W. Hoyt, Gavin Nunns, and Christopher J. Crick. Life-signs determination model for warfighter physiological status monitoring. Technical report, U.S. Army and GCAS Inc., Sept 2004.

[3] H.S. Carvalho, W.B. Heinzelman, A.L. Murphy, and C.J.N. Coelho. A general data fusion architecture. *Proceedings of the Sixth International Conference of Information Fusion*, 2:1465–1472, 2003.

[4] D. Ceer. Pervasive medical devices: less invasive, more productive. *IEEE Pervasive Computing*, 5(2):85–87, April-June 2006.

[5] Chung-Min Chen, Hira Agrawal, Munir Cochinwala, and David Rosenbluth. Stream query processing for healthcare bio-sensor applications. In *ICDE '04: Proceedings of the 20th International Conference on Data Engineering*, page 791. IEEE Computer Society, 2004.

[6] Sriram Cherukuri, Krishna K. Venkatasubramanian, and Sandeep K. S. Gupta. BioSec: A biometric based approach for securing communication in wireless networks of biosensors implanted in the human body.

In *Proceedings of the 2003 International Conference on Parallel Processing Workshops*, page 432, Los Alamitos, CA, USA, 2003. IEEE Computer Society.

[7] Ahyoung Choi and Woontack Woo. Context based physiological signal analysis in a ubiquitous VR environment. In Dongpyo Hong and Seokhee Jeon, editors, *ISUVR*, volume 260 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2007.

[8] FDA. FDA's human factors program. As viewed April 2008. http://www.fda.gov/cdrh/humanfactors.

[9] MIT Laboratory for Computational Physiology. Integrating Data, Models, and Reasoning in Critical Care project, A Bioengineering Research Partnership. Project web site, as viewed March 2008. http://mimic.mit.edu/index.html.

[10] Liliana Grajales and IonV. Nicolaescu. Wearable multisensor heart rate monitor. In *BSN '06:Proceedingsof the International Workshop on Wearable and Implantable Body Sensor Networks*, pages 154–157. IEEE Computer Society, April 2006.

[11] Daniel Halperin, Thomas S. Heydt-Benjamin, KevinFu, TadayoshiKohno, and William H. Maisel. Security and privacy for implantable medical devices. *IEEE Pervasive Computing*, 7(1):30–39, Jan.-March 2008.

[12] HIPAA. As viewed April 2008. http://www.hipaa.org.

[13] University of Oxford Invensys UTC, Department Engineering Science. Self Validating sensor project at University of Oxford. Project web site, as viewed March 2008. http://seva.eng.ox.ac.uk/self validation.html.

[14] Evangelos Bekiaris Ioannis G. Damousis, Dimitrios Tzovaras. Unobtrusive multimodal biometric authentication: The HUMABIO project concept. *EURASIPJournal on Advances in Signal Processing*, 2008.

[15] David Jea, Jason Liu, Thomas Schmid, and Mani B Srivastava. Hassle free fitness monitoring. In *Proceedings of the 2nd International Workshop on Systems and Networking Support for Healthcare and Assisted Living Environments (HealthNet)*, Jun. 2008.

[16] Andrew D. Jurik and Alfred C. Weaver. Remote medical monitoring. *Computer*, 41(4):96–99, 2008.

[17] Boris P. Kovatchev, Christopher King, Marc Breton, and Stacey Anderson. Clinical assessment and mathematical modeling of the accuracy of continuous glucose sensors (cgs). In *EMBS '06: Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 30, pages 71–74, Sep. 2006. *Communications Review*, 11(3):12–30, 2007.

[18] Prajakta Kulkarni and Yusuf Öztürk. Requirements and design spaces of mobile medical care. *SIGMOBILE Mobile Computing*

[19] I. Maglogiannis, E. Zafiropoulos, A. Platis, and C. Lambrinoudakis. Risk analysis of a patient monitoring system using bayesian network modeling. *J. of Biomedical Informatics*, 39(6):637–647, 2006.

[20] I. Mohomed, A. Misra, M. Ebling, and W. Jerome. Harmoni: Context-aware filtering of sensor data for continuous remote health monitoring. *PerCom 2008: Proceedings of the Sixth Annual IEEE International Conference on Pervasive Computing and Communications*, pages 248–251, March 2008.

[21] Joanna Alicja Muras, Vinny Cahill, and Emma Katherine Stokes. A taxonomy of pervasive healthcare systems. *Proceedings of the Pervasive Health Conference and Workshops*, pages 1–10, 2006.

[22] John O. Donoghue, John Herbert, and David Sammon. Patient sensors: A data quality perspective. In *Proceedings of the 6th International Conference on Smart Homes and Health Telematics*, pages 54–61, 2008.

[23] Department of Electrical Engineering and Berkeley Computer Science, University of California. Center for Biomedical Informatics in Critical Care (C-BICC) project at UC Berkeley. Project web site, as viewed March 2008. http://www.eecs.berkeley.edu/Research/ Projects/Data/102178.html.

[24] U.S. Department of Health and Office of the National Coordinator for Health Information Technology Human Services. Remote Monitoring, detailed use case. Detailed use case document published on March 21, 2008. www.hhs.gov/healthit/usecases/documents/ RMonDetailed.pdf.

[25] Christian Peter, Eric Ebert, and Helmut Beikirch. Awearable multi-sensor system for mobile acquisition of emotion-related physiological data. In *ACII '05: Proceedings of the 1st International Conference on Affective Computing and Intelligent Interaction*, Lecture Notes in Computer Science, pages 691–698. Springer, 2005.

[26] StephaneLo Presti, Michael Butler, Michael Leuschel, and Chris Booth. Atrust analysis methodology for pervasive computing systems. In *Proceedings of the 7th International Workshop on Trustin Agent Societies*, volume 3577 of *LNCS*, pages 129–143. Springer, 2004.

[27] StephaneLo Presti, Michael Butler, Michael Leuschel, Colin Snook, and PhillipTurner. Formal modelling and verification of trust in a pervasive application. Technical report, Trusted Software Agents and Services for Pervasive Information Environments, University of Southampton, June 2004. Available as project deliverable at url=http://eprints.ecs.soton.ac.uk/10183/1/TSAS -WP4-01 v1.pdf.

[28] Stephane Lo Presti, Mark Cusack, Chris Booth, David Allsopp, Mike Kirton, Nick Exon, and Patrick Beautement. Trust issues in pervasive environments. Technical report, Trusted Software Agents and Services for Pervasive Information Environments project, University of Southampton, Sept 2003. Available as Project deliverable at http://eprints.ecs.soton.ac.uk/10183/1/TSAS -WP2-01 v1. pdf.

[29] Maxime Raya, Panagiotis (Panos) Papadimitratos, Virgil Gligor, and Jean-Pierre Hubaux. On data-centric trust establishment in ephemeral ad hoc networks. In *INFOCOM '08: Proceedings of the 27th Conference on Computer Communications*, pages 1238– 1246. IEEE Computer Society, April 2008.

[30] Jatinder Singh, Jean Bacon, and Ken Moody. Dynamic trust domains for secure, private, technology-assisted living. In *ARES '07: Proceedings of the The Second International Conference on Availability, Reliability and Security*, pages 27–34. IEEE Computer Society, 2007.

[31] Nesime Tatbul, Mark Buller, Reed Hoyt, Steve Mullen, and Stan Zdonik. Confidence-based data management for personal area sensor networks. In *DMSN '04: Proceeedings of the 1st international workshop on Data management for sensor networks*, pages 24–31. ACM, 2004.

[32] Trusted Computing Group (TCG). Project web site, as viewed April 2008. https://www.trustedcomputinggroup.org/home.

[33] Surapa Thiemjarus, Benny Lo, and Guang-ZhongYang. Context aware sensing -what's the significance? In *Perspective in Pervasive Computing*, pages 163–170, October 2005.

[34] Upkar Varshney. Pervasive healthcare and wireless health monitoring. *Mobile Networks and Applications*, 12(2–3):113–127, 2007.