A Trust Evaluation Framework for Sensor Readings in Body Area Sensor Networks

Vinh Bui, Richard Verhoeven, Johan Lukkien, Rafal Kocielnik
Department of Mathematics and Computer Science
Eindhoven University of Technology
{t.v.bui, p.h.f.m.verhoeven, j.j.lukkien, r.d.kocielnik}@tue.nl

ABSTRACT

This paper addresses a framework to evaluate trustworthiness of a Body Area Sensor Networks (BASN), in particular, of sensor readings. We show that such trustworthiness is to be interpreted with respect to a certain statement or goal; its evaluation is based on quality aspects derived from observations and opinions from others. We examine relevant quality aspects of sensor readings which correspond to potential deviating behaviors of sensors. We then look at how to derive such qualities from observations taking uncertainty into the evaluation as well as decay over time. We develop an extension of subjective logic for this purpose and we show how we can compute quality properties without storing long time series. We then demonstrate this for two examples, including Galvanic Skin Response (GSR) and Electrocardiography (ECG) sensed data.

1. INTRODUCTION

We consider a monitoring and reasoning framework for evaluating trustworthiness of collected data in a Body Area Sensor Network (BASN) based on *Subjective Logic* [10, 14]. A BASN consists of sensor nodes mounted on or implanted in the human body together with a more powerful device which we call the *body hub*. This hub is capable of storing data and running application-specific components; it controls the BASN and acts as a single access point. In current technology, the hub is typically a smart phone.

A BASN is a subsystem of a larger system; typically, it connects to a back-end through the Internet for the purpose of data exchange and (re)configuration. In some designs there is little intelligence at the user side of the network except for bridging intermittent connectivity. However, in our vision a BASN can operate independently under control of its owner, including storing and managing data. Today initial explorations of BASNs are mostly in the domain of self-monitoring for well-being and sports achievements thus avoiding strict medical regulations, though we can already see a development in which health monitoring, initiated by either a medical specialist or the user himself, becomes part of the management of users' health. An important question then is to what extent a stakeholder (user, specialist) can rely on a BASN, i.e., how he determines its trustworthiness.

Trustworthiness, in general, refers to a relation among entities, where one relies on the other. At the user-system interaction level, this is commonly based on a high transparency and a clear interaction, complemented with reliable and correct operation. When used inside the system, trust and trustworthiness refer to relations between entities in the system itself. Trust and a trust model form the basis of system decisions (e.g., whether or not an operating system accepts an application, how it asserts the quality of outcomes or decides to take an action). In current systems such trust is mostly implicit in system operation like checking a certificate or asking a users' consent. In our work we choose to relate the trust to the subject of the decision. For example, one may trust a simple ECG sensor to deliver a heart rate service but not a detailed ECG. This also suggests that in the evaluation of such trust there are objective elements (facts like measurements, or the opinion of another entity on the subject matter) and subjective elements (the value a trustor assigns to such a fact). Overall, trust is built from some compositions of properties, typically determined through monitoring, and is related to the service that is expected. Following [3], trust is therefore defined as the degree to which a trustor has a justifiable belief that, in a given context, a trustee will live up to a given set of statements about its behavior.

In [3], a trust management model was presented along with a component-based architecture to maintain the BASN as a trustworthy platform under changes in applications such as dynamic extension and configuration of the BASN applications. The model allows to evaluate and monitor the trustworthiness at component, application, and system levels. In the current work we address three additional aspects, applied to trustworthiness of sensed data. First, we examine in more detail how to derive an opinion (which leads to a level of trust) from observation. Second, we examine the role of uncertainty in this, i.e., how to deal with trust-based decisions when there is no prior knowledge. We also propose a method to decay trust over time. Third, we propose a way to combine opinions that contain such uncertainty. For this we use the formalism of subjective logic.

The rest of the paper is organized as follows. First we examine the nature of the expected erroneous behaviors of sensor readings (whether they are malfunctioning or exhibiting malicious behavior) as observed by the body hub in Section

BODYNETS 2013, September 30-October 02, Boston, United States

DOI 10.4108/icst.bodynets.2013.253677

2. Based on these typical errors we define a means to evaluate the trust. We explain the operation of subjective logic for this purpose and extend it with new methods to obtain opinions from observation in Section 3. We demonstrate the resulting scheme for two properties of sensor readings for the examples of GSR and ECG sensors in Section 4. Section 5 gives an overview of the related work in the literature. Finally, Section 6 gives the conclusion and future work.

2. PRACTICAL ISSUES WITH SENSORS

2.1 Quality and physical aspects

A sensor is attached to an analog-digital-convertor (ADC) of a micro-controller, which might store the readings to flash or transmit them wirelessly. Depending on the battery level, the supply voltage to the ADC differs, such that the sensor readings are affected. The ADC driver affects the timing characteristics with respect to jitter and sampling period accuracy, such that noise filters (like a 50 Hz filter for ECG) do not work properly. For high sampling rates, the microcontroller might be insufficient to process, store or transmit the collected readings. The antenna configuration affects the communication range that results in connectivity problems.

The Galvanic Skin Response (GSR) is a measurement of electrical conductance of the skin that varies with moisture level [4]. The sweat glands, responsible for moisture level of the skin, are activated when an individual experiences emotional arousal. Real-life measurements of GSR pose a number of challenges related to artifacts removal and evaluation of trustworthiness of the signal. Artifacts in case of GSR are recordings which do not stem from emotion-induced changes in sweat gland activity. These may result from the recording procedure or from physiological responses in systems other than the electrodermal one [1]. Figure 1 depicts some of the most common recording artifacts experienced in real-life settings. The picture shows an example of a GSR data measurement from the DTI-2 wristband device [13] with a 10 Hz sampling frequency.

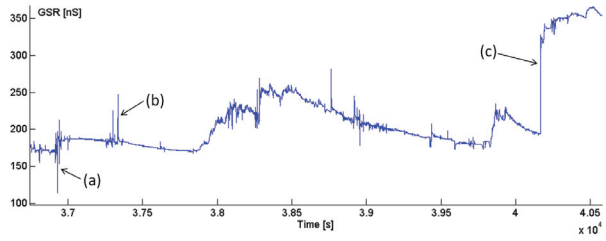

Figure 1: Examples of artifacts of GSR data measurement: (a) loss of skin contact; (b) motion artifacts; (c) repositioning of the electrodes.

Recording artifacts may stem from properties and sensitivity of the device that determine the minimal and maximal values of the GSR levels that can be reliably measured. Recordings that are too low may result from lack of connectivity with the skin (see (a) in Figure 1); too high values can be a result of an unexpected response to a strong stimulus that exceeds the measurement limits. Other types of artifacts can result from, so called, calibration marks. These often appear as jumps in the recording, followed by an exponential adaptation to the new baseline. Similar jumps can also result from repositioning of the electrodes on the skin (see (c) in Figure 1).

Physiologically based artifacts may result from skin movements directly beneath electrodes, but also from muscular activity even quite far from the recording site itself (see (b) in Figure 1). Additional artifacts may arise from pressure and stretching of the skin and from changes in skin blood flow. The GSR recordings can also be affected by a strong respiratory activity, gross physical movement or participant's speech activity.

We consider the GSR signal around these artifacts as less trustworthy; we want to develop methods to bring this into our trust model.

2.2 Malicious activities

The proper operation of the body hub and the sensors can be affected by malicious activities [6, 9]. A denial-of-service attack can target the operation of the body hub by sending excessive data, block the communication channel, or deplete the battery of the sensor by sending malicious requests. To pretend certain body behavior, a sequence of sensor data can be replayed by a malicious sensor or tool. In an on-off attack, a sensor alternates between periods of reporting correct and incorrect information, which is often harder to detect since it might be legitimate. A temperature or light sensor that alternates between positions in shadow and in direct sunlight results in an on-off behavior.

2.3 Abnormal readings

For most sensors, there exists a typical behavior. For a temperature sensor, the readings should be fairly stable when operated in a home or office, while for an ECG sensor, a periodic peak should occur. The described problems result in measurements which are inconsistent with the typical behavior, for example deviating from temperature measurements, but the behavior might also be the result of unknown phenomena for which no countermeasures exist. To detect abnormal readings, a generic specification of the expected behavior is required, for example, based on the maximum variance within a certain interval (for temperature), the expected variance within an interval (for ECG) or the expected frequencies (in the frequency domain).

3. TRUST EVALUATION SCHEME

3.1 Preliminaries

We use the definition of trust defined above for evaluating the trustworthiness of sensors, which are considered as trustees while the body hub and applications running on the body hub are trustors. Applying this definition to the relation between the body hub and sensors, the behavior and statements about it refer to the sensor operation, the sensed data, and properties thereof. The trust is then a measure of how the body hub judges trustworthiness of the sensor readings, while the justification of this pertains to the computation of the trust value, e.g., based on monitoring, data consistency, reputation or other means.

We consider the approach for trust evaluation based on subjective logic, in which the trust evaluation process follows certain evaluation steps to build *opinions* on quality attributes of the trustee. Opinions are the basis of subjective logic and they are used to express beliefs. Quality attributes may contain direct observations regarding the

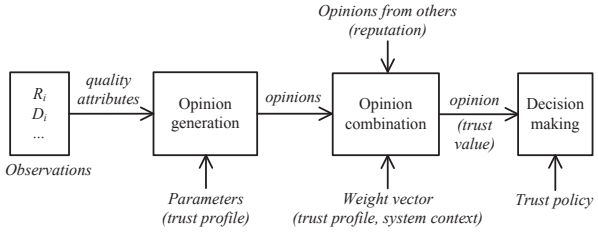

Figure 2: The trust evaluation scheme with two steps of the opinion generation and opinion combination.

technical performance of the sensor (e.g., availability, data integrity) or *recommendations* from other entities. Figure 2 shows the trust evaluation scheme proposed with two steps of the opinion generation and the opinion combination.

- *Opinion generation*: an opinion of each quality attribute is generated based on the direct observation of the trustee, e.g., the body hub observes quality attributes of the sensor readings. Since each sensor can have a different set of quality attributes and even a quality attribute can have different properties (e.g., the body temperature should be stable within a certain interval while an ECG should not), a particular generation function is implemented for each property or attribute. The opinion generation function can also be customized through a set of parameters, which are specified in the *trust profile* of the trustor. The values of these parameters might be different according to a given context.
- *Opinion combination*: the opinions of quality attributes are combined according to statements about the trustee's behavior or an expectation of the trustor. Optionally, opinions of other trustees or trustors (reputation), which are related to the given context, can be considered for the trust evaluation. Using the opinions of others helps to evaluate the trustworthiness more precisely and to detect more complex abnormal sensor readings. The combination is based on a weight factor (called *important rate* in [14]) of the attributes; this is similar to our trust evaluation method proposed in [3]. The weight factor is determined in the trustor's trust profile or is based on the system context.

An outcome of the final evaluation is an overall opinion or a vector of opinions, which are compared to a threshold for making a decision. The threshold and guidelines for decision making are specified in the *trust policy*.

We ignore typical security problems of message authentication and key distribution. There are many approaches and solutions in the literature for such problems.

3.2 Trust evaluation using subjective logic

Within opinions we want to distinguish trust, distrust and uncertainty. The latter captures a lack of knowledge; a high uncertainty means that we need to fall back to a blind trust. This is captured precisely within subjective logic. Subjective logic is a logic that operates on subjective beliefs about the world. It consists of a belief model called "opinion" and a set of binary operations for combining opinions. It is suited for our purpose for three reasons: 1) it supports reasoning from the perspective of a subject rather than aiming at an objective outcome, equal for all subjects; 2) it has mechanisms to combine such opinions; and 3) it allows us to take uncertainty into account and into this composition. Formally, an opinion denoted by $\omega_x^A = (b, d, u, a)$ expresses the relying party A's belief in the truth of statement x. For example, the opinion of the truth of statement "Sensor S is reliable to provide sensed data with quality q" can be interpreted as trust in S within the scope of the action or behavior "provide sensed data with quality q." Here, b, d, and u represent belief, disbelief, and uncertainty respectively, where $b, d, u \in [0, 1]$ and $b + d + u = 1$. Uncertainty is caused by the lack of evidence to support either belief or disbelief. The parameter $a \in [0, 1]$ is called the base rate, and represents the belief when no evidence is known; in addition it is used for computing an opinion's expectation value defined as $E(\omega_x^A) = b + au$.

Subjective logic presents several operators to combine opinions (see [10] for details). For example, *consensus* ($\oplus$) combines opinions of different entities on the same entity; *discount* ($\otimes$) generates an opinion of a recommendation or a chain of recommendations; *conjunction* ($\wedge$) deduces a combined opinion on the conjunction of two distinct opinions. An important one for our work is the *opinion generator*, used to generate opinions from real world observations on a certain statement. Let p be the number positive observations about the statement, n the number of negative observations, $\varepsilon \geq 1$ a parameter controlling the rate of loss of certainty, which can be used to tune the use of uncertainty in the model for the requirements of different scenarios, then:

$$b = \frac{p}{p+n+\varepsilon};\ d = \frac{n}{p+n+\varepsilon};\ u = \frac{\varepsilon}{p+n+\varepsilon} \qquad (1)$$

This definition of opinion generator has limitations: it does not take time into account and when evaluated over a large number of values it stabilizes. We will discuss these limitations and our proposed generation functions (e.g. addressing a limited history of values) in Section 3.3.

In addition to the above operators, the *adding operator* ($\sum$) is introduced in [14], which is used to aggregate opinions over a number of different sub-opinions. Let $QA = \{qa_1, qa_2, ..., qa_n\}$ be a set of attributes that may influence an entity's opinion on a statement. $W = \{w_1, w_2, ..., w_n\}$ ($\sum_{i=1}^{n} w_i = 1$) is a weight vector defining importance rate of attributes in QA, according to the entity's considerations. $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ is a set of opinions about the quality attributes. $\omega_i = (b_i, d_i, u_i)$ is the opinion on qa_i. The base rate $a_i = a$ for a given constant a. The aggregated opinion $\omega_\sum = (b_\sum, d_\sum, u_\sum)$, in which $b_\sum$ is defined as

$$b_\sum = \sum_{i=1}^{n} w_i \cdot b_i \qquad (2)$$

Similarly, $d_\sum$ and $u_\sum$ are defined.

3.3 Opinion generation

The opinion that trustor A (e.g., an application) has on sensor s_i is denoted by $\omega_i^A = (b, d, u)$ (a is a given constant). To simplify the notation, here we use b instead of b_i^A (similar to d and u). The corresponding trust value is defined as the opinion's expectation value: $tv_i^A = b + au$; this trust value is in the range $[0, 1]$. The opinion about each quality attribute of s_i is computed based on a set of direct observations about such quality attribute or other evidence. In

the scope of this paper, we consider two quality attributes: sampling reliability (R_i) and data integrity (D_i). We will develop the opinion generation functions (generator) for these quality attributes. The corresponding opinions are denoted by $\omega^A_{R_i} = (b_R, d_R, u_R)$ and $\omega^A_{D_i} = (b_D, d_D, u_D)$, respectively.

3.3.1 Sampling reliability

The body hub prescribes the sampling period ($\overline{p}_i$) at which sensor s_i should collect samples. Due to practical issues mentioned in Section 2.1, the actual sampling period might deviate from the requested period. Given a series of sampling times τ^n_i for sensor s_i with $n \geq 0$ (timestamps put on the data by s_i), we define $p^n_i = \tau^n_i - \tau^{n-1}_i$ to represent the actual sample period for sample n. The accuracy of sample n is then defined as:

$$\Delta p^n_i = |p^n_i - \overline{p}_i| \tag{3}$$

When the application requires a relative accuracy of α, the number of successful and failed samples are collected in rs_i and rf_i, respectively, according to the following formula:

$$\begin{cases} if \;\; \Delta p^n_i \leq \alpha \cdot \overline{p}_i & rs_i \leftarrow rs_i + 1 \\ else & rf_i \leftarrow rf_i + 1 \end{cases} \tag{4}$$

In (4), relative accuracy is used to make a decision and the complete history of sensor s_i is represented by the values rs_i and rf_i. With minor adjustments, the decision can be based on absolute accuracy, and the values of rs_i and rf_i can updated to reduce the effect of a long history. For example, the following formulas will be used.

$$\begin{cases} if \;\; \Delta p^n_i \leq \alpha \cdot \overline{p}_i & rs_i \leftarrow rs_i + 1;\; rf_i \leftarrow \gamma \cdot rf_i \\ else & rf_i \leftarrow rf_i + 1;\; rs_i \leftarrow \delta \cdot rs_i \end{cases} \tag{5}$$

where γ and δ are in $[0, 1]$. The γ and δ values indicate the level of pessimism and optimism and determine how fast disbelief or belief are lost when faced with success or failure, respectively. With fluctuating successes and failures, both rs and rf remain small and will converge to $\frac{1}{1-\delta}$ and $\frac{1}{1-\gamma}$, respectively, such that the uncertainty factor ε plays a more important role.

When success and failure are non-boolean values in the range $[0, 1]$ (e.g. based on the accuracy), the above formulas can be captured with the following single formula.

$$\begin{cases} rs_i \leftarrow \delta^f \cdot rs_i + s \\ rf_i \leftarrow \gamma^s \cdot rf_i + f \end{cases} \tag{6}$$

where $s = \alpha^{\Delta p^n_i / \overline{p}_i}$ and $f = 1 - s$; $s, f \in [0, 1]$. For $(s, f) = (1, 0)$ and $(s, f) = (0, 1)$, it results in the previous formulas.

The opinion $\omega^A_{R_i}$ of the sampling reliability attribute is computed from the number of successful and failed samples, according to the generation operator (see Equation (1)).

$$b_R = \frac{rs_i}{rs_i + rf_i + \varepsilon};\; d_R = \frac{rf_i}{rs_i + rf_i + \varepsilon};\; u_R = \frac{\varepsilon}{rs_i + rf_i + \varepsilon} \tag{7}$$

3.3.2 Data integrity

As indicated in Section 2.3, the data samples collected by a sensor will often follow a certain expected behavior. To determine the quality attribute for data integrity, the body hub uses the expected behavior for the samples recorded by sensor s_i, based on previous samples. Let $data_i$ represent the data stream of samples from sensor s_i, where $data^n_i$ represents the value of data sample n, for $n \geq 0$. For a stable or slowly changing signal, the expected value $expect^n_i$ will be close to the average value of the last N samples:

$$expect^n_i = \frac{\sum_{k=n-N}^{n-1} data^k_i}{N} \text{ , for } n \geq N \tag{8}$$

The number of previous samples (N) used to compute the average is a parameter. Based on this expected value the data integrity of each sample is determined. Similar to Section 3.3.1, the number of successful and failed samples are collected in is_i and if_i, according to the following formula:

$$\begin{cases} if \;\; |expect^n_i - data^n_i| < \beta & is_i \leftarrow is_i + 1 \\ else & if_i \leftarrow if_i + 1 \end{cases} \tag{9}$$

where β represents the maximum allowed absolute deviation from the expected value.

The opinion $\omega^A_{D_i}$ of the data integrity is then computed by considering is_i and if_i, using the generation operator. This is similar to the calculation in Equation (7).

The above quality attribute is relevant for sensors that are naturally stable (e.g., the body temperature), where the parameters N and β are specific for the type of sensor or the application. For sensor data that are naturally fluctuating (e.g., ECG) a different quality attribute needs to be used although it can still be used to detect extreme outliers.

3.4 Opinion combination

Multiple opinions or quality attributes, transitive opinions, or opinions from others are combined to an overall opinion (or a vector of opinions). The opinions or recommendations from others, in a similar context, can be considered to increase the accuracy of the evaluation process or to detect complex abnormal behaviors of the sensors. For example, the correlation between the data received from the internal accelerometers of the body hub (e.g., a smart phone) and the data received from external accelerometer sensor can be taken into account.

The combination operators mentioned in Section 3.2 are used in the opinion combination process. For example, we consider the combination of two opinions about the sampling reliability and data integrity attributes, $\omega^A_{R_i}$ and $\omega^A_{D_i}$ respectively. The overall opinion, $\Omega^A_i = (b_\sum, d_\sum, u_\sum, a)$, is determined according to the adding operator (see Equation (2)); for example $b_\sum$ is defined as follows.

$$b_\sum = w_R \cdot b_R + w_D \cdot b_D \tag{10}$$

where w_R and w_D represent the weights for reliability and integrity attributes, respectively. Again, the weight vector can be specified for each set of quality attributes in the *trust profile* or it can be adapted based on the given context.

To illustrate the opinion combination, we describe an example of the stress monitoring application (A), which monitors the stress level from the heart-rate data. The heart-rate data is provided by the ECG sensor (s). Let $\omega^A_s = (0.8, 0.1, 0.1)$ (the base rate $a = 0.5$) be the opinion of A about the data of s. B is the epilepsy monitoring application that detects epileptic seizures based on the heart rate. When B is installed on the system, B uses the heart-rate data of s. Assume B has referral trust in A, denoted by $\omega^B_A = (0.6, 0.1, 0.3)$. In certain preconditions [11] it can be concluded from (B trusts A) and (A trusts s) that (B trusts s). B's indirect trust in s can then be derived by discounting A's trust in s with B's trust in A. The derived trust is denoted by $\omega^{B:A}_s = (b^{B:A}_s, d^{B:A}_s, u^{B:A}_s)$. By using the discount

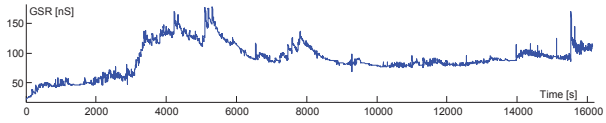

Figure 3: The GSR data measurement from the DTI-2 wristband with 2 Hz sampling frequency.

operator ($\otimes$), we have $\omega_s^{B:A} \equiv \omega_A^B \otimes \omega_s^A = (0.48, 0.06, 0.46)$. The effect of discounting is to increase uncertainty, i.e. to reduce the confidence in the expected value.

The application B then decides to evaluate the trustworthiness of s based on its direct observation about s. Assume that B has the opinion about s, denoted by $\omega_s^B = (0.4, 0.2, 0.4)$. The opinion $\omega_s^{A \diamond B} = (b_s^{A \diamond B}, d_s^{A \diamond B}, u_s^{A \diamond B})$ is then called the consensus between ω_s^A and ω_s^B, denoting the trust that an imaginary application $[A, B]$ would have in s. By using the consensus operator ($\oplus$), we have $\omega_s^{A \diamond B} \equiv \omega_s^A \oplus \omega_s^B = (0.78, 0.13, 0.09)$. The consensus operator reduces uncertainty, i.e. to increase the confidence in the expected value.

4. EXAMPLES

We demonstrate the trust evaluation according to two quality properties (sampling reliability and data integrity) of sensor readings for the examples of GSR and ECG data. These data are collected from the sensors by using the VITRUVIUS body sensor platform [2].

4.1 An example of GSR data

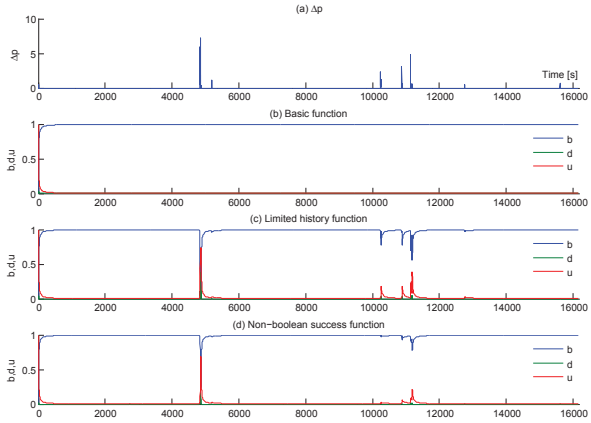

Figure 4: The opinion $\omega_R = (b, d, u)$ about the GSR data according to the sampling reliability attribute.

Figure 3 presents an example of the GSR data measurement from the DTI-2 wristband [13] with 2 Hz sampling frequency. In the first evaluation, opinions about GSR data according to the sampling reliability attribute are generated by using three different functions: the basic function (see Equation (4)), the limited history function (see Equation (5)), and the non-boolean success function (see Equation (6)). The following parameters of the generation functions are chosen: $\varepsilon = 10, \alpha = 0.5, \gamma = 0.5, \delta = 0.5$, and the expected sampling period $\bar{p} = 0.5$ (because the sampling frequency is 2 Hz). The accuracy of samples in term of the sample period (Δp) and the opinions $\omega_R = (b, d, u)$ are computed and shown in Figure 4.

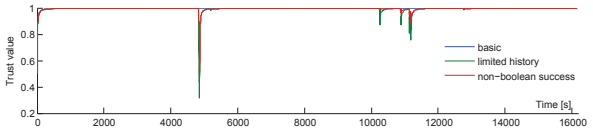

Figure 5: The trust values of sampling reliability for GSR data according to the basic, limited history, and non-boolean success functions.

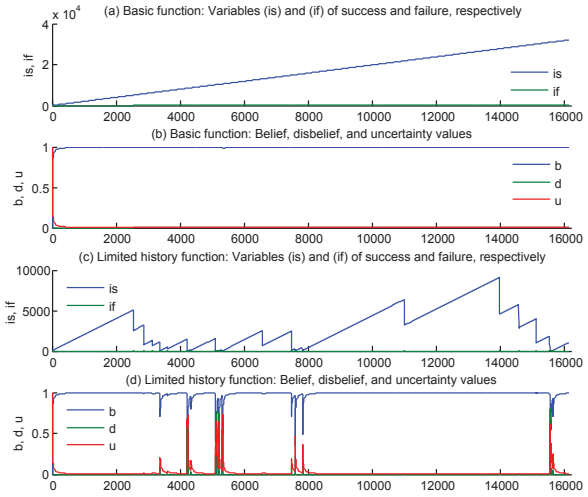

Figure 6: The opinion $\omega_D = (b, d, u)$ about the GSR data according to the data integrity attribute.

From the opinions (b, d, u) the corresponding trust values ($tv = b + au$, where the base rate $a = 0.5$) are computed. These trust values are shown in Figure 5.

In the second evaluation, the opinions are generated according to the data integrity attribute by using the basic function (see Equation (9)) and the limited history function (see Equation (5)). The following parameters are chosen: $\beta = 16$ and $N = 100$. The number of successful and failed samples (is and if) and the opinions $\omega_D = (b, d, u)$ are computed and shown in Figure 6. From the opinions the trust values are computed for the cases of the basic function and the limited history function. The corresponding trust values are shown in Figure 7.

In the above evaluations for both quality attributes, the limited history and non-boolean success functions are better to adapt to the quality of the sensor reading. Since the basic function uses the complete history of the sensor reading and the number of failed samples are dominated by the number of successful samples in such example, the values of (b, d, u) are mostly constant. The belief value b is mostly equal to 1 (though the belief decreases slightly when there are failed samples, but it is not visible on the figure because of the low resolution). In case of the limited history (or non-boolean success) function, the variables is and if in Figure 6.(c) do not reflect the success and failure occurrences. These variables are updated and controlled by the parameters γ and

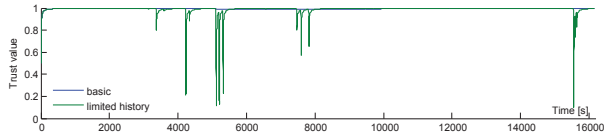

Figure 7: The trust values of data integrity for GSR data according to the basic and limited history functions.

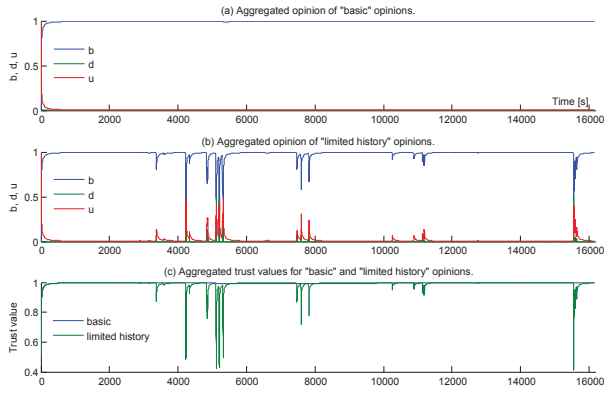

Figure 8: The overall opinions and trust values, based on combining the opinions on the sampling reliability and data integrity.

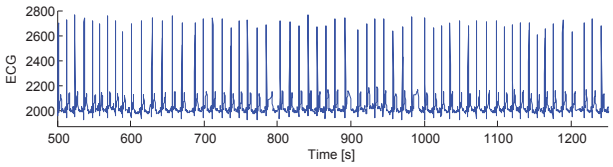

Figure 9: The ECG data measurement from the Shimmer sensor with 100 Hz sampling frequency.

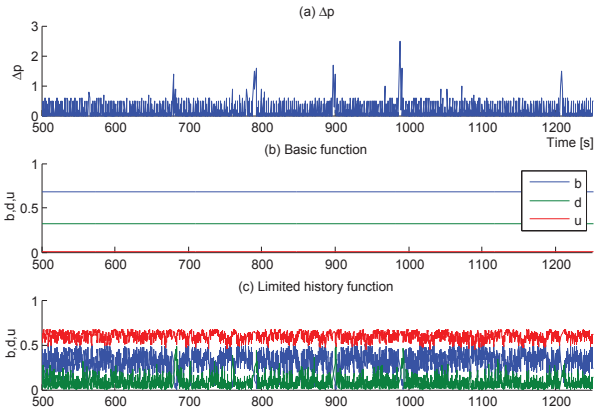

Figure 10: The trust evaluation for ECG data using the sampling reliability attribute.

δ to reduce the effect of a long history. For example, the value of is decreases to $\delta \cdot is$ when there is a failure (i.e. the trustor forgets part of the success). The value of δ determines whether the trustor is optimistic or pessimistic.

We note that these functions do not provide an effective evaluation against all possible malicious activities or abnormal readings. For example, to detect a replay attack, an algorithm is required to determine the similarity between recent and historical sensor readings. When the similarity is expressed as a (non-boolean) success value, it can be used as input for an opinion generators.

We demonstrate the opinion combination by using the adding operator ($\sum$) to aggregate the overall opinion over two opinions (ω_R and ω_D). The combination is based on the importance rate of the quality attributes. We assume that the GSR data is used by a stress monitoring application to detect stress levels of a user. From this application perspective the sampling reliability (qa_R) is less important than the data integrity (qa_D), and the corresponding weight vector of these properties is $W = (w_R, w_D) = (0.35, 0.65)$. Figure 8 plots the overall opinions and the corresponding trust values, where the opinions for sampling reliability and data integrity are combined for the basic and limited history functions. The overall trust values provide the application with a better evaluation of the trustworthiness of the data. For example, in Figure 8.(c), the trust value of the GSR data is decreased at time 11000, which is caused by a reduced belief in the sampling reliability. However, the reduced trust is not visible in Figure 7.

4.2 An example of ECG data

Figure 9 plots the ECG data measurement from the Shimmer ECG sensor [12] with 100 Hz sampling frequency. Similar to the above example of the GSR data, we evaluate the trustworthiness of the ECG data by considering the sampling reliability attribute with two opinion generation functions: the basic and the limited history functions. The following parameters are chosen: $\varepsilon = 10, \alpha = 0.5, \gamma = 0.65, \delta = 0.65$, and the expected sampling period $\bar{p} = 0.01$ (because the sampling frequency is 100 Hz). The Δp and the corresponding values of belief, disbelief, and uncertainty (b, d, u) are computed and shown in Figure 10. The basic function results in a similar behavior (the values of (b, d, u) are mostly constant) as in the example of the GSR data. The belief and disbelief values are around 0.7 and 0.3, respectively, which indicates alternations between successes and failures. In case of the limited history function, those alternations result in low values for rs and rf, such that the uncertainty factor ε has a more dominant role. As shown in Figure 10.(c), the uncertainty gets a higher value than the belief and disbelief.

The corresponding trust values computed from the opinions are shown in Figure 11.

5. RELATED WORK

In wireless sensor networks (WSNs), trust management is specifically useful to manage the trustworthiness of interactions among network entities [15, 7]. The authors in [15] presented a security framework with trust management that is used to secure sensor networks. A distributed trust model enabling recommendation-based trust and trust-based recommendation has been proposed, to build reasonable trust relationships among network entities. In [8] the authors proposed a trust management model for evaluating the trust value of beacon nodes that provides the location information

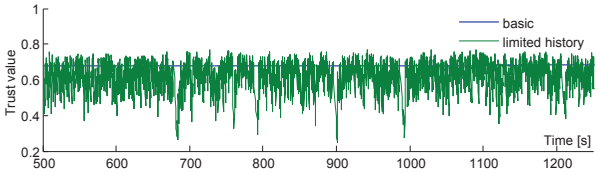

Figure 11: The trust values of ECG data according to the basic and the limited history functions.

for a BASN. The model is based on a distributed reputation-based trust model, in which each beacon node stores its past direct interaction experiences (i.e. the number of correct and incorrect location broadcasts from each neighboring beacon node) locally. Each beacon node evaluates the reputation of its neighbors by combining its observed past evidences and the first-hand evidence. The BASN then uses the reputation values received from beacon nodes to compute the overall reputation (trustworthiness value) of each beacon node and the location information from trusted beacon nodes will be used to estimate the current location of the BASN.

In [9] the authors claimed that most of the trust management models proposed for MANETs and WSNs are not suitable for BASNs since they do not consider the unique operational and security requirements of BASNs. For example, the trust management protocols in [5] compute trust in a fully distributed manner, in which each sensor node monitors the behaviors and manages the trust records of other nodes. These mechanisms incur high costs on sensor nodes in terms of processing power, memory, bandwidth, and energy consumption. The authors also described their attack-resistant and lightweight trust management scheme named ReTrust. The trust is established based on both the direct observation and recommendations from other entities according to the number of successful and failed interactions between nodes in the network. The authors however did not describe how to compute these successful and failed numbers. In our proposed framework, the opinion concept of subjective logic is used as the basis for trust evaluation, which takes into account the uncertainty factor. Subjective logic also provides several operators for combining both direct and indirect trust values.

6. CONCLUSION

This paper investigates the nature of observed sensor errors in a body sensor network. It proposes a framework to evaluate the trustworthiness of sensor readings from quality attributes (with sampling reliability and data integrity as examples) without storing long time series. The trust evaluation is performed through two steps: the opinion generation and the opinion combination. The extension of subjective logic is developed for this purpose. The trust evaluation was demonstrated with two examples of the GSR and ECG sensed data. It was shown that the opinion generation function with the limited history is better to adapt to the quality of the sensor readings. In addition, the function's parameters give a trustor the ability to select optimistic or pessimistic behavior according to the system context. The combination of opinions helps to evaluate the trustworthiness more precisely and to detect the complex sensor errors.

Current and future work is on extending the set of quality attributes thus specifying abnormal readings more precisely for different types of sensors. Moreover, cause-detection and control mechanisms will be investigated. Feedback loops using these mechanisms aid to maintain the trustworthiness.

Acknowledgment

We would like to thank IOP GenCom and the VITRUVIUS project for the financial support for this work.

7. REFERENCES

[1] W. Boucsein. *Electrodermal activity*. Springer, 2011.

[2] V. Bui, R. Verhoeven, and J. Lukkien. A body sensor platform for concurrent applications. In *IEEE Int. Conf. on Consumer Electronics - Berlin (ICCE-Berlin)*, pages 38–42, 2012.

[3] V. T. Bui, J. J. Lukkien, and R. Verhoeven. Toward a trust management model for a configurable body sensor platform. In *Proc. 6th Int. Conf. on Body Area Networks (BodyNets 2011)*, pages 23–26, 2011.

[4] C. W. Darrow. The rationale for treating the change in galvanic skin response as a change in conductance. *Psychophysiology*, 1(1):31–38, 1964.

[5] S. Ganeriwal, L. K. Balzano, and M. B. Srivastava. Reputation-based framework for high integrity sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 4(3):1–37, June 2008.

[6] S. Ganeriwal and M. B. Srivastava. Trustworthy sensor networks: Issues and challenges. Technical report, UC Los Angeles: Center for Embedded Network Sensing, 2004.

[7] L. Gomez, A. Laube, and A. Sorniotti. Trustworthiness assessment of wireless sensor data for business applications. In *Int. Conf. on Advanced Info. Networking and Applications*, pages 355–362, 2009.

[8] Z. S. Han Yu and C. Leung. Towards trust-aware health monitoring body area sensor networks. *Int. Journal of Information Technology*, 16, 2010.

[9] D. He, C. Chen, S. Chan, J. Bu, and A. V. Vasilakos. Retrust: Attack-resistant and lightweight trust management for medical sensor networks. *IEEE Trans. on Info. Tech. in Biomedicine*, 16(4):623–632, 2012.

[10] A. Jøsang. A logic for uncertain probabilities. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9(3):279–311, June 2001.

[11] A. Jøsang, R. Hayward, and S. Pope. Trust network analysis with subjective logic. In *Proc. 29th Australasian Computer Science Conference*, ACSC '06, pages 85–94, Australia, 2006.

[12] Shimmer research. http://shimmer-research.com/.

[13] J. Westerink, M. Ouwerkerk, G.-J. de Vries, S. De Waele, J. van den Eerenbeemd, and M. van Boven. Emotion measurement platform for daily life situations. In *3rd Int. Conf. on Affective Computing and Intelligent Interaction*, pages 1–6, 2009.

[14] Z. Yan and C. Prehofer. Autonomic trust management for a component-based software system. *IEEE Trans. on Dependable and Secure Comp.*, 8(6):810–823, 2011.

[15] Z. Yao, D. Kim, I. Lee, K. Kim, and J. Jang. A security framework with trust management for sensor networks. In *1st Int. Conf. SecureComm*, pages 190–198, sept. 2005.