Trustworthiness Assessment of Wireless Sensor Data for Business Applications

Laurent Gomez
SAP Research
805, Docteur Maurice Donat
06250 Mougins, France
laurent.gomez@sap.com

Annett Laube
SAP Research
805, Docteur Maurice Donat
06250 Mougins, France
annett.laube@sap.com

Alessandro Sorniotti
Institut Eurécom
2229, Route des Crêtes
06560 Valbonne, France
sorniott@eurecom.fr

ABSTRACT

Nowadays, Wireless Sensor Networks are mature enough to be used by Business Applications. These applications rely on trustworthy sensor data to control business processes. We propose an approach to assess the trustworthiness of sensor data during its lifecycle from acquisition on the node, over processing and to routing to the Business Application. We rely on the Subjective logic framework to compute the probability that sensor data are trustworthy enough to be used by an application. With the definition of new operators for the subjective logic, we develop a trust model, that allows to detect erroneous sensor data which are originated either unintentionally by malfunctioning of sensor nodes or intentionally by attackers.

General Terms

WSN, Trust Assessment, Subjective Logic

1. INTRODUCTION

Nowadays, Wireless Sensor Network technology is mature enough to be used by Business Applications. An increasing number of applications are developed in several business domains, reaching from defense, over public security, manufacturing and traffic control to health care [**?**]. The particular interests lay in the ability of Wireless Sensor Networks (WSNs) to control and monitor different physical environments.

Business Applications (BAs) have a strong need to assess the trustworthiness of the data delivered by WSN. When WSNs are integrated into business processes, the delivered data can influence severely the decisions in the applications and the taken actions in the real world. In the worst case, the life of people can be endangered. A good example is the remote patient monitoring: a patient is equipped with a Body Sensor Network (BSN) **??** which monitors body performance (e.g. heart rate, body temperature, SPO2, etc.) and activities (e.g. walking, running, falling). Erroneous sensor data can lead to a wrong therapy for this person. If an emergency stays undetected, it can result in the dead of the person.

Erroneous or non-trustworthy sensor data can have 2 different reasons: *Intentional misbehavior* and *unintentional errors*.

Unintentional errors of the sensor data are caused by malfunction of the hardware (broken or obstructed sensors), malposition of the node (untied or incorrectly attached node) or exhausted batteries.

Intentional misbehavior is caused by attackers, exploiting security vulnerabilities of WSNs, e.g. in the routing protocols. [**?**] Security for WSN has often to be balanced with the requirements of energy saving and the limited resources (memory, CPU) that are available. But, implementating 'cheaper' security solutions, like symmetric key algorithms, can open more exploitable flaws. Related to the capability of easy deployment and mobility, WSN node are often easily accessible and rarely tamper-resistant. Hijacking of nodes and extraction of cryptographic material is quite easy and gives the attacker the possibility to add malicious nodes or inject bogus data into the network.

Instead of hardening the security in WSN, assessing the trustworthiness of sensor data is an alternative solution. The goal of the trustworthiness assessment is to assist business applications in decision making whether they can rely on the data or not. Trustworthiness is the probability that a sensor data really corresponds to the measurement in the physical world. This approach has two main advantages: (i) it allows business applications to separate erroneous sensor data from trustworthy ones, and (ii) it supports energy optimisation in WSNs.

Energy saving can be achieved when non-trustworthy sensor data is filtered out already on the sensor or within the network considering that data transmission is the most energy consuming task in a WSN **??**.

In this paper, we propose a framework for the trustworthiness assessment of sensor data, from the acquisition on the node, to their delivery to business application, including any intermediary routing or processing. The assessment aims at identifying erroneous sensor data caused intentionally or unintentionally.

The paper is organised as follows. In Section 2 we discuss the state of the art. In order to propose our trust model in Section 4, we introduce the notion of sensor data life cycle in Section 3. Section 5 is dedicated to the use of subjective logic for our model. In this section, we define new operatorsin order to support the trustworthiness assessment at acquisition, processing and routing time. At last, we evaluate the feasibility of our approach in Section 6. Futur work and conclusion are discussed in Section 7.

2. RELATED WORK

According to [**?**], sensor node data can be compromised in two different ways: (i) unintentionally or (ii) intentionally. Unintentional errors are cause by malfunction of hardware, misposition of the nodes or battery exhaustion and decay. In literature, several approaches can be found for sensor node failure detection (see Subsection 2.1).

The deliberated exploitation of node vulnerabilities can intentionally cause erroraous sensor date. Many security mechanism rely on the assupmtion that sensor nodes are tamper-resistant devices []. Due to the nodes resource restriction, the support of strong security mechanism on sensor node is hardly envisioned. Eventhough security mechanism are tailored to those constraints, it is still quite easy capture node and to get the embedded cryptographic material [**?**]. Based on those assumption, several attacks are possible: the injection of malicious nodes with proper cryptographic material, the interception and forgery of sensor data on the wireless channel, etc. The impact of the exploitation of such threats on business applications can be severe consequences.

In [18], the authors aims at identifying compromised nodes. They develop an alert reasoning algorithm where each node evaluates its security estimation on its neighbors. This approach is only based on authentication and analysis of data delivered by sensor nodes. Furthermore, the authors do not study influence of compromised sensor data involved in further processing.

Other authors (discuses in Subsection subsec:RepSystems) use distributed reputation and trust systems to detect unreliable sensor node. We consider also approach for mobile ahhoc networks (MANETs) which address similar problems.

2.1 Failure Detection

Within a WSN, sensor nodes are prone to different kinds of failure [16]. : crash, omission, timing, value or arbitrary. Crash or omission imply none response from the sensor to sensor data request. Timing refers to timeout during request processing. Value failure deals with delivering incorrect value due to malfunctioning or compromised sensor nodes. Finally, arbitrary failures include all the types of failures that cannot be classified in previously described categories [13, 12]. In order to improve resilience to sensor node failure, a few approaches propose failure detection mechanisms. In sensor node failure detection, nodes can either detect its own failure [8] (e.g. based on battery exhaustion), or its neighbors failure [7].

2.2 Reputation and Trust Systems

Reputation systems have been developed in order to identify compromised nodes, based on the behavior. Reputation is based on a collection of evidence of good and bad behavior undertaken by other entities. In [4], the authors capitalize on Bayesian formulation of reputation representation, updates, integration and trust evolution in a Reputation based Framework for Sensor Network (RFSN). The latter addresses bad mouthing and ballot stuffing attacks. Thus in this approach, only good behaving nodes can get access to others nodes information. In [2], the authors also propose a reputation system based on Bayesian approach. They clearly distinguish the reputation from trust in sensor nodes. The former represents the opinion formed by a node on another nodes and the latter the opinion formed by a node about how honest another nodes are. In this approach, each node is in charge of maintaining its reputation and trust rating on its nodes of interest (e.g. the ones that it is interacting with).

At the contrary, in trust systems, the objective is to evaluate the probability exceptation that a given event occured. In the case of sensor data, this event woudl be that the sensor data really reflects the actual physical environment context. In [14], the authors propose a trust-and-clustering based framework based on public key authentication for mobile ad hoc wireless networks. They define a decentralised trust model where each nodes monitor and rate each other with quantitative trust values. The goal is thus to discover and isolate dishonest nodes. A chain of trust is established among nodes similar to Pretty Good Privacy (PGP). Any nodes can sign another's public key with its own private key. The authors developed a trust-and-clustering based public key authentication mechanism supported by new security operations on public key certification and update of a trust table. In [19], the authors propose a trust based framework for secure aggregation in wireless sensor network based on Bayesian model and beta distribution probability. Based on subjective logic, this framework computes an opinion which encompasses belief and uncertainty on the aggregation of sensor data with the consensus operator. Still those approaches are restricted to the trustworthiness evaluation of origin and value of data. Besides, in [19], the authors use the consensus operator, which computes a fused opinion of several entities about a same event. Consensus then appears not to be the most appropriated operator (see section **??**).

Nevertheless, all existing approaches barely address erroneous data processing. Apart [**?**] which considers data aggregation, they all restrict themselves to detection of raw bogus sensor data. In addition, those approaches evaluate the reliability of data based on their origin or value. But none of them target the confidence assessment of data processing (e.g. filtering, aggregation, fusion) based on potentially bogus data. {FIX1}

3. SENSOR DATA LIFE CYCLE

As identified in Section 1, business applications need to evaluate the trusthworthiness of sensor data in order to distinguish between erroneous and trustworthy sensor data. To the best of our knowledge (see Section **??**), existing approaches base their confidence on the raw sensor data and take only sensor data values and/or origin into account. When it comes to data routing or processing, none of those

[1] FIX. **ALR:** Do not forget to mention BARAK-2008

Figure 1: Sensor Data Life Cycle

approaches target the trustworthiness of sensor data.

In order to address this problem, we formalise the life cycle of sensor data, from its acquisition on the sensor nodes, to their delivery to business applications, including intermediary processing or routing from entity to entity (e.g. nodes, middleware, business application). 1.[{FIX2}] In this life cycle, we identify three states for sensor data: (i) raw, (ii) routed, (iii) processed. A sensor node firstly produces a **raw** sensor data which is transmitted to a radio/processing board attached to the node. Usually a bunch of sensor nodes are attached to the same radio/processing board **??**. Nevertheless, a sensor data is **raw** as soon as it has been acquired (sensed) by a node without any additional routing or processing. With processing, we understand any data manipulation such as filtering, fusion or aggregation. As soon as a sensor data is sent to other node in the WSN, it is considered as **routed**. The data has been delivered from an entity (e.g. node, forwarding node, middleware) to another entity (e.g. forwarding node, business applications).

In the remaining of this paper, we will note e_i $i \in [0, n]$ the set of entities on which the sensor data is routed to.[{FIX3}] e_0 is the production node [{FIX4}] , and e_n is a business applications. Last but not least, a sensor data can be the result of a processing (e.g. fusion, aggregation) of multiple sensor data. Sensor data can be processed several time on the same entity.

4. TRUST MODEL FOR SENSOR DATA

Based on the sensor data life cycle described in Section 3, we define now our trust model. Trust [11, 5] is commonly defined as the propability expectation of an entity (e.g. node, sink, middleware, applications) that an event occurs or that a proposition becomes true.

In Section 3, we identified three states in the sensor data life cycle: raw, routed and processed. Following this definition, we have three possible propositions P_x: (i) P_{raw}: *"the raw sensor data is trustworthy enough to be used"* , (ii) P_{routed}: *"the routed sensor data is trustworthy enough to be used"* and (iii) $P_{processed}$: : *"the processed sensor data is trustworthy enough to be used"*. The trust assessment of each of those three propositions is impacted by different parameters. For example, the trustworthyness of raw sensor data is impacted by type of measurement, whereas routed data trustworthiness is impacted by the trust relationship between the source and destination nodes[{FIX5}].

We formalise trust along the three sensor data states as follows:

DEFINITION 1. ***(Trust)***
Let e be an entity. Trust $\psi(P_x, e)$ is e's expectation probability that P_x is true where $\psi(P_x, e) \in [0, 1]$.

[2]FIX. **ALR:** Remove Ei form figure!
[3]FIX. **ALR:** not correct, while nothing is routed to e0
[4]FIX. **ALR:** Think about terminology
[5]FIX. **ALR:** better depending on the used routing protocol?

Figure 2: Sensor Data

In the remaining of this section, we formalise the trust of raw, routed and processed sensor data, and identify the specific parameters for the trust assessment along the sensor data life cycle.

4.1 Trust Assessment of Raw Sensor Data

As depicted in Figure 2, a sensor data is a composition of a type (e.g. ambient temperature, pulse) and a list of attributes (e.g. accuracy, value, origin). We formalize raw sensor data as follows

DEFINITION 2. ***(Sensor Data Attribute)***
Let a sensor data attribute be a pair
$a =< a_{type}, v >$ where
a_{type} is the attribute type and v its value.

DEFINITION 3. ***(Sensor Data)***
Let a piece of sensor data s be a pair $< s_{type}, < a_i >_{i=0}^{n} >$ where
s_{type} is the sensor data type and
$< a_i >_{i=0}^{n} =< a_{ti}, a_{vi} >_{i=0}^{n}$ is the list of its n attributes where a_{ti} and a_{iv} are type and value of attribute a_i and such as for each i=1,..,n, it exists only one sensor data attribute a_i of type a_{ti}.

We have for example patient's body temperature represented as follows:

```
s = < bodytemperature,
     < <value, "37.5">,
       <metric, "celsius">,
       <typeofmeasurement, "behind ears">,
       <origin, "sensor1234">,
       <accuracy, "+-0.5">
    >
```

We consider that each sensor attributes have an impact on trust assessment of raw sensor data, eventhough the influence of metric can be neglected in comparison to the origin of the data. We note $P_{(a_i,e)}$ the impact of the attribute a_i on sensor data trust assessment, and α_i its weight. The trust assessment of a raw sensor data is then defined as follows:

DEFINITION 4. ***(Trust Assessment of Raw Sensor Data)***
$\psi(P_s, e) = \psi_{acquisition}(\psi(P_{a_i}, e), \alpha_i)$
where i=1,..,n
and α_i is the weight of the a_i attribute on $\psi(P_s, e)$.

4.2 Trust Assessment of Routed Sensor Data

We note the process of sensor data routing s from an entity e to an entity f as $s\ |^{e \to f}$. Each entity e has a confidence, $\psi(P_f, e)$, in the entity f delivering data. The trust assessment of the delivered sensor data from entity f to e is expressed as follows:

DEFINITION 5. ***(Trust Assessment of Routed Sensor Data)***
$\psi(P_{s|f\to e}, e) = \psi_{delivery}(\psi(P_s, f), \psi(P_f, e))$

4.3 Trust Assessment of Processed Sensor Data

Numerous sensor data processing are available along the life cycle from on-the-fly average to aggregation including fusion of sensor data, mainly depending on resource capabilities of the processing entity. Sensor data processing is the result a computation over a set of sensor data $< s_i > i = 0, \ldots, n$. Likely the trust assessment of raw sensor data, we weight each sensor data of $< s_i > i = 0, \ldots, n$. Their impact on the sensor data processing can vary. To that purpose, we weight each sensor data with α_i.

DEFINITION 6. ***(Trust Assesment of Sensor Data Processing)***
$\psi(P_{(s_i)_{i=0}^n}, e) = \psi(\psi(P_{s_i}, e), \alpha_i)_{i=0}^n$
where i=1,..,n the number of aggregated sensor data and α_i *is the weight of* s_i *sensor data in* $\psi(P_{(s_i)_{i=0}^n}, e)$.

$\psi(P_{(s_i)_{i=0}^n}, e)$ *is called trustworthiness evaluation of sensor data aggregation* $\bigoplus(s_i)_{i=0}^n$.

5. A SUBJECTIVE LOGIC APPROACH

In section 4, we model trust of sensor data from acquisition from nodes to delivery to business applications including aggregation and fusion. In this section, we implement our trust model mapping trust assessment to subjective logic opinion.

5.1 Subjective Logic

Subjective logic is a theoretical framework based on Dempster-Schafer theory of evidence [15]. In subjective logic, we manipulate opinions about proposition P. An opinion is represented by the 4-tuple *(b,d,u,a)* when a represents the a *priori* probability in absence of opinion. As we only consider binary state space for P, we set a to $\frac{1}{2}$. Respectively, b, d and u represent the belief that P is true, the belief that P is false, and the uncertainty is the amount of belief that is no committed to the truth or falsehood of PŠs. The range of those four value is [0,1] where b+d+u=1. The opinion of A about P is defined as ω_P^A = b+a.u. Moreover, subjective logic framework provides a set of logical operator for combining opinions such as conjunction, disjunction and negation, in addition to non-traditional operators for consensus or discount of opinions. .

Applied to trust evaluation of sensor data, it consists of determining opinion in the following proposition : *'a sensor data is trustworthy enough to be used'*. In addition, subjective logic allows to represent the uncertainty with respect to sensor data measurement (e.g. quality of service, accuracy of node). Josang's model [10] is thus suitable for measuring uncertainty with respect to sensor data. Moreover, data processing benefits from subjective logic operators for combining opinions on collected sensor data.[FIX6]

[6]FIX. **LG:** motivate the use of subjective logic for sensor data

Table 1: Opinion Determination

	Subjective	*Measurable*
Attribute	Reputation of Origin Trust	Accuracy Type of measurement Value
Entity	Reputation Trust	Communication protocol Entity credentials

5.2 Opinion Policy

5.3 Trusted Sensor Data Attribute and Entity

Relying on our trust model introduced in section 4 and capitalizing on subjective logic, we define $\psi(P_a, e)$, the trust evaluation of sensor data attribute a by an entity e, as ω_a^e. Similarly, we define $\psi(P_f, e)$, the trust evaluation of an entity f by e , as ω_f^e.

Those opinions on sensor data attributes and entity are to be determined based on a combination of subjective and measurable criteria. Following Covington et al's [3] approach, we distinguish subjective aspect (e.g. node or entity reputation) from measurable aspects (e.g accuracy, freshness). The subjective aspects of an opinion are based on the past experience with a given entity (e.g. node, sink), while the measurable aspects are derived from elements which characterize a sensor data or an entity.

Table 1 proposes some subjective and measurable aspects for opinion determination of attributes and entities. Reputation of node can be seen as subjective element of origin attribute of sensor data. Credential of an entity and its used communication protocol are measurable elements which support the determination of trustworthiness of an entity.

For the sake of readability, s denotes a subjective aspect, and m a measurable aspect. In [3], the authors propose the following combination in order to determine an opinion ω based on those two parameters:

$$\omega = (b, d, u, a) \; where \begin{cases} b = s \cdot m \\ d = s \cdot (1 - m) \\ u = 1 - m \end{cases}$$

Belief is then defined as a combination of subjective and measurable aspects whereas uncertainty is defined as the opposite of subjective aspect. Based on this combination of subjective and measurable aspects, sensor data attribute and entity are determined.

5.4 Trusted Sensor Data Acquisition

When applying subjective logic to our trust model, we define $\psi(P_s, e)$, trust evaluation of sensor data s of an entity e, as ω_s^e. Based on $\psi(P_s, e)$ formulation in section **??**, we suggest the definition of as *combine* operator which computes ω_s^e based on opinions on its sensor data attributes, $\omega_{a_i}^e$. Contrary to the *consensus* operator [17] which fuses opinions of different entities about a same proposition, $\psi_{acquisition}$ operator aims at combining opinions of a single entity about different propositions P_{a_i}, related to P_s, and leveraging the

influence of the most weighted opinions, $\psi(P_{a_i}, e)$. We then propose the following mapping between $\psi_{acquisition}$ and the *combine* operator:

$$\begin{aligned} \psi(P_s, e) &= \psi_{acquisition}(\psi(P_{a_i}, e), \alpha_i) \\ \omega_s^e &= combine(\omega_{a_i}^e, \alpha_i)_{i=0}^n \\ \omega_s^e &= combine^{\alpha_i}((\omega_{a_i}^e))_{i=0}^n \end{aligned}$$

As described in section 5.1, an opinion is a 4-tuple {b, d, u, a} encompassing belief, disbelief, uncertainty and atomicity (set to $\frac{1}{2}$) of a given proposition trustworthiness. We then have

$$\omega_s^e = (b, d, u, a) \ where$$

$$\begin{cases} b = combineb^{\alpha_i}((b_{a_i}))_{i=0}^n \\ d = 1 - b - u \\ u = min(1 - b, combineu^{\alpha_i}((u_{a_i})_{i=0}^n) \end{cases}$$

combineb aims at smoothly increasing belief of combined opinions. u is defined in such manner that the influence of uncertainty on combination of opinion is minimized in comparison with belief. Additionally, we respect the constraint on $b + d + u = 1$.

combineb and *combineu* functions are defined in the following sections.

5.4.1 *combineb* function

We first define the *combineb* function for two beliefs, and then extend it for more than two beliefs. As regards combination of two beliefs b_e and b_f, it consists of a smooth increase of their maximum, depending on the distance $|b_e - b_f|$. This increase is to be exponentially proportional to their maximum belief and to their distance. The *combineb* operator has then to fulfill the following requirements:

- **RE1**:

  $$\forall b_e, b_f \in [0, 1], 1 \geq combineb(b_e, b_f) \geq max(b_e, b_f)$$

- **RE2**: $combineb(b_e, b_f)$ is proportional to the distance $| b_e - b_f |$.
- **RE3**: $combineb(b_e, b_f)$ is exponentially proportional to $\max(b_e, b_f)$.

With requirement **RE1**, we express the fact that the combination of two beliefs always results in an increase. In case of $\min(b_e, b_f)$=0, $combineb(b_e, b_f)$ is equal to the lower bound, $\max(b_e, b_f)$. Besides, $combineb(b_e, b_f)$ is up to 1. **RE2** reflects the fact that the closer the $\min(b_e, b_f)$ is to $\max(b_e, b_f)$, the bigger the combination acceleration has to be. In other words, we tend to reward the combination of belief with are close each others. Finally, **RE3** is to reward the combination of high beliefs. We prefer the combination of two beliefs 0.5 and 0.9 rather than 0.5 and 0.5, even if they are close each others. Figure 3 depicts the fact that a combination of two low beliefs leads to almost no belief increases. On the contrary, the combination with two strong beliefs leads to a quick increase of combined belief. In addition, the combination of two high beliefs is to be bigger than combination of high and low beliefs.

]\[t]

Figure 3: *combineb* **Evolution**

Figure 4: *combineb* **Operator**

Based on **RE1**, **RE2** and **RE3**, we define combination between two beliefs as follows:

DEFINITION 7. ***(Belief Combination)***
Let b_e and b_f be agent's beliefs about two distinct propositions a and b. Let $combineb(b_e, b_f)$ be the belief such that:
$combineb(b_e, b_f) = \min(1, \max(b_e, b_f) + \varepsilon(b_e, b_f))$
where $\varepsilon(b_e, b_f) = (b_e \cdot b_f)^{(2 - b_e - b_f)}$.

$combineb(b_e, b_f)$ is called the combination of b_e and b_f representing the agents' beliefs about the combination of a and b being true.

In figure 5, we depict the evolution of $combineb(b_e, b_f)$ for five values of the maximum between b_e and b_f. We clearly demonstrate our three requirements. When the maximum between the beliefs equals to 0.1, the increase of $combineb(b_e, b_f)$ is smaller than the one with the maximum equals to 0.9 (**RE1** and **RE3**). Moreover $combineb(b_e, b_f)$ increases progressively while $|b_e - b_f|$ tends to zero (**RE2**).

As far combination of more than two beliefs is considered, we define belief combination as follows:

$$combineb(b_{a_i})_{i=0}^n =$$

$$\begin{cases} for \quad n = 0, \omega_{a_0} \\ for \quad n = 1, combineb(b_{a_0}, b_{a_1}) \\ otherwise \quad , combineb(b_{a_n}^e, combineb(b_{a_i}^e)_{i=0}^{n-1}) \end{cases}$$

5.4.2 *combineu* function

With respect to uncertainty combination, we reduce combined uncertainty by computing an average of uncertainty as follows:

$$combineu^{\alpha_i}(u_{a_i})_{i=0}^n = \frac{\sum_{i=0}^n (u_{a_i} \cdot \alpha_i)}{\sum_{i=0}^n (\alpha_i)}$$

5.4.3 *combine* operator

In Figure 5, we illustrate the evolution of the combination of two opinions, ω_e and ω_f, equally weighted. We set three fixed values of belief and uncertainty of ω_f. For curve (a), (b) and (c), we set ω_f to 0.375 ($b_f = 0.25$ and $u_f = 0.25$), 0.875 ($b_f = 0.75$ and $u_f = 0.25$) and 0.625 ($b_f = 0.25$ and

Figure 5: Combination Operator

$u_f = 0.75$) respectively. With (a), (b) and (c), we clearly demonstrate then that combination of opinions smoothly increase up to 1. With (a) and (c), we show the impact of high uncertainty of combination of opinions. Additionally, comparing (a) and (b), we leverage combined belief in comparison to uncertainty.

5.5 Trusted Sensor Data Delivery

We define the trustworthiness in a sensor data delivery by a entity f to e as follows:

$$\begin{aligned} \psi(P_{s|f\rightarrow e}, e) &= \psi_{delivery}(\psi(P_s, f), \psi(P_f, e)) \\ \omega^e_{s|f\rightarrow e} &= discount(\omega^f_s, \omega^e_f) \end{aligned}$$

The *discount* operator perfectly fits, by definition, to compute $\omega^e_{s|f\rightarrow e}$ based on ω^f_s and ω^e_f. Dedicated to transitive trust computation, the *discount* operator enables any entity e to form an opinion on a proposition P_s by discounting opinion ω^f_s and ω^e_f.

5.6 Trusted Sensor Data Fusion

In order to determine trustworthiness of sensor data fusion, we decide to reuse the *combine* operator. As defined in section 5.4, the *combine* operator combines opinions of a entity e in a single opinion. In the case of fusion of data, we combine opinions of a single entity on different propositions used to infer on a single proposition.

$$\begin{aligned} \psi(P_{\bigotimes(s_i)^n_{i=0}}, e) &= \psi_{\bigotimes}(\psi(P_{s_i}, e), \alpha_i)^n_{i=0} \\ \omega^e_{\bigotimes(s_i)^n_{i=0}} &= combine(\omega^e_{s_i}, \alpha_i)^n_{i=0} \\ \omega^e_{\bigotimes(s_i)^n_{i=0}} &= combine^{\alpha_i}(\omega^e_{s_i})^n_{i=0} \end{aligned}$$

5.7 Trusted Sensor Data Aggregation

We propose to compute an average on opinion on the sensor data of the aggregation. This average computation has to involve weight of aggregated sensor data, as expressed in the following definition of the *average* operator.

$$\begin{aligned} \psi(P_{\bigoplus(s_i)^n_{i=0}}, e) &= \psi_{\bigoplus}(\psi(P_{s_i}, e), \alpha_i)^n_{i=0} \\ \omega^e_{\bigoplus(s_i)^n_{i=0}} &= average(\omega^e_{s_i}, \alpha_i)^n_{i=0} \end{aligned}$$

We then define the *average* operator as such:

$$average(\omega^e_{s_i}, \alpha_i)^n_{i=0} \begin{cases} b_{average} = \frac{\sum^n_{i=0}(b_{P_{s_i}}.\alpha_i)}{\sum^n_{i=0}(\alpha_i)} \\ d_{average} = \frac{\sum^n_{i=0}(d_{P_{s_i}}.\alpha_i)}{\sum^n_{i=0}(\alpha_i)} \\ u_{average} = \frac{\sum^n_{i=0}(u_{P_{s_i}}.\alpha_i)}{\sum^n_{i=0}(\alpha_i)} \end{cases}$$

In summary, in this section, we implement our model trust with subjective logic. The notion of opinion is then used to assess the trustworthiness of sensor data. Including belief and uncertainty on the fact that *"the sensor data is trustworthy enough to be used"*, subjective logic framework supports the combination of opinions with several operators (e.g. discount, consensus). In order to determine the opinion of a single entity about different interdependent propositions, we introduce the *combine* operator. The latter is used to determine opinions of acquired and fuse data. In addition, we adapt *average* operator defined by subjective logic in order to consider weight of opinions. At last, we re-use the *discount* operator at sensor data delivery time.

Figure 7: Trust Assessment Illustration

6. EVALUATION

6.1 A remote healthcare monitoring scenarion

In order to illustrate the implementation of our trust model with subjective logic, we introduce a remote health care monitoring example where patients activities (e.g. patient's hydration, awakening schedule) or physiological information (e.g. blood pressure, pulse, body temperature) are monitored 24 hours a day via a Body Sensor Network[6]. The latter is defined as a combination of any kind of wearable[9] and implantable[1] sensors. The BSN is connected to a Medical Emergency Response Center (MERC) through patients PDA. The latter is in charge of detecting any irregularities in patient health condition, and of aggregating physiological patient information. The PDA can for example trigger an alert in case of emergency to the MERC, that contacts the closest physician to patient for a home visit. In addition to visiting, physician can request for aggregation of physiological information for a better diagnostic. Figure 6 outlines trust assessment at acquisition, aggregation and fusion of physiological information from the PDA to the MERC.

Mapping to the life cycle illustrated by Figure 1, we detail the trustworthiness evaluation of pulse and blood pressure from nodes to application in Figure 7. Applying the defined operator at acquisition, aggregation, fusion and delivery steps, we demonstrate confidence evolution in sensor data and processing. For sake of readability, in Figure 7, we note opinion as such $pe - \{b, d, u\}$, where pe is expectation probability, b belief, d disbelief and u uncertainty. In this example, each sensor attributes or element of fusion and aggregation are equally weighted.

Our implementation then homogenizes trust evaluation of sensor data within their life cycle. As depicted in Figure 7, we use the same metric from acquisition to delivery of sensor data. Moreover, we consider any sensor data attributes for trustworthiness evaluation, and do not restrict it to origin or value. At last, we use our *combine* operator in order to assess trustworthiness of data at acquisition and fusion time.

6.2 Herd control scenario

In this section, we can speak about the experiment and the collected data by TU/e. On the field, farmers detect the following events: node detached from the cow, low battery, ... How can we infer that a cow lost her node ...

Show some results with the data received.

Figure 6: Remote Health care Monitoring

7. CONCLUSION

In this paper, we proposed a novel trust model for sensor data during their entire life cycle. Capitalizing on subjective logic, we implement our model and design new operators on opinions for the combination and aggregation of opinions. Opinion on data is then used by application for further decision-making. To the best of our knowledge, our work is the first in the field to introduce a complete trust evaluation of sensor data during their life cycle. The proposed implementation of our trust model then address drawbacks introduced by existing approaches: heterogeneity, the fact that they only consider origin or value in trust evaluation, and the lack of confidence determination in fusion of information.

In the future, we plan to investigate on the establishment of similarity between sensor data based on ontology for contextual information. The idea is to support business application in the acquisition of sensor data which are not directly available in WSN. For example, a business application can request pulse from a WSN delivering heart rate data. If pulse and heart rate data are defined as similar, it would enable the middleware to convert heart rate into pulse information. The latter would be mapped with a confidence value depending on heart rate and similarity relation trustworthiness.

8. REFERENCES

[1] D. Bradley. Implantable chips. *in the Reactive Reports. Chemistry Webmagazine*, October 2004.

[2] S. Buchegger and J.-Y. L. Boudec. A robust reputation system for p2p and mobile ad-hoc networks. *in the Proceedings of the Second Workshop on the Economics of Peer-to-Peer Systems.*, 2004.

[3] M. Covington, M. Ahamd, I. Essa1, and H. Venkateswaran. Parametrized authentication. *in ESORICS*, 2004.

[4] S. Ganeriwal and M. Srivastava. Reputation-based framework for high integrity sensor networks. *in ACM ESAS Workshop*, pages 66–77, 2004.

[5] J. Golbeck, J. Hendler, and B. Parsia. Trust networks on the Semantic Web. *in the Proceedings of Cooperative Intelligent Agents*, 2003.

[6] Y. Guang-Zhong. *Body Sensor Networks*. Springer, 2006.

[7] G. Gupta and M. Younis. Fault-tolerant clustering of wireless sensor networks. *in IEEE INFOCOM*, 2005.

[8] S. Harte and A. Rahman. Fault tolerance in sensor networks using self-diagnosing sensor nodes. *in the Proceedings of the IEEE International Workshop on Intelligent Environment*, 2005.

[9] R. Jafari, F. Dabiri, P. Brisk, and M. Sarrafzadeh. Adaptive and fault tolerant medical vest for life-critical medical monitoring. *in the Proceedings of the ACM Symposium on Applied computing*, pages 272–279, 2005.

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

[11] A. Josang and S. Pope. Semantic constraints for trust transitivity. *in the Proceedings of APCCM*, 2005.

[12] C. Koo. Broadcast in radio networks tolerating byzantine adversarial behavior. *in the Proceedings of the twenty-third annual ACM symposium on Principles of distributed computing*, pages 275–282, 2004.

[13] L. Lamport, R. Shostak, and M. Pease. The byzantine generals problem. *in the ACM Transactions on Programming Language and Systems*, 4(3):382–401, July 1982.

[14] E. H. Ngai and M. Lyu. Trust andclustering-based authentication services in mobile ad hoc networks. *in the Proceedings of ICDCS Workshops*, pages 582–587, 2004.

[15] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, NJ, 1976.

[16] A. Tanenbaum and M. V. Steen. *Distributed Systems: Principles and Paradigms*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2001.

[17] D. Yu and D. Frincke. Alert confidence fusion in intrusion detection systems with extended dempster-shafer theory. *in the Proceedings of the Southeast Regional Conference*, 43:142–147, 2005.

[18] Q. Zhang, T. Yu, and P. Ning. A framework for identifying compromised nodes in sensor networks. *in Proceedings of the IEEE SecureComm*, August 2006.

[19] W. Zhang, S. Das, and Y. Liu. A trust based framework for secure data aggregation on wireless sensor networks. *in the Proceedings of the IEEE SECON*, September 2006.