We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

23

Probabilistic modelling and recursive bayesian estimation of trust in wireless sensor networks

Mohammad Momani[1] and Subhash Challa[2]
[1]*University of Technology Sydney, Australia mmomani@eng.uts.edu.au*
[2]*NICTA, VRL, University of Melbourne, Australia, Subhash.Challa@nicta.com.au*

1. Introduction

Wireless Sensor Networks closely resemble a human behaviour model, in which a number of nodes that have just met are able to communicate with each other based on mutual trust levels developed over a period of time. WSNs are characterised by their performance of an additional function to the traditional functions of an ad-hoc network, which is monitoring events and reporting data and, as such, the sensed data represent the core component of trust-modelling in this research.

The trust-modelling problem in wireless networks is characterised by uncertainty. It is a decision problem under uncertainty and the only coherent way to deal with uncertainty is through probability. There are several frameworks for reasoning under uncertainty, but it is well accepted that the probabilistic paradigm is the theoretically sound framework for solving a decision problem involving uncertainty. Some of the trust models introduced for sensor networks employ probabilistic solutions mixed with ad-hoc approaches. None of them produces a full probabilistic answer to the problem. Each node's reliability is an unknown quantity. The ensuing decision problems concern is which nodes are to be trusted. It is these decision problems; regarding when to terminate nodes, that motivate research in trust models.

We look at applying trust evaluation to WSNs, providing continuous data in the form of a new reputation system we call GTRSSN: *Gaussian Trust and Reputation System for Sensor Networks*. It has been argued that previous studies on WSNs focused on the trust associated with the routing and the successful performance of a sensor node in some predetermined task. This resulted in looking at binary events. The trustworthiness and reliability of the nodes of a WSN, when the sensed data are continuous, has not been addressed. Our main contribution is therefore the introduction of a statistical approach; a theoretically sound Bayesian probabilistic approach for modelling trust in WSNs in the case of continuous sensor data; that is, we derive a Bayesian probabilistic reputation system and trust model for WSNs, as presented in our work in (Momani et al., 2007a) and (Momani et al., 2007b).

2. Node Misbehaviour Classification

The main idea behind reputation and trust-based systems is to discover and exclude misbehaving nodes and to minimise the damage caused by inside attackers. Node misbehaviour can be classified in two categories: communication misbehaviour and data misinforming. Most of the researchers classify node misbehaviour in the same way they model trust: from the communication point of view. However, as discussed so far, WSNs are deployed to sense events and report data, so the node misbehaviour diagram presented in (Srinivasan et al., 2007) is extended by introducing a new branch addressing sensor data misbehaviour; misinforming, as a second category of nodes' misbehaviour classification in WSNs, as illustrated below in Figure 1, to reflect the way trust is being modelled here.

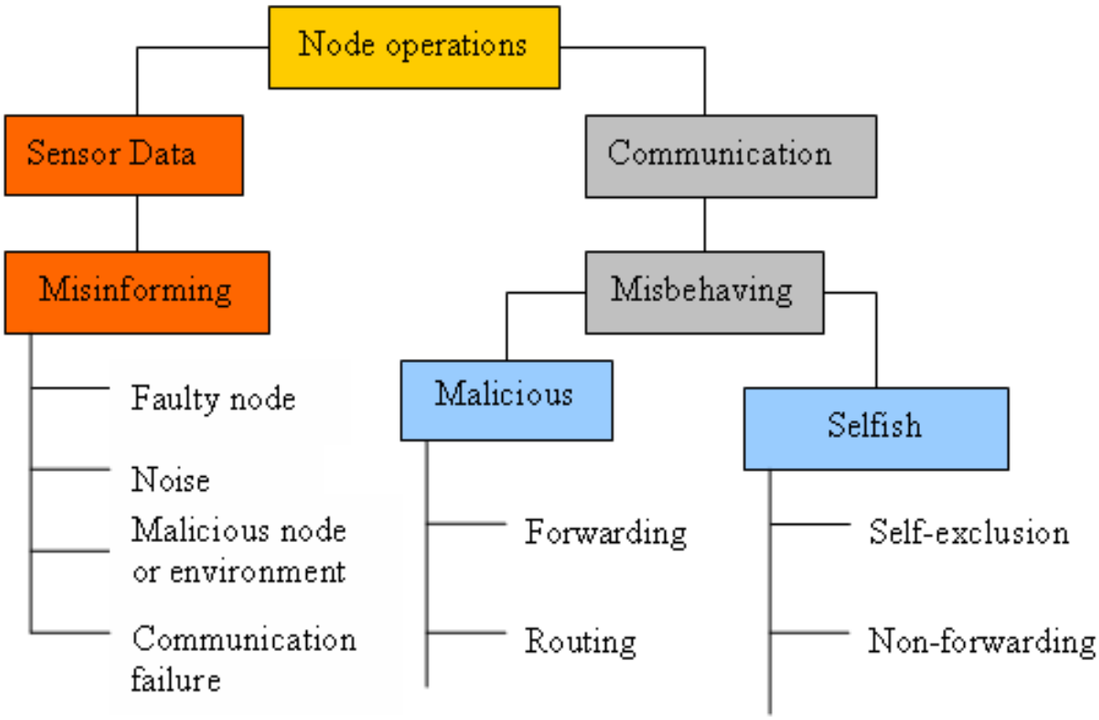


Fig. 1. Node misbehaviour classification

As can be seen from the diagram in Figure 1, the new branch dealing with sensor data includes the misinforming behaviour of a sensor node. This can be caused due to a faulty node, a node that is damaged or has expired, or due to a noise, as sensor data are not without noise, a malicious node or environment. The node might have been captured or the environment is malfunctioning or there might have been a communication failure, or there has been interference or the communication between nodes is cut off for some reason. The communication misbehaviour classification is due to the node being malicious, an intruder attacking and damaging the network, or the node is selfish, trying to save resources for later usage. Further detailed information regarding the node misbehaviour communication branch is provided in (Srinivasan et al., 2006).

3. Modelling Trust

Initially, the primary focus of the research on trust in WSNs was on whether a node will detect appropriately, will or will not report the detected event(s), and will route information. The uncertainty in these actions warranted the development of reputation systems and corresponding trust models. Modelling trust in general is the process of representing the trustworthiness of one node in the opinion of another node, that is, how much one node trusts every other node in the surrounding area, and it has been the focus of many researchers from different domains. In other words, trust-modelling is simply the mathematical representation of a node's opinion of another node in a network. Figure 2 below shows the two main sources for trust formation in WSNs: the observation of the behaviour of the surrounding nodes, direct trust and the recommendation from other nodes, indirect trust.

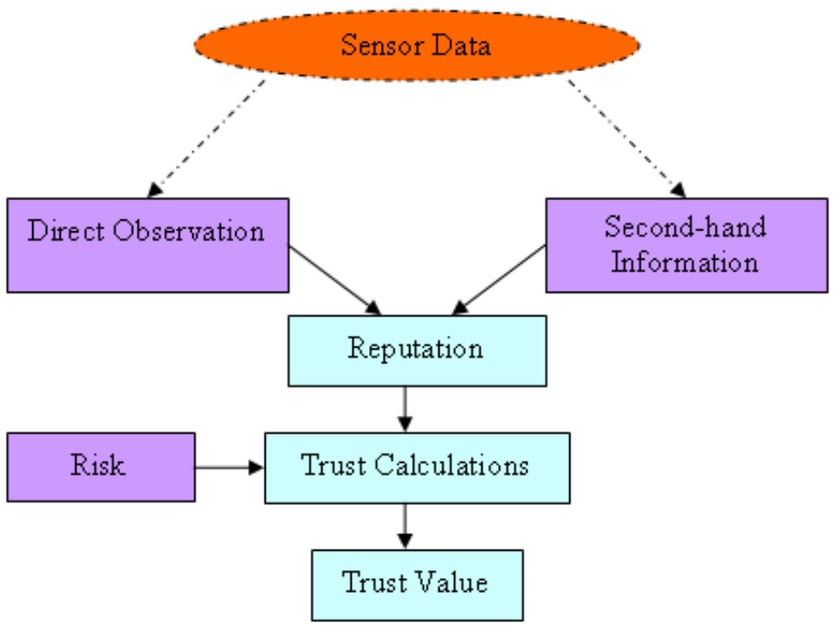


Fig. 2. Trust computational model for WSN

3.1. Direct Observations

A node will observe a neighbouring node's behaviour and build a reputation for that node based on the observed data. The neighbouring node's transactions data are direct observations referred to as *first-hand information*. By their nature, the considered events are binary, and the mathematical trust models developed for WSNs are for binary transactions. We argue that the problem of assessing a reputation based on observed data is a statistical problem. Some trust models make use of this observation and introduce a probabilistic modelling. For example, the reputation-based framework for high integrity sensor networks (RFSN) trust model presented in (Ganeriwal & Srivastava, 2004) by Ganeriwal and Srivastava uses a Bayesian updating scheme known as the *Beta Reputation System* for assessing and updating the nodes' reputations. The Beta reputation system was introduced by Josang and Ismail (Jøsang & Ismail, 2002), who used the Beta distribution to model binary statistical events.

3.2. Second-hand Information

A second source of information in trust-modelling is information provided by other nodes. This source of information is referred to as *second-hand information*. It consists of information gathered by nodes as first-hand information and converted into an assessment. Due to the limitations of a WSN, the second-hand information is summarised before being shared. For example, the RFSN in (Ganeriwal & Srivastava, 2004) uses the Beta probability model and share the values of the parameters of the probability distributions as second-hand information. This shared information is not hard data for the node receiving the information. A proper way is required to incorporate this new information into the trust model by combining it with observed data. While some trust models build reputations purely on the basis of observations, most of them attempt to use the second-hand information. The reasons are obvious from a statistical point of view. But the interest is also motivated by the desire to speed up the assessment of reputations. Due to the asymmetric transactions in a network, some nodes may not have enough observations about all neighbouring nodes.

Using shared information improves the efficiency and speed of reputation assessment, however, combining the two sources of information is handled differently by different trust models. For example, the RFSN uses the Dempster-Shafer Belief Theory. The Belief Theory is a framework for reasoning under uncertainty that differs from the probabilistic framework. The discussion of the fundamental differences between these two theories is beyond the scope of this research. Although the two approaches can be joined in some cases, they differ in their philosophies on how to treat uncertainty. The RFSN uses both of them in the same problem. We propose a probabilistic treatment of trust, and apply it to the case of continuous sensor data.

Although a reputation system is designed to reduce the harmful effect of an unreliable or malicious node, such a system can be used by a malicious node to harm the network. Systems such as the RFSN in (Ganeriwal & Srivastava, 2004) and the distributed reputation-based beacon trust system (DRBTS) in (Srinivasan et al., 2006) are confronted with the issue of what second-hand information is allowed to be shared. For example, some prohibit negative second-hand information to be shared, in order to reduce the risk of a negative campaign by malicious nodes. Our proposed model incorporates all of the second-hand information. To resolve the issue of the validity of the information source, the information is modulated using the reputation of the source. This probabilistic approach rigorously answers the question of how to combine the two types of data in the exercise of assessing reputations in a sensor network. It is based on work undertaken in modelling *Expert Opinion* (Lindley & Singpurwalla, 1986; Morris, 1971; West, 1984). Expert opinions are used whenever few data are observed. The expert opinion is second-hand information that is merged with hard data according to the laws of probability. Information provided by knowledgeable sources is known as “expert opinion” in the statistical literature. These opinions are modulated by existing knowledge about the experts themselves, to provide a calibrated answer.

4. The Beta Reputation System

The Beta Reputation System was proposed by Josang and Ismail in (Jøsang & Ismail, 2002) to derive reputation ratings in the context of e-commerce. It was presented as a flexible system with foundations in the theory of statistics, and is based on the Beta probability

density function. The Beta distribution can be used in the probability modelling of binary events. Let θ be a random variable representing a binary event, $\theta = 0; 1$, and p the probability that the event occurs, $\theta = 1$. Then the Beta-family of probability distributions, a continuous family of functions indexed by two parameters a and β, can be used to represent the probability density distribution of p, noted as $Beta(a, \beta)$, as shown in equation (1):

$$f(p \mid \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} \tag{1}$$

where $0 \leq p \leq 1$; $a > 0$; $\beta > 0$. If the number of outcomes where there are r occurrences and s non-occurrences of the event is observed, then using a Bayesian probabilistic argument, the probability density function of p can be expressed as a Beta distribution, where $a = r + 1$ and $\beta = s + 1$. This probabilistic mechanism is applied to model the reputation of an entity using events of completion of a task by the assessed entity. The reputation system counts the number r of successful transactions, and the number s of failed transactions, and applies the Beta probability model. This provides for an easily updatable system, since it is easy to update both r and s in the model. Each new transaction results either in r or s being augmented by 1.

For the RFSN (Ganeriwal & Srivastava, 2004) Ganeriwal and Srivastava used the work of Josang and Ismail presented in (Jøsang & Ismail, 2002), in their trust model for WSNs. For each node n_j, a reputation R_{ij} can be carried by a neighbouring node n_i. The reputation is embodied in the Beta model and carried by two parameters a_{ij} and β_{ij}. a_{ij} represents the number of successful transactions node n_i had with n_j, and β_{ij} represents the number of unsuccessful transactions. The reputation of node n_j maintained by node n_i is $R_{ij} = Beta(a_{ij} + 1, \beta_{ij} + 1)$. The trust is defined as the expected value of the reputation, as shown in equation (2):

$$T_{ij} = E(R_{ij}) = E(Beta(\alpha_{ij} + 1, \beta_{ij} + 1)) = \frac{\alpha_{ij} + 1}{\alpha_{ij} + \beta_{ij} + 2} \tag{2}$$

Second-hand information is presented to node n_i by another neighbouring node n_k. Node n_i receives the reputation of node n_j by node n_k, R_{kj}, in the form of the two parameters a_{kj} and β_{kj}. Using this new information, node n_i combines it with its current assessment R_{ij} to obtain a new reputation R_{ij}^{new}, as given in equation (3):

$$R_{ij}^{new} = Beta(\alpha_{ij}^{new}, \beta_{ij}^{new}) \tag{3}$$

where

$$\alpha_{ij}^{new} = \alpha_{ij} + \frac{2\alpha_{ik}\alpha_{kj}}{(\beta_{ik} + 2)(\alpha_{kj} + \beta_{kj} + 2)(2\alpha_{ik})} \tag{4}$$

$$\beta_{ij}^{new} = \beta_{ij} + \frac{2\alpha_{ik}\beta_{kj}}{(\beta_{ik} + 2)(\alpha_{kj} + \beta_{kj} + 2)(2\alpha_{ik})} \tag{5}$$

Note that node n_i uses its reputation of node n_k in the combination process. The authors of the RFSN defined how their trust model can be used in practice. They brought out some important points concerning the way information is to be used to avoid two major

problems: (i) data incest, and (ii) a game theoretic set-up. Some researchers (Agah et al., 2004; Liu et al., 2004) have looked into the game theory aspect, which is no doubt inherent in a problem with malicious nodes in the network. However, a game theory solution might be difficult to obtain, in view of the large number of nodes. The RFSN forces the WSNs protocols into an exchange of information that limits any game aspect. The effectiveness of the notion of reputation and trust resides in the assumption that the majority of nodes in any neighbourhood is trustworthy, therefore creating a resilience of the system. Trust assessment is used to flush out the bad nodes. In combining information, the authors of the RFSN followed the approach of (Jøsang & Ismail, 2002), by mapping the problem into a Dempster-Shafer belief theory model (Shafer, 1976), solving it using the concept of belief discounting, and conducting a reverse mapping from belief theory to probability. In our work we find it unnecessary to use the Belief theory. Rather, probability theory, and the ensuing work on expert opinion provide a way to combine the two types of information.

5. Expert Opinion Theory

The use of expert opinion has received much attention in the statistical literature. It allows for the formal incorporation of informed knowledge into a statistical analysis. Expert opinion, or informed judgement, is often available in the form of vendor information, engineering knowledge, manufacturer's knowledge, or simply an opinion formed over time. It is often a subjective opinion based on knowledge. Its main departure from hard data is that it cannot be claimed as objectively observed data. Nevertheless, it is often valuable information that has been formed over the course of time. In our case, reputation is offered to neighbouring nodes as an opinion. The node making the assessment has not observed that reputation, and therefore treats it as an opinion. Early work to formalise ad-hoc procedures for the use of expert opinion includes (Dalkey & Helmer, 1963; Morris, 1971). Morris (Morris, 1974) recognised the importance of treating the expert opinion as data, stating the general principle on which subsequent work was based. The topic was further enlarged by the Bayesian statistical community to the problem of reconciliation prior information from different sources (Dawid, 1987; French, 1980; Genest & Schervish, 1985; Lindley et al., 1979), a topic that dated back to Winkler (Winkler, 1968). Lindley (Lindley, 1983) highlighted the theory in the statistical arena, with others following with work on reliability (Aboura & Robinson, 1995; Mazzuchi & Soyer, 1993; Singpurwalla, 1988), on maintenance optimization (Aboura, 1995; Mazzuchi & Soyer, 1996; Van Nortwijk et al., 1992) and on nuclear safety (Cooke, 1994).

The probabilistic approach adopted in the elicitation and use of expert opinion considers the opinion given by the expert as data and treats it according to the laws of probability. If θ is a random variable, and μ represents an opinion from an expert about θ, then $P(\theta \mid \mu)$ obtains, using Bayes' theorem as discussed in appendix A, the following formula, as shown in equation (6):

$$P(\theta \mid \mu) = \frac{P(\mu \mid \theta)P(\theta)}{P(\mu)} \tag{6}$$

$$P(\mu) = \int_{\theta} P(\mu \mid \theta)P(\theta)d\theta \tag{7}$$

- $P(\mu \mid \theta)$ is the likelihood function, and represents the analyst model of the expert's input
- $P(\theta)$ is the distribution that represents any prior knowledge the analyst may have about the quantity of interest
- $P(\mu)$ is the normalising constant

Bayes' theorem inverses the probability, so that the evidence μ highlights the value of θ that is most likely. The likelihood function $L(\theta) = P(\mu \mid \theta)$ refers to where the expert opinion is modelled. As an example, consider the reliability scenario of (Aboura & Robinson, 1995). In it, an expert provides reliability estimates for a device or machine. The work was undertaken in the context of maintenance optimisation.

Figure 3 shows the expert's input along the unknown reliability curve that the analyst wants to estimate. Each assessment by the expert is about the reliability as a time t_i, in the form of a value $0 < r_i < 1$. If the expert was perfect, and assuming that the reliability at time t_i is $e^{-\lambda t_i^{\beta}}$, then

$$r_i = e^{-\lambda t_i^{\beta}} \tag{8}$$

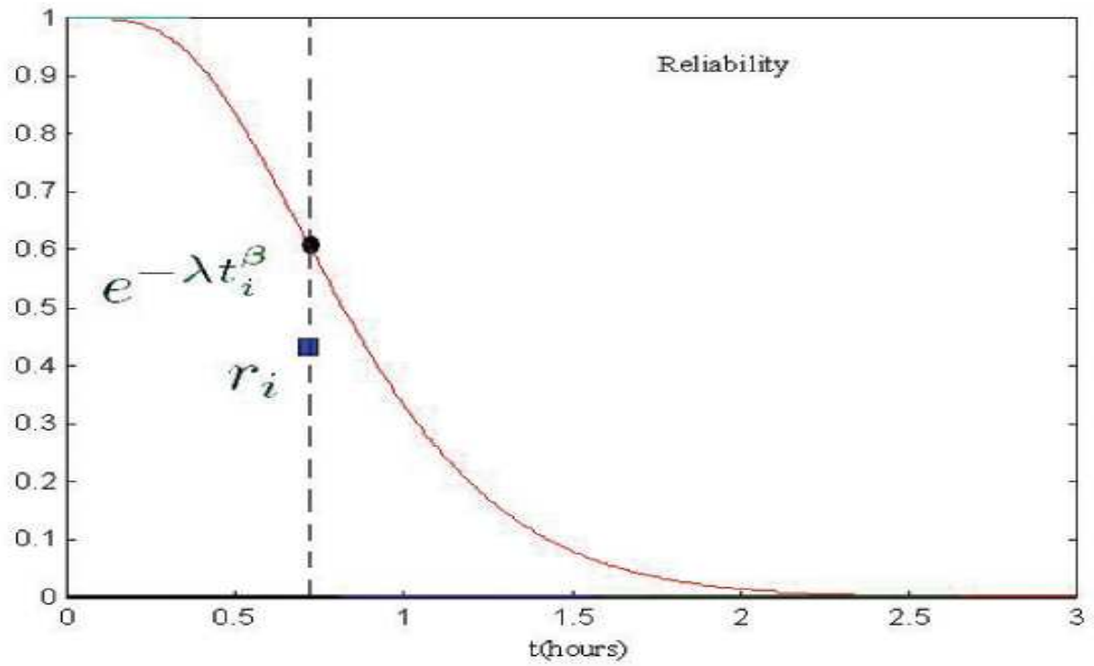


Fig. 3. Expert opinion r_i for reliability at time t_i

However, it will not necessarily be the case, and a probability distribution is needed to model the input. That probability distribution is the likelihood function, in this case

$$L(\lambda, \beta) = P(r_i \mid \lambda, \beta) \tag{9}$$

The authors of (Aboura & Robinson, 1995) modelled it using a Beta distribution, such that

$$E(r_i \mid \lambda, \beta) = \alpha_i + \sigma_i e^{-\lambda t_i^{\beta}} \tag{10}$$

where σ_i and α_i are inflation and bias, respectively, carried by the expert about the reliability at time t_i. These two values reflect the analyst's modulation of the expert opinion. To model several correlated inputs, a Dirichlet model is used. Once the likelihood function is built, then it can be used to combine the actual expert opinion with any existing knowledge about the random variable of interest. The analyst may not only have prior knowledge but also some observed data y about a random variable of interest, θ. Bayes' theorem is applied to combine the three sources of information, as shown in equation (11):

$$P(\theta \mid y, \mu) = \frac{P(y \mid \theta, \mu) P(\mu \mid \theta) P(\theta)}{P(y, \mu)} \tag{11}$$

One often writes, $P(\theta \mid y, \mu) \propto P(y \mid \theta, \mu)\ P(\mu \mid \theta)\ P(\theta)$, the denominator being a normalising constant that does not affect the combination occurring in the numerator. This seemingly simple operation can effectively combine many sources of information. We use it to model the reputation of a node when opinions about that node are provided by other nodes.

6. GTRSSN: Gaussian Trust and Reputation System for Wireless Sensor Networks

Taking into consideration the above discussion, let us assume that the wireless sensor network shown in Figure 4 consists of N nodes $(n_1, n_2, ..., n_N)$, and the corresponding matrix $\Gamma = [\Gamma_{i,j}]$ is given as follows:

$$\Gamma = [\Gamma_{i,j}] = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

If node n_i is connected to node n_j, then $\Gamma_{i,j} = \Gamma_{j,i} = 1$, otherwise it is equal to (0). Let (X) be a field variable monitored in the environment where the WSN is deployed. This variable, might represent temperature, chemical component or atmospheric value, is detected and estimated by the sensor nodes and it is assumed to be of a continuous nature. The nodes are synchronised and can report at discrete times t = 0, 1, 2, …., k.

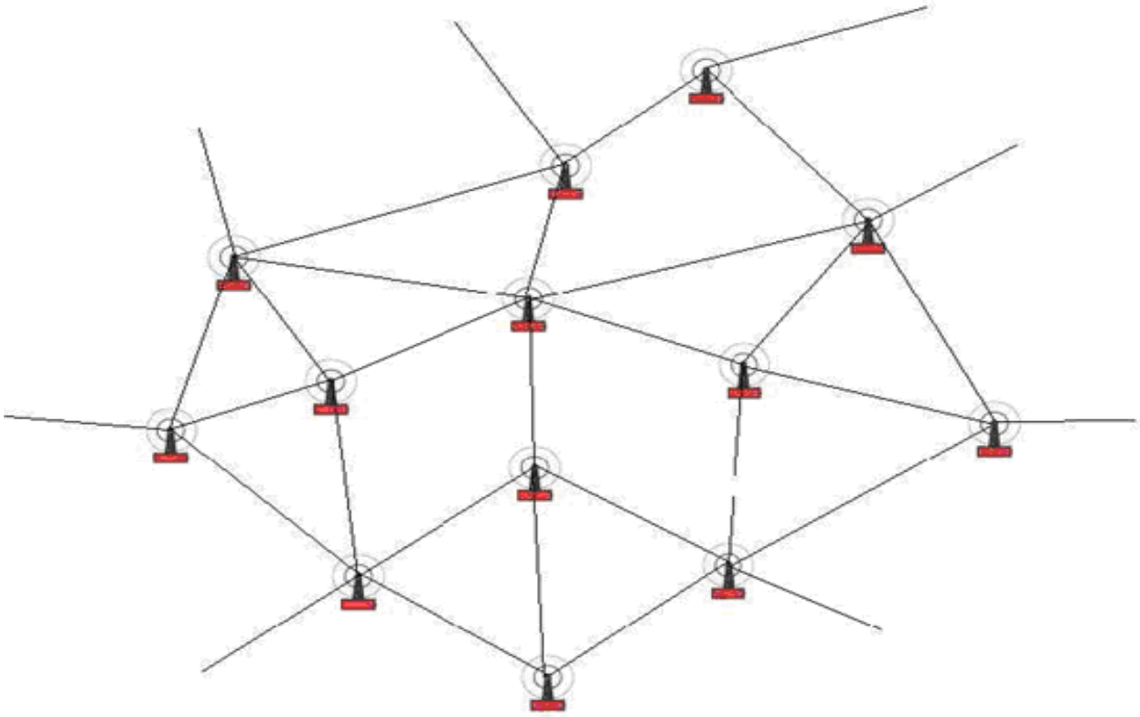


Fig. 4. Network of wireless sensor nodes

The random variable $(X_{n_i} = X_i)$ is the sensed value reported by node n_i, $i = 1, \ldots, N$. $x_i(t)$ is the realisation of that random variable at time t. Each node n_i, $i = 1, \ldots\ldots, N$ has a time series $(x_i(t))$. These time series are most likely different, as nodes are requested to provide readings at different times, depending on the sources of the requests. It could also be that the nodes provide such readings when triggered by particular events. We assume that each time a node provides a reading, its one-hop neighbours that route its report see that report, and can evaluate the reported value. For example, if node n_j reports $x_j(t)$ at some time t, then node n_i, such that $\Gamma_{i,j} = 1$, obtains a copy of that report for routing purposes, and has its own assessment $x_i(t)$ of the sensed variable. Let $y_{i,j}(t) = x_j(t)\text{-}x_i(t)$. From the node n_i perspective, $X_i(t)$ is known, and $Y_{i,j}(t) = X_j(t) - X_i(t)$ represents the error that node n_j commits in reporting the sensed field value $X_j(t)$ at time t. $Y_{i,j}(t)$ is a random variable modelled as a Normal (Gaussian), as shown in equation (12):

$$Y_{i,j}(t) \sim N(\theta_{i,j}, \tau^2) \tag{12}$$

where τ is assumed to be known (error variance), and it is the same for all nodes. If we let $\overline{y}_{i,j}$ be the mean of the observed error, as observed by n_i about n_j reporting, as in equation (13):

$$\overline{y}_{i,j} = \sum_{t=1}^{k} y_{i,j}(t)/k \tag{13}$$

then

$$(\theta_{i,j} \mid y_{i,j}) \sim N(\overline{y}_{i,j}, \tau^2 / k) \tag{14}$$

where $y_{i,j} = \{y_{i,j}(t)$, for all t values at which a report is issued by n_j and routed through $n_i\}$. This is a well-known straightforward Bayesian updating where a diffuse prior is used.

We let $\mu_{i,j} = \overline{y}_{i,j}$ and $\sigma_{i,j}^2 = \tau^2 / k$. Recall that k is node-dependent. It is the number of reports issued by node n_j and routed through n_i, and differs from node to node. We define the reputation as the probability density function, as in equation (15):

$$R_{i,j} = N(\mu_{i,j}, \sigma_{i,j}^2) \tag{15}$$

where $\mu_{i,j} = \overline{y}_{i,j}$ and $\sigma_{i,j}^2 = \tau^2 / k$ are the equivalent of a_{ij} and β_{ij} in RFSN (Ganeriwal & Srivastava, 2004).

Trust is defined differently, since we want it to remain between (0) and (1), a convention that seems to be unanimous among researchers, except for the occasional translation to the scale [-1, 1]. In our trust model, we define the trust to be the probability, as shown in equations (16) and (17):

$$T_{i,j} = \text{Prob}\{|\theta_{i,j}| < \varepsilon\} \tag{16}$$

$$\begin{aligned} T_{i,j} &= \text{Prob}\{-\varepsilon < \theta_{i,j} < +\varepsilon\} \\ &= \phi\left(\frac{\varepsilon - \overline{y}_{i,j}}{\tau / \sqrt{k}}\right) - \phi\left(\frac{-\varepsilon - \overline{y}_{i,j}}{\tau / \sqrt{k}}\right) \\ &= \phi\left(\frac{\varepsilon - \mu_{i,j}}{\sigma}\right) - \phi\left(\frac{-\varepsilon - \mu_{i,j}}{\sigma}\right) \end{aligned} \tag{17}$$

where ϕ is the cumulative probability distribution (cdf) of the Normal $N(0, 1)$. As shown in Figure 5, the area under the *Gaussian* curve $N(\mu_{i,j}, \sigma_{i,j}^2)$ within the interval $[-\varepsilon, +\varepsilon]$ is the trust value. The bigger the error θ_{ij} is, meaning its mean shifting to the right or left of 0, and the more spread that error is, the lower the trust value is. Each node n_i maintains a line of reputation assessments composed of T_{ij} for each j, such that $\Gamma_{i,j} \neq 0$ (one-hop connection). T_{ij} is updated for each time period t for which data is received from some connecting node j. The filled areas in Figure 5 represent the Gaussian Trust T_{ij} in two cases.

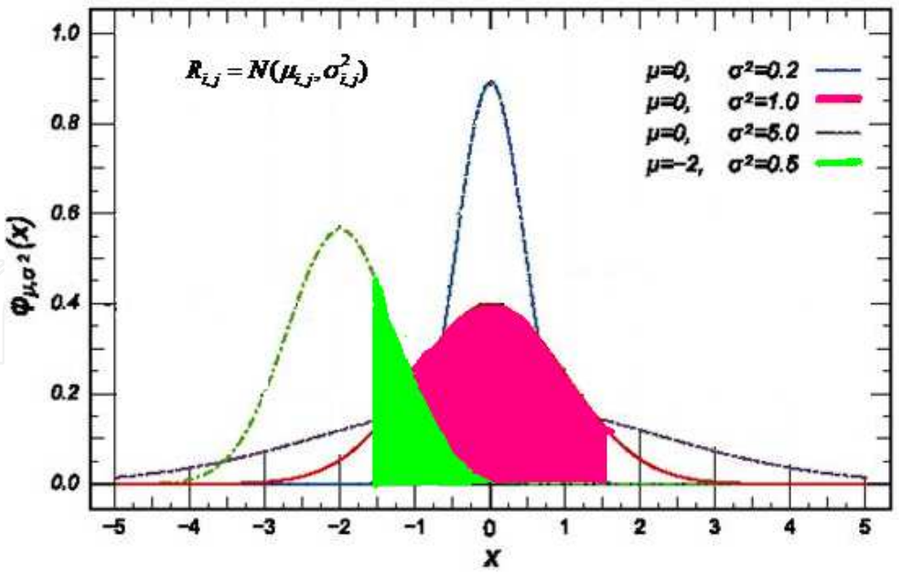


Fig. 5. Normal (Gaussian) distribution example

In addition to data observed in form of $y_{i,j} = \{y_{i,j}(t)\}$, for all t values at which a report is issued by n_j and routed through n_i}, node n_i uses *second-hand information* in the form of $(\mu_{l_s,j}, \sigma_{l_s,j})$, s = 1, …, m, from the m nodes connected to n_j and n_i, as shown in Figure 6, below. This is an "expert opinion", that is, soft information from external sources. Each of these m nodes has observed node n_j reports and produced assessments of its error in the form of $(\mu_{l_s,j}, \sigma_{l_s,j})$, s = 1,…, m, and consequently $T_{ls,j}$, s = 1, …, m. In using the expert opinion theory, one needs to modulate it. Node n_i uses its own assessment of the nodes $n_{l_1}, \ldots, n_{l_m}$, in the form of $(\mu_{i,l_s}, \sigma_{i,l_s})$, s = 1, … , m, and consequently $T_{i,ls}$, s = 1, …, m.

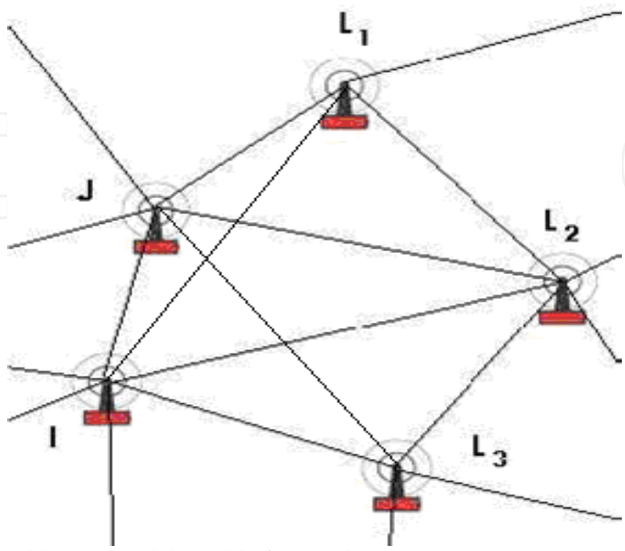


Fig. 6. Nodes that provide second-hand information

Using Bayes' theorem, the probability distribution of $\theta_{i,j}$ is obtained using the observed data along with the second-hand modulated information, as shown in equation (18):

$$P(\theta_{i,j} \mid y_{i,j},(\mu_{l_1,j},\sigma_{l_1,j}),...,(\mu_{l_m,j},\sigma_{l_m,j}) \\ ,(\mu_{i,l_1},\sigma_{i,l_1}),...,(\mu_{i,l_m},\sigma_{i,l_m})) \tag{18}$$

and it is proportional to the product of three terms shown in equations (19), (20) and (21):

$$P(y_{i,j} \mid \theta_{i,j},(\mu_{l_1,j},\sigma_{l_1,j}),...,(\mu_{l_m,j},\sigma_{l_m,j}), \\ (\mu_{i,l_1},\sigma_{i,l_1}),...,(\mu_{i,l_m},\sigma_{i,l_m})) \tag{19}$$

$$P((\mu_{l_1,j},\sigma_{l_1,j}),...,(\mu_{l_m,j},\sigma_{l_m,j}) \mid \theta_{i,j},(\mu_{i,l_1},\sigma_{i,l_1}),...,(\mu_{i,l_m},\sigma_{i,l_m})) \tag{20}$$

and

$$P(\theta_{i,j} \mid (\mu_{i,l_1},\sigma_{i,l_1}),...,(\mu_{i,l_m},\sigma_{i,l_m})) \tag{21}$$

The first term, equation (19), reduces to $P(y_{i,j} \mid \theta_{i,j})$ through conditional independence, and is equal to the product of the likelihoods

$$\prod_{t=1}^{k} N(\theta_{i,j},\tau^2) \tag{22}$$

The third term, equation (21), also reduces to $P(\theta_{i,j})$, due to the conditional independence of $\theta_{i,j}$ from $(\mu_{i,l_1},\sigma_{i,l_1}),...,(\mu_{i,l_m},\sigma_{i,l_m})$, and it represents the prior distribution of $\theta_{i,j}$ which we model as a diffuse prior $N(0,\infty)$.

The second term, equation (20), models the use of the second-hand information. This term requires some elaboration and can be reduced to the product of equation (23) through conditional independence arguments.

$$\prod_{s=1}^{m} P((\mu_{l_s,j},\sigma_{l_s,j}) \mid \theta_{i,j},(\mu_{i,l_s},\sigma_{i,l_s})) \tag{23}$$

To derive $P((\mu_{l_s,j},\sigma_{l_s,j}) \mid \theta_{i,j},(\mu_{i,l_s},\sigma_{i,l_s}))$ for each $s = 1, \ldots, m$, we observe the following:
for some t's,

$$\theta_{i,j} = x_j(t) - x_i(t) \tag{24}$$

and for some t's

$$\theta_{l,j} = x_j(t) - x_l(t) \tag{25}$$

and, if all t's were the same, then

$$\theta_{i,j} = x_j(t) - x_i(t) = (x_j - x_l) + (x_l - x_i) = \theta_{l,j} + \theta_{i,l} \tag{26}$$

But not all *t*'s are the same, so all data are not used for all assessments. We inspire ourselves from this relationship to model the expert opinion likelihood. We assume that

$$\theta_{l,j} \sim \theta_{i,j} - \theta_{i,l} \tag{27}$$

$$\mu_{l,j} \sim \theta_{i,j} - \mu_{i,l} \tag{28}$$

and we model

$$\mu_{l,j} \sim N(\theta_{i,j} - \mu_{i,l}, var) \tag{29}$$

where we choose *var* to be inversely related to node's n_i assessment of the reputation of node n_l, that is

$$var = \left(\frac{1}{T_{i,l}} - 1\right)\psi \tag{30}$$

where ψ is a model parameter.

$$\mu_{l,j} \sim N(\theta_{i,j} - \mu_{i,l}, \left(\frac{1}{T_{i,l}} - 1\right)\psi) \tag{31}$$

leads to equation (32):

$$\prod_{s=1}^{m} P((\mu_{l_s,j}, \sigma_{l_s,j}) \mid \theta_{i,j}, (\mu_{i,l_s}, \sigma_{i,l_s})) = \prod_{s=1}^{m} N(\theta_{i,j} - \mu_{i,l_s}, \left(\frac{1}{T_{i,l_s}} - 1\right)\psi) \tag{32}$$

and consequently proves that equation (33)

$$P(\theta_{i,j} \mid y_{i,j}, (\mu_{l_1,j}, \sigma_{l_1,j}), ..., (\mu_{l_m,j}, \sigma_{l_m,j}), (\mu_{i,l_1}, \sigma_{i,l_1}), ..., (\mu_{i,l_m}, \sigma_{i,l_m})) \tag{33}$$

is a Normal (Gaussian) distribution with mean and variance as shown in equations (34) and (35) respectively:

$$\mu_{i,j}^{new} = \frac{\sum_{s=1}^{m} \frac{(\mu_{l_s,j} + \mu_{i,l_s})}{\left(\frac{1}{T_{i,l_s}} - 1\right)\psi} + (k\overline{y} / \tau^2)}{\sum_{s=1}^{m} \frac{1}{\left(\frac{1}{T_{i,l_s}} - 1\right)\psi} + (k / \tau^2)} \tag{34}$$

$$\sigma_{i,j}^{2\ new} = \frac{1}{\sum_{s=1}^{m} \frac{1}{\left(\frac{1}{T_{i,l_s}} - 1\right)\psi} + (k/\tau^2)} \tag{35}$$

These values $(\mu_{i,j}^{new}, \sigma_{i,j}^{new})$, along with $(\mu_{i,j}, \sigma_{i,j})$, are easily updatable values that represent the continuous Gaussian version of the $(\alpha_{i,j}, \beta_{i,j})$ and $(\alpha_{i,j}^{new}, \beta_{i,j}^{new})$ of the binary approach in (Ganeriwal & Srivastava, 2004), as derived from the approach in (Jøsang & Ismail, 2002). The solution presented is simple and easily computed, keeping in mind that the solution applies to networks with limited computational power. In the binary work, $(\alpha_{i,j}, \beta_{i,j})$ are obtained through a Bayesian approach, while $(\alpha_{i,j}^{new}, \beta_{i,j}^{new})$ are obtained through the combination approach of Belief functions. The Gaussian solution provides a full probabilistic approach in the case of continuous sensor data.

Some would object to the use of a diffuse prior, which, in effect, forces a null prior trust value, regardless of the ε value. A way to remedy to this is to start with a $N(\mu_0, \sigma_0^2)$ prior distribution for all θ_{ij}, such that the prior trust is (1/2). This choice not only answers the diffuse prior issue, but also allows the choice of the parameters involved. ε can be determined: given μ_0 and σ_0, μ_0 is most likely to be set to (0). Therefore, σ_0 and ε determine each other. Once one is set, the other is automatically deducted. Note that the prior is really node-dependent, making our definition of trust, and therefore **ε**, node-dependent. In practice, it is most likely that all priors are tuned to the same values so that the prior trusts are started at some level, say (1/2), with a proper prior $\theta_{i,j}$, as shown in equation (36):

$$\theta_{i,j} \sim N(\mu_0, \sigma_0^2) \tag{36}$$

The reputation parameters $\mu_{i,j}$ and $\sigma_{i,j}^2$ are presented in equations (37) and (38):

$$\mu_{i,j} = \frac{(\mu_0/\sigma_0^2) + (k\overline{y}_{i,j}/\tau^2)}{(1/\sigma_0^2) + (k/\tau^2)} \tag{37}$$

$$\sigma_{i,j}^2 = \frac{1}{(1/\sigma_0^2) + (k/\tau^2)} \tag{38}$$

and the updated values are presented in equations (39) and (40) respectively:

$$\mu_{i,j}^{new} = \frac{(\mu_0 / \sigma_0^2) + \sum_{s=1}^{m} \frac{(\mu_{l_s,j} + \mu_{i,l_s})}{\left(\frac{1}{T_{i,l_s}} - 1\right)\psi} + (k\bar{y}_{i,j} / \tau^2)}{(1/\sigma_0^2) + \sum_{s=1}^{m} \frac{1}{\left(\frac{1}{T_{i,l_s}} - 1\right)\psi} + (k/\tau^2)} \tag{39}$$

$$\sigma_{i,j}^{2\ new} = \frac{1}{(1/\sigma_0^2) + \sum_{s=1}^{m} \frac{1}{\left(\frac{1}{T_{i,l_s}} - 1\right)\psi} + (k/\tau^2)} \tag{40}$$

Once $\mu_{i,j}^{new}$ and $\sigma_{i,j}^{2\ new}$ are formulated, the new trust value $T_{i,j}^{new}$ will be presented as shown in equation (41):

$$T_{i,j}^{new} = \phi\left(\frac{\varepsilon - \mu_{i,j}^{new}}{\sigma_{i,j}^{new}}\right) - \phi\left(\frac{-\varepsilon - \mu_{i,j}^{new}}{\sigma_{i,j}^{new}}\right) \tag{41}$$

We call this trust and reputation system (GTRSSN), which stands for Gaussian Trust and Reputation System for Sensor Networks. It can be seen as an extension of the concepts of RFSN and DRBTS for sensor data and it introduces a full probabilistic approach to the combination of information in the reputation assessment.

7. Simulation Results

To verify the theory introduced in this chapter, several simulation experiments in different scenarios were developed. The results from the simulations conducted on the network shown in Figure 7, for one scenario, where only a random region from the network is selected to report data on every time series, are presented in this section. In all simulation experiments, the trust relationship between four nodes (1, 6, 7 and 13) in a sub-network of the fifteen-node network shown in Figure 7 is calculated.

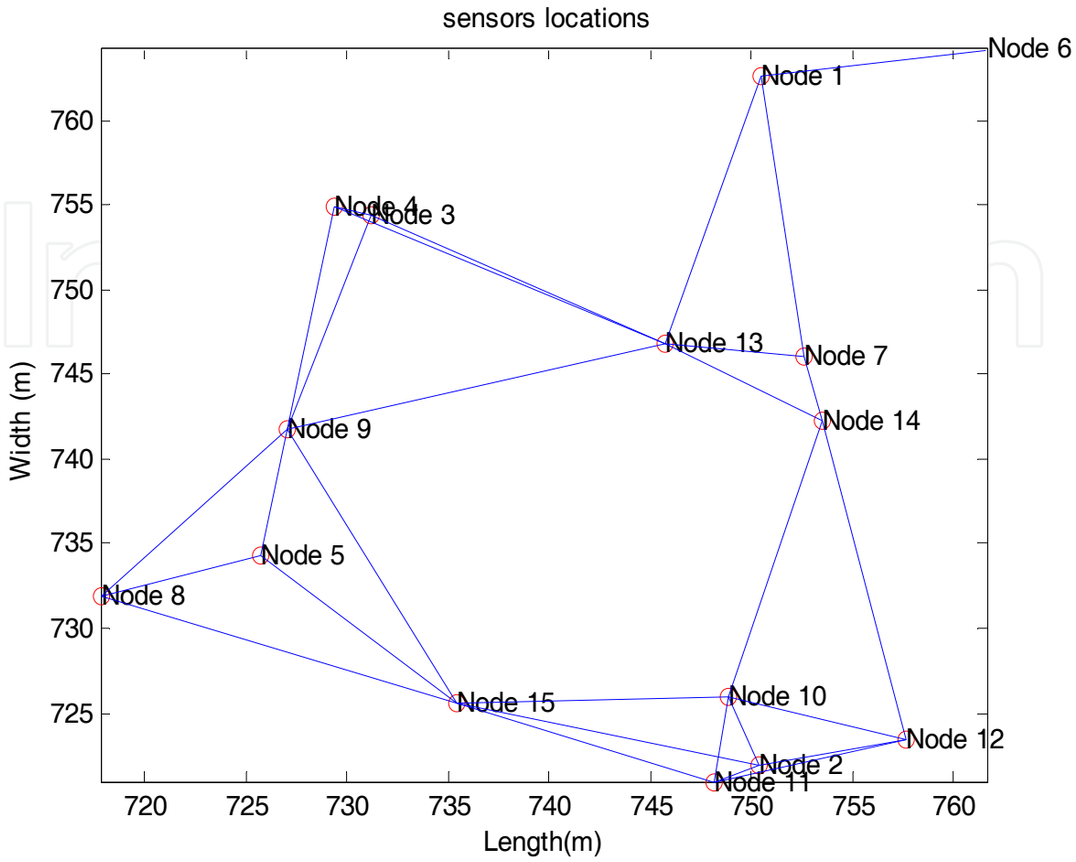


Fig. 7. Wireless Sensor Network Diagram

In this scenario and as stated before, it is assumed that, at each time slot a group of nodes are selected to report their sensed data, and when one node is sending its own reading to a specific node in the group, all the surrounding nodes connected to the sending node hear the reported value and start to send the output of that reading as a second-hand information to the receiving node regarding the sending node. The output of that reading between the sending and the receiving nodes is regarded as the direct observation, as discussed before. In other words, and in the case of selected sub-network, when node (7) is sending its reading to node (1), nodes (6) and (13) hear the reported data, use it to find the trust between them and node (7) and report that trust to node (1) as second-hand information about node (7). Node (1), at the same time, uses the reading reported directly from node (7) to calculate the direct trust between node (1) and node (7).

7.1. No faulty or malicious nodes are present in the network

At the beginning, it is assumed that all nodes are working properly, that no faulty or malicious nodes exist in the network, and report the sensed event (temperature) with minimum error. Figure 8 below presents the result of the simulation and shows the trust value between node (1) and the other nodes (6, 7 and 13). At first node (1) assesses node (13) based on the direct interactions only between the two nodes, without second-hand

information, and then node (1) assesses node (13) based on the direct information between the two nodes and the second-hand information received from node (7) about node (13), with second-hand information. Node (1) performs the same assessment procedure for all nodes directly connected to it.

It can be seen from Figure 8 that trust values between node (1) and nodes (7) and (13) are slightly different but they eventually all converge to the value of one. The trust value between node (1) and node (6) is the same in both cases, with and without second-hand information as there is no second-hand information for node (6). Node (6) is not connected to any other node other than node (1).

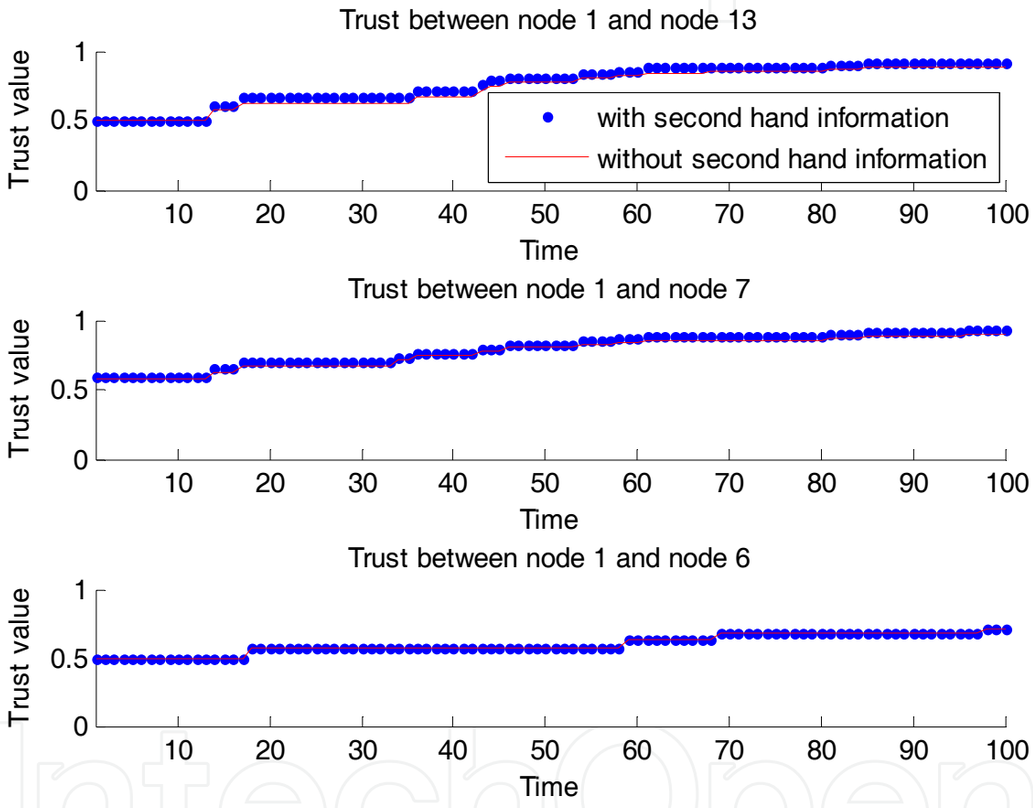


Fig. 8. All nodes are normal

7.2. Node (13) is Faulty or Malicious

In another experiment, the same network was simulated, but with the introduction of a significant error in node (13) readings, that is, node (13) is faulty or malicious. Simulation results are shown in Figure 9, below and, as can be seen from Figure 9, the trust value between node (1) and node (13) dropped to almost zero for both cases, with and without second-hand information, which means node (7) is assessing node (13) as a faulty or malicious node. The situation for node (6) is not affected, as there is no connection between node (6) and node (13). The interesting result here is that the trust value between node (1)

and node (7) is not affected in either case even though there is a connection between node (7) and node (13). Node (13) is faulty, and one would think that it could harm the reputation of node (7), but that was not the case, which proves that the modulation in the approach makes the reputation system robust to bad-mouthing attacks.

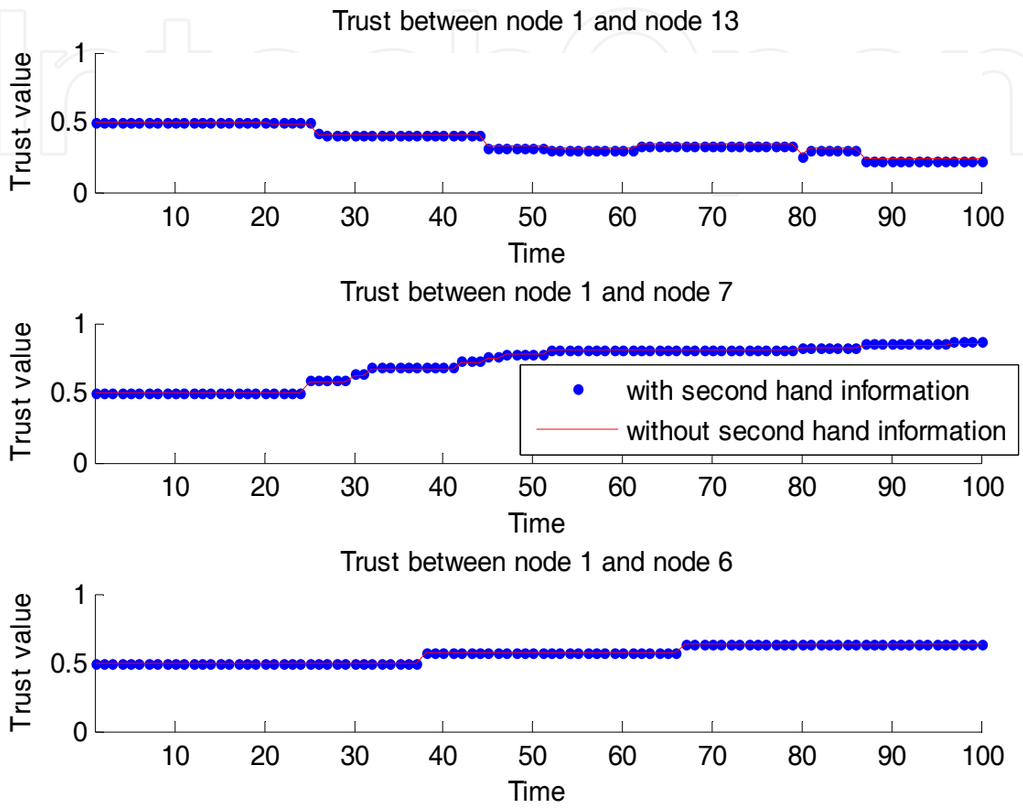


Fig. 9. Node (13) is faulty

7.3. Node (7) and Node (13) are Faulty

In this simulation experiment, it has been assumed that node (7) and node (13) are faulty. The results of the simulation are presented in Figure 5.10, showing that the trust values for both nodes (7) and (13) are dropping to zero in both cases. Node (6) is assumed reliable and the trust value associated with it is the same in both cases, as there is no connection between node (6) and the other faulty nodes, (7) or (13), to affect that trust value.

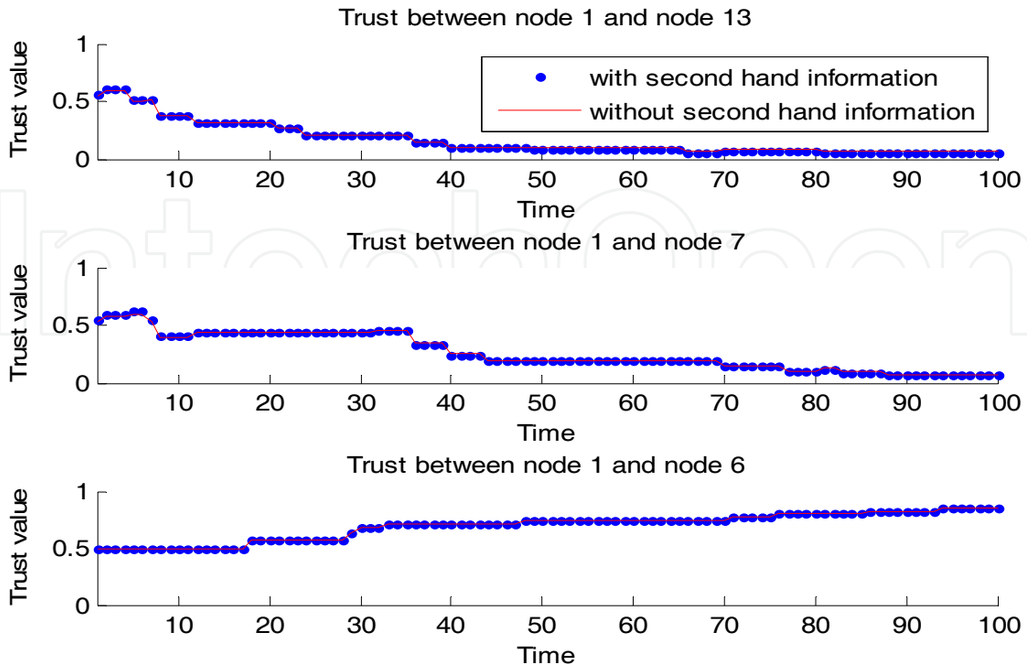


Fig. 10. Node (7) and node (13) are faulty

7.4. Node (6) is Faulty or Malicious

The simulation results presented in Figure 11 below show that when node (6) is faulty or malicious, nothing almost will change in the trust values between node (1) and either of nodes (7) and (13), as there is no direct or indirect connection between them. In other words, when node (6) is faulty, node (1) will discover that, as it has a direct connection with node (6) and the direct trust with node (6) will be affected. As there is no indirect trust for node (6), both trust values will stay on the initial trust value or will decrease to the value of zero.

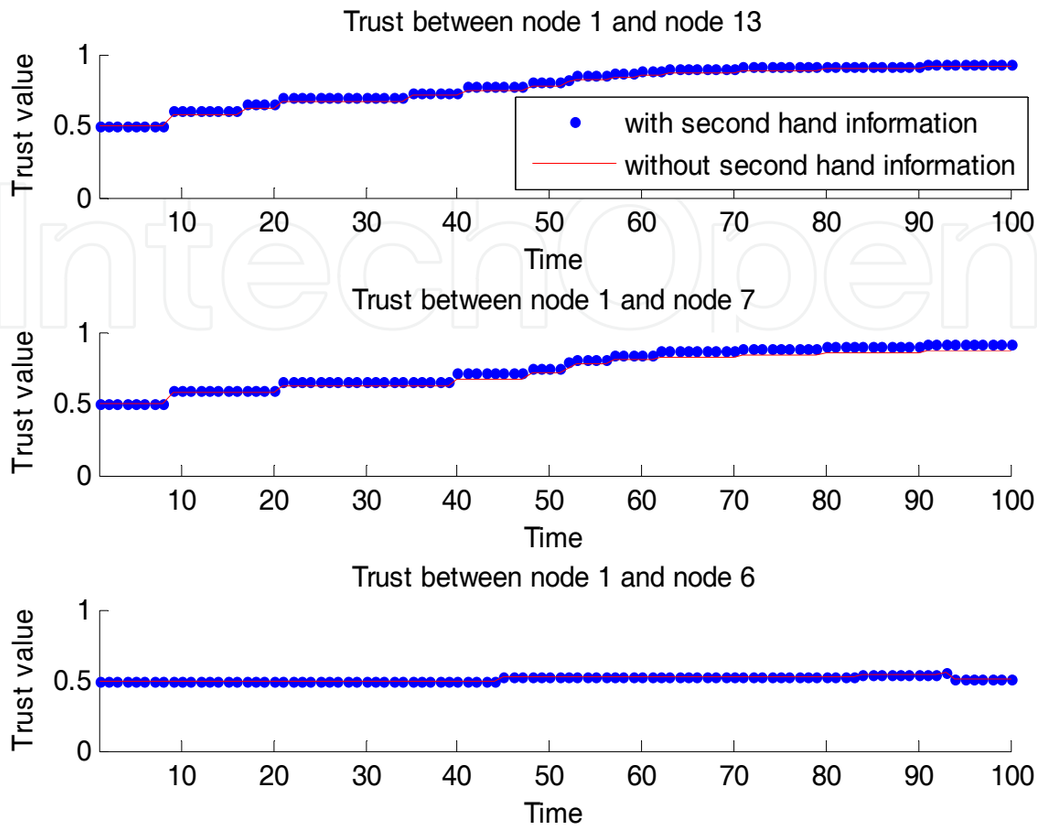


Fig. 11. Node (6) is faulty

7.5. Node (1) is Faulty or Malicious

It is assumed in this experiment that node (1) is faulty or malicious. Node (1) is the main node in the sub-network and is acting as the receiving node, and all the simulations show the trust relationship between node (1) and all the other nodes connected to it. As can be seen from Figure 5.12, the direct trust value for both nodes (7) and (13), is declining to the value of zero, as node 1 is faulty. That will leave the two nodes (7) and (13) to assess each other indirectly, which is a very interesting case again, as both nodes (7) and (13) are now assessing node (1) as a faulty node, so the indirect trust value for both nodes are slowly converging to the value of one. The trust value for node (6) is set to the initial value (0.5) and will decrease on both values to zero, as there is no second-hand information available to node (6) and node (1) is a faulty node.

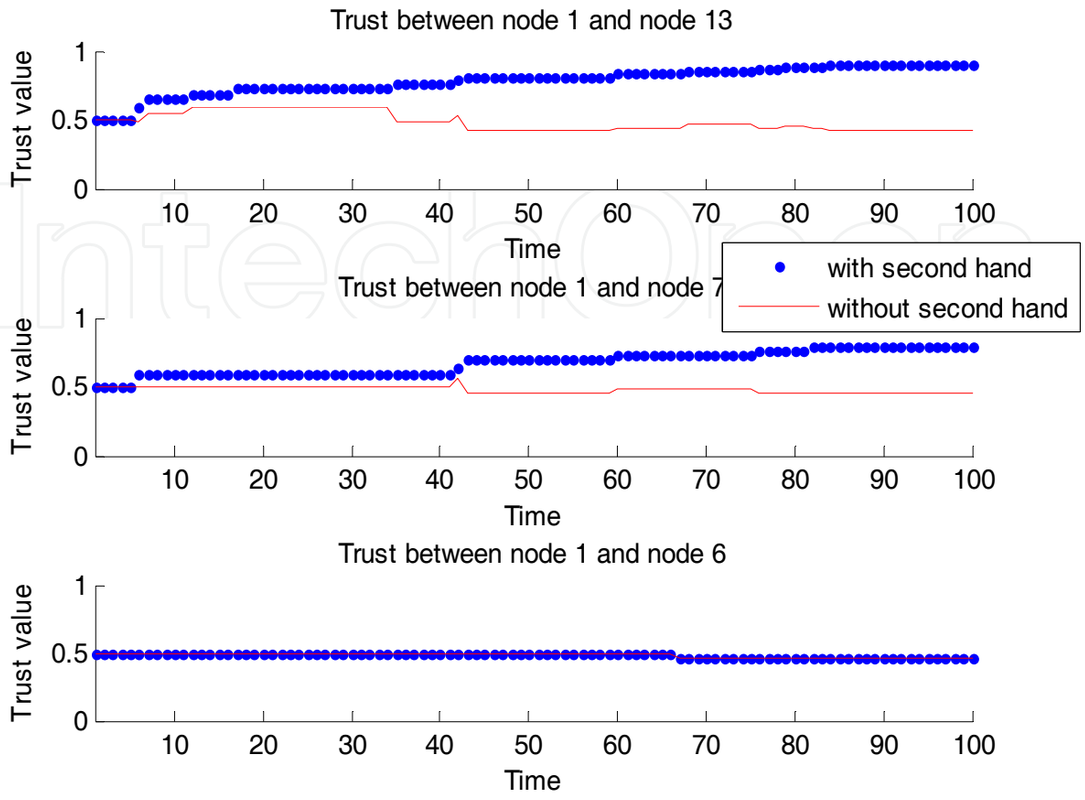


Fig. 12. Node (1) is a malicious node

The last example shows precisely the reason the trust system is instituted. It allows the classification of nodes into separate sets according to their trustworthiness. In the last example, it is known that node (1) is faulty, since it is a simulation exercise. The results should clearly indicate to the network that node (1) is faulty. However, it could also be the case that the nodes (7) and (13) are malicious. The trust system works on the assumption that a majority of nodes in a neighbourhood are reliable. This principle helps purge the system of bad elements. In this case, at this point, it is observed that the developed trust system is effective in distinguishing among nodes.

8. Conclusion

It has been argued that the trust-modelling problem is characterised by uncertainty, and the only coherent way to deal with uncertainty is through probability. Even though some of the trust models introduced for sensor networks employ probabilistic solutions mixed with ad-hoc approaches, none of them produces a full probabilistic answer to the problem. In this chapter we introduced a theoretically sound Bayesian probabilistic approach for calculating trust and reputation systems in WSNs. We introduced a new Gaussian Trust and Reputation System for Sensor Networks (GTRSSN), which we believe is a breakthrough in modelling trust in WSNs, as previous studies in WSNs focused on the trust associated with the routing

and the successful performance of a sensor node in some predetermined task, that is, looking at binary events to model trust and the trustworthiness and reliability of the nodes of a WSN when the sensed data is continuous has not been addressed before. Having said that, introducing the sensor data as a major component of trust leads to the modification of node misbehaviour classification, the trust computational model and the way first-hand and second-hand information is formulated. These issues have been presented in this chapter. Also, a brief summary about the Beta reputation system and the expert opinion theory has been presented. A very detailed GTRSSN, which is the significant contribution of this research, has also been presented, with some simulation results. The simulation results show the implications of sensor data for the direct and indirect trust relationship between nodes, which helps to distinguish among nodes and purge the bad nodes from the network.

9. References

Aboura, K. & Robinson, N. I. (1995). Optimal Replacement in a Renewal Process, CSIRO.

Aboura, K. (1995). Bayesian Adaptive Maintenance Plans using Initial Expert Reliability Estimates, CSIRO.

Agah, A.; Das, S. K. & Basu, K. (2004). *A Game Theory based Approach for Security in Wireless Sensor Networks*. IEEE International Conference on Performance, Computing, and Communications, Phoenix, Arizona.

Cooke, R. (1994). Uncertainty in Dispersion and Deposition in Accident Consequence Modelling Assessed with Performance-based Expert judgement. *Reliability Engineering and System Safety* 45: 35-46.

Dalkey, N. C. & Helmer, O. (1963). An Experimental Application of Delphi Method to the Use of Experts. *Management Science* 9(3): 458-467.

Dawid, A. P. (1987). The Difficulty about Conjunction. *The Statistician* 36: 91-97.

French, S. (1980). Updating of Belief in the Light of Someone else's Opinion. *Journal of Royal Statistical Society* 143: 43-48.

Ganeriwal, S. & Srivastava, M. B. (2004). *Reputation-based Framework for High Integrity Sensor Networks*. The 2nd ACM Workshop on Security of Ad-hoc and Sensor Networks Washington DC, USA

Genest, C. & Schervish, M. J. (1985). Modelling Expert Judgments for Bayesian Updating. *The Annals of Statistics* 13(3): 1198-1212.

Jøsang, A. & Ismail, R. (2002). The Beta Reputation System. *The 15th Bled Electronic Commerce Conference*. Bled, Slovenia.

Lindley, D. V. & Singpurwalla, N. D. (1986). Reliability (and fault tree) Analysis using Expert Opinions. *Journal of the American Statistical Association* 81: 87-90.

Lindley, D. V.; Tversky, A. & Brown, R. V. (1979). On the Reconciliation of Probability Assessments. *Journal of Royal Statistical Society* 142: 146-180.

Lindley, D. V. (1983). Reconciliation of Probability Distributions. *Operations Research Journal* 31: 866-880.

Liu, Y.; Comaniciu, C. & Man, H. (2004). *A Bayesian game approach for intrusion detection in wireless ad hoc networks*. Proc. 2006 workshop on Game theory for communications and networks, ACM Int. Conf., Pissa, Italy.

Mazzuchi, T. A. & Soyer, R. (1993). A Bayes Method for Assessing Product Reliability during Development Testing. *IEEE Transactions on Reliability* 42(3): 503-510.

Mazzuchi, T. A. & Soyer, R. (1996). Adaptive Bayesian Replacement Strategies. *Bayesian Statistics* **5**: 667-674.

Momani, M.; Challa, S. & Aboura, K. (2007a). Modelling Trust in Wireless Sensor Networks from the Sensor Reliability Prospective. *Innovative Algorithms and Techniques in Automation, Industrial Electronics and Telecommunications*. Sobh, T., Elleithy, K., Mahmood, A. and Karim, M., Springer Netherlands.

Momani, M.; Aboura, K. & Challa, S. (2007b). *RBATMWSN: Recursive Bayesian Approach to Trust Management in Wireless Sensor Networks*. The Third International Conference on Intelligent Sensors, Sensor Networks and Information, Melbourne, Australia, IEEE.

Morris, P. A. (1971). Bayesian Expert Resolution. *Department of Engineering-Economic Systems*, Stanford University. Ph.D.

Morris, P. A. (1974). Decision Analysis Eexpert Use. *Management Science* 20: 1233-1241.

Shafer, G. (1976). *A Mathematical Theory of Evidence*, Princeton University Press.

Singpurwalla, N. D. (1988). An Interactive PC-Based Procedure for Reliability Assessment Incorporating Expert Opinion and Survival Data. *Journal of the American Statistical Association* **83**(401): 43-51.

Srinivasan, A.; Teitelbaum, J.; Liang, H.; Wu, J. & Cardei, M. (2007). Reputation and Trust-based Systems for Ad-hoc and Sensor Networks. *Algorithms and Protocols for Wireless Ad-hoc and Sensor Networks*. Boukerche, A., Wiley & Sons.

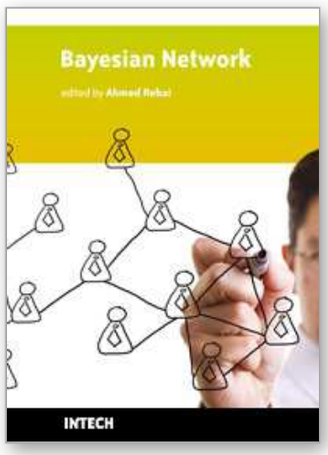
Srinivasan, A.; Teitelbaum, J. & Wu, J. (2006). DRBTS: Distributed Reputation-based Beacon Trust System. *The 2nd IEEE International Symposium on Dependable, Autonomic and Secure Computing (DASC '06)*.

Van Nortwijk, J. M.; Dekker, R.; Cooke, R. & Mazzuchi, T. A. (1992). Expert Judgement in Maintenance Optimization. *IEEE Transactions on Reliability* 41(3): 427-432.

West, M. (1984). Bayesian aggregation. *Journal of the Royal Staistical Society* 147(4): 600-607.

Winkler, R. L. (1968). The Concensus of Subjective Probability Distributions. *Management Science* 15: B61-B75.

IntechOpen



Bayesian Network

Edited by Ahmed Rebai

ISBN 978-953-307-124-4
Hard cover, 432 pages
Publisher Sciyo
Published online 18, August, 2010
Published in print edition August, 2010

Bayesian networks are a very general and powerful tool that can be used for a large number of problems involving uncertainty: reasoning, learning, planning and perception. They provide a language that supports efficient algorithms for the automatic construction of expert systems in several different contexts. The range of applications of Bayesian networks currently extends over almost all fields including engineering, biology and medicine, information and communication technologies and finance. This book is a collection of original contributions to the methodology and applications of Bayesian networks. It contains recent developments in the field and illustrates, on a sample of applications, the power of Bayesian networks in dealing the modeling of complex systems. Readers that are not familiar with this tool, but have some technical background, will find in this book all necessary theoretical and practical information on how to use and implement Bayesian networks in their own work. There is no doubt that this book constitutes a valuable resource for engineers, researchers, students and all those who are interested in discovering and experiencing the potential of this major tool of the century.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen