Providing trusted data for industrial wireless sensor networks

Yu and He *EURASIP Journal on Wireless Communications and Networking* (2018) 2018:289
https://doi.org/10.1186/s13638-018-1307-y

EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

Providing trusted data for industrial wireless sensor networks

Shuyan Yu[1*] and Jinyuan He[2]

Abstract

The deployment of wireless sensor networks, or WSNs, in industrial domains has attracted much attention over the past few years. An increasing number of applications have been developed such as for condition monitoring in the railway industry. Nevertheless, compared with traditional WSNs, the industrial environment is harsher, noisier, and more complex, which poses a higher requirement for the network security especially in terms of data trustiness and which further deters WSN practical integration in industrial applications. The main contribution of this research is to partially address the security issues by means of providing trusted data for industrial WSNs. To this end, a negative binomial distribution-based trust scheme combined with the D–S belief theory and a noise filter method is proposed and designed for industrial WSNs. In this paper, we first discuss the trust theory in WSNs and the disadvantages of traditional trust schemes for industrial applications, then analyze and evaluate the proposed method, and finally compare the performance of our method with some classic trust schemes. Through simulation tests about temperature readings of a factory workshop, it shows that the proposed method can improve the data trustiness, reliability, and robustness in the trust evaluation process under industrial environments and ensure the security of the network.

Keywords: Trusted data, Industrial wireless sensor networks, Noise filter, Negative binomial distribution

1 Introduction

Industrial wireless sensor networks, or IWSNs, have received more and more attention in recent years [1]. IWSNs consist of a certain number of small sensors and several base stations or data sinks [2]. IWSNs are mainly used for collecting and transmitting data from field devices. With limited computing abilities and storage capacities, these battery-powered small device sensors shown in Fig. 1 [3] are usually equipped with sense unit and signal transmission unit. The basic operations of such networks are periodic sensing, data gathering, and data transmission by individual sensor nodes to the data sink via intermediate nodes. Sensor nodes in IWSNs are resource constraint devices since their processing capabilities, power supply, memory capacity, and bandwidth have stringent constraints. But due to their low cost and high scalability, partially with the help of cloud computing [4, 5], IWSNs have been used in a wide variety of real industrial applications ranging from nuclear plant facility management, supply and demand energy management, industrial process control to conditioning monitoring in the railway industry as is presented in Fig. 2 [3]. According to [3], sensor devices are attached to the object being monitored such as tracks, bridges, or train mechanics with one or more sensors mounted on a sensor board; the sensor nodes communicate with the base station using a wireless transmission protocol; the base station collates data and transmits data to the control center server possibly through satellites or GPRS; and the sensor nodes may communicate directly with the server rather than via the base station, or the user accesses the data directly via the base station.

The environment of IWSNs is extremely complex with strict requirements such as speed and reliability [6], and IWSN device nodes are often deployed in unattended or even hostile industrial areas; therefore data trustiness and security must be taken into consideration when the networks are being designed. Further, lack of physical security makes sensor nodes easy to be compromised by

*Correspondence: shuyanyu1231@qq.com
[1]College of Management and Information, Zhejiang Post and Telecommunication College, Shaoxing, China
Full list of author information is available at the end of the article

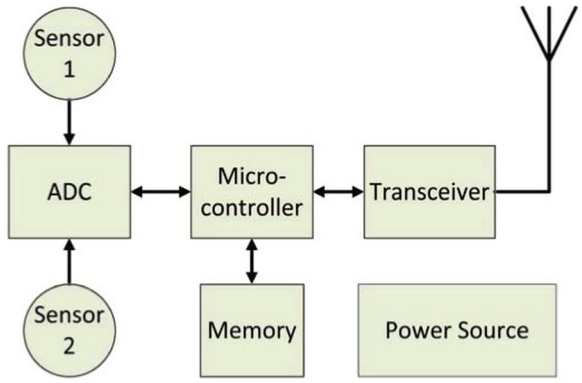

Fig. 1 Components of a sensor node

intruders who will later attack the whole network. If the compromised nodes or unreliable data sources cannot be identified in time, secret information may be revealed and the whole network could be under the control of the adversaries [7]. Besides, individual nodes are not always being honest in their interactions with others and may not provide trust or reliable information for their peers. Thus, the existence of unreliable sources in the network will deteriorate the accuracy as well as the system performance [8], which later threats the full functioning of IWSNs. For example, a body sensor network can remotely monitor the vital information and activities of a patient, but untrusted data might lead to a wrong therapy or even death of this patient.

In order to provide data accuracy and security for the network, trust theory [9–16] has been gradually studied by researchers. Through the evaluation and storing the trust values of sensor nodes in WSNs, it can compute how much data from those nodes can be trusted when they are doing a certain job such as packet delivery and routing response.

In this research, we propose a trust scheme to provide trusted data for IWSNs, which uses the negative binomial distribution as a trust computation model. Considering the noise under industrial environments, a noise filter method is designed and combined with the proposed scheme. The organization of this study is as follows. Section 2 discusses the trust theory in WSNs and the disadvantages of traditional trust schemes for industrial applications, Section 3 analyzes and evaluates the proposed method, Section 4 shows the simulation tests, and Section 5 concludes this work.

2 Related work

Traditional security solutions such as cryptography and intrusion detection have been successfully applied in the computer networks, but when dealing with the internal compromised nodes, these methods are not so effective. The reason is that compromised nodes still have the access to the cryptographic keys that are used to secure the communication links within the network. Additionally, a compromised node pretending to be an authorized one cannot be detected by using cryptographic primitives. Thus, compromised nodes can pretend to be a legitimate one from a cryptographic standpoint while undertaking malicious actions.

To deal with the untrusted source from the malicious nodes, through the evaluation and storing the trust values of WSN nodes, it is possible to know how much those nodes can be trusted when they are undertaking a certain task. Trust schemes, usually defined as a node's belief in the reliability of another one's behaviors or actions, have been studied and proposed as an alternative to traditional security solutions. A typical trust scheme architecture is presented in Fig. 3 [15], and trust properties are shown in Fig. 4 [16].

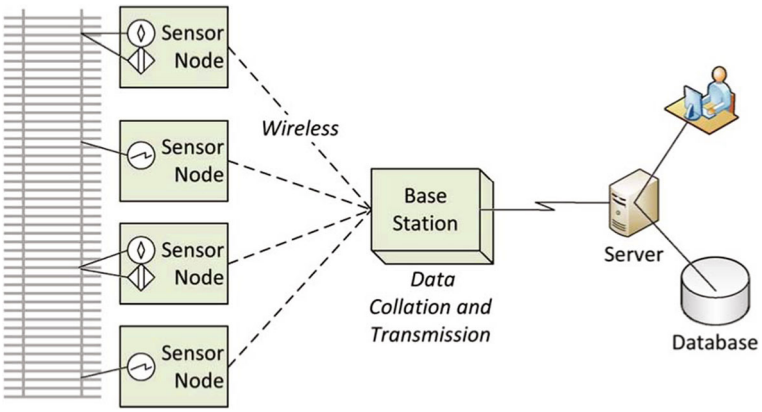

Fig. 2 IWSN application of conditioning monitoring in the railway industry

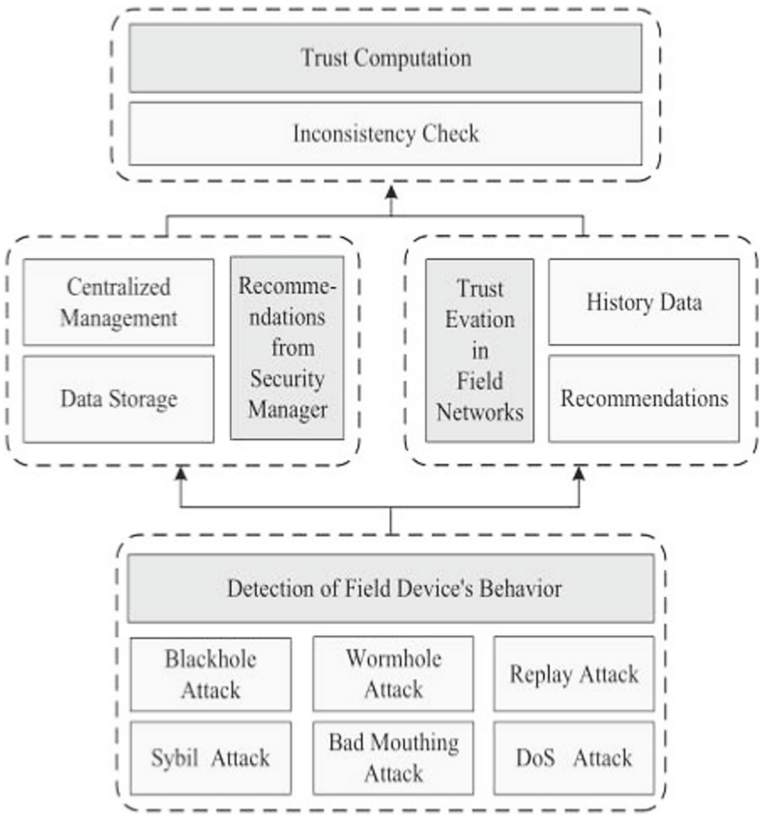

Fig. 3 A typical trust scheme architecture

Among the trust schemes, statistics-based models such as Bayesian theory models [17–24] have received wide attention, based on which many trust models have been proposed by researchers in the past several years. The basis of trust mechanism is that its calculation is either directly based on the historical behaviors of participating nodes or indirectly based on the recommendations from other nodes.

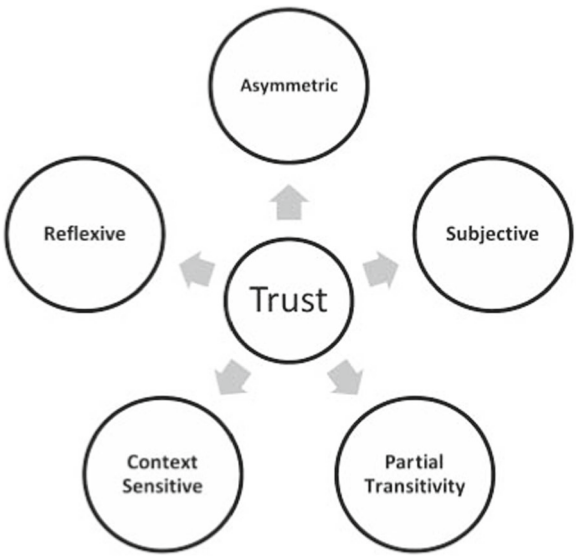

Fig. 4 Trust properties

Generally, Bayesian theory fundamentally conforms to the procedure of trust evaluation. Bayesian theory-based trust system attempts to discover the behavior patterns through historical actions [13]. In Bayesian theory, it first calculates the prior probability of an event, then applies the prior probability into the binomial distribution, and finally modifies or updates the probability by using a posterior inference according to the relevant evidences.

In reputation-based framework for high integrity sensor networks (RFSN) [10], a representative application of binomial distribution-based trust scheme in WSNs, each sensor holds trust metrics representing past behavior of other nodes in order to predict these nodes' future behavior. RFSN uses a completely decentralized method and can run on each sensor node. Nodes in RFSN only interact with other nodes within the wireless communication range; therefore, they only maintain the trust of nodes within the neighborhood.

In RFSN, a transaction is defined as two nodes exchanging information or participating in a collaborative process. Based on the trust metrics built for other nodes by the behavior monitoring mechanism, a sensor node can treat

them as *cooperative* or *non-cooperative* and evaluate the trustworthiness of these nodes. In the practical application, trust in RFSN is defined as the probability that a node will cooperate. In [10], let Θ represent the probability that a certain node will cooperate, and a prior distribution that denotes the probability that a node would cooperate with another one is defined by

$$P(\Theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\Theta^{\alpha-1}(1-\Theta)^{\beta-1} \tag{1}$$

where $0 \le \Theta \le 1, \alpha \ge 0$, and $\beta \ge 0$. Θ can be used as the success probability in Bernoulli observations. For example, let $T \in [0,1]$ be the node *i's* rating for node *j* in one transaction, then

$$P(T|\Theta) = \Theta^{T}(1-\Theta)^{1-T} \tag{2}$$

When the transaction is complete, the posterior of Θ is defined by

$$\begin{aligned} P(\Theta|T) =& \frac{P(T|\Theta)P(\Theta)}{\int P(T|\Theta)P(\Theta)d\Theta} \\ & \sim \text{beta}(\alpha+T, \beta+1-T) \end{aligned} \tag{3}$$

Then, the mathematical expectation of Θ is defined by

$$E(\Theta) = \frac{\alpha+T}{\alpha+T+\beta+1-T} \tag{4}$$

In Eq. (4) or Eq. (5), $E(\Theta)$ can be regarded as the trust value a sensor node in the practical application of WSNs. The higher the value of $E(\Theta)$, the more trusted the sensor node becomes.

After *n* transactions, the mathematical expectation of Θ is defined by

$$E(\Theta) = \frac{\alpha+nT}{\alpha+\beta+n} \tag{5}$$

And the two trust parameters become

$$\alpha = \alpha + nT, \beta = \beta + n \times (1-T) \tag{6}$$

In Eqs. (4), (5), and (6), the trust parameter α and β can be interpreted respectively as the observed number of positive outcomes and the observed number of negative outcomes regarding a certain transaction. For example, in a packet delay transaction, node *i* asks node *j* to transmit its data packets; after 10 requests from node *i*, node *j* has successfully transmitted 5 packets and failed 4 packets, then the trust parameters about node *j* can be expressed as $\alpha+ = 5$ and $\beta+ = 4$ which are observed and recorded by node *i*.

3 The proposed scheme

3.1 Trust computation model

The proposed scheme uses the negative binomial distribution as the trust computation model, which is presented as follows.

In a sequence of independent Bernoulli trials with success probability ρ, let Z denote the number of failures until the *r*th success. The random Z is called the negative binomial random variable with parameters ρ and *s*. Its probability mass function is defined by

$$P(Z=s|r,\rho) = \binom{r+s-1}{s}\rho^{r}(1-\rho)^{s} \tag{7}$$

for $s = 0, 1, 2 \ldots$ and $0 < \rho < 1$.

This is the probability of observing *s* failures before the *r*th success. In Eq. (7), ρ is the binomial success probability and its conjugate prior distribution is beta distribution. Then, the posterior of ρ is defined by

$$\begin{aligned} P(\rho|Z) =& \frac{P(Z|\rho)P(\rho)}{\int P(Z|\rho)P(\rho)d\rho} \\ =& \frac{\Gamma(\alpha+\beta+r+s)}{\Gamma(\alpha+r)\Gamma(\beta+s)}\rho^{\alpha+r-1}(1-\rho)^{\beta+s-1} \end{aligned} \tag{8}$$

It indicates that the posterior $P(\rho\ |Z)$ has a beta distribution with parameters $\alpha+r$ and $\beta+s$. Then, the expectation of ρ is defined by

$$E(\rho) = \frac{\alpha+r}{\alpha+\beta+r+s} \tag{9}$$

Traditional trust schemes are not suitable for IWSNs. For example, in Eq. (4), trust parameters α and β are limited with adding 1 after each transaction, so neighboring nodes have to keep observing the observed node so as to compute and record its trust values. Under some IWSN applications, nodes usually need not to keep tracking a certain event, rather they could be put into sleep state for some time when there is no sensing tasks to follow so that their energy can be saved. In this case, Eq. (4) is not applicable. In contrast, in Eq. (9), increment of trust parameters can be set to a certain number according to the characteristics of specific IWSNs, and after $r+s$ transactions, neighboring nodes can compute once the trust of that node.

In addition, as is shown in Fig. 5, trust from the third parties should be added as indirect references. The indirect method can be mapped into *D-S* belief theory [25]. Assume *j* obtains the trust of *i* through *h*. Let $\left(\alpha_i^h, \beta_i^h\right)$ denote such indirect trust. *j* has the past trust records about *i* and *h*, which is represented by (α_i, β_i) and (α_h, β_h) respectively. After combining with indirect reputation, trust parameters are defined by

$$\alpha_i^{'} = \alpha_i + \frac{2\alpha_h\alpha_i^h}{(\beta_h+2)+\left(\alpha_i^h+\beta_i^h+2\right)+2\alpha_h} \tag{10}$$

$$\beta_i^{'} = \beta_i + \frac{2\alpha_h\beta_i^h}{(\beta_h+2)+\left(\alpha_i^h+\beta_i^h+2\right)+2\alpha_h} \tag{11}$$

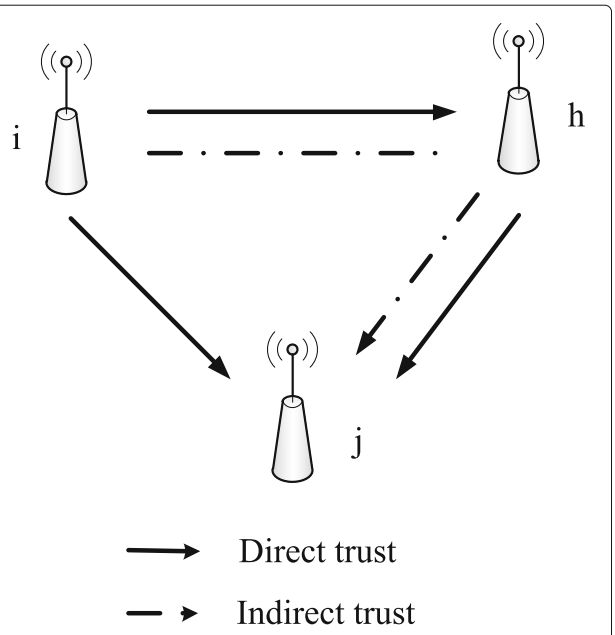

Fig. 5 Direct trust and indirect trust in sensor networks

where $\alpha_i^{'}$ and $\beta_i^{'}$ are the combined trust parameters about node i respectively.

3.2 Noise filter method

Due to the harshness of industrial environment, noises accompanied by radio interference and node temporary error often occur during the deployment of IWSNs. In this case, the device nodes are unable to record all the actual observations whether they are positive or negative, and the trust parameters α and β become the minimum number of observed successes and failures respectively in the real situation. Obviously, noise filter is important for the trust computation of IWSNs. Based on the mean-value theorem of definite integral, the noise filter method designed for IWSNs is presented below.

Let ξ and $f(\xi)$ denote the probability of an event and the corresponding probability density function respectively. According to the mean-value theorem of definite integral, the mean value of $f(\xi)$ is $\frac{\int_0^1 f(\xi)d\xi}{1-0} = 1$. Because $f(\xi)$ has a mean value of 1 and both the increase and the decrease from 1 are counted twice by $|f(\xi) - 1|$, combined with Eqs. (9), (10), and (11), the noise filter denoted by NF is defined by

$$\int_0^1 |\frac{\Gamma\left(\alpha^{'} + \beta^{'} + r + s\right)}{2\Gamma(\alpha^{'} + r)\Gamma(\beta^{'} + s)} \rho^{\alpha^{'}+r-1}(1-\rho)^{\beta^{'}+s-1} - 1|d\rho \tag{12}$$

Then, the expectation of ρ with noise filter is defined by

$$E(\rho_{NF}) = NF \times \frac{\alpha^{'} + r}{\alpha^{'} + \beta^{'} + r + s} \tag{13}$$

4 Results and discussion

To test the performance of the proposed trust scheme, NS−2 is used for the simulation and RFSN as a classic trust scheme that utilized the binomial distribution-based trust is selected for comparison. A cluster-based sensor network is formed to monitor the temperature readings of a factory workshop. Tests in this section consist of three different scenarios, and the parameters are set as follows:

- One hundred twenty sensor nodes are randomly deployed in a rectangular region, and a base station is located in the center.
- Sensor nodes are divided into 8 clusters with 15 nodes in each cluster.
- The temperature reading of each normal sensor node is within (25, 30).
- Assume that there are compromised nodes in the network and their data readings are within (40-45).
- Suppose the NIC of each sensor node is in promiscuous mode so that it can overhear the data packets from its nearby neighbors.
- To simulate the noise, 5% temperature readings from normal sensor nodes is not within (25, 30).
- In each simulation, the base station launches 2000 queries to collect temperature readings from the monitored region.

4.1 Test 1

In this part, changes in proportion of compromised nodes are considered to test the convergence time of the trust schemes. It is desirable that the convergence time should be as fast as possible.

We can notice that in Fig. 6, as the proportion of compromised nodes goes up, convergence time of the two trust schemes increases. For a given proportion value,

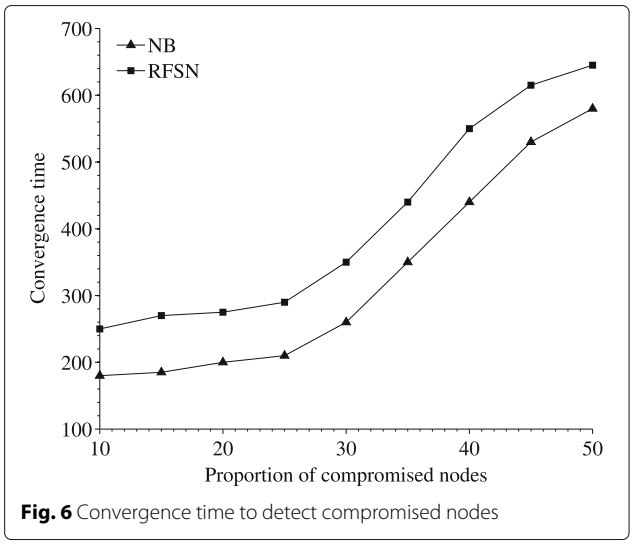

Fig. 6 Convergence time to detect compromised nodes

convergence time of the negative binomial is shorter, e.g., when the proportion is up to 50%, the convergence time is about 585 and 650 for both schemes respectively. It indicates that the *NB* scheme uses less time to detect the malicious nodes and data from the *NB* scheme can be more trusted.

4.2 Test 2

In this part, the success rate of service request attempts that are launched by compromised nodes and answered by normal nodes is tested between the two trust schemes. The lower the rate is, the more reliable the scheme becomes.

In Fig. 7, as the running time increases, the success rate of the compromised nodes drops gradually in both schemes and it comes close to zero at about 1000 and 1200 s for both schemes respectively. The rate drops faster under the *NB* scheme which indicates that it is more effective in checking compromised nodes and that data from *NB* scheme is more reliable.

4.3 Test 3

Trust robustness in both schemes is tested in Figs. 8 and 9 when the compromised nodes are increasing gradually in the network.

In Figs. 8 and 9, compromised nodes increase at a rate of 5% in every 100 query. Before reaching at about the 800 th query (equivalent to about 40% compromised nodes in the network), trust in both schemes can effectively detect the compromised nodes and minimize their influence on the network. From the 800th query, the average trust value of legitimate nodes begins to go downward and the average trust value of compromised nodes start to go upward. This trend continues until to about the 1500th query in Fig. 8 and 1550th query in Fig. 9 (equivalent to about 75% / 77.5% compromised nodes in the network) when the average trust values of both kinds of nodes reach the same point. After that, the average trust value of compromised nodes exceeds that of the legitimate nodes, which makes the whole network compromised and unreliable.

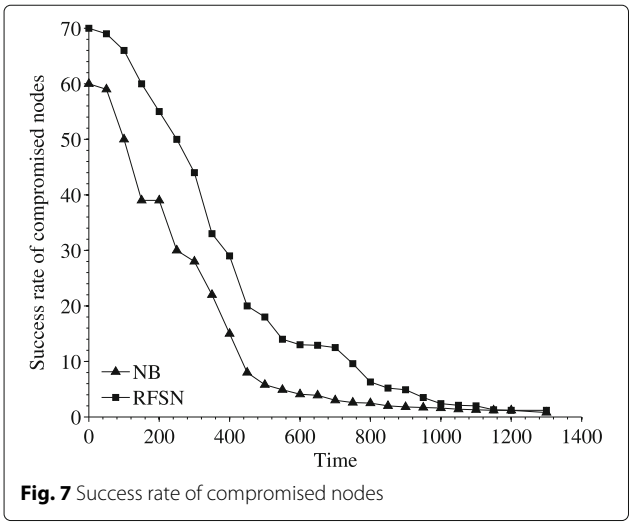

Fig. 7 Success rate of compromised nodes

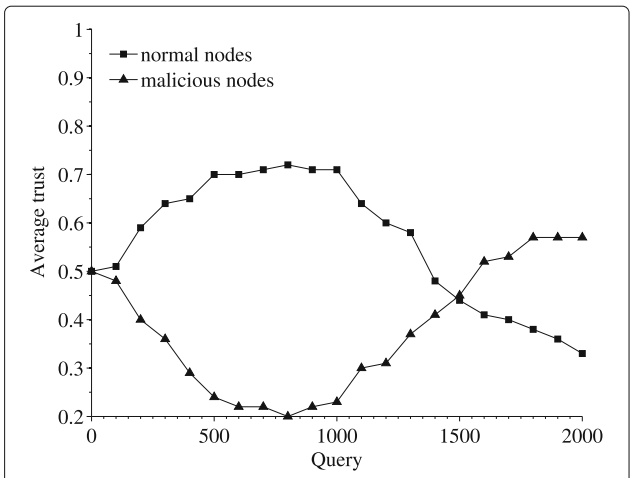

Fig. 8 Average trust with 5% increasing rate of compromised nodes in every 100 query in RFSN

Figures 8 and 9 also indicate that the *NB* scheme is more robust under the attack of compromised nodes and they also show that when the compromised nodes outnumber the legitimate nodes in the both schemes, the network becomes vulnerable and easy to be attacked.

5 Conclusions

Although cryptography primitives can provide the capability to tackle the attacks from the external networks, they cannot address the problem caused by the internal compromised devices, which results in untrusted data in

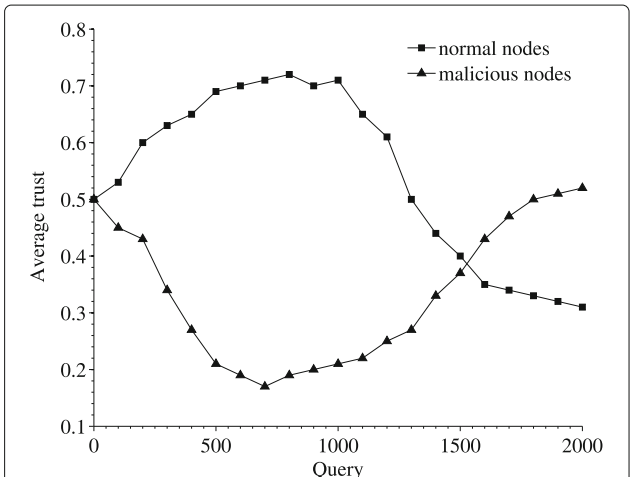

Fig. 9 Average trust with 5% increasing rate of compromised nodes in every 100 query in NB

the network. In this article, a trust scheme with noise filter is proposed to provide trusted data for IWSNs, and the simulations show that the proposed method can improve the data trustiness, reliability, and robustness under industrial environments. In our future work, we will continue to study the noise filter algorithms and refine the trust granularity so as to further improve the data trustiness for IWSNs. Additionally, the behavior monitoring algorithm among sensor nodes will also be studied to enhance the observing accuracy.

Abbreviations

D–S: Dempster–Shafer; GPRS: General packet radio service; IWSN: Industrial wireless sensor network; NB: Negative binomial; NIC: Network interface card; RFSN: Reputation-based framework for high integrity sensor networks; WSN: Wireless sensor network

Acknowledgments

The authors gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

Funding

This work is partially supported by Zhejiang Academy of Education Planning and Research under the grant no. 2017SCG384.

Availability of data and materials

The data sets in the simulation tests are assumed to be the temperature readings of a factory workshop, and the data readings are randomly generated within each value interval; therefore, interested researchers can generate their own random data within the same three value intervals as presented in our simulation tests.

Authors' contributions

SY conducted the experiments and wrote the first draft of the paper. JH helped to revise the paper and polished the paper. Both authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]College of Management and Information, Zhejiang Post and Telecommunication College, Shaoxing, China. [2]Institute of Sustainable Industrial and Liveable Cities, Victoria University, Victoria, Australia.

Received: 22 May 2018 Accepted: 22 November 2018
Published online: 17 December 2018

References

1. B.C. Villaverde, S. Rea, D. Pesch, InRout-A QoS aware route selection algorithm for industrial wireless sensor networks. Ad Hoc Netw. **10**(3), 458–478 (2012)
2. D. Ioannis, Interactive multimedia installation art development using recycled input and sensing devices. Int. J. Arts Technol. **9**(2), 108–125 (2016)
3. V.J. Hodge, S. O'Keefe, M. Weeks, A. Moulds, Wireless sensor networks for conditioning monitoring in the railway industry: a survey. IEEE Trans. Intell. Trans. Syst. **16**(3), 1088–1106 (2015)
4. Z. Xia, Y. Zhu, X. Sun, Z. Qin, K. Ren, Towards privacy-preserving content-based image retrieval in cloud computing. IEEE Trans. Comput. **6**(1), 276–286 (2018)
5. Z. Xia, N.N. Xiong, A.V. Vasilakos, X. Sun, EPCBIR: an efficient and privacy-preserving content-based image retrieval scheme in cloud computing. Inf. Sci. **387**, 195–204 (2017)
6. C. Pei, Y. Xiao, W. Liang, X. Han, Trade-off of security and performance of lightweight block ciphers in Industrial Wireless Sensor Networks. EURASIP. J. Wirel. Commun. Netw. **2018**(1), 117–135 (2018)
7. R. Feng, X. Xu, X. Zhou, J. Wan, A trust evaluation algorithm for wireless sensor networks based on node behaviors and d-s evidence theory. Sensors. **11**(2), 1345–1360 (2011)
8. W. Li, S. Saruwatari, M. Bandai, T. Watanabe, Discussions on trade-offs in data aggregation in wireless sensor networks. Comput. Syst. Sci. Eng. **29**(1), 51–63 (2014)
9. S. Seo, J.W. Kim, J.D. Kim, J.M. Chung, Reconfiguration time and complexity minimized trust-based clustering scheme for MANETs. Eurasip J. Wirel. Commun. Netw. **2017**(1), 155–121 (2017)
10. S. Ganeriwal, M.B. Srivastava, Reputation based framework for high integrity sensor networks. ACM Trans. Sens. Netw. **4**(3), 15–37 (2008)
11. Y. Wang, M. Zhang, W. Shu, An emerging intelligent optimization algorithm based on trust sensing model for wireless sensor networks. Eurasip J. Wirel. Commun. Netw. **2018**(1), 145–154 (2018)
12. G. Amudha, P. Narayanasamy, Distributed location and trust based replica detection in wireless sensor networks. Wirel. Pers. Commun. **102**(4), 3303–3321 (2018)
13. F. Wang, F. Wang, B. Huang, L.T. Yang, SONR: a reliable reputation system of self-organized network. J. Netw. Comput. Appl. **35**(3), 914–926 (2012)
14. T. Zhang, L. Yan, Y. Yang, Trust evaluation method for clustered wireless sensor networks based on cloud model. Wirel. Netw. **24**(3), 777–797 (2018)
15. J. Duan, D. Yang, S. Zhang, J. Zhao, M. Gidlund, in *the 39th Annual Conference of the IEEE Industrial Electronics Society, IECON 2013*. A trust management scheme for industrial wireless sensor networks (Vienna, 2013), pp. 5576–5581
16. O. Khalid, S.U. Khan, S.A. Madani, K. Hayat, M.I. Khan, Comparative study of trust and reputation systems for wireless sensor networks. Security Comm. Netw. **6**(6), 669–688 (2013)
17. M.A.A. Kappel, D.C. Knupp, R.P. Domingos, I.N. Bastos, Analysis of hydrogen permeation in metals by means of a new anomalous diffusion model and Bayesian inference. Comput. Mater. Continua. **49-50**(1), 13–29 (2015)
18. G. Jayaprakash, M.P. Muthuraj, Prediction of compressive strength of various SCC mixes using relevance vector machine. Comput. Mater. Continua. **54**(1), 83–102 (2018)
19. G. D'Angelo, S. Rampone, F. Palmieri, Developing a trust model for pervasive computing based on Apriori association rules learning and Bayesian classification. Soft Comput. **21**(21), 6297–6315 (2017)
20. S. Yoon, Y. Yu, Extended virtual in-situ calibration method in building systems using Bayesian inference. Autom. Constr. **73**, 20–30 (2017)
21. W. Meng, W. Li, Y. Xiang, K.K.R. Choo, A bayesian inference-based detection mechanism to defend medical smartphone networks against insider attacks. J. Netw. Comput. Appl. **78**, 162–169 (2017)
22. E.T. Chancey, J. P. Bliss, Y. Yamani, H. Hah, Trust and the compliance-reliance paradigm: the effects of risk, error bias, and reliability on trust and dependence. Hum. Factors. **59**(3), 333–345 (2017)
23. V.S. Janani, M.S.K. Manikandan, Efficient trust management with Bayesian-Evidence theorem to secure public key infrastructure-based mobile ad hoc networks. Eurasip J. Wirel. Commun. Netw. **2018**(1), 25–52 (2018)
24. V.S. Janani, M.S.K. Manikandan, Mobility aware clustering scheme with bayesian-evidence trust management for public key infrastructure in ad hoc networks. Wirel. Pers. Commun. **99**(1), 371–401 (2018)
25. Z. Jiao, H. Gong, Y. Wang, A D–S evidence theory-based relay protection system hidden failures detection method in smart grid. IEEE Trans. Smart Grid. **9**(3), 2118–2126 (2018)