Approach to Anomaly Detection in Self-Organized Decentralized Wireless Sensor Network for Air Pollution Monitoring

Alexey Meleshko¹, Vasily Desnitsky^{1,}, and Igor Kotenko¹*

¹St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 199178, 39, 14-th Linia, St. Petersburg, Russia.

Abstract. The paper reveals the essence and features of the proposed approach to detecting anomalies in a self-organizing decentralized wireless sensor network (WSN). As a basis for detecting anomalies, the used WSN is intended to monitor atmospheric air pollution near industrial facilities and human life objects. The distinctive features of such a network are the decentralized nature of its structure and services, the autonomy and mobility of the network nodes, as well as the possibility of non-deterministic physical movement of nodes in space. The spontaneous nature of the dynamic formation of the network topology as well as the assignment of roles and private monitoring functions between the available network nodes determines such networks are subject to attacks that exploit the properties of network decentralization and its self-organization. The proposed approach to detecting anomalies is based on the collection and analysis of data from sensors and is designed to increase the security of self-organizing decentralized WSN by identifying anomalies that are critical in the context of the monitoring purposes.

1 Introduction

Currently various cyber-physical systems are gaining more and more developed and aimed at tracking pollution characteristics of the environment. In particular, highly specialized wireless sensor networks with sensor nodes collecting readings and transmitting them to coordinating nodes for subsequent processing and analysis are becoming widespread [1]. Due to flexibility and scalability requirements, in such networks data is transmitted over wireless communication channels, considering principles of self-organization of the network as capabilities of its dynamic rebuilding. Besides the self-organization, there is a need to ensure the decentralization of such networks, i.e. to set and dynamically redistribute target and supporting functions between nodes in the process of the network functioning.

The main contribution of this paper is the developed approach to detecting anomalies in a self-organizing decentralized wireless sensor network for air pollution monitoring. A laboratory oriented hardware/software prototype of a network and an application layer

* Corresponding author: desnitsky@comsec.spb.ru

protocol for decentralized interaction of network devices have been developed on the basis of a self-organizing ZigBee network layer protocol. In addition, a common model of attacking actions aimed at exploiting the properties of decentralization and self-organization of the network has been built. Experiments performed on the constructed hardware/software prototype demonstrate the applicability of the proposed approach.

The key feature of obtained results, which distinguishes it from existing works in the field, is that the proposed approach is focused on the identification and detection of attacks being specific to wireless sensor networks with the self-organizing decentralized control. At the same time, timely detection of compromises of mechanisms for assigning and redistributing the roles in the WSN, including identification of attacks of unauthorized substitution of the node role, will improve the quality indicators of the detection of false network nodes.

The constructed application protocol for decentralized control of the WSN also represents an element of the novelty. In particular, this protocol allows initiating joint, group interaction of nodes in the network. Such interaction, first, ensures the decentralization of the monitoring functions, and, second, supports uninterrupted communication of the nodes in conditions of their spontaneous switching on/off, changing their operation mode, their physical location, other parameters and the network topology.

The rest of the paper is structured as follows. Section 2 presents a review of the related work on the security of WSN monitoring systems. Section 3 describes the requirements. Section 4 provides a description of the WSN prototype for monitoring air pollution, describes the hardware and software components, as well as presents the application protocol for transferring data between participants in network. Section 5 comprises the attack model and results of its analysis. Section 6 comprises fragments of the software/hardware implementation, experiments and analysis of the results obtained.

2 Related works

Avci, et al. describe a decentralized method for real-time detection of WSN damage [2]. Their method is based on 1D convolutional neural networks. For each wireless network node, an individual neural network is trained so that each neural network processes only the local data of the node and works directly with the raw data. This approach does not require data transmission and synchronization, and, according to the authors, it is minimally resource-intensive, since 1D convolutional neural networks combine the tasks of feature extraction and classification into a single training block.

Rajasegarar, et al. considers the problem of decentralized detection of anomalies in a WSN [3]. The proposed approach is built on cluster-based anomaly detection. All nodes perform a clustering operation on their own local data and then send sufficient cluster statistics to their parent node. The parent node merges the resulting set of clusters with its own clusters and creates merged clusters. On their basis, the parent node distinguishes abnormal clusters. The assessment of the normality and non-normality of a cluster is determined by the methods of k nearest neighbors with the parameter $k = 4$. The authors performed practical experiments that showed the applicability of the approach. In comparison with the centralized detection the detection accuracy is a bit lower, but it is at an acceptable level.

Luo and Nagarajan describe a distributed approach to detecting anomalies in WSN for the Internet of Things on the basis of an autoencoder neural network [4]. The proposed algorithm consists of two parts. The first part is based on local anomaly detection on the network sensors, while the second one is focused on anomaly detection by using an IoT cloud. That is, anomalies can be detected locally on sensors in a completely distributed manner without the need for communication with any other sensors or the cloud. Besides a

more complex training task – detection on data from the entire network or a group of network sensors – can be performed in the cloud. Paper [5] is devoted to the description of a decentralized intrusion detection system (IDS). The main feature of this IDS is that its detection functions are distributed over several network nodes. This IDS has been tested on several well-known attacks on WSN, such as Message delay, Repetition, Wormhole, Jamming, Data alteration attacks. The modeling results show that the proposed IDS is effective and sufficiently accurate in detecting various types of the modeled attacks.

Huang and Yuen propose a scheme for decentralizing the data identification process in a WSN [6]. In fact, this scheme represents a protocol for decentralized information processing in WSN, and it is designed to remove the disadvantages of the centralization, such as the transfer of large amounts of data at a high frequency and a large computing load on the central node of the network. The proposed scheme is aimed at reducing the load on the communication channel of the nodes, as well as at reducing the load on the base station. The proposed method can be widely used for online identification and for monitoring the state of constructions by using WSN. Safaei, et al. propose an approach to the detection of anomalies and strongly noisy data of WSN (outliers). The approach is based on the use of time series and an adaptive Bayesian network [7]. The proposed local outlier detection is a decentralized noise detection algorithm that runs on each sensor node individually, and an adaptive Bayesian network is used as a classification algorithm to predict and identify outliers at each node.

Kohno, et al. consider a problem of secure decentralized data transmission, as well as protection against host hijacking attacks [8]. The node hijacking refers to an attack aimed at stealing or cracking secret keys used for data encryption and node authentication. The authors propose a method for transmitting information from nodes that is resistant to hijacking attacks. The declared security is ensured by using the Skipjack symmetric encryption algorithm. The main advantage of the method proposed is the use of the Secret Sharing Scheme, which is based on solving the Lagrange polynomial.

The analysis of the related work showed that the existing papers in the field aimed mainly either at the development of decentralized security systems that are distributed among several nodes of the WSN, or at the creation of secure protocols for decentralizing WSN. At the same time, issues of modeling and evaluation of attacks that exploit the properties of decentralization and self-organization, which have great potential, are currently rather poorly covered in existing works. The purpose of this paper is to study and develop means for detecting anomalies associated with attacks on the decentralization and self-organization mechanisms of the network, taking into account the specificity and nature of such attacks.

3 Requirements to the monitoring

The main object of the research is a wireless sensor network for monitoring air pollution. Such a network includes sensor nodes that are designed to get readings of the characteristics describing the state of the atmospheric air, including humidity, percentage of carbon dioxide or carbon monoxide, etc. Besides the network has a coordinator node, which performs the functions of collecting data from nodes and their processing.

For air pollution monitoring wireless sensor network the following requirements can be identified as the main ones:

- Requirements for decentralization of the network. They are conditioned by the lack of a single center for network management and decision-making on its monitoring. Due to the spontaneity of the network topology formation, the delegating the responsibilities of a single control center to a certain node turns out to be a rather difficult and even practically unattainable task.

- Requirements for self-organization of network nodes. At the network level of the interaction these allow forming automatically network topologies of nodes in a most natural way for the given situation. The feasibility of these requirements at the network level makes it possible to ensure the decentralization of the network at the level of the application protocol.
- Requirements for the target characteristics and monitoring indicators. These requirements determine types of attacks that should be taken into account in the monitoring process as well as the most critical attack types. The requirements also specify target values of the indicators of precision, completeness and fl-measures that determine the quality of the monitoring process.
- Rapidness and feasibility Requirements. These ones expressed in the desired ratio between the resources of network nodes, on the one side, and the required functionality and monitoring characteristics, on the other side.
- Software and hardware compatibility requirements. They specify a list of types of electronic components, network protocols, sketches and software [9].

The specific numerical threshold values used in the requirements are refined for a concrete specification of software and hardware components and WSN functionality.

4 Prototyping and decentralized control protocol

To ensure the possibility of obtaining datasets and simulating various modes of operation of the WSN for monitoring atmospheric air pollution and attack effects, a software/hardware prototype of such a network has been developed. Single-board computers Raspberry Pi Model B (RPI) were used to implement functions for data analysis and other computing procedures. XBee radio communication modules were used to ensure the interconnection of nodes with each other. Arduino controllers were applied to implement data preprocessing on nodes that do not have RPI. Besides sensors designed to control parameters of the air such as temperature and humidity sensors, gas sensors were used. The prototype architecture is shown in Figure 1.

The Python is used as a programming language for implementation of the interaction of sensors and XBee modules with RPI on each of the nodes. To interact with XBee, a special Python library is used, namely digi-xbee. Besides the RPi.GPIO module is used for interconnection with sensors. For each node, the work is performed as follows – using the program installed on the RPi, the sensor readings are got, then the correctness of the readings is checked and transferred through the XBee module for the subsequent analysis. Since the prototype network must support decentralization, in addition to processing its own data, each node in the RPi receives and processes information from other nodes in the network according to its assigned role.

For the interaction of the components of WSN prototype for monitoring of atmospheric air pollution with each other, as well as for the possibility of maintaining the property of decentralization, a decentralized control protocol was developed. Since XBee modules are used to transfer information, the basis is the ZigBee network protocol. And the property of decentralization itself is implemented at the software level of the nodes, that is, at the application level. The application-level decentralization protocol maintains a mechanism for dividing all network nodes into specific roles.

In Figure 2 shows a diagram of the decentralized control protocol. Within the framework of the protocol, the network is specified in the form of a graph of states and transitions. Each state defines a set of actions that should be performed in accordance with this protocol, after that the network transfers to another state. The transition from one state to another is carried out when some logical condition is triggered.

Fig. 1. The architecture of the prototype of WSN.

Fig. 2. State/transition graph for the decentralized control protocol.

In Figure 2 network states are depicted by blocks S₁ - S₅. The beginning and the end of the protocol script are indicated by the Start and Fin blocks, respectively. Table 1 shows the actions performed in each of the network states.

Table 1. States and actions of WSN.

State	Actions
S ₁	Session formation: – establishment of the initial composition of participants (nodes) – negotiation of initial session security parameters, incl. encryption
S ₂	Address and broadcast data transmission by participants
S ₃	Updating session participants: – cross-poll of participants (formation of the current list of available nodes) – adding new nodes from the waiting queue – making a decision on the need or absence of the need to redistribute the roles of the nodes
S ₄	Role reassignment
S ₅	Updating network security settings
S ₆	Forced end of the session

This approach to decentralizing network operation has several advantages, such as dynamically rebuilding node roles, reducing the load on hardware resources due to the distribution of functions, the absence of a single point of failure of the WSN and, as a consequence, increasing the security of the network. The increase in security is due to the fact that a potential attacker does not know in advance the role of each node at a certain point in time. Additionally, it is possible to implement parallel data processing on different nodes with one role. This mechanism allows increasing the effectiveness of the data monitoring functions.

5 Attack model

The developed model of attacks on a wireless sensor network for monitoring atmospheric air pollution covers the results of the classification of possible attacking effects by the type of object that is being affected. Also it covers results of the analysis of possible targets of the attacker. Table 2 shows the possible types of targets that are influenced and their characteristics.

Table 2. Classification of attacks on WSN by the type of object that is affected.

Attack target	Characteristic
Attacks on hardware elements	Impacts on <ul style="list-style-type: none">– sensors– microcontrollers– single board computers– wireless interfaces– peripheral equipment
Attacks on software elements	Modification of the code and its settings, data substitution, and other effects on <ul style="list-style-type: none">– microcontroller sketches– software components
Attacks on decentralized management protocol	Impacts at the network level, including violations of addressing mechanisms, routing, ensuring the timeliness of data delivery, ensuring adequate channel bandwidth, ensuring fault tolerance and availability of data and services Application-level impacts, including violations of data confidentiality and integrity mechanisms as well as data source authenticity
Attacks on physical environment, i.e. object of monitoring	Indirect attacks on network sensors by means of mechanical, chemical, radio-electronic and other influences

The main goals of a potential attacker trying to compromise the target network are the following. These are an unauthorized retrieval of data on the network structure, parameters and composition of nodes, dynamic characteristics of node activity, user and business data. Besides it includes modification and substitution of data, code, settings; modification of the hardware of the nodes, incl. false node injection; tampering with network management processes; interference in the monitoring process. The developed attack model maps the possible targets of the attacker and the objects of the attack to possible scenarios. Also it ranks them according to the degree of criticality and feasibility.

6 Experiments and discussion

To demonstrate the work of the prototype, and to confirm the feasibility of the proposed approach, an experiment was performed to model an attack on this prototype and to detect

it. A man-in-the-middle attack was modeled, its essence is to connect a malicious host to a network and then eavesdrop on network data.

The modeled attack is aimed at violating confidentiality and integrity of data. The attack consists in introducing a node controlled by an attacker into the network under the guise of a legitimate network expansion. The attack is modeled by exploiting the self-organizing properties of the network as a result of its legitimate extensibility. The attack is illustrated in Figure 3.

Fig. 3. Modeled attack exploiting self-organization.

Figure 3 demonstrates that a new node appears in the WSN, and through this node data is retransmitted from other nodes. In the process of retransmission, an attacker is able to eavesdrop on data on a node and modify it. This attack is detected through the use of intelligent data analysis, a number of supervised machine learning methods, including the k-nearest neighbors method, support vector machine, etc. The feasibility of detection based on the analysis of data from sensors has been experimentally confirmed.

The full-scale modeling of the attack on the WSN prototype made it possible to confirm its feasibility in practice. The feasibility of this attack is due to the specificity of the implementation of the properties of self-organization and decentralization of such networks as ZigBee, SigFox, etc., namely, the mechanism for identifying nodes of the same network. However, in this case, it is possible to detect an attack by analyzing transmitted data from all nodes. The prototype is suitable for modeling attacks on WSN and makes it possible to evaluate such attacks. In the future it is planned to expand the list of investigated attacks and conduct their modeling in order to study the possibility of their detection.

7 Conclusion

An approach to detect anomalies in WSN data for monitoring atmospheric air pollution has been proposed. It is based on a decentralized search for anomalies in relation to the considered type of WSN. The approach includes intelligent data analysis by using unsupervised machine learning techniques. Intelligent data analysis is of a group nature and is fulfilled in a decentralized manner through using the collected data and computing power

from several network nodes. Within the framework of the constructed attack model, the attacks that are most critical for WSN are identified. As part of the practical part, a software/hardware prototype of a fragment of an air pollution monitoring network and a decentralized management protocol for it were developed. The performed experiments confirmed the feasibility of the developed approach and the achievement of the monitoring quality indicators.

Further improvement of the developed prototype is planned as well as the expansion of the analyzed types of attacks and anomalies. In addition, a formal specification and validation of the constructed distributed communication protocol is assumed.

This work is performed by a partial financial support of Russian Foundation for Basic Research (project 19-07-00953).

References

1. A. Meleshko, V. Desnitsky, I. Kotenko, *Machine learning based approach to detection of anomalous data from sensors in cyber-physical water supply systems*, IOP Conf. Series: Materials Science and Engineering, IOP Publishing, v. **709**, n. 033034, pp. 1-7, doi: 10.1088/1757-899X/709/3/033034 (2019)
2. O. Avcı, O. Abdeljaber, S. Kiranyaz, M. Hussein, D.J. Inman, *Wireless and real-time structural damage detection: A novel decentralized method for wireless sensor networks*, Journal of Sound and Vibration, v. **424**, pp. 158–172, doi:10.1016/j.jsv.2018.03.008 (2018)
3. S. Rajasegarar, C. Leckie, M. Palaniswami, J. C. Bezdek, *Distributed Anomaly Detection in Wireless Sensor Networks*, 10th IEEE Singapore International Conference on Communication Systems, pp. 1-5, doi: 10.1109/ICCS.2006.301508 (2006)
4. T. Luo, S. G. Nagarajan, *Distributed Anomaly Detection Using Autoencoder Neural Networks in WSN for IoT*, 2018 IEEE International Conference on Communications (ICC), pp. 1-6, doi: 10.1109/ICC.2018.8422402 (2018)
5. A.R. Silva, H.T. Martins, P.S. Rocha, A.F. Loureiro, B. Ruiz, H.C. Wong, *Decentralized intrusion detection in wireless sensor networks*, Proceedings of the 1st ACM international workshop on Quality of service & security in wireless and mobile networks (Q2SWinet '05), pp. 16–23. doi: <https://doi.org/10.1145/1089761.1089765> (2005)
6. K. Huang, K. Yuen, *Online dual-rate decentralized structural identification for wireless sensor networks*, Structural Control and Health Monitoring, pp 1-18, doi:10.1002/stc.2453 (2019)
7. M. Safaei, A.S. Ismail, H. Chizari, M. Driss, W. Boulila, S. Asadi, M. Safaei, *Standalone noise and anomaly detection in wireless sensor networks: A novel time-series and adaptive Bayesian-network-based approach*, Software: Practice and Experience, pp. 1-19, doi:10.1002/spe.2785 (2020)
8. E. Kohno, T. Ohta, Y. Kakuda, *Secure decentralized data transfer against node capture attacks for wireless sensor networks*, International Symposium on Autonomous Decentralized Systems, pp. 1-6, doi: 10.1109/ISADS.2009.5207366 (2009)
9. V. Desnitsky, A. Chechulin, I. Kotenko, D. Levshun, M. Kolomeec, *Combined Design Technique for Secure Embedded Devices Exemplified by a Perimeter Protection System*, SPIIRAS Proceedings, v. **48**, pp.5-31, doi: 10.15622/sp.48.1 (2016)