Detection of Malicious Node in RPL-Based Internet of Things through Provenance

Sabah Suhail, Shashi Raj Pandey, Choong Seon Hong
Department of Computer Engineering, Kyung Hee University, Yongin, 446-701 Korea
Email: sabah, shashiraj, cshong@khu.ac.kr

Abstract

In the Internet of Things (IoT), resource-constrained things are connected to the Internet via IPv6 and 6LoWPAN networks. Such interconnection is provided by the Routing Protocol for Low-Power and Lossy Networks (RPL). However, the data transportation using RPL protocol is subjected to various attacks due to the interconnection of unattended things with untrusted Internet. For instance, the data generated by sensors are vulnerable to attacks (including data forgery, data disruption or data drop) and therefore, the accurate and reliable information cannot be assured in the decision-making process. Provenance can be used to keep track of data traversal. In this paper, we have used provenance to trace packet path along with packet data rate (PDR) to evaluate the performance of the network at each forwarding node in network topology. We have evaluated the proposed approach by computing PDR and throughput with reference to the position of malicious node in the network.

Keywords*—*RPL, LLN, 6LoWPAN, IoT, provenance, PDR

I. Introduction

IoT is deployed in numerous pervasive application areas, for instance, environmental monitoring, energy management, health-care system, industrial automation, surveillance and military [1]. To enable the interconnection of resource-constrained things with the Internet, RPL routing protocol has been standardized for constrained environments such as 6LoWPAN networks. However, RPL implementations do not enable secure operation modes and are vulnerable to various attacks, for example, packet drop attack. To identify the source of an attack, it is important to find the source causing the data loss or network interruption.

Provenance can be used to keep track of data source and the actions performed by the participating entities during the data propagation and processing [2]. Provenance has been used extensively in various application areas including databases, scientific workflows, distributed systems, and networks [3]. However, the use of provenance in IoT domain still requires attention. We have discussed the integration of provenance in IoT in [4], [5] and [6]. Due to the resource-constrained nature of sensor devices, the inclusion of provenance in IoT requires a couple a of constraints, for instance, storage, energy, processing [2]. The main contribution of this paper are:

- We propose a provenance-enabled scheme for RPL-based IoT environment.
- We use Packet Delivery Ratio (PDR) as provenance information at each forwarding node in the packet path.
- We evaluate the proposed scheme in terms of PDR and throughput.

The rest of the paper is organized as follows: Section II presents the RPL overview. Section III discusses the system model including network, provenance and attacker models. Section IV explains the working of the provenance scheme. Section V presents the experimental evaluation results. Finally, we conclude the paper with future research directions in Section VI.

II. RPL Overview: Topology formation and packet forwarding

RPL is a gradient-based proactive routing protocol for low power and lossy network (LLN) that builds directed acyclic graphs (DAGs) based on routing metrics and constraints. In order to construct the DODAG, the nodes exchange control messages[1]. First, the root node multi-cast DIO messages. Neighbors of root node receive DIO message and use this information to compute their rank, join DODAG, and choose a preferred parent. If a node can not receive a DIO after a predefined time interval, it may explicitly solicit the DIO messages from the neighbor nodes through DIS message. DAO is also uni-casted by the child nodes to their respective parent node in case of point-to-multipoint (P2M) and point-to-point (P2P) connectivity. For topology maintenance, the DIOs are sent periodically controlled by the trickle timer.

To send a packet in RPL network, the source (or child) node forwards the data packet to its preferred parent node. The forwarding nodes keep on forwarding the data packet based on routing information derived from DAG topology formation.

III. System Model

In this section, we discuss about the main components of our network model, provenance model and the attacker model.

A. Network Model

The network is modeled as a graph $G(N_f, R)$ where N_f represents the set of forwarding nodes (or the parent nodes) responsible for forwarding the data packet based on routing

[1] DODAG Information Solicitation (DIS), DODAG Destination Advertisement Object (DAO) and DODAG Information Object (DIO)

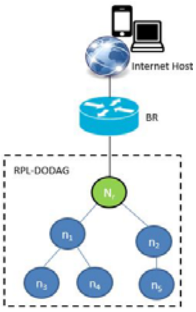


Fig. 1: Interconnection of RPL-connected things to the Internet through the Border Router(BR).

information R. The routing information R in maintained by the Routing Table (RT). [2]

B. Provenance Model

The provenance data (P_{data}) consists of PDR at N_f. To compute the average PDR, we measure the number of sent packets from all the nodes to the sink node and divide it by the number of successfully received packets at the sink node.

$$\begin{aligned} \mathcal{A}veragePDR = (TotalPacketsReceived/ \\ TotalPacketsSent) * 100 \end{aligned} \tag{1}$$

Thus, at any given forwarding node, the provenance data can be defined as:

$$\mathcal{P}_{data} = PDR@N_f \tag{2}$$

C. Attacker Model

In this paper, we assume that the malicious node can only impersonate as forwarding node N_m. Whenever it receives a data packet d from a legitimate forwarding node, it may drop some of the data packets randomly, thus affecting the PDR at subsequent forwarding nodes.

IV. PDR Provenance Scheme

In this section, we discuss the provenance scheme that employs the use of PDR at each forwarding node to identify the packet loss or packet drop (maliciously or accidentally).

A. Provenance Embedding

In order to embed the provenance as PDR in RT, we compute PDR at each forwarding node and store it in RT against each route entry. To illustrate the working of provenance embedding let us consider an example (topology shown in Fig.2. The data generating node n_1 send data to its preferred parent n_2. n_2 will compute the PDR (as in eq.1) and insert it the RT against the route entry $\langle n_2, n_1 \rangle$. Similarly, n_2 forward the data packet to its preferred parent n_3 and update its RT by computing PDR for route entry $\langle n_3, n_2 \rangle$. The RT @ n_3 is shown in Fig.2.

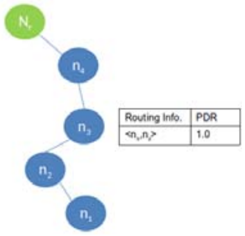


Fig. 2: Example: Provenance information as PDR in the RT @ n_3

TABLE I: Network parameters used in simulation

Parameter	Value
Network layer	RPL
Simulation time [4]	600 s
Number of forwarding nodes	8
Packet size (excluding header)	50 bytes
Data rate	1packet/5sec
τ	0.8 ~1.0

B. Provenance Verification

To verify the provenance information embedded as PDR, the root node can check the RT at each node after some interval I. In that case, if PDR is less than the minimum PDR threshold τ [3] then it will mark the immediate child node as malicious node. Another reason for low PDR can be any form of network discrepancy (congestion or link loss). We will consider that case as a part of our future work.

V. Simulation

We run our experiments on Cooja that is a Contiki based simulator [7]. For our simulations, we use Tmote sky as things. Other parameters are mentioned in Table I.

A. Benign Scenario

To evaluate the working of the proposed scheme for IoT devices, we consider 1 data generating node (node 10), 8 forwarding nodes (node 2 to node 9) and 1 sink node (node 1). The data packet generated by node 10 keeps on traversing the packet path (9→8→7→6→5→4→3→2→1) based on the routing information in RT. RT also stores information about PDR against each route entry. Fig. shows that PDR is 1 for case of no malicious node.

B. Malicious Scenario

We consider a malicious node N_m. We place the N_m at two different positions (as shown in Fig.4b and 4c). We have analyzed that if N_m is placed near the source node then it may drop packets initially on the packet path. Thus, causing a significant drop down in the PDR at many nodes making the packet drop attack obvious. On the other hand, if the N_m

[2] Each node maintains a RT that contains routing information about its child node(s).

[3] The value of τ can be decided on the basis of data rate depending on application.

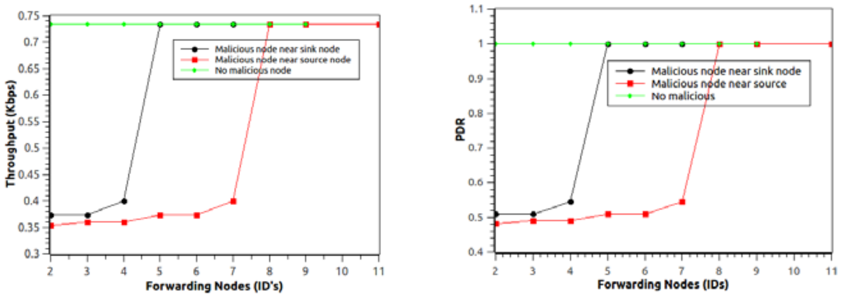
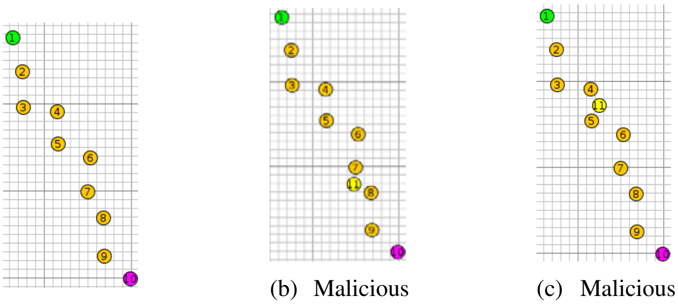


Fig. 3: Throughput and PDR at forwarding nodes.

is placed near the sink node, then the PDR is dropped out at only a few nodes, thus making it hard to detect the packet drop attack. Moreover, the forwarding nodes near the sink are prone to packet drop issues due to traffic load at them. Hence, we can say that the computation of PDR will not only help to detect the packet drop attack but also identify the other network discrepancies, for instance, congestion, link loss etc.

We have computed average PDR at each forwarding node (shown in Fig. . It can be concluded that if the malicious node is near the source node then many of the forwarding nodes are affected, however, if a malicious node is near sink node then the PDR abnormality can be depicted by a few nodes. Also, the identification of a malicious node is usually based on the node soon before the node where PDR is dropped (for example, in scenario 4b PDR starts to drop at node 7 while in scenario 4c PDR drops at node 4). Similarly, we have computed average throughput at each forwarding node as follows:

$$\mathcal{T}hroughput = (TotalPacketsReceived * Packetsize(bits)) \\ /Simulation\ time(sec) \qquad (3)$$



(a) Benign nodes.

(b) Malicious node near source node.

(c) Malicious node near sink node.

Fig. 4: Cooja screenshot of an RPL network showing network configurations for benign nodes and malicious node

VI. Conclusion

In this paper, we have presented a provenance-based scheme for detection of malicious nodes in RPL-connected networks. We have computed PDR at each forwarding node in the packet route and add it as a provenance information in the routing table maintained by each node. Based on the data rate we have identified the minimum threshold of PDR. If PDR at any of the forwarding node is below that certain threshold then we can identify the malicious node. In future, we will other network parameters also to identify the disruption in the network. Also, we will incorporate other factors that cause the drop down of PDR.

Acknowledgment

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (No.2015-0-00557, Resilient/Fault-Tolerant Autonomic Networking Based on Physicality, Relationship and Service Semantic of IoT Devices) *Dr. CS Hong is the corresponding author.

References

[1] Sabah Suhail, Choon Seong Hong, Zuhaib Uddin Ahmad, Faheem Zafar, and Abid Khan. Introducing secure provenance in iot: Requirements and challenges. In *Secure Internet of Things (SIoT), 2016 International Workshop on*, pages 39–46. IEEE, 2016.

[2] Sabah Suhail, Choon Seong Hong, Lodhi Ali, Faheem Zafar, Abid Khan, and Faisal Bashir. Data trustworthiness in iot. In *ICOIN*. IEEE, 2018.

[3] Faheem Zafar, Abid Khan, Saba Suhail, Idrees Ahmed, Khizar Hameed, Hayat Mohammad Khan, Farhana Jabeen, and Adeel Anjum. Trustworthy data: A survey, taxonomy and future trends of secure provenance schemes. *Journal of Network and Computer Applications*, 94:50–68, 2017.

[4] Sabah Suhail, Shashi Raj Pandey, Choong Seon Hong, and M Ali Lodhi. Trustworthy data communication in vanet. , pages 1089–1091, 2017.

[5] Sabah Suhail, Choong Seon Hong, Faheem Zafar, and Adeel Anjum. Are the recommendations from recommender system trustworthy? , pages 1060–1062, 2017.

[6] Sabah Suhail and Choong Seon Hong. A secure provenance-aware model for internet of things. , pages 1154–1156, 2016.

[7] Adam Dunkels, Bjorn Gronvall, and Thiemo Voigt. Contiki-a lightweight and flexible operating system for tiny networked sensors. In *Local Computer Networks, 2004. 29th Annual IEEE International Conference on*, pages 455–462. IEEE, 2004.