Detection of Selective Forwarding Attack 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 through IPv6 and 6LoWPAN networks. The Routing Protocol for Low-Power and Lossy Networks (RPL) has enabled such interconnection. However, the data transportation using RPL is vulnerable to various attacks due to the interconnection of unattended things with the untrusted Internet. For instance, the data generated by sensors are vulnerable to attacks (selective forwarding attack) and therefore, the error-free and reliable information cannot be assured in the decision-making process. Provenance can be used to keep track of data acquisition and data traversal. In this paper we use provenance to evaluate the performance of the network by computing packet delivery ratio (PDR) at each forwarding node in the packet path. Moreover, for investigating the faulty nodes, we maintain the count for received packets from respective child nodes in the routing table at each parent node. We have evaluated the proposed approach in terms of provenance size and provenance generation time.

Keywords—IoT, Provenance, PDR, RPL, Selective forwarding attack, Anomaly-based detection, 6LoWPAN

I. Introduction

IoT is deployed in various application areas, for instance, environmental monitoring, energy management, health-care system, industrial automation, surveillance and military [1]. To enable communication between the resource-constrained things with the Internet, RPL routing protocol has been standardized for constrained environments (6LoWPAN networks). However, secure operation modes are usually not enabled by RPL implementations making it vulnerable to various attacks, for example, selective forwarding attack. In selective forwarding attack [2], an attacker forwards only selected packets. For instance, an attacker could forward only routing or control messages and drop all other packets to disconnect nodes from the network. 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 [3]. Provenance has been used extensively in various application areas including databases, scientific workflows, distributed systems, and networks [4]. However, the use of provenance in IoT domain still requires considering constraints, for instance, storage, energy, processing [1]. We have discussed in detail about the integration of provenance in IoT in [5], [6], [7], and [8]. The main contribution of this paper are:

- We propose a provenance-enabled scheme to identify selective forwarding attack in RPL-based IoT environment.
- To identify anomaly in the network, we compute PDR as provenance information at each forwarding node and embed it in the payload. For further investigation of malicious node, we insert the packet count received by the child node in RT of the parent node.
- We evaluate the proposed scheme in terms of provenance size and provenance generation time.

The rest of the paper is organized as follows. Section II presents the RPL overview. Section III discusses the system model including network, data, provenance and attacker models. Section IV explains the working of the provenance scheme. Section V presents the 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 constructs a directed acyclic graphs (DAGs) by utilizing routing metrics and constraints. To construct the DODAG, the nodes exchange control messages[1]. The root node multicast DIO messages. Upon reception, the neighboring nodes compute their rank, join DODAG, and choose a preferred parent. If a node can not receive a DIO after a specific time interval, it solicit the DIO messages from the neighbor nodes via DIS message. DAO is also unicasted by the child nodes to their respective parent node to establish point-to-multipoint (P2M) and point-to-point (P2P) connectivity. For overall topology maintenance, the DIOs are sent aperiodically based on the trickle timer. To send a packet, the source node forwards the data packet to its preferred parent node. The forwarding nodes based on routing information keep on forwarding the data packet until packet arrives at sink node for further data processing and analysis.

III. System Model

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

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

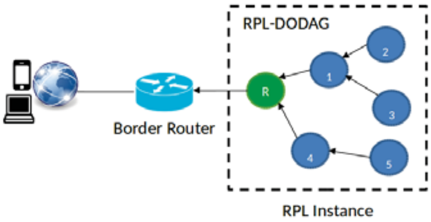


Fig. 1: Interconnection of RPL-connected things to the Internet through the Border Router(BR).

A. Network Model

The network is modeled as a graph $G(N, R)$ where N represents the set of nodes (including source and forwarding) responsible for generating and forwarding the data packet based on routing information R maintained by the Routing Table (RT). [2]

B. Data Model

A data packet consists of the following entities: i) a unique sequence number, ii) payload, iii) PDR.

C. Provenance Model

The provenance data (P_{data}) consists of PDR computed at forwarding node (n_f) with respect to number of packets sent by the source node (n_s). The average PDR can be computed as:

$$\mathcal{A}vg.\ PDR = (packet_r\ by\ n_f/packet_s\ by\ n_s). \qquad (1)$$

Thus, at any n_f, the P_{data} can be defined as:

$$\mathcal{P}_{data} = payload \leftarrow PDR@n_f. \qquad (2)$$

Also, each n_f stores the total number of packets received (p_r) by its child node at its RT.

$$\mathcal{RT}@n_f \leftarrow p_r. \qquad (3)$$

D. Attacker Model

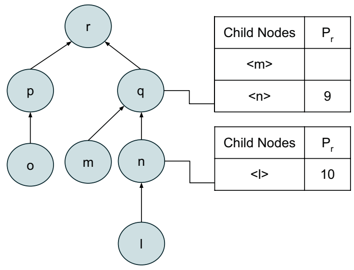
In this paper, we assume that the malicious node (M) can only impersonate as a forwarding node. Whenever it receives a data packet from a legitimate forwarding node, it may drop some of the data packets randomly, thus affecting the PDR at subsequent forwarding nodes.

IV. Provenance Scheme

In this section, we discuss the provenance encoding and decoding scheme that compute *PDR* at each forwarding node in the packet path to identify the anomaly (high packet drop rate) occurred either by the malicious nodes or by network disruptions. We have adopted anomaly-based detection mechanism that tries to detect anomalies in the system by determining the ordinary behavior and using it as a baseline. Therefore, any deviations from that baseline are considered an anomaly.

A. Provenance Embedding

The process of provenance embedding consists of two main steps. Firstly, upon packet reception, the forwarding node (n_f) computes PDR with respect to the number of packets sent by the source node (n_s). It then embeds the PDR in the payload. Secondly, n_f inserts the total number of received packets ($packet_r$) from its immediate child node in its routing table (RT) against respective child node entries. Let us consider an example scenario to illustrate the working of the proposed scheme. l as a source node sends data to its preferred parent n. n will compute the PDR (as in eq.1) and inserts it the payload. n computes PDR and embeds it in the payload. It also inserts the count of $packet_r$ in the RT against the routing entry of its child node l and then forwards the packet to q. Similarly, the next node in the packet path i.e., q updates the payload by inserting its PDR and updates its RT entries with the count of $packet_r$ from its child node n. This process continues until the packet arrives at the root node r.



Child Nodes	P_r
<m>	
<n>	9

Child Nodes	P_r
<l>	10

Fig. 2: Example showing working of provenance encoding.

B. Provenance Decoding

The root node (n_r) verify the P_{data} in two incremental steps. *Step 1:* it extracts the PDR from the payload and compare it with the minimum baseline PDR threshold (τ)[3]. If PDR is less than τ, then it leads to some suspicious anomaly in the network and hence, n_r proceed to next step. *Step 2:* n_r checks the RT of the n_f in the packet path[4] and recompute the PDR with respect to the p_s to identify the M. n_r may also carry out the data trustworthy assessment process after some predefined interval I. Other reasons affecting the PDR can be any form of network discrepancy (congestion or link loss). We may consider those factors as a part of our future work.

V. Simulation

We perform simulations on Cooja (Contiki-based simulator) using Tmote sky as *things*. Other simulation parameters are given in Table I.

[2] Each node maintains 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.

[4] r has acquired the knowledge of the packet path during topology formation

TABLE I: Network parameters used in simulation

Parameter	Value
Network layer	RPL
Simulation time (excluding topology convergence time)	600 s
Packet size (excluding header)	50 bytes
Data rate	1packet/10sec
minimum baseline PDR threshold (τ)	1.0

A. Performance Metrics

1) Provenance Generation Time: Provenance generation time can be defined as the amount of time required by each forwarding node to compute and embed provenance and then forward the data packet to the next neighboring node.

2) Provenance Size: Provenance size can be defined as the amount of extra meta-data included in the payload as provenance data.

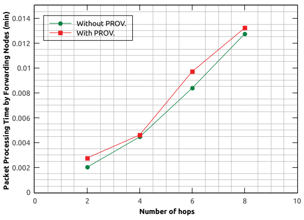


Fig. 3: Comparison: Packet Processing Time (min).

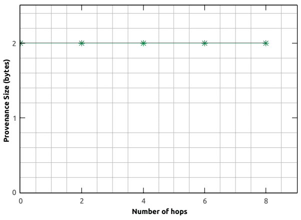


Fig. 4: Provenance Size at different hops.

Fig. 3 shows the packet processing time in case of only RPL and provenance-based RPL. It can be seen that the provenance generation time (min) is almost negligible. Fig. 4 shows the provenance size which is 2 bytes as we are only embedding the PDR in the payload as P_{data}. We setup a simulation environment consisting of 8-hop network with node ID 3 marked as malicious node that performs selective forwarding attack. We set natural packet loss as 1%. Fig. 5 shows that the packet loss rate becomes more obvious as the packet drop increases from 3% to 7%.

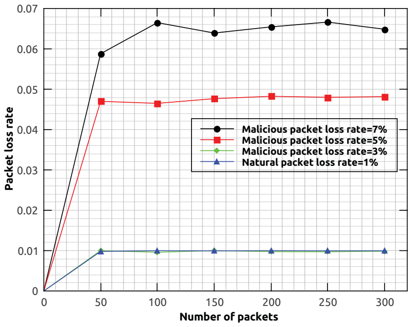


Fig. 5: Packet Loss Rate.

VI. Conclusion

In this paper, we have presented a provenance-based scheme for detection of malicious nodes performing selective forwarding attack in RPL-connected networks. To identify the malicious node that is performing selective forwarding attack, we have computed PDR at each forwarding node in the packet route and add it as a provenance information in the payload. We have also added the total number of packets received by the forwarding node in the routing table against its respective child node route entry. Based on the received packet count at the forwarding node, we compute PDR and identify the faulty node. If the PDR of the forwarding node is below baseline threshold (for instance, less than natural link loss rate=1%) then we can detect the anomaly in the network. In the future, we will extend the proposed scheme to identify other attacks in the RPL network.

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] Chris Karlof and David Wagner. Secure routing in wireless sensor networks: Attacks and countermeasures. In *Sensor Network Protocols and Applications, 2003. Proceedings of the First IEEE. 2003 IEEE International Workshop on*, pages 113–127. IEEE, 2003.

[3] Sabah Suhail, Choong Seon Hong, M Ali Lodhi, Faheem Zafar, Abid Khan, and Faisal Bashir. Data trustworthiness in iot. In *Information Networking (ICOIN), 2018 International Conference on*, pages 414–419. IEEE, 2018.

[4] 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.

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

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

[7] Sabah Suhail, Shashi Raj Pandey, and Choong Seon Hong. Detection of malicious node in rpl-based internet of things through provenance. *한국정보과학회 학술발표논문집*, pages 1171–1173, 2018.

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