
Peer reviewed version

Link to published version (if available): 10.1109/WF-IoT.2016.7845515

Link to publication record in Explore Bristol Research

PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via IEEE at http://ieeexplore.ieee.org/document/7845515/. Please refer to any applicable terms of use of the publisher.

**University of Bristol - Explore Bristol Research**

**General rights**

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/user-guides/explore-bristol-research/ebr-terms/
Traffic Forensics for IPv6-Based Wireless Sensor Networks and the Internet of Things

Vijay Kumar, George Oikonomou, and Theo Tryfonas
Faculty of Engineering, University of Bristol, Bristol, UK
Email: vk12122@my.bristol.ac.uk; {g.oikonomou; theo.tryfonas } @ bristol.ac.uk

Abstract—Research and standardisation efforts in the fields of Wireless Sensor Networks (WSNs) and the Internet of Things (IoT) are leading towards the adoption of TCP/IP for deployments of networks of severely constrained smart embedded objects. As a result, wireless sensors can now be uniquely identified by an IPv6 address and thus be directly connected to and reachable from the internet. This has a series of advantages but also exposes sensor deployments to new security vulnerabilities. Should a deployment be compromised, post-incident analysis can provide information about the nature of the attack by inspecting the network’s state and traffic during the time period prior, during and after the attack. In this paper we adopt traffic forensic techniques in order to achieve post-hoc detection of attacks against availability in IPv6-based Low-Power Wireless Personal Area Networks. To this end, we first implement an attack which exploits inherent vulnerabilities of the IPv6 Routing Protocol for Low-Power and Lossy Networks (RPL). Subsequently, we present an automated method to detect and analyse this attack by examining network packet captures.

Index Terms—6LoWPAN Forensics, Traffic Forensics, Wireless Sensor Networks

I. INTRODUCTION

The development of standards-compliant TCP/IP stacks for embedded devices has been among the key enablers for the development of Internet of Things applications. The Institute of Electrical and Electronics Engineers (IEEE)’s 802.15.4 standard for low-power wireless communications is among the key building blocks of WSN deployments. IPv6 over Low Power Wireless Personal Area Networks (6LoWPAN) [1] and related Internet Engineering Task Force (IETF) specifications [2], [3] have made it possible to use IPv6 in networks of networked embedded smart objects. For those networks, the RPL [4] is the de-facto standard for routing.

The adoption of protocols of the TCP/IP family has made WSNs vulnerable to new security threats. For instance, the Internet Control Message Protocol Version 6 (ICMPv6) is used to perform key functions in IPv6 networks, one of which is Neighbor Discovery (ND). IPv6 ND’s security vulnerabilities have been previously documented [5], [6], while a detailed overview of 6LoWPAN security threats and countermeasures is presented in [7].

Additionally, IPv6 Routing Protocol for Low-Power and Lossy Networks (RPL) itself has vulnerabilities that have been discussed by the research and standardisation communities: It is susceptible to selective forwarding, wormhole and sinkhole attacks [8]–[11], which can result in loss of data integrity, availability and confidentiality. The RPL specification defines some security counter-measures aiming to achieve confidentiality, integrity and replay protection [4]. Possible solutions have also been contributed by the academic community, such as VeRA [8]. However, these mechanisms are not widely adopted yet. For instance, the RPL implementation in the Contiki open source operating system for the IoT does not support any of them. Among the possible causes are problems with RPL’s complexity and implementability, some of which have been previously documented [12]. Additionally, some of RPL’s security services rely on the Advanced Encryption Standard (AES). Due to processing constraints of nodes forming 6LoWPANs, there is no well-established method for key negotiation and agreement. Recent research has demonstrated the feasibility of Elliptic Curve-based approaches [13].

In case of an incident, post-hoc analysis can provide information about the nature of the attack as well as the mechanisms used to implement it. The Contiki embedded Operating System for the Internet of Things (IoT) provides mature support for several of the aforementioned standards and specifications and it is this a very suitable platform for research in this area.

In this context, this paper’s contributions are the following:

- We implement and demonstrate the feasibility of an attack against RPL. In doing so, we bring out the necessity for forensic readiness of IPv6-connected WSNs.
- Using this attack as a use-case, we present an automated method for the post-hoc incident detection and analysis through an examination of 6LoWPAN traffic captures.

II. RELATED WORK

Previous work on forensic analysis of WSN traffic has predominantly relied on powerful observer nodes forming part of a WSN deployment. For instance, a network of investigator nodes has been proposed as a solution for digital investigations of wormhole attacks [14]. Those observers are responsible for capturing sensor node behaviour and of forwarding this information to the network’s base station. The same work proposes a set of algorithms to analyse evidence, in order to identify collaborating malicious nodes and to reconstruct wormhole attack scenarios.

In a similar fashion, it has been demonstrated that a digital forensic readiness layer can be added over a pre-deployed IEEE 802.15.4 WSN [15], by the addition of powerful forensic nodes. They capture all WSN data plane traffic and maintain frame authenticity and integrity. This work mainly focuses on
the reduction of time and cost involved in performing a digital investigation and demonstrates the ability to collect evidence without any modification to an existing network.

Powerful observer nodes with the ability to analyse traffic and detect attacks have been adopted by the work documented in [16]. Observers can detect various patterns, such as wormhole, black-hole, sinkhole and sybil attacks. Only illegitimate behaviour is forwarded to the network’s base station, thus reducing communication overheads.

In a different approach without observers, a remote live forensic protection framework has been proposed aiming to prevent the execution of illegitimate software on sensor devices [17]. Nodes are capable of notifying their peers about an intrusion as soon as they get tampered with. The framework uses sand-boxing to restrict a running application’s memory access to within a legitimate memory space. It also proposes techniques aiming to prevent the execution of malicious code by validating software authenticity.

Foren6 is a recent research effort aiming to provide diagnostic and debugging capability for 6LoWPANs [18]. It is a passive monitoring tool capable of collecting information from multiple, potentially mobile sniffers. Foren6 stores a history of network state and topology changes, called versions, and provides the ability to navigate through the entire history in a post-hoc fashion through a network visualiser. Its current version primarily focuses on network debugging and diagnostics but lacks automated incident detection.

SVELTE [11] is an IDS for 6LoWPAN / RPL networks that adopts a hybrid centralised - distributed approach. The centralised component is executed on the network’s Border Router and uses three modules. The distributed component runs on all network nodes and also uses three modules which complement the centralised component.

It has been demonstrated that compressed sensing techniques can be employed in order to overhear encrypted wireless transmissions, detect the traffic’s periodic components and ultimately reveal the type of application deployed in the network [19]. In this work, the authors are discussing the attacker’s side, but similar principles could be adopted to conduct traffic analysis for forensic purposes.

The IETF is currently undertaking standardisation efforts in the field of Intrusion Detection and Defence in RPL networks. A recently published Internet Draft presents a classification of Intrusion Detection System (IDS) architectures which could be applicable to RPL deployments [20]. The document subsequently discusses data source location, collection frequency and intrusion response in a context of RPL networks. This work in progress is still at a very early stage.

III. RPL VERSION AND RANK EXPLOITATION

RPL is a Distance-Vector routing protocol and perceives the network as a tree-like structure called a Destination Oriented Directed Acyclic Graph (DODAG). Data traffic normally flows from members of the network towards the DODAG root (upwards). Data can also travel downwards, on a path from nodes closer to the root towards nodes further away. According to the RFC, support for this type of data flow is optional.

RPL control plane packets are transported inside ICMPv6 datagrams. DODAG Information Solicitation (DIS) messages are probes, DODAG Information Object (DIO) messages advertise the presence of a DODAG and are used to form ‘upwards’ paths. Lastly Destination Advertisement Object (DAO) messages are used to construct downward paths. The sender of a DAO message can optionally request an acknowledgement (DAO-ACK).

DODAG formation and maintenance is based on a series of criteria, such as the DODAG version, a node’s rank and link metrics. Nodes with a lower rank are perceived to be closer to the root than nodes with higher ranks. Therefore the minimum rank across the entire DODAG corresponds to its root. Link metrics are determined by Objective Functions (OFs), which are not defined as part of the RPL specification. OFs are also used to calculate node ranks.

The basic principle of a rank exploitation attack is illustrated in Fig. 1. The attacker transmits malicious RPL DIO messages advertising a low rank. These DIOs lead the node under attack to believe that the malicious node can provide a better path towards the DODAG root and selects it as its parent. As a result, the compromised node and all its children in the tree lose connectivity to the rest of the network. If adequate protective mechanisms are not in place, such as encryption of application layer traffic, the same attack can also compromise data confidentiality and integrity.
IV. USE-CASE AND TRAFFIC CAPTURE IMPLEMENTATION

To conduct our investigations, we simulate a 6LoWPAN with the Cooja simulator, which is distributed as part of the Contiki OS\(^1\). Cooja simulates IEEE 802.15.4 networks and includes a plugin which can export all network traffic as a standard packet capture file (PCAP), suitable for processing and analysis with Wireshark\(^2\) and similar tools. However, it is not normal to assume that capturing all network traffic would be feasible in a real deployment. Therefore, inside the simulated network we position network sniffers which capture traffic within their vicinity. We subsequently merge the captured PCAP files and use them to conduct our analysis. To confirm the validity of the merged PCAP file, we compared it with the one exported by Cooja and verified that the merged capture is a subset of the entire network traffic. In other words, we verify that the merge does not introduce any erroneous traffic, all frames in the merged PCAP file are also encountered in the one exported by Cooja.

A sample network setup is presented in Fig. 2. Node 1 is the DODAG root, node 2 is the attacker, nodes 4 and 24-27 are passive sniffers and the remainder are normal RPL nodes.

To implement the attack with Contiki, we use as a starting point the standard code used for DODAG root nodes or RPL router nodes and make the following changes:

1) We suppress DIS messages. This prevents the malicious node from probing for existing networks. This is not strictly required.

2) We listen for incoming DIO messages as normal. In doing so, the malicious node keeps a record of the network’s current DODAG version [4], as advertised by the root and other nodes.

The attacker subsequently has two options: It can either join the network as a normal node or it can impersonate a Border Router. In the former case, the attacker will behave like a normal node, but it will pretend to offer a very good path to the DODAG root by advertising a very low rank. In the latter case, it ignores DIO messages and therefore does not join an existing network.

The attacker then assumes normal operation, advertising a network of identical parameters as the ones used by the existing deployment in exactly the same fashion as a legitimate node would do.

The attack is illustrated in Fig. 3. Node 2 attacks nodes 8, 20 and 22 causes a total of 6 nodes to get partitioned out of the network, with the compromised network topology displayed in Fig. 3b.

The success of the rank exploitation attack depends on ranks advertised by nodes as well as on path metrics between nodes. Consider nodes 2, 11 and 20 in Fig. 3. In order for the attack from node 2 against node 20 to work, node 2’s rank plus the path cost between nodes 20 and 2 must be lower than node 11’s rank plus the cost of the path between nodes 20 and 11. If this is the case, node 20 is led to believe that node 2 offers a better path to the DODAG root and selects it as its new parent. All traffic passing through node 20 is forwarded to node 2 and this constitutes an availability breach. Furthermore, if application layer traffic is not encrypted, there is also the possibility of confidentiality and integrity breaches, depending on node 2’s behaviour.

V. TRAFFIC ANALYSIS AND INCIDENT DETECTION

Post-incident detection and analysis of an attack play an important role in determining the root cause of a problem and the various events and entities involved in the attack. Determining the problem is useful for improving network security and identifying existing network vulnerabilities. After the attack, we are provided with a packet capture which includes network traffic before, during and after the incident. Information available to the investigator is that an incident has taken place, the identifiers of compromised nodes (in our case IPv6 addresses) and the nature of the compromise (in our case that the nodes in question suddenly became unreachable over the network). The investigator aims to determine the exact steps that led to the incident.

\(^1\)http://www.contiki-os.org

\(^2\)http://www.wireshark.org/
For instance, assume that node X was reachable until time $T_0$ and then at some point after time $T_1$ it became unreachable. We look for suspicious activity between times $T_0$ and $T_1$ which could have influenced X’s connectivity. For example, we observe that node X was sending DAO messages to node Y until time $T_0$. At some point between $T_0$ and $T_1$ node X stopped sending DAOs to Y and started sending DAOs to the malicious node (M) instead. M was not visible in the capture before $T_0$ but it appeared between $T_0$ and $T_1$. This could be part of normal network activity, whereby M would be a node that recently joined the network, or it could be a sign of malicious activity.

To detect and confirm the presence of a malicious node, we need behavioural analysis incorporating a temporal element. Below, we provide the algorithm for detecting network anomalies:

1) We scan the packet capture for DAO messages. DAO messages are sent by the node to its parent to inform about their own presence and to advertise a ‘downward’ path towards their DODAG children. Since the capture provides information about the network state before the incident, we are aware of node parent-child relationships. Any changes in the parent-child or network topology could be detected by inspecting the source and destination of DAO messages.

2) Changes in the network topology are normally triggered by DIO messages, which are sent periodically by nodes and which typically have a link-local broadcast IPv6 destination and therefore get transmitted as broadcast frames at the link layer. Nodes listen for DIOs and use the information therein to select a parent that minimizes the cost of the path towards the DODAG root. Following the event mentioned in the previous step, we look for a DIO message which may have been sent by a malicious node. However, there is a possibility that the network topology change is genuine and their is no malicious intent. The presence of malicious activity is further confirmed by the following two steps.

3) Since the lost node has chosen a new parent, it may have decreased its rank. Node ranks are advertised inside RPL structures. Once loading a PCAP file is complete, we go through all captured datagrams and we use the Foren6 internal data representations to extract information contained inside DIS, DIO and DAO messages, to detect events of interest and to generate our reports. The same data structures also help us record the paths of application layer traffic. Packets analysed by Foren6 are further filtered in order to make sure we have no duplicates.

4) As discussed previously, the malicious node may itself be acting either as a normal node or it may be impersonating a border router. In the latter case, it will not be sending out DIS and DAO messages. This information can also be retrieved by examining the packet capture.

5) Among all parent change occurrences, we record parent IPv6 addresses and the number of parent switches related to this node. Based on the assumption that a malicious node is not present in the network right from the start, we search for DIOs sent by these IPv6 addresses, which advertise a very low rank and which were not present during the early time windows included in the capture.

6) Nodes present in the network will normally transmit or forward application layer traffic in either direction. If application layer traffic was present in the network between times $T_0$ and $T_1$, the packet capture can reveal the location at which this traffic got dropped.

7) If multiple nodes become unresponsive at approximately the same point in time, we can co-relate the events that led to the incident. For example, if multiple nodes became unresponsive after switching to the same new parent, this increases the likelihood that this parent was acting maliciously.

VI. IMPLEMENTATION AND EVALUATION

We implemented the traffic capturing with a tool developed in-house and the merging using mergecap, which is distributed as part of Wireshark. Subsequently, to extract events of interest from the merged capture and to detect the presence of malicious nodes, we extended the open source Foren6 toolkit [18]. Foren6 parses PCAP files and stores information about captured datagrams in very detailed data structures. Once loading a PCAP file is complete, we go through all captured datagrams and we use the Foren6 internal data representations to extract information contained inside DIS, DIO and DAO messages, to detect events of interest and to generate our reports. The same data structures also help us record the paths of application layer traffic. Packets analysed by Foren6 are further filtered in order to make sure we have no duplicates.

When we encounter any message that signifies a parent change, we extract its source and destination IPv6 addresses and verify whether it was a consequence of a prior message, based on the algorithm discussed in the previous section. A snippet of the tool’s output is presented in Fig. 4. Observe that the last row corresponds to a node which i) first started sending DIO messages long after the capture’s start time and ii) advertises a very low rank. The second column lists the total number of devices which selected the node as their parent. A second snippet showing rank decreases and parent changes for an individual node is displayed in Fig. 5.

By conducting traffic analysis, we can identify events of potential interest and present them to the forensic investigator. The tool outputs the following information:

3http://cetic.github.io/foren6/
1) All nodes, their parents and a timestamp of every parent change, ordered by timestamp.
2) Rank decrease details of all nodes involved in parent changes.
3) A list of all nodes in the network and whether they were the source of DIS, DIO or DAO messages.
4) A list of all nodes acting as the new parent during a parent switch event, alongside a count of nodes selecting them as a parent, the timestamp of their first outgoing DIO and the rank advertised therein.
5) The ranks advertised by nodes acting as parents.

By presenting this information in a human-readable form and by flagging events of interest, the investigator can apply his/her domain expertise to less mundane tasks in a more time-efficient manner.

A. Evaluation

To evaluate the effectiveness of our work, we ran two experiments in networks consisting 1 legitimate Border Router, 20 normal nodes, a single attacker and a host of sniffers. In the first experiment the attacker impersonated an alternative Border Router whereas in the second it acted as normal node. For this pilot evaluation, we positioned sniffers in such a way as to capture all network traffic. Testing our mechanism with partial traffic captures is part of our future work.

For the first experiment, the malicious host was not sending out DAO messages. With the output from our tool, identifying sighting the malicious node is straightforward based on information numbered 1) and 4) in the previous sub-section.

In the second experiment, whereby the malicious node acts as a normal router, detection is more complicated. In this scenario, it behaves in every respect as a legitimate node, sending out DAO messages and generally operating in a stealthy fashion. Detection is achieved based on sudden connectivity loss events for other nodes, abrupt rank decreases, DIOs advertising very low ranks, and lost application layer traffic.

Fig. 6 illustrates the speed of the analysis using our tool for PCAP captures of increasing sizes. The numbers were obtained on a standard desktop PC. The X axis corresponds to the number of datagrams in the PCAP, while the Y axis displays the total duration of the analysis in seconds. The values shown include the time needed for Foren6 to parse the PCAP as well as the time it took for our extensions to analyse the traffic, detect events of interest and generate the report.

B. Countermeasures

The version and rank exploitation attack can be addressed through the following countermeasures:

1) Enabling DAO-ACK messages. DAO-ACK messages are sent by a DAO recipient in response to a unicast DIO [4]. Enabling DAO-ACK messages will assist in identifying malicious activity and can also be used as part of an intrusion detection system.
2) White lists of legitimate IPv6 addresses, maintained by every network node in order to avoid communication with malicious nodes. This approach has several drawbacks, primarily related to the potential high churn rate of those networks which would make the integrity of such white lists very difficult to maintain. Additionally, this approach would not scale well with networks containing large numbers of nodes [10].
3) Enabling RPL’s authenticated mode, whereby nodes need to authenticate themselves before joining a network as (non-routing) hosts. In order to join as a router, a node has to obtain a second key from a key authority. This mode of secure operation is briefly discussed in the RPL specification [4]. However, according to the same document, authenticated mode cannot be implemented with symmetric keys and “RPL supports only symmetric algorithms: authenticated mode is included for the benefit of potential future cryptographic primitives”.

VII. CONCLUSION AND FUTURE WORK

In this work we have presented a traffic analysis tool which can identify attacks against RPL in 6LoWPANs, flag events of interest and present the analysis results to an investigator in a human readable format. As part of our future work we aim to improve the algorithm’s accuracy. Additionally, we plan to run further experiments in order to investigate the accuracy of our approach when the traffic captures provide reduced network.
coverage. Part of this task will be an attempt to reconstruct the network topology based on incomplete information.

This work is complementary to our RAM extraction and carving tool for devices used in those networks [21]. By combining an analysis of network activity with an analysis of the RAM contents of network nodes, we can reveal useful information about the events that led to a security incident. As part of our future work we plan to integrate those two works into a single toolkit.

REFERENCES