Detecting Incidents in Wireless Mesh Networks using Flow and Routing Information

Lothar Braun[1], Alexander Klein[1], Leon Aaron Kaplan[2], Georg Carle[1]

[1] Lehrstuhl für Netzarchitekturen und Netzdienste, Technische Universität München
[2] CERT.at - Computer Emergency Response Team Austria

Abstract. The topology in wireless multi-hop networks can change frequently due to characteristics of the shared medium, mobility of the users, or miss-behaving and malicious nodes. Frequent topology changes typically lead to inconsistencies in the network topology since routing protocols can only cope with certain topology change rate. These incidents may even lead to a temporary collapse of the network if countermeasures are not applied in time. Therefore, an early detection of incidents is necessary to achieve a high availability of these networks. In this work, we propose a monitoring infrastructure for wireless ad-hoc networks which is able to detect incidents by evaluating flow and routing information.

1 Introduction

Wireless community ad-hoc and mesh networks that provide network connectivity to their users can be found throughout many cities in Germany [1], Austria [2], Greece [3], Italy [4] and Spain [5]. Many of these networks deploy the OLSR [6], BGP [7], or the B.A.T.M.A.N [8] routing protocols to build a reliable and stable topology. The applied routing protocols are designed to operate in the context of frequent link breaks which are usually the consequence of interference, high noise and packet loss and mobility of the users.

While these networks are large, ranging from several hundreds to several thousand devices, their user group is commonly very decentralized: It can be regarded as a distributed community of individuals, which also provides and administrates the deployed networking devices. These user devices play an active role in the network since their communication behavior, e.g. forwarding and dissemination of routing information, affects the topology in the network. A single miss-behaving device, miss-behaving either due to technical problems or intended malicious behavior, can have a large impact on the topology and therefore service availability and service quality. A malicious user could try to modify the topology of the network in several ways in order to route network traffic through himself or to partition the network by injecting and forwarding manipulated routing messages.

Detecting and mitigation such incidents as quickly as possible can help to increase the network stability. A carefully designed monitoring framework can

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4

help to improve the situation awareness and can help to gain a better understanding of the behavior of the network which is necessary to identify, classify and pinpoint the source of network anomalies.

In this paper, we present our ideas on how to build such a monitoring infrastructure for routing and flow data that can be used for early network failure detection and attack identification. The introduced infrastructure is dedicated for wireless multi-hop networks. It consists of two basic components: The first component is designed for mapping the global topology of ad-hoc routing tree by collecting and analysing routing information from multiple observation points. The second component aims at monitoring and analysing the traffic that can be observed on multiple nodes. Our goal is to combine traffic information and routing topology information in order to find network anomalies.

Section 2 discusses our monitoring goals and outlines expected challenges. Our architecture for a distributed monitoring architecture and our proposed analysis methods are discussed in Section 3.

2 Goals and Challenges

Our work aims at providing a complete view of the wireless network that employs routing and flow measurement information to model the current state of the network. This global network view is built by aggregating data from different data sources such as OLSR routing tables and flow measurements gathered from various observation points. The main challenge regarding the determination of the topology is to filter out inconsistent information. Due to the long topology control interval of OLSR and the lossy nature of wireless links, it is likely that the topology information at an observation point is based on outdated information since the topology usually changes frequently. Moreover, incorrect selection of multipoint relays can prevent nodes from receiving the latest routing information which has to be considered during the calculation of the actual topology.

The limited hardware resources of wireless devices in terms of memory and computational power have to be taken into account when deploying flow monitoring tools. In general, the user community is using heterogeneous hardware with different hardware resources. Thus, the monitoring infrastructure has to be designed such that it can be deployed on the hardware with the highest hardware constraints. As a result, the functionality of the flow monitoring tool should be reduced to the absolute necessary. However, the collected routing and flow information has to be forwarded to a central node with more powerful hardware for evaluation which increases the traffic load in the network. Therefore, the current network utilization has to be considered to avoid congestion.

Anomaly detection represents challenging task in wireless networks since the topology of these networks changes frequently. For this reason, training data based on long term monitoring is needed to infer the normal network state. Furthermore, metrics for anomaly detection have to be defined to detect operational relevant changes. In addition, these changes have to be classified to distinguish necessary changes due to variation in the link quality and malicious changes.

2.1 Challenges concerning OLSR

OLSR is one of the most popular routing protocols for wireless mesh and ad-hoc networks due to its efficiency in terms of routing overhead and its capability to react quickly to topology changes [9]. This capability is achieved by periodic transmission of hello messages for detecting changes in the two hop neighborhood and topology control messages for changes further away from the node. Both mechanisms work very well until the topology changes in the network become too frequent.

If the topology changes become too frequent, nodes start to disseminate already outdated information which results in inconsistencies in the routing table of other nodes [10]. Thus, the collection of topology information has to be done with respect to possible inconsistencies. Each node in the network has a limited point of view even if a link-state routing protocol like OLSR is applied since it is heavily dependent from the information that it receives from other nodes. Therefore, the question arises whether the global routing information can be derived from a single router. A detailed evaluation of received routing messages can help to minimize inconsistencies in the global routing information if previously received messages are taken into account. Such mechanisms enable the detection of malicious routing messages, e.g. by detecting that a node has unlikely changes in its neighborhood.

The dissemination of routing information of the OLSR protocol is strongly affected by the calculated MPR set. Unreliable links and frequent topology changes lead to errors in the MPR set calculation which limits the dissemination of routing information [10]. For this reason, a distributed approach is recommended to gather as much routing information as possible to detect such inconsistencies and to initiate counteractive measures.

2.2 Challenges concerning Flow measurement

Wireless mesh networks are usually built from constrained limited hardware devices. Deployed routing devices have to perform OLSR updates, and have to forward user traffic . Adding one more component for flow monitoring may not infer with these previous tasks, as flow monitoring is a non-important task for the devices. Hence, generating flow data needs to be as simple as possible in order to ensure low system load.

All flow information needs to be exported from the wireless devices towards a central collector that can analyse the data. This export needs to either employ a side channel, e.g. a wired network connection, or needs to be performed over the wireless link. Employing the wireless link adds more traffic to the link, thus reducing the available bandwidth for user traffic. One challenge is therefore to reduce the amount of exported flow traffic.

Furthermore, flow monitoring always raises questions concerning privacy issues. These privacy issues get even more severe, as soon as flow data is transmitted over the wireless link that are accessed by all users. In order to cope with these problems, anonymisation techniques and end-to-end encryption needs to be employed in the flow monitoring and analysis architecture.

3 Approach

Our approach aims at providing a global view on the current network state. By observing and analysing the collected data over time, we hope to find emerging network problems such as network attacks. A key component of such flow and routing information analysis is a measurement infrastructure which allows to collect measurement data from several observation points. Any distributed measurement approach must consider synchronising the individual nodes' clocks. We propose NTP as a reasonable well established protocol giving us good enough granularity. Several components within the network need to collect, process and export network state information. These components will be described in the following subsections.

3.1 OLSR Monitoring

In order to detect problems in the self-organizing routing tree, routing information from different observation points need to be collected. As routing problems can result in a break-down of the communication abilities over the wireless interfaces, special considerations concerning the availability of routing information in the event of routing problems need to be taken into account: Export of OLSR routing information should even be possible in the event of routing or other network problems. Devices that are able to communicate over an independent side-channel, e.g. a wired internet connection, are good candidates for transmitting their OLSR network view.

The views of several measurement points can then be collected at a central analysing component. This component can correlate the different views and is able to check for inconsistencies that can be related to network troubles.

3.2 Flow Monitoring

The second data source we would like to consider is flow data, which describes the user data that is transmitted over the mesh network. Using this data source, we can observe which traffic patterns are routed over which device. By correlating the OLSR routing view with the monitored flow information, we can potentially detect problems like inconsistencies, overloaded links and devices or other kinds of anomalies.

The flow monitoring process needs to be deployed on the mesh devices, which is typically built from low-cost hardware with little processing power and memory. Packet capturing is one of the processes that is known to be a performance bottleneck, when the observed traffic exceeds the available processing resources. As our flow measurement setup is built on Linux, we can use known capturing optimization such as PF_RING [11] to reduce the capturing costs.

Flows can then be generated using a standard flow generation architecture, and exported using an flow export protocol such as IPFIX [12]. This general architecture needs to be adapted and extended to cope with the requirements in the constrained mesh networking environment:

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4

Sampling: Sampling needs to be performed in order to reduce the number of packets that need to be processed during flow generation, and can also help to reduce the number of generated flows. Our setup eliminates the choice of several of these algorithm: As we want to correlate flow data from different observation points, we may not employ randomized sampling. Hence, we can only use techniques that sample the same subset of traffic, e.g. hash based sampling algorithms that select packets based on packet properties. Several flow and connection-based sampling algorithms have shown to be yield good sampling results while preserving much security related information [13].

Anonymisation and Data Security: Privacy considerations are always important if sensible information about communication is collected and stored. The deployed anonymisation techniques need to be the same over all observation points. Well-known techniques for prefix preserving IP address anonymisation can be used to gain a certain degree of anonymisation. As flow data might be transmitted over the wireless link, any participant is potentially able to read communication information. In order to circumvent this eavesdropping, end-to-end encryption between the flow exporting process and flow collecting process need to be deployed.

Flow Data Compression: If flow information is transmitted over the wireless links, flow export will consume a potential high amount of bandwidth. Keeping this bandwidth as small as possible in order to minimize this impact is therefore desired. Previous evaluations have shown that compression of flow information can result in very high compression ratios [14]. However, compression algorithms can put a significant load on a constrained device. Choosing the right compression algorithm which yields a good compression to computing load ratio is therefore an important task.

3.3 Data Storage and Analysis

Data from both sources can be collected in a central database, which is responsible for calculating the global view. In a first step, OLSRs' Dijkstra algorithm for any operation can be calculated for each reporting node. These different views can be combined into a single global view, as long as no routing inconsistencies are detected, e.g. conflicting routes. Long-term topology maps can be used to build a model on routes and routing changes which can occur during normal operations. This model can be used to detect extreme changes on the topology that can point administrative relevant incidents.

Different metrics from the graph theory domain, such as the average graph diameter or node degree, can be used to describe a model of the normal network state. Graph analysis can be used to detect cycles or other types of routing problems. The dominator tree algorithm [15] can be used to find single points of failures in a path. Malicious or broken routing views can be detected by comparing the different views of multiple routing measurement points.

Anomaly detection can also be used to analyse traffic flow information: Traffic flows on the wireless mesh network should move along the routing topology that has been created by OLSR. Using multi-variate anomaly detection methods,

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4

which are able to correlate data from multiple observation points, a model that includes both routing metrics as well as flow metrics can be created. Changes in the correlation structure can be detected, e.g. if the measured traffic flows are no longer observed at the nodes that should observe these flow according to the global routing view.

4 Conclusion and Outlook

This paper presented our idea on a distributed monitoring architecture for wireless ad-hoc and mesh networks. Our architecture aims at correlating routing and traffic measurement data into a global network state view. Building on this network view, we aim at detecting technical problems and security incidents that have negative impact on the network service quality. After building the necessary network measurement tools, we hope to deploy our architecture in a real-world environment in the Funkfeuer project [2].

References

1. Freifunk: http://www.freifunk.net/ (April 2011)
2. Funkfeuer: http://www.funkfeuer.at/ (April 2011)
3. Athens Wireless Metropolitan Network: http://www.awmn.net/ (April 2011)
4. Ninux: http://www.ninux.org/ (April 2011)
5. GUIFI.net: http://www.guifi.net/ (April 2011)
6. Clausen, T., Jacquet, P.: Optimized Link State Routing Protocol (OLSR). RFC 3626 (Experimental) (October 2003)
7. Rekhter, Y., Li, T., Hares, S.: A Border Gateway Protocol 4 (BGP-4). RFC 4271 (Draft Standard) (January 2006)
8. Neumann, A., Aichele, C., Lindner, M., Wunderlich, S.: Better approach to mobile ad-hoc networking (b.a.t.m.a.n.). http://tools.ietf.org/html/draft-wunderlich-openmesh-manet-routing-00 (April 2008)
9. Viennot, L., Jacquet, P., Clausen, T.H.: Analyzing control traffic overhead versus mobility and data traffic activity in mobile ad-hoc network protocols. ACM Wireless Networks Journal (Winet) **10**(4) (2004) 447–455
10. Klein, A.: Performance comparison and evaluation of AODV, OLSR, and SBR in mobile ad-hoc networks. In: Proc. 3rd International Symposium on Wireless Pervasive Computing ISWPC 2008. (May 2008) 571–575
11. Deri, L.: Improving passive packet capture: Beyond device polling. In: Proceedings of the Fourth International System Administration and Network Engineering Conference (SANE 2004). (September 2004)
12. Claise, B.: Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information. RFC 5101 (January 2008)
13. Braun, L., Münz, G., Carle, G.: Packet sampling for worm and botnet detection in tcp connections. In: Proceedings of IEEE/IFIP Network Operations and Management Symposium (NOMS 2010). (2010)
14. G. Münz and L. Braun: Lossless compression for ip flow information export (ipfix). https://tools.ietf.org/html/draft-muenz-ipfix-compression-00 (July 2008)
15. Lengauer, T., Endre, R., Jan, T.: A fast algorithm for finding dominators in a flowgraph. ACM Transactions on Programming Languages and Systems **1** (1979) 121–141

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4