Anomaly detection and mitigation for disaster area networks *

Jordi Cucurull, Mikael Asplund, Simin Nadjm-Tehrani

Department of Computer and Information Science, Linköping University
SE-581 83 Linköping, Sweden
[jordi.cucurull,mikael.asplund,simin.nadjm-tehrani]@liu.se

Abstract. One of the most challenging applications of wireless networking are in disaster area networks where lack of infrastructure, limited energy resources, need for common operational picture and thereby reliable dissemination are prevalent. In this paper we address anomaly detection in intermittently connected mobile ad hoc networks in which there is little or no knowledge about the actors on the scene, and opportunistic contacts together with a store-and-forward mechanism are used to overcome temporary partitions. The approach uses a statistical method for detecting anomalies when running a manycast protocol for dissemination of important messages to k receivers. Simulation of the random walk gossip (RWG) protocol combined with detection and mitigation mechanisms is used to illustrate that resilience can be built into a network in a fully distributed and attack-agnostic manner, at a modest cost in terms of drop in delivery ratio and additional transmissions. The approach is evaluated with attacks by adversaries that behave in a similar manner to fair nodes when invoking protocol actions.

1 Introduction

Disaster area networks are created through spontaneous mobilisation of ad hoc communication when the existing infrastructure is wiped out or severely overloaded. In such environments, in addition to local establishments of cellular networks and satellite connections there is a potential for hastily formed networks (HFN) [1] of wireless devices connecting via 802.11 or similar technologies. One of the major needs in a disaster area is timely dissemination of information destined for a large group of actors. However, due to the nature of the multi-party engagements, and massive engagement of volunteers there is little room for establishment of mutual trust infrastructures. Any dissemination protocols destined for such environments require to function in a chaotic context where the node identifiers or even the number of involved nodes cannot be assumed.

The physical aspects of above networks can be characterised by intermittent connectivity, leading to networks in which existence of partitions is a norm. This creates a variation of mobile ad hoc networks (MANET) with no contemporaneous routes among the nodes, also referred to as intermittently connected

* The original publication is available at www.springerlink.com

MANET (IC-MANET). Experience from the Katrina HFN [2] shows that even in disaster environments there are security threats – actors who try to disrupt operations or use the information for own benefit. However, the application of security measures in these environments is far from trivial. This paper addresses security issues that impact dissemination of messages, and thereby focuses on the availability of the dissemination services in IC-MANET.

We study the impact of intrusions on a dissemination protocol, Random Walk Gossip (RWG) that is designed to run in IC-MANET in a disaster area network [3]. This algorithm is intended to disseminate important messages to any k receivers, thereby a manycast algorithm that does not rely on knowledge about the network nodes. Its main characteristics are that it overcomes lack of information with opportunistic contacts and uses a store-and-forward mechanism to overcome network partitions. Moreover, to act in an energy-efficient manner it will try to avoid excessive transmissions by decentralised control of number of custodians for a given message.

In such a scenario the adversary has no choice other than behaving in a way that resembles the rest of the nodes in the network, and we further assume that the adversary too needs to be efficient in its use of energy and bandwidth resources. In fact the adversary may not act normally, but seen from a restricted view, being the local knowledge at each mobile node, it follows fragments of the protocol specification.

Intrusion detection mechanisms are intended to identify malicious activity targeted at the resources of a monitored system, broadly classified as misuse or anomaly detection. The former requires the formulation of misuse constraints, which are extremely complex when the adversary behaves within the boundaries of the protocol specifications and when at the same time it must be suitable for different environments, i.e. dynamic load, partition sizes, varying local densities and connectivity changes. In the IC-MANET context anomaly detection is a suitable approach to intrusion detection while misuse detection is less appropriate. First, the fact that the adversary behaves in a similar way to the fair nodes makes formulation of misuse constraints hard if not impossible. Second, even if we can formulate a set of rules for undesirable packet sequences, these will hardly work in all nodes due to dynamic load and partition changes.

Our approach builds on a learning based anomaly detection technique that uses only normal data in the learning phase. While this might be a limitation of the approach, since there is no guarantee that attacks are not present in the early deployment phase in the scenario, we believe that the efficiency of the technique will outweigh its disadvantages. Another major problem in highly unpredictable and partitionable networks is what to do when an attack is suspected. If the network is generally chaotic and the goal is to maintain dissemination levels then it is less relevant to exactly identify adversary nodes and try to isolate or ignore them. We therefore suggest mitigation approaches that enable each node to adjust its own behaviour thereby reducing the effect of the suspected attack.

The threat model that we consider is that the adversary tries (1) to drain the network resources, both at node level (battery life) and network level (available

bandwidth), thereby reducing the dissemination flows, and (2) acts as an absorbing patch which reduces some message dissemination in its vicinity, acting as a grey hole at certain times and locations. Clearly this threat model creates a challenging type of adversary to detect.

Our detection and mitigation approach has been evaluated in a simulation setting where an implementation of the RWG algorithm is running with a disaster mobility model adapted from earlier work [4]. The evaluations indicate that the approach indeed creates a resistance to attacks that match the above threat model, and show that the network recovers from the attack when it is of a transient nature. Moreover, our approach dampens the effect of the attacks on the network resources by preserving the overall overhead in the network compared to the non-mitigated case, whilst not losing the delivery goals significantly. These results are obtained despite the fact that the classical metrics used for evaluation of intrusion detection do not show good results. The paper discusses why the network performance metrics are more useful in IC-MANET clarifying the impact of partitions, traffic load and the store-and-forward approach.

The contributions of the paper are as follows:

- Presentation of a scalable approach to anomaly detection and mitigation in partitionable ad-hoc networks with scarce resources that run a given dissemination protocol suitable for these environments. The detection algorithm is scalable since it is fully distributed and efficient. It is a statistical mechanism reminiscent of the chi-square technique [5]. It has been adapted to the specific RWG protocol by selection of appropriate (general) features.
- Illustration of the approach using a simulation platform, and specifically analysing why the performance based metrics outperform the classic detection rate and false positive rate metrics in such disaster area networks.

2 Related work

Yang *et al.* [6] describe the major security challenges in MANET and some of the existing solutions. Among the identified challenges are the lack of a well-defined place to deploy the security solutions, the vulnerability of the information contained within the devices, the fact of communicating within a shared medium, bandwidth, computation and energy resource constraints, the dynamic network topology, and the wireless channel characteristics (e.g. interference and other effects leading to unreliability). It is also stated that a complete security solution should integrate prevention, detection and reaction components. Prevention typically evolves around establishment of trust, often based on cryptographic techniques. However, trust is not easy to achieve in such scenarios [1] and cryptographic techniques, as studied in Prasithsangaree and Krishnamurthy [7], usually are computationally too expensive. Farrell and Cahill [8], in the context of delay-tolerant networks, also mention the lack of cryptographic key management schemes as an open issue.

Orthogonal to the preventive perspective we need to consider the role of intrusion detection in IC-MANET. Several approaches to intrusion detection

have already been proposed for the MANET Ad hoc On Demand Distance Vector (AODV) and Optimised Link State Routing (OLSR) protocols. However, to our knowledge no earlier works address multicast protocols, and specifically not those suitable to run in a partitioned MANET. Some authors propose that specifying, distributing and updating the signatures of the attacks is not feasible [9] in these environments. Instead, anomaly detection is easier to apply since the normality model can be created and updated in each node from its own observations. Hence, abnormal behaviours in the specific context of a particular node, even if they are within the boundaries of the protocol specifications, can be detected. Garcia-Teodoro *et al.* [10] present an extensive review of the most popular techniques used in anomaly detection, which can be roughly classified into statistical based, knowledge based, and machine learning based. The most significant challenges of anomaly detection are also mentioned, namely low detection efficiency and high false positive rate, low throughput and high cost, absence of appropriate metrics and assessment methodologies for evaluating and comparing different techniques, protection of the intrusion detection mechanism from attacks, and analysis of ciphered data. In our work we confirm that the classic metrics used in wired or fully connected wireless networks are not appropriate in IC-MANET. We believe comparisons on common simulation platforms (as long as the authors make their code accessible to other researchers) is a first step for comparative evaluation.

Although anomaly detection for IC-MANET has to be geared towards protocols that in fact manage the challenges of multiple partitions – what we aim to address in this paper – we would like to name a few precursors for anomaly detection in MANET. Nakayama *et al.* [11] propose an adaptive method to detect attacks on the AODV routing protocol. The method is based on the calculation of projection distances using multidimensional statistics and the Principal Component Analysis (PCA) to determine the axis that expresses the most relevant data correlations. Cabrera *et al.* [12] propose a method to detect attacks on AODV and OLSR protocols. The method consists of three hierarchical and distributed intrusion detection modules based on pattern recognition and classifier fusion. Liu *et al.* [9] too present a method to detect attacks on the AODV routing protocol. The method is based on the combination of two data mining approaches over data obtained from the MAC and network layers. The technique allows the identification of the MAC address of the attacker, although it can be easily spoofed. Later, a Bayesian method is used to correlate local and global alerts. It also proposes a collaborative decision mechanism among nodes in the radio transmission range. These approaches are not applicable to our problem area. First due to the manycast nature of dissemination and secondly due to the intermittent connectivity in disaster area networks.

Among the few works that address intrusion detection in IC-MANET, there is work by Chuah *et al.* [13] proposing a rule-based misuse detection mechanism targeted towards delay-tolerant networks. It builds on a history-based routing protocol, detects attacker nodes, and mitigates their effects by keeping them on a black list, i.e. through isolation. Other security-related work in IC-MANET is

concerned with prevention, e.g. Scalavino *et al.* [14] who propose an authorisation scheme to protect the data exchanged in disaster events.

A main challenge of anomaly detection in MANET [9] is that most of the approaches do not succeed with localisation of the attack sources. There is no prior trust relationship among the network nodes, and network operation relies on altruism and cooperation of neighbour nodes. Also the fact that nodes fail to respond, e.g. through network congestion, link failure, or topology changes, can be confused with intrusions [15], producing high false positive rates.

The metrics used on intrusion detection in MANET are usually based on the accounting for detection rate and false positive rate typically presented as ROC curves. In most cases these metrics reflect the number of attacks detected, but sometimes they show the number of attackers detected [13]. The detection delay is not usually taken into account, but there are a few exceptions [16, 17, 13]. There are approaches that quantify the impact of the detectors, such as network overhead or the CPU speed-up [18], or data delivery ratio to evaluate the impact of attack response [13]. In this paper we also advocate the use of delivery ratio and total transmissions (as an indicator of overhead) for evaluation purposes.

Finally, response and mitigation of attacks is one of the topics that has not been considered much in the intrusion detection literature for wireless ad-hoc environments. Some MANET works [16, 19] just mention it and propose certain actions, but do not really apply it. There are a few exceptions in which mitigation is really applied [13, 20]. In addition to the network performance metrics to show the benefits of the approaches, the work by Wang *et al.* [20] also proposes metrics to decide whether it is worth to enable the mitigation or not. This is an interesting direction that should be explored within IC-MANET too, but we postpone it to future works.

3 Protocol description and threat model

3.1 Protocol description

The Random Walk Gossip (RWG) is a message dissemination protocol for intermittently connected networks that has been presented and evaluated in a previous publication [3]. Here we will try to provide just as much information as needed to understand the threat model that we have used.

The protocol is designed to cope with the challenges faced in disaster area networks such as intermittent connectivity, scarcity of bandwidth and energy, as well as unknown and unpredictable network topologies with partitions. RWG is a manycast protocol, which means that a message is intended to reach a given number k of nodes. When k nodes have been reached, the message is k-delivered and does not need to be propagated anymore, thus not wasting energy. RWG is based on a store-and-forward mechanism, i.e. each node keeps the messages to forward in a local buffer until they have been delivered. This mechanism prevents the loss of messages because of network partitions.

When a message is sent in a connected part of the network, it will perform a random walk over the nodes, until all the nodes in the partition are informed of

this message. This walk is controlled by a three-way packet exchange shown in Figure 1. First a Request to Forward (REQF), that includes the message payload, is sent by the current custodian of the message (indicated as grey in the picture). The neighbouring nodes that hear the REQF reply with an acknowledgement packet (ACK). The custodian chooses one of these nodes randomly and sends an OK to Forward (OKTF) to this node indicating that it will be the next custodian. The other nodes will retain the message without actively disseminating it. They will keep the message as *inactive* until it expires. Partitions can be overcome by the movement of nodes. Thus, new uninformed nodes will be informed by some node that keeps the message as *inactive* and restarts to disseminate. This process will continue as long as no more uninformed nodes remain in the network or the message is k-delivered.

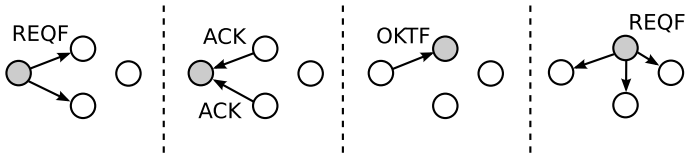

Fig. 1: Random Walk Gossip

All the packet types share the same header structure. In order to keep track of which nodes have seen a given message, each packet header contains a bit vector called the *informed* vector. When a node receives the message it produces a hash of its own address and puts a 1 in the bit vector in the field corresponding to the hash. This allows the protocol to know when a message is k-delivered, and to tell the potential future recipients of the message how far the message has reached towards its dissemination goal (summing the number of 1's indicates the current known local knowledge on this). The *informed* vector also indicates which nodes have received the message. If a node A hears a new neighbour B, then A will go through the messages stored in its buffer to see if B has not yet been informed of any of the messages, in which case those messages will be reactivated and sent to node B (and other uninformed nodes in the vicinity).

Finally, when a node realises that a message is k-delivered it sends a Be Silent (BS) packet to its vicinity. This packet will cause all receiving nodes to also realise that the message is k-delivered and thus remove it from their buffers. No new BS packets are sent upon the reception of a BS packet.

3.2 Threat model

Routing and dissemination protocols for MANET are usually based on cooperation of the network nodes and assume fair play. RWG is not an exception and an attacker can take advantage of it. There are many ways a malicious node can attempt to disrupt the dissemination activity in a network. This paper focuses

on the mitigation of low-cost attacks which are consistent with the protocol specification. We study how a disrupting node will try to impact the dissemination and resource usage of the network without investing too much of its own energy resources. Recall that the only packet type with a payload is the REQF packet. This is also the one that claims more in terms of transmission energy and bandwidth. Using forged inexpensive ACKs three aspects of the protocol operation can be attacked:

- **Discovery of new nodes**: RWG discovers new nodes in the vicinity by listening to the packets the node receives. Each time a packet is received the messages stored in the local buffer are checked to see if they have been already delivered to the sender of that packet. If that is not the case they are forwarded to the node. An attacker can exploit this mechanism by sending many ACK packets with different fake sender addresses and create a high number of message retransmissions. The fake addresses are randomly generated and there is no mechanism to prevent their usage.
- **Delivery status propagation**: The propagation of the delivery status of the messages is done through the *informed* vector included in the sent packet headers. An attacker can manipulate these vectors and take advantage of the other nodes to propagate them using ACK packets.
- **Selection of custodians for a given message**: When a message is forwarded to a group of nodes, they answer with an ACK packet. RWG uses these ACK packets to randomly choose one of the nodes as the next custodian of the message. An attacker could exploit this mechanism to be elected as the next custodian by answering with several ACKs increasing the probability of being chosen.

It is assumed that the adversaries will have a complete knowledge of the protocol and that will act according to the RWG specifications. Though our anomaly detection algorithm is oblivious to the attack patterns, we will later use two specific instances of attacks (see Section 5.2) based on exploiting some of the operations described for the purpose of evaluation.

4 Anomaly detection and mitigation

Anomaly detection is based on the construction of a model that represents the normal behaviour of a system and which is used to determine abnormal situations. Since MANET are usually operated by resource constrained devices a statistical-based approach has been selected as an anomaly detector since it has a smaller footprint than other techniques.

4.1 Detection algorithm

The anomaly detector we propose represents normality as a vector of numerical values called features. The algorithm is based on a distance function $D(x_i)$ that

calculates sums of squared differences between a given observation x_i of the system (which contains F features) and the normality model $\bar{x}$ to decide if the observation is anomalous or not (see Eq. 1). An observation is obtained and evaluated each time a packet is received. According to the central limit theorem, if the number of variables is large enough, then the calculated sums of squared differences will follow a normal distribution. Hence a threshold (T_1), based on the statistical three-sigma rule (also called 68-95-99.7 rule) [21], is introduced to determine if the distance measured is outside of the values considered normal. The work flow of the system has two differentiated parts.

$$D(x_i) = \sum_{j=1}^{F} (x_{i,j} - \bar{x_j})^2 \tag{1}$$

1. **Training**: In this part in which the normality model of the system is created, only observations of the normal behaviour of the system are used. The model consists of a vector ($\bar{x}$) with the average value of each feature, two vectors (max, min) with the maximum and minimum values expected for each feature under normal conditions, and a threshold (T_1) that states which is the maximum distance observed from the average field of normality. The normality model is created during a period of time that includes two consecutive steps that comprise N and M observations respectively.
   (a) **Calculation of average, maximum and minimum values**: During a period of time with a number of N observations, the average ($\bar{x}$), maximum (max), and minimum (min) vectors are calculated. The last two vectors are used for normalisation, i.e. to keep all the features from normal observations within the same range of values. Normalisation is also applied to $\bar{x}$.

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{2}$$

   (b) **Calculation of the threshold**: During a period of time, and for a number of M observations, the distance $D(x_i)$ between an observation x_i and the calculated average $\bar{x}$ is measured. T_1 (see Eq. 3, 4, and 5) is defined as the mean of the distances calculated (μ) plus three times their standard deviation (σ). According to the three-sigma rule the range $[\mu - 3\sigma, \mu + 3\sigma]$ should cover 99.7% of the distances of the normal observations evaluated. Note that just the upper limit is used, because evaluations with small distances are not considered anomalous.

$$\mu = \frac{1}{M} \sum_{i=1}^{M} D(x_i) \tag{3}$$

$$\sigma = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (\mu - D(x_i))^2} \tag{4}$$

$$T_1 = \mu + 3\sigma \tag{5}$$

2. **Testing**: During this step the detector is fed with observations of the system behaviour that can be anomalous. The detector decides whether an observation x_i is anomalous by calculating the distance $D(x_i)$ from $\bar{x}$, which determines how far is the current observation from the normal behaviour of the system, and compares it with T_1. If $D(x_i) > T_1$ the observation is categorised as anomalous, and if $D(x_i) \leq T_1$ it is categorised as normal.

4.2 Features

The features of an anomaly detector are the variables which are believed to characterise the behaviour of the monitored system. Our approach uses features at the routing layer and most of them are statistical.

- **Packet rates**: Number of packets of each type received during the last I_1 seconds. There are four of these features, one for each packet type.
- **Packet distances**: Distance, measured in number of packets received, between the reception of two specific types of packets. E.g., number of packets received between the reception of a REQF and the next ACK. There are sixteen of these features that cover all the possible packet type combinations.
- **Packet rate differences**: Relative difference in the packet rates calculated for each type of packet. There are six features, one for each relevant combination.
- **Number of different source addresses**: Number of different source addresses counted in the packets received during the last I_2 seconds.
- **Packet ratios**: Quotient of the number of packets received of a specific type compared to another packet type among the last I_3 packets received. There are three of these features: ACK/REQF, ACK/OKTF, ACK/BS.
- **Summation of informed vectors**: Summation of all the positions of the *informed* vectors received in the last I_4 packets.

Because the evaluation is carried out each time a packet is received, the features that provide information about the last packets received are implemented as sliding windows over the intervals I_1, I_2, I_3, and I_4.

4.3 Alert aggregation

Statistical anomaly detection requires a certain time to detect an anomaly within the system. As alerts cannot be mapped to the specific packets causing the attacks, the alarms must be raised after an interval of suspicion. This is the reason why the alerts raised by the detector are processed and aggregated during an interval I_a of aggregation.

In each of these periods the number of packets evaluated and the number of alerts registered are counted. Then, an alarm is raised if the number of alerts within that period exceeds a certain threshold (T_2). The threshold is a tunable parameter of the system which is defined in terms of proportion of alerts registered over the number of packets evaluated during I_a.

4.4 Mitigation

When an alarm is raised in a node the mitigation scheme is locally enabled. As it will be explained in Section 5.3, a careful RWG operational mode is proposed to cover the possible attacks that fall within the threat model defined. Since it is not clear whether an attack is transient, continuous or intermittent, we need to decide how long a mitigation should take place. In this paper we have simply evaluated a mitigation that takes place over a constant interval I_m ($> I_a$). This prevents the system from disabling the mitigation too early as a consequence of the beneficial effects of the mitigation instead of the finalisation of the attack.

5 Evaluation

This section evaluates the detection and mitigation approach applied to RWG in a disaster area scenario against the threat model described in Section 3.2.

5.1 Simulation setup

The performance of the approach has been evaluated using the Network Simulator 3 (ns-3) with an implementation of the detection and mitigation mechanisms embedded in the RWG protocol implementation at the network layer.

The disaster area scenario includes the mobility traces from Aschenbruck *et al.* [4], based on a large training manoeuvre in preparation of the FIFA world cup in Germany in 2006. The original traces include 150 mobile nodes. To induce partitions and create an intermittently connected network we have selected 25 of the nodes, chosen uniformly over the locations in the area, while maintaining the trace for that node. This creates a similar network with lower density. Five other nodes have been chosen as attackers while acting in the same network, again with a uniform distribution among the fair nodes. The attacker nodes do not create normal traffic (data) in the network, but produce packets that are compatible with the protocol specification as described in section 5.2. This is aligned to the threat model defined, where the attacker spend the minimum energy possible. All the nodes use the 802.11a protocol, at 6Mb/s data rate with a radio range of 24 meters. The speed of the nodes varied in the range 1-2 m/s in an area of 200m x 350m. The load is generated by introducing in the network a total of 15 messages to disseminate every second from randomly chosen nodes. Each message has set to be delivered to minimum number of 10 nodes ($k = 10$).

The simulation time for each run is 3000 seconds. The first 200 seconds are discarded (not used for anomaly detection) due to start-up time for the protocol. The following 1400 seconds are used for training the system (half of them for calculating $\bar{x}$, min and max vectors and the rest for the threshold), and the last 1400 seconds are used for evaluation. Each simulation is repeated 10 times with different sets of traces and all the results shown are averages over these 10 runs. The alert aggregation window (I_a) is chosen as 10 seconds (unless otherwise stated). The selected threshold (T_2) for the alert aggregation process is set up to

30%. The mitigation period (I_m), during which a mitigation remains enabled is set up to 200 seconds. The intervals used to calculate the features (I_1, I_2, I_3, and I_4) are set up to 5 seconds, 10 seconds, 50 packets, and 100 packets, respectively.

5.2 Generated attacks

To show the effectiveness of the detection and mitigation approach, two attacks that fall into the threat model described in Section 3.2 have been implemented.

– **Draining attack**: It makes the nodes around the attacker to transmit more packets than usual in order to drain their batteries and waste some bandwidth. The effect, that exploits the RWG node discovery mechanism, is achieved by regularly sending ACK packets with different fake identities. As it is depicted on Fig. 2 the affected neighbours (A and B affected by C in the example) respond to each ACK by sending all the messages stored in their buffers ($m1$, $m2$, $m3$) which are in *inactive* state, since the identity announced in the ACK (n_F) is completely new and it seems to come from a not yet informed node. The attack is cheap since just one inexpensive ACK packet issued by the attacker may reach several nodes which can answer with several possibly expensive REQF packets that, besides, induce other responses to them (3 REQF, 3 ACK and 3 OKTF in the example).

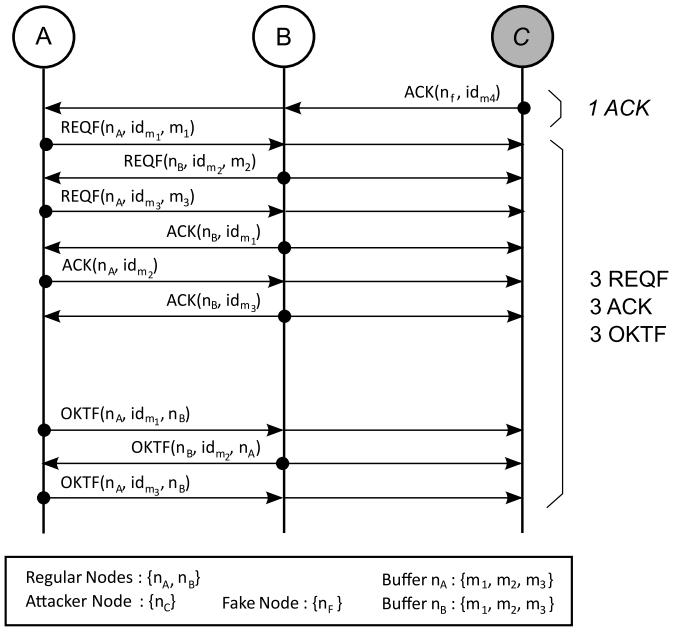

Fig. 2: Draining attack

– **Grey hole attack**: This attack, which exploits the propagation of the message delivery status, consists of making the nodes around the attacker to believe that the messages they disseminate have already reached k nodes as required. This makes the fair nodes to execute the mechanisms for removing the message, thus resulting in a reduction of network message k-delivery ratio. As can be seen in Fig. 3 the attacker answers the REQF packets received with an ACK that contains a forged *informed* vector (see values within parenthesis in the example). The vector is modified to include $k-1$ bits set to 1. Hence, when another fair ACK is received the node which has sent the REQF considers that the message has been disseminated to k nodes and issues a BS packet. Note that the attacker does not directly set the number of bits of the *informed* vector to k in order to go unnoticed.

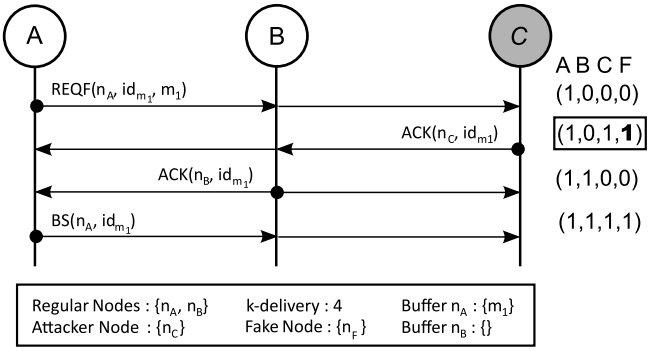

Fig. 3: Grey hole attack

In both cases the adversaries do not participate in the normal operation of the network, but can listen and send packets as any other node. Both of the attacks are tested in both continuous and transient modes. The continuous mode enables the attack during 2/3 of the detection time. From the 2067th time step in our tests, until the end of the simulation. While the transient mode enables the attack during a certain interval of the simulation, from 2200 to 2400 seconds in our tests. The former shows the effects of a persistent attack, while the latter shows the effects of an attack that disappears shortly after.

These attacks have indeed a significant impact on network performance. Beginning with the draining attack, it is performed by 5 nodes each sending 10 ACK packets/second with different identities. Each of these ACKs produces around 15 direct responses as REQF packets issued by the victims. The impact of the continuous draining attack can be seen on Fig. 4, where a huge and sharp increase of the network packet transmissions can be observed soon after the attack. Note that a peak, with around 150% higher packet transmission rate, is registered during the first 100 seconds of the attack. Later this rate is reduced

and stabilised to around a 90% higher rate compared to the no attack case. This is due to the fact that just at the beginning of the attack there are more inactive messages ready to be forwarded in the buffers of the fair nodes.

The grey hole attack, whose goal is to reduce the chances of successful dissemination of messages, is performed by 5 nodes each one answering to all the REQF packets they receive with forged ACK packets. The impact of the continuous grey hole attack can be seen in Fig. 6, which depicts how the message k-delivery rate, in comparison with the messages introduced into the network, suddenly drops to a 10% of the normal rate (which in this scenario is around 10 messages/second) just after the beginning of the attack.

5.3 Implemented mitigations

In a highly unpredictable environment with pockets of connectivity, we need to act in a way that works with unknown node IDs and "fuzzy" normality. Instead of suspecting individual nodes and isolating them (which is very difficult to do accurately) as for example in the work by Wang *et al.* [20], our approach is based on the adjustment of the protocol behaviour in the own node to a careful mode. In this mode the performance can slightly decrease, but the impact of the attacks is strongly reduced. The new operational mode responds to the threats described in Section 3.2 and it is generic enough to provide a unified response to them.

For the attacks that target the RWG mechanisms for discovery of new nodes and selection of custodians, the mitigation consists of ignoring cheap packets (ACK, OKTF, and BS) with "fake" identities. Of course, in the normal operation of the protocol none of the nodes have knowledge to distinguish good and fake identities. We propose that we have a chance of recognising such nodes if we add a low overhead mechanism to the protocol, namely creating a list of known nodes during the periods in which the mitigation is not enabled. This can be effectively done if a list is updated with identities of nodes that have sent REQF messages. This addition to the protocol is not wasteful of energy (given that transmission energy is the dominant factor) but uses up some storage at each node. We also expect a slight increase in the latency for detection of new nodes in the vicinity.

For the attacks that target the RWG mechanism for propagation of delivery status, the solution consists of going into a "suspicious mode". In this mode we restrict the update of the delivery information from the ACK packets received (i.e. do not set zeros to ones in the bit vector). More specifically, when the mitigation is enabled, the *informed* vectors of the messages contained in the node's local buffer are only updated from the *informed* vectors of the REQF, OKTF and BS packets. If an ACK is received the local *informed* vectors are just updated for the position that corresponds to the sender of the ACK, but the *informed* vector contained within the ACK packet is ignored. This mitigation imposes a heavier burden on the network resources. The information regarding the number of deliveries of each message is propagated slower than usual and

the message is kept in the network for a longer time than needed increasing the transmission of packets around a 25%.

Obviously, the application of these techniques reduces the performance of the network if enabled indefinitely (we lose some of the strengths of RWG). This is the reason why they are not an integrated part of the protocol specification. Instead, it is best to apply them just during an interval (I_m) after the detection of an attack. Further studies should show what are the optimal intervals to select for I_m in a given network environment.

5.4 Evaluation metrics

Given the chaotic nature of the scenario we would not be able to use the classic detection rate (DR) and false positive rate (FPR) metrics for evaluation. This is due to the fact that the success of the approach is not measurable with those metrics neither on a per node basis nor on a network wide (average) basis. The locality of the attackers, the nature of the partitions, and the mobility of the nodes, all affect the results so that there are no meaningful homogeneous outcomes using these metrics. However, we will come back to them and analyse the above locality aspects in Section 5.6. Our main evaluation metrics for detection and mitigation are instead:

- **K-Delivery Rate (KDR)**: Depending on the connectivity of the network, the message load, and the dynamics only a proportion of the messages sent are finally k-delivered. Thus, a good metric to evaluate the possible effects of an attack and its mitigation is the number of messages which are in fact k-delivered over the interval of study.
- **Packet Transmission Rate (PTR)**: Another relevant metric is the number of packets transmitted during the interval of study. Besides being an indicator of the usage of bandwidth as a resource, the PTR is an indicator of the energy spent by the nodes, since the more transmissions the more energy is consumed.

5.5 Detection and mitigation results

The detection approach proposed in Section 4 has been tested with the two attacks and two combinations described in Section 5.2 (continuous and transient). In the following, whenever an attack is sensed the anomaly detector enables both mitigations at the same time (ignores ACK packets with possible bogus IDs, and does not update the informed vector on ACK packets received). The I_m interval is selected as 200 seconds.

Fig. 4 shows the effect of applying the detection and mitigation to the continuous draining attack. When the detection and mitigation mechanism is disabled the PTR in the network is around 90% higher than the normal rate as a result of the attack (except during the initial peak which is higher). However, when the mechanism is enabled, the PTR increases, but as soon as the attack becomes detected in most of the nodes, the mitigation actions are taken and the

attack impact is reduced. Fig. 5 shows the transient draining attack, which as in the previous case increases the PTR with the same proportions. Nonetheless, in this case an initial peak of the PTR with mitigation is noticeable since the PTR in the simulation without attack is also increasing. In this case it is worth mentioning that after the attack, the number of packets sent gradually returns to the normality as the mitigation is disabled within I_m of detecting in each node. In both cases the detection delay observed is about 10-30 seconds after the beginning of the attack for nodes close to the attackers.

Fig. 6 shows the effect of applying the detection and mitigation to the continuous grey hole attack. When the mechanism is disabled and the attack starts the KDR drastically drops to around 10% of the normal rate. With the mechanism enabled the KDR also drops, but not so low, and after a certain period it stabilises to values slightly below the values without an attack . In Fig. 7 the impact of the transient grey hole attack is shown, which as in the previous case drastically decreases the delivery ratio. The detection and mitigation responds similarly, but in this case it can be observed that the mechanism helps to a fast recovery once the attack has ended. With this attack the detection delay is longer and highly dependant on each node. The nodes close to the attackers show a detection delay around 10-60 seconds after the beginning of the attack for both continuous and transient modes. It is worth to say that this is a complex attack to mitigate since once the *informed* vector is sent there is a contagious impact on the other partitions while the mitigation is not enabled everywhere (since the detection is not strong enough in some places with the same threshold everywhere).

The results shown demonstrate that the approach is successful in creating a resistance to the attacks that conform to the given threat model despite the difficulties that the complexity of IC-MANET bring to the picture.

5.6 Locality and classic metrics

The most usual evaluation metrics for measuring the anomaly detection performance are the Detection Rate (DR) [1] and the False Positive Rate (FPR) [2]. In this section we show why these metrics are less meaningful in IC-MANET. We begin by discussing how to apply the metrics in the evaluation. To calculate these metrics we need to determine whether during the period of alarm the node was under attack. In intermittently connected networks the concept of being under attack for a particular node is not clear. Just considering a fixed attack interval for all the nodes is meaningless, since attacks do not always take place in a well-determined time interval and confined space. Nodes can be isolated from the attackers during some periods or can be too far from the attackers to be significantly affected by them. Besides, some attacks can be propagated further even if their source has stopped attacking, such as some types of flooding attacks. Hence, our attempt to account for the classic DR and FPR metrics,

[1] $DR = TP/(TP+FN)$ where TP stands for true positives and FN for false negatives
[2] $FPR = FP/(FP+TN)$ where TN stands for true negatives

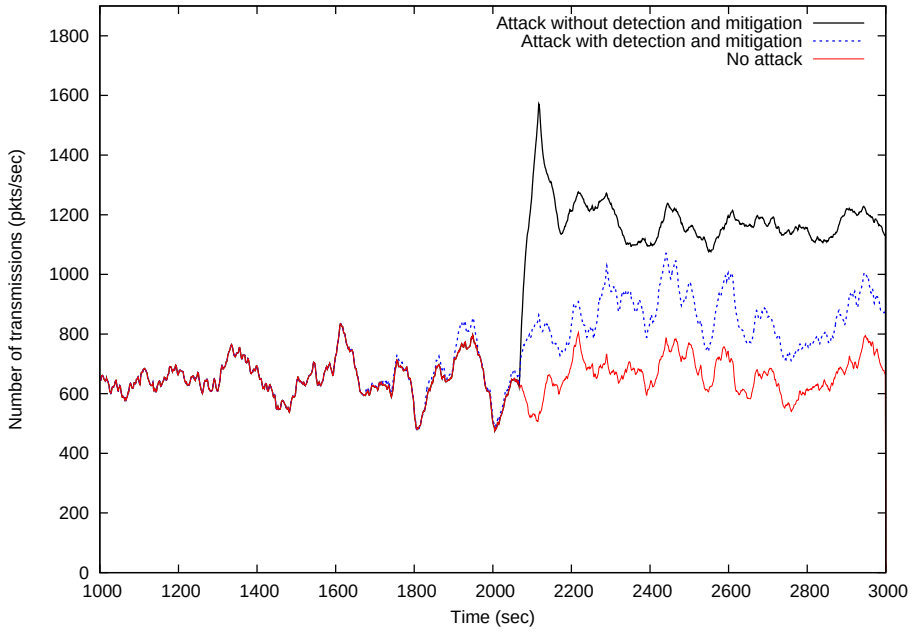

Fig. 4: Draining continuous attack

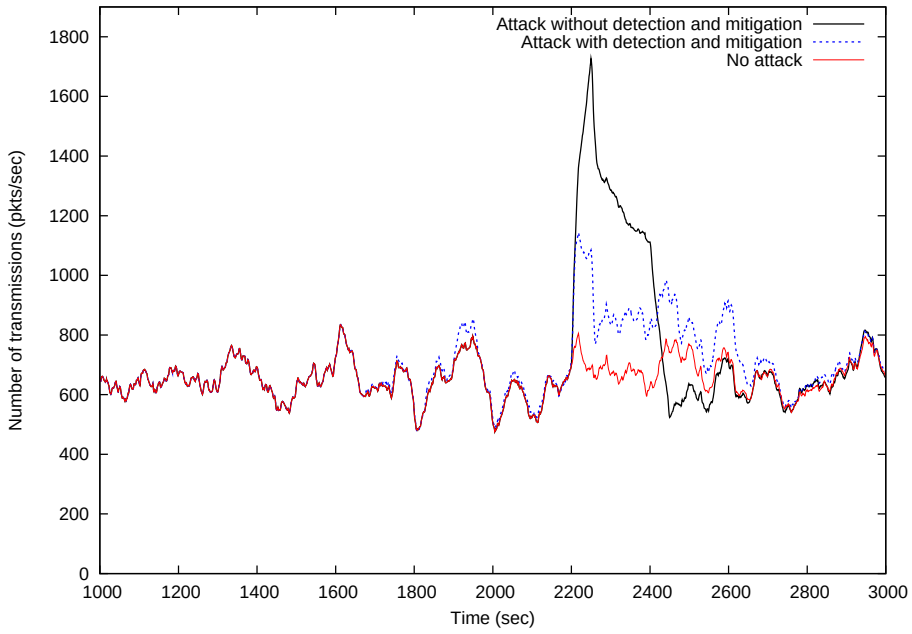

Fig. 5: Draining transient attack

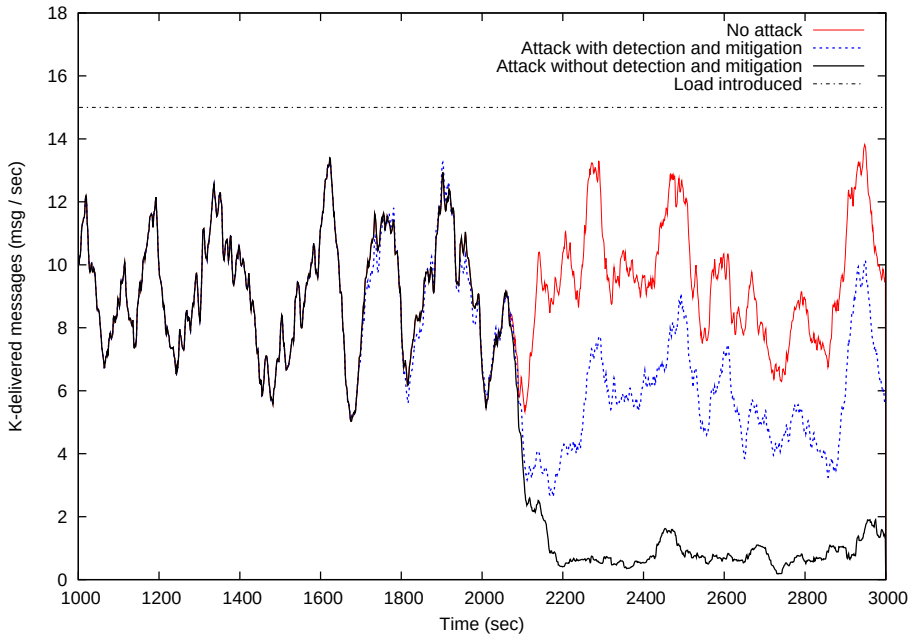

Fig. 6: Grey hole continuous attack

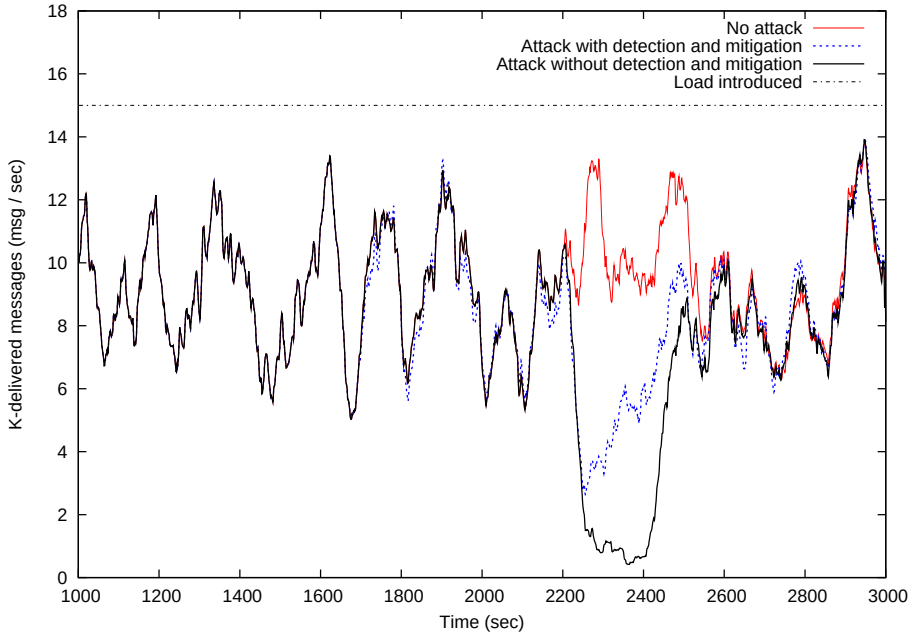

Fig. 7: Grey hole transient attack

is based on tagging the packets sent by the attacker and the packets sent in response to them. Then, in each aggregation interval I_a(see Section 4.3) a node has been considered as being under attack if at least one of the tagged packets has been received in that interval.

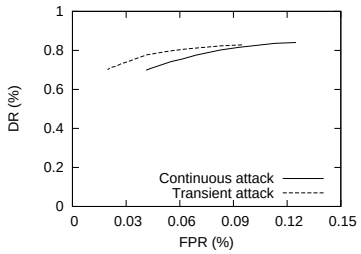

Fig. 8: ROC drain attack

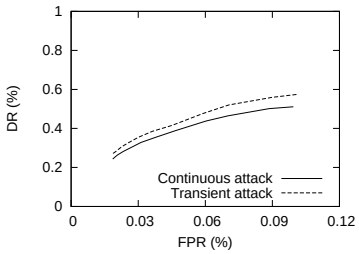

Fig. 9: ROC grey hole attack

The network wide average results obtained in terms of DR and FPR, by using different T_2 values have been depicted on Fig. 8 and 9. These numbers are computed by averaging the performance of all 25 anomaly detectors over the entire test interval. The curves demonstrate that in highly partitioned networks with very different conditions it is not feasible to analyse the results of the detection mechanism on an aggregate basis using these metrics. While the earlier results were convincing about the success of the approach, these curves show mediocre results overall.

We have observed that the traffic flow, the type of attack, and the number of attackers in each partition produce very different detection rates. The network topology in our disaster area is composed of eight partitions more or less stable along the whole simulation, with moving nodes acting as "bridges" over the partitions. Analysing the results node by node we have confirmed that the parameter with more influence over the detection performance is the proximity of the adversaries to the fair nodes. Table 1 show the best, worst, and average DR and FPR, for the continuous draining and grey hole attacks. Results in each column are categorised into different classes. Each class (different rows as described in column 1) shows the results aggregated for partitions that have similar number of adversaries, i.e. partions with no adversaries, partitions with 1 adversary, and so on. There are around 1/3 of the fair nodes in each class. The results, calculated with the alert aggregation threshold T_2 at 5%, demonstrate that the less the partition is affected by attacks the worse is the performance of the detection. That is, the classes with zero and one adversary are the ones that reduce the average detection performance. Note that despite having partitions with no adversaries, some attacks are received by sporadic contacts with other partitions.

Another aspect which has been observed is that in the transient cases the false positive rate is a bit lower than in the continuous cases. The reason is that the

# Adversaries per partition	Draining Attack						Grey Hole Attack					
	Average		Best		Worst		Average		Best		Worst	
	DR	FPR	DR	FPR	DR	FPR	DR	FPR	DR	FPR	DR	FPR
2	94%	6%	95%	6%	93%	8%	63%	8%	70%	5%	60%	10%
1	90%	5%	97%	3%	85%	7%	44%	4%	55%	2%	40%	7%
0	58%	5%	93%	4%	45%	8%	29%	6%	66%	3%	11%	9%

Table 1: Detection performance for the continuous attacks

attacks are always detected with a small delay, but the alarm also persists when the attack is finished. Since the attack is not continuously received uniformly by all the nodes, because of their mobility, there are some gaps during which the alarms are enabled and counted as false positives. The continuous attacks are longer and present more of these gaps. This shows, once again, the complexity of the performance accounting using these metrics.

6 Conclusions

In this article we have presented a holistic anomaly detection and mitigation approach for dissemination protocols for intermittently connected networks. The approach has been integrated and evaluated in the Random Walk Gossip dissemination protocol applied within a disaster area scenario.

We have adopted a statistical-based detector algorithm to combat the typical resource constraints associated with the devices with respect to CPU power used for learning and detection. The threat model for which the approach has been validated focuses on making a big impact on fair nodes with little invested energy by the adversary. Moreover, the adversary behaviour is so similar to the normal behaviour that is hard to distinguish the attacks by creation of constraints, signatures or rules. So this environment is indeed a challenging environment.

Taking into account this threat model we have had to add a mitigation mode to the basic protocol operation. When in this mode, small modifications in the protocol create a chance of deciding when the own behaviour has to be changed due to a suspected attack. This is different from earlier works where identification of the culprit and individual isolation or specific treatment is the response. The integrated protocol can of course be run in the original no-mitigation mode when no attacks are expected and then no protection is provided either. Hence, the added detection-mitigation algorithm can be seen as an enhancement of an earlier protocol that works in a fair-play scenario. We believe this way of thinking can be generalised and applied in other dissemination protocols too.

Furthermore, our approach assumes full knowledge of the adversary about the protocol and even the anomaly detection scheme. The adversary cannot easily adapt to avoid detection by the algorithm due to the unpredictability of what learning has accomplished in the normality model. This is a simple and powerful aspect of our scheme.

The evaluation of the approach has demonstrated its effectiveness by showing resistance to the attacks using network performance metrics. In two attack modes, transient and continuous, we have shown that mitigation brings back the network to performance levels close to pre-attack scenarios. The analysis has also highlighted the complexity of using the classic metrics, detection rate and false positive rate, in highly partitioned networks. These metrics are not appropriate to measure the detection performance on a global basis in highly partitioned networks.

Future work includes identifying the applicability of the methods to more attack types, an intermittent version of the current attacks, and the addition of new threat models. It is also interesting to explore which parts of this resilience to attacks can be beneficially integrated into the dissemination algorithm. Current work includes the addition of two new components to the detection-mitigation loop. First, a diagnosis element that runs in parallel with a general (early) mitigation. This would be useful to adapting the mitigation without pinpointing attacker nodes. Second, an adaptive component that decides when and how to end a given mitigation phase, and a return to the less careful mode.

Another aspect in which more research is required is the study of impact of mitigation actions. When a node enables the mitigation, in some cases this may change the behaviour of the system and can be detected as an anomaly creating a recursive chain of alarms among the nodes. This is a complex problem because the behaviour of the system can be affected by the mitigation actions applied by all the nodes.

Acknowledgements

This work was supported by a grant from the Swedish Civil Contingencies Agency (MSB) and the national Graduate school in computer science (CUGS).

References

1. Denning, P.J.: Hastily formed networks. Communications of the ACM **49**(4) (2006) 15–20
2. Steckler, B., Bradford, B.L., Urrea, S.: Hastily formed networks for complex humanitarian disasters after action report and lessons learned from the naval postgraduate school's response to hurricane katrina. Technical report, Naval Postgraduate School (2005)
3. Asplund, M., Nadjm-Tehrani, S.: A partition-tolerant manycast algorithm for disaster area networks. IEEE Symposium on Reliable Distributed Systems (2009) 156–165
4. Aschenbruck, N., Gerhards-Padilla, E., Gerharz, M., Frank, M., Martini, P.: Modelling mobility in disaster area scenarios. In: MSWiM '07: Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems, New York, NY, USA, ACM (2007) 4–12
5. Ye, N., Chen, Q.: An anomaly detection technique based on a chi-square statistic for detecting intrusions into information systems. Quality and Reliability Engineering International **17**(2) (2001) 105–112 John Wiley & Sons.

6. Yang, H., Luo, H., Ye, F., Lu, S., Zhang, L.: Security in mobile ad hoc networks: challenges and solutions. IEEE Wireless Communications **11**(1) (2004) 38 – 47
7. Prasithsangaree, P., Krishnamurthy, P.: On a framework for energy-efficient security protocols in wireless networks. Computer Communications **27**(17) (2004) 1716 – 1729 Elsevier.
8. Farrell, S., Cahill, V.: Security considerations in space and delay tolerant networks. In: Second IEEE International Conference on Space Mission Challenges for Information Technology, Washington, DC, USA, IEEE (2006) 29–38
9. Liu, Y., Li, Y., Man, H., Jiang, W.: A hybrid data mining anomaly detection technique in ad hoc networks. International Journal of Wireless and Mobile Computing **2**(1) (2007) 37–46 Inderscience.
10. García-Teodoro, P., Díaz-Verdejo, J., Maciá-Fernández, G., Vázquez, E.: Anomaly-based network intrusion detection: Techniques, systems and challenges. Computers & Security **28**(1-2) (2009) 18–28 Elsevier.
11. Nakayama, H., Kurosawa, S., Jamalipour, A., Nemoto, Y., Kato, N.: A dynamic anomaly detection scheme for AODV-based mobile ad hoc networks. IEEE Transactions on Vehicular Technology **58**(5) (2009) 2471–2481
12. Cabrera, J.B., Gutirrez, C., Mehra, R.K.: Ensemble methods for anomaly detection and distributed intrusion detection in mobile ad-hoc networks. Information Fusion **9**(1) (2008) 96–119 Elsevier.
13. Chuah, M., Yang, P., Han, J.: A ferry-based intrusion detection scheme for sparsely connected ad hoc networks. In: Fourth Annual International Conference on Mobile and Ubiquitous Systems: Networking & Services, IEEE (2007) 1–8
14. Scalavino, E., Russello, G., Ball, R., Gowadia, V., Lupu, E.C.: An opportunistic authority evaluation scheme for data security in crisis management scenarios. In: ASIACCS '10: Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security, New York, NY, USA, ACM (2010) 157–168
15. Thamilarasu, G., Balasubramanian, A., Mishra, S., Sridhar, R.: A cross-layer based intrusion detection approach for wireless ad hoc networks. In: IEEE International Conference on Mobile Adhoc and Sensor Systems Conference, IEEE (2005) 854–861
16. Sun, B., Wu, K., Pooch, U.W.: Zone-based intrusion detection for ad hoc networks. International Journal of Ad Hoc & Sensor Wireless Networks. Old City Publishing (2004)
17. Tseng, C.H., Wang, S.H., Ko, C., Levitt, K.: DEMEM: Distributed evidence-driven message exchange intrusion detection model for MANET. In: Recent Advances in Intrusion Detection. Volume 4219 of LNCS., Springer (2006) 249–271
18. Huang, Y.a., Lee, W.: A cooperative intrusion detection system for ad hoc networks. In: SASN '03: Proceedings of the 1st ACM workshop on Security of ad hoc and sensor networks, New York, NY, USA, ACM (2003) 135–147
19. Deodhar, A., Gujarathi, R.: A cluster based intrusion detection system for mobile ad hoc networks. Technical report, Virginia Polytechnic Institute & State University
20. Wang, S.H., Tseng, C.H., Levitt, K., Bishop, M.: Cost-sensitive intrusion responses for mobile ad hoc networks. In: Recent Advances in Intrusion Detection. Volume 4637 of LNCS., Springer (2007) 127–145
21. Moore, D.S., Cabe, G.P.M.: Introduction to the practice of statistics. 5th edn. W. H. Freeman (2005)