Social Insect-based Sybil Attack Detection in Mobile Ad-hoc Networks

Parisa Memarmoshrefi
Institute for Computer Science, Telematics Group
University of Göttingen, Goldschmidtstrasse 7,
37077 Göttingen, Germany
memarmoshrefi @cs.uni-goettingen.de

Hang Zhang
Institute for Computer Science, Telematics Group
University of Göttingen, Goldschmidtstrasse 7,
37077 Göttingen, Germany
hang.zhang @cs.uni-goettingen.de

Dieter Hogrefe
Institute for Computer Science, Telematics Group
University of Göttingen, Goldschmidtstrasse 7,
37077 Göttingen, Germany
hogrefe @cs.uni-goettingen.de

ABSTRACT

Due to the characteristics of mobile ad-hoc networks (MANETs) such as distributed self-organizing nature and multi-hops communication, bio-inspired approaches proposed to address challenges in these networks. Like in all communication networks, one of the major challenges in MANETs is security provision in these environments. As the traditional approaches such as Public Key Infrastructure (PKI) which are based on central authority are not suitable for dynamic, distributed and cooperative ad-hoc networks, one of the fundamental security threats in MANETs is Identity-based attack.

In this work we propose a scheme inspired of the social life of ants, to discriminant the sever Sybil attacker and their related shadow/hidden identities in a self-organized authentication mechanism. For this aim, we adapt the idea of discrimination between mate and non-mate members, in real ants' life, based on aggression concept. We also present our simulation results and discuss some open research issues.

Categories and Subject Descriptors

C.2.2 [Computer-Communication Networks]: Network Protocols

General terms

Security, Performance, Design.

Keywords

Insect-based aggression, Sybil attack, Ad-hoc Networks.

1. INTRODUCTION

Mobile ad-hoc networks (MANETs) are complex distributed systems, consisting of wireless mobile nodes that are communicating in multi-hops fashion, without any infrastructure. Joining of nodes freely and dynamically into the network allows people and devices to connect to the network in areas where no pre-existing communication network exists. As these kind of self-organized networks are built on the cooperation between nodes, so they pose important challenges for security provision. One of the fundamental security threats is Identity-based attack. Due to the lack of centralized identity management in MANETs, providing a unique, distinct and persistence Identity per node is a challenging and vital task. In self-organized authentication mechanisms the objective of the identification protocol is: in case of having two parties *A* (verifier) and *B* (claimant), B is able to prove itself as an authentic party to A; then A will complete the authentication process by verifying the correctness of a message which shows that B is in possession of the corresponding keying materials.

In authentication mechanism one-to-one mapping between entity (node) and its identity and its keying materials is important. According to [1] a node is termed as Sybil attack if it manages to create and control more than one identity for a single entity. This malicious behavior leads to serious threats for varieties of network functionalities such as authentication process, multipath routing, in reputation and trust-based security mechanisms. In wireless sensor networks, Sybil nodes report incorrect sensed measurements, which might affect significantly the data aggregation results [2]. In vehicular ad-hoc networks, Sybil attackers are able to provide fake impression of traffic congestion in order to divert the traffic [3]. Therefore, detection and prevention of Sybil attacker is strongly desirable in self-organized ad-hoc networks.

Characteristics of bio-inspired approaches made them suitable for addressing the networking challenges. These characteristics are such as adaptive to environmental changes, resiliency to failure, ability to learn the new conditions, successful and collaborative operation to form the global intelligence, self-organization etc.[4].

In this work we propose a scheme inspired of the social life of ants, which discriminant Sybil identities in a self-organized authentication mechanism. In nature, the stability of animal societies is based on the discriminating among individuals and rejection of non-group/mate members. Ants are good models for studying recognition mechanisms, because they are efficient in discriminating "friends" (nest-mates) form "foes" (non-nest-mates) [5]. Ants have a Cuticular Hydrocarbon (CH) profile in which multicomponent cues are encoded. The CH profile plays a role as a chemical signature and is unique to each colony members due to their shared diet. Ants, using their antenna, detect the CH of other and use it as a detector for mate or non-mate members. Any additional components in CH, which is not habituated to previously, will elicit aggressive behavior. In [5] it is indicated that aggressive behavior is caused only based on the additional chemical of the CH profile and not on odor similarity. This

BICT 2014, December 01-03, Boston, United States

DOI 10.4108/icst.bict.2014.258041

chemical reveals foes and makes ants to be capable of detecting intruder rather than recognizing the nest-mates. Our security scheme is based on this phenomenon. We consider mobile ad hoc network as a colony of ants and inspired from ants' life, we propose a model to discriminant the Sybil attackers in the network who aim to defeat the self-organized authentication process.

The rest of the work is organized as follows: Section 2 presents an overview about the related work. In section 3 we present our Sybil attack detection approach. In sections 4 we introduce the simulation parameters and evaluate the results of the implementation. Finally, section 5 concludes the work and presents the directions for the future work.

2. RELATED WORK

The Sybil attack taxonomy is presented in [2], as a malicious node who illegitimately has multiple identities. In direct communication, the Sybil node communicates directly with the legitimate nodes. However, in indirect communication no legitimate node can communicate directly with the Sybil nodes. Instead malicious node claims to be able to reach the Sybil identities. Furthermore, by fabricated identities, the attacker can create arbitrary Sybil identities; while by stolen identities, an attacker cannot fabricate new identities. In this case the attacker assigns other legitimate identities to its Sybil nodes. Another dimension of Sybil attacker model is simultaneity of the Sybil identities. The attacker may have all Sybil identities participating in the network at the same time. So, it can cycle through, in order to pretend they are different and separate nodes. Alternately, the attacker can present their identities once or simulates leaving and joining nodes. While using a new identity and discarding the previously used identities, the attacker tries to hide the history and consequently the relation between its Sybil identities.

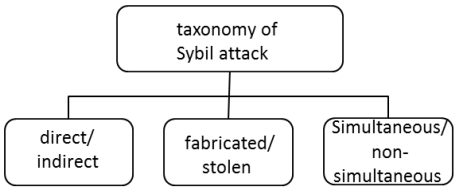


Figure 1. Taxonomy of Sybil attack

In order to prevent Sybil attack, it is the mission of authentication mechanism that provides a one-to-one mapping between entity (node) and its identity and corresponding keying materials. A classification of authentication protocol for mobile ad hoc networks is presented in [6]. The author categorized the existing authentication protocols based on: a) role of the nodes in the network with respect to authentication process; b) type of the credentials; c) the time when credentials are established.

In [7] a classification of authentication protocols for MANETs is presented correspond to the role of the node in the network. a) In central Certificate Authority (CA), known also as Trusted Third Party (TTP), only CA has the role of assigning a unique identity (e.g. Public key) to an entity and issue a certificate (e.g. Public key Certificate) for it. Although, this authentication mechanism is resistance to Sybil attack; however, it is not appropriate for mobile ad hoc network due to topology changes in these environment and furthermore could be a single point of attack. b) In distribute CA system, a set of *n* nodes in the system have the role of the Certification Authority and any *k* nodes out of these *n* nodes collectively can provide CA services. However, most of the distributed CA involves with the generating and verifying cryptographic identities; but the one-by-one binding of these cryptographic identities to the corresponding nodes (e.g. IP) is overlooked. Therefore this leads them to be vulnerable to the fatal Sybil attack [8]. c) In self-CA system, every node in the networks has the role of the CA; and they are able to generate their owing keying materials and issue public key certificate for own and for others based on their knowledge. Via a certificate, the binding of a node's identity to its corresponding public key is proven by a digital signature of the issuer. Each node maintains a local certificate repository, and performs the public key authentication via chain of certificates. Authors in [9],[10] and [11] proposed self-organized authentication mechanism for ad-hoc networks; however they are not resistant to the Sybil attack. Self-CA system is mostly based on existence or creation of the web of trust.

Authors in [6] categorized authentication protocols based on type of credentials. Credentials are information required for authentication. They are based on two types: identity-based credentials or context-based credentials. In the first class, node's possession of a unique identifier (e.g. symmetric, asymmetric cryptographic key) proves the authenticity of the node with high certainty. While in latter category, unique contextual attribute is used in order to identify the claimant node with high certainty. In general, these credentials can be behavioral or physical. In behavioral-based contextual credentials, identifying and authenticating the supplicant are based on its pattern of behavior. The authentication process is done by monitoring the behavioral pattern of the supplicant with respect to certain functionality and by classifying it according to its performance. In physical-characteristic based, credential is the unique physical characteristic of the supplicants, like its GPS (Global Position System) location or RSSI (Received Signal Strength Indication) or SNR (Signal to Noise Ratio).

[12] proposed a Sybil attack detection scheme based on the Received Signal Strength. However, this it incurred extra overhead due to periodic localization of the nodes]. If identity-based credentials assure that node *B* (claimant) possesses the identity and keying materials it claims, node *A* (verifier) could authenticate it successfully.

In self-CA, the concept of trust in authentication process, which also called identity trust, is the trust value to the credentials. In our previous work [13], we proposed a self-authentication mechanism where each node creates its own public/private key pairs and issues certificate for its direct neighbor nodes and signs it with its private key. To all created certificates in the initialization phase of the network, a trust value is assigned. This value is equal to a threshold value which represents the uncertainty of the correctness of the certificates at the networking setup phase. In this self-organized authentication mechanism, public key authentication is performed through a chain of certificates. The model aimed at finding the most trustworthy certificate chain for a PK authentication; however the model is still prone to the Sybil attacker.

3. PROPOSED SYBIL IDENTITIES DETECTION

When a node, S, wants to authenticate the public key of a destination node located out of S's radio range, (node D in figure 2), a chain of valid certificates from S to D is needed. Every certificate in the chain is verified by the public key of the previous certificate in the chain (Figure 2).

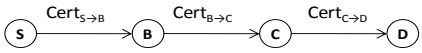


Figure 2. PK Certificate chain: $\{Cert_{S\to B}, Cert_{B\to C}, Cert_{C\to D}\}$

Upon getting the certificate chain, source node calculates the trustworthy of the destination's PK via reported certificate chain as follows:

$$t_{SD} = \prod_{k=1}^{n} t_k$$

t_k is the trust value between two directly connected nodes on the certificate chain.
However, the Sybil attacker tries to deceive other nodes so that to make them believe in a false certificate.

3.1 Attack Model

We assume that a Sybil node creates different fake identities with the corresponding credentials (public/private key pairs) known to the Sybil node. We call these fake identities shadow identities. Sybil node uses these shadow/hidden identities to issue fake certificates. In figure 3, node B is a Sybil node with two shadow identities B1 and B2.

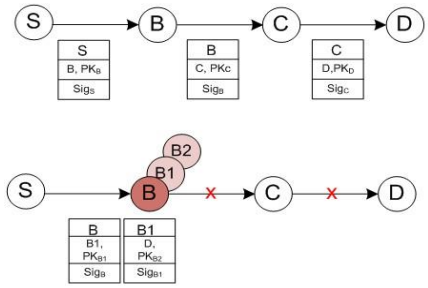


Figure 3 Malicious node with two hidden identities

We assume, a Sybil node always keeps one of its shadow identities specifically for impersonating the requested destination, D. In this example node B, reserves identity B2 for this aim; and PK_{B1} will be assigned to the destination. Node B tries to deceive source node to make it believe that PK_{B1} is the public key of the destination. It uses other shadow identities B2 to issue fake certificates and insert them into the certificate chains. In this case, the certificate chain received from D to S is (Figure 3):

$$\{cert_{S\to B}, cert_{B\to B2}, cert_{B2\to B1}\}$$

3.2 Social Insect-based Detection

In a real social ant colony, ants are called "walking chemical factories" [5], because their recognition is based on chemical cues [5]. Their nest mate and non-nest mate discrimination is also chemical mediated and is based on CH profiles. It is determined by genetic and environmental factors such as diet, nest materials and physical contact with colony members [14]. The genetic and odor dissimilarity lead to cause the aggressive behaviors and consequently non-nest mate detection. The aggression value of each ant, encountering other insects, will be increased; and consequently it leads to represent a special behavior toward non-nest mates.

In order to make an aggression bioassay, biologists in [5] carried out some test on real ants for duration of time. Each observed behavior of ants was given a score indicating increasing aggression level (al): 0 = investigate with antenna, 1 = open mandibles, 2 = bite, 3 = ready to spray formic acid. For each aggression test, the overall aggression index (AI) is calculated using following formula:

$$AI = \frac{\sum_{i=1}^{3} al_i \times t_i}{T}$$

where al_i and t_i is the aggression level and the corresponding duration of the *i*-th behavior and T is the total interaction time. Borrowed form ants' CH profile and their bioassay aggression, in next section, we define an aggression metric for mobile nodes in order to detect the malicious node in the network.

3.3 Aggression Metric in Mobile Nodes

In order to design a social-insect based detection, we need an unsupervised mechanism that let the nodes in the network to detect the malicious Sybil attackers. As we assumed, the Sybil node creates several fake local IDs and applies them as a proposed credentials for the requested destination; therefore, any detected group of nodes that their behaviors are similar to each other but far from the norm of the society (non-mate members), would irritate the aggression factor of the mate members. For this non-mate detection, we apply a cluster analysis method based on hierarchical agglomerative approach. Figure 4 represents the flow chart of the designed model.

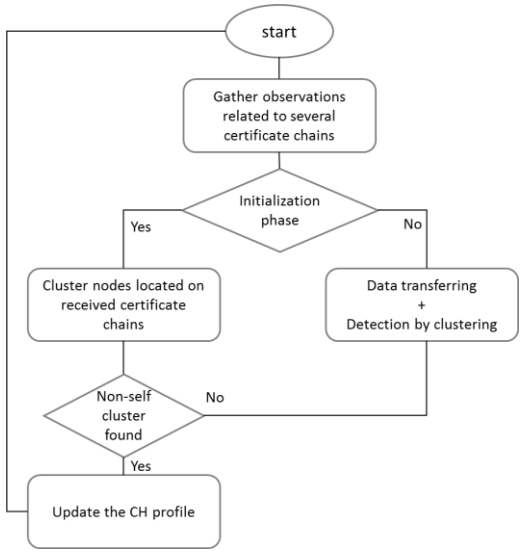


Figure 4. Flow chart of the approach

In the initialization phase each node needs to gather some information about its environment. For this phase we assume all nodes are belongs to one colony and therefore, aggression value is

set to 0; indicating all nodes are self-mate members. On the other hand, at the beginning stage of the network establishment, nodes don't have any prior-experience about the certificate transferring process; so, the pheromone values initially are considered as the threshold value. This represents the uncertain situations of the network in this phase. In initialization period, nodes starts requesting the public key of some target destination nodes in the network. For each PK request, source node sends out some number of forward ants (FA) toward destination. When a FA reaches to the destination D, it will be transformed to backward ant (BA) and the BA retraces exactly the path of FA back to the source. Through returning back, BA carries the PK certificates that intermediate nodes added to it, including the destination's certificate. Once source node receives a BA, it recovers the public key of the destination reported trough each certificate chain (Table 1). After receiving all the BAs or the PK discovery timeout occurs, node S starts the analysis phase.

3.4 Calculating Dissimilarity of CH Profiles by Node Clustering

Through received certificate chains, in analysis part, source node starts to cluster nodes by exploiting the behavior relationships between nodes. For this aim, following parameters are defined for each individual node to create the comparison objects. These metrics then will be considered as a part of the node's CH profile updating.

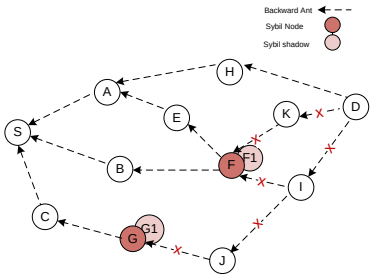


Figure 5. Source node receives different certificate chains

Table 1. Certificate chain tables of node S

Certificate chains	Reported PK for D
SAH	PK_D
SBF	PK_{f1}
SAEF	PK_{f1}
SCG	PK_{G1}

- Cln: it is a value that shows a node belongs to how many different colonies/groups of reported PKs. If a node belongs to two certificate chains which are reported different PKs for the same destination, D, then this value is 2 (e.g. node A in table 1).
- Dist: this value represents the average distance of a node to destination, in certificate chains.
- Soc: this metric demonstrates how social a node is; i.e. how often a node belongs to different colonies/groups. Three states are defined for this metric:
  - 0, means node has been seen in different colonies;
  - 0.5, means nodes has been seen in none of the colonies;
  - 1, means node always seen in one colony.
- Ph: this value is the average pheromone/trust value of the chain that node belongs to.

3.4.1 *Node Clustering*

In Initialization phase each node (e.g. node S) proactively makes PK requests for a random set of target destinations (in table 2, *m* destinations). For each destination, received certificate chains carried by backward ants will be investigated by source node; and the required metrics will be exploited from all received certificate chains. We assume source node makes the PK request for each destination node sequentially. Information corresponding to each destination partially make the objects (certificate chains observations) needed for clustering process. The clustering process will be performed iteratively. Each PK request for a specific target destination is considered as iteration. In each iteration, clustering objects are made up of the metrics related to the PK request of the current target destination together with all other previous destinations. Table 2 show the objects composed of *m* destinations in initial phase. In order to perform node clustering, we choose one unsupervised cluster analysis method [15]: Hierarchical Cluster Analysis (HCA) which is aimed to build a hierarchy of clusters. Based on how to build the hierarchy, there are agglomerative HCA and divisive HCA. The former is a "bottom up" approach. It starts with each object and aggregates the nearest two objects into one cluster. This merger is repeated until all the objects/clusters become in one cluster. On the contrary, divisive HCA begins with one cluster and splits the cluster recursively as one moves down the hierarchy. Considering the complexity, we use agglomerative hierarchical clustering in our work to analysis the information gathered by the ants.

For this aim upon receiving certificate chains, source node retrieves the required information and saves it in form of the objects table (Table 2); and calculates the similarity and dissimilarity between every pair of objects gathered for his current PK request. This similarity is calculated based on the Euclidean distance by using the following formula:

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2}$$

In the next step, source node groups the observed objects into a binary, hierarchical cluster tree, by linking the pairs of objects that are in close proximity.

Table 2. All received certificate chains corresponding to m different destinations form the clustering objects in node S

Node S	Objects\Certificate Chains Observation												
	Distination 1				Destination 2				…	Destination m			
	Cln	Dis	Soc	Ph	Cln	Dis	Soc	Ph		Cln	Dis	Soc	Ph
Node-1	0	0	0.5	0	0	0	0.5	0		0	0	**0.5**	0
Node-2	0	0	0.5	0	0	0	0.5	0		0	0	0.5	0
…	0	0	0.5	0	0	0	0.5	0		0	0	0.5	0
Node-n	0	0	0.5	0	0	0	0.5	0		0	0	0.5	0

Complete linkage clustering is applied to calculate the distance between clusters. It uses the distance information generated in previous step, in order to determine the proximity of objects to each other. Objects are paired into binary clusters; the newly formed clusters are grouped into larger clusters until a hierarchical tree is formed. Figure 6 shows a dendrogram, which is a tree diagram graph for visualizing the arrangement of nodes, clustered by hierarchical clustering calculation.

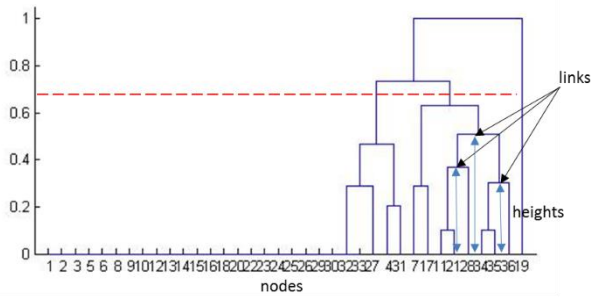


Figure 6. Cluster tree (dendrogram) of 30 nodes with 6 malicious Sybil nodes. Each Sybil node has one shadow identity.

The distance between objects is indicated by the height of the inverse U form edge.

3.4.2 Aggression-based Sybil Discrimination

We consider, in the artificial Cuticular Hydrocarbon (CH) profile, each node maintains two overall values representing for the availability and reliability of the certificates. It consists of two following values:

Pheromone: inspired from food availability in real ants life, this value represents how a node willingly participate in certificate chains and fulfilling the authentication process correctly.

Aggression: inspired from mate and non-mate discrimination in real ants' life this value shows whether a node is Sybil attacker (non-mate) or not.

In this work we assume that local pheromone updating is deactivated. As this work is aimed to make Sybil node discrimination, therefore we focus on aggression value updating in each nodes.

Each node in the network has an aggression vector in that it maintains its aggression level towards other nodes who are clustered differently. We give this value a number [0,1]; 0 means no aggression and 1 means complete aggression. In initialization phase of the network, where nodes have no information about the behavior of other participants, the aggression value is set to 0. By each PK request, the aggression value of a node *n* will be updated as follows:

$$Agg_n = \frac{\sum_{i=1}^{N} Agg_{in}}{N}$$

where Agg_{in} is the aggression value of node *i* toward node *n* and Agg_n is the average aggression value of all nodes in the network toward node n. *N* is the number of nodes in the network including the shadow identities of the Sybil nodes.

Furthermore, the overall aggression index for node *n*, updating after requesting the PK for all *m* destinations is the average of the Agg_n over all destinations:

$$AI = \frac{\sum_{d=1}^{m} Agg_{nd}}{m}$$

This value helps the source node to determine the correctness of the binding between credential of a requested node, node's PK, and the node itself.

4. EVALUATION

4.1 Simulation Parameters

We implemented our approach in QualNet [16]. We consider a network of 30 nodes that move in an open area of 1500 ×1500 m^2, and we assume that there is no obstacle that could influence the mobility or signal propagation of the nodes. All 30 nodes are randomly distributed in the area. Random Waypoint Mobility Model is chosen for implementing the node movement, since it is flexible and it could provide quite realistic mobility patterns. We keep the minimum speed to 0, and use two different maximum speeds: 0 m/s and 5 m/s. The pause time is 30s and the total duration of each experiment is 100s. Some other simulation parameters are show in Table 3.

In this set of experiments, we basically test our approach with static and movement (maximum nodes' speed: 5 m/s) two cases. For each speed, we use 5 different random initial topologies. For each topology, we increase the malicious nodes' number from 1 up to 6. In each single scenario, there are 5 source nodes. Each of them sends out ants to discover 4 public keys of 4 different destination nodes in sequence.

Table 3. Some other simulation parameters

Parameter	Value
Hello interval	1s
Chain discovery timeout	1s
Allowed hello loss	2
Number of forward ants	5
Number of Sybil hidden ids	1
Intermediate nodes reply	No
Buffer size in packet	100
Certificate size	512 bytes

Inside QualNet, we use ALGLIB [17] to apply the agglomerative hierarchical cluster analysis. We consider each node (including the hidden ids from the malicious nodes) as an object. Based on the received certificate chains, every source node sets Cln, Dist, Soc and Ph, 4 metrics to describe the behavior of each object. Then, these metrics are used for clustering the objects through ALGLIB functions. ALGLIB merges the objects into a hierarchical tree based on the complete linkage, which is set by default. The final clustering results are output from QualNet, and we analysis these data in Matlab [18].

4.2 Experimental results

After merging the object into binary clusters, the source node needs to determine where to cut the hierarchical tree into different clusters containing mate and non-mate nodes (dashed red line in figure 6).

Correct aggression value updating is depends on the correct partitioning of the nodes based on their observed behaviors. In order to find the appropriate cut-point, we consider the inconsistency coefficient value. Inconsistency coefficient value compares the height of a link in cluster tree with the average height of the links below it (Figure 6).

The hierarchical cluster tree is to dividing the nodes into distinct and separated clusters based on the nodes' behavior. Consequently, the inconsistency coefficient of links identifies these divisions where the similarity between objects are changed.

In this work we made some experiments on finding the best point to cut the hierarchical cluster tree so that nodes with the same trend in behavior be grouped in the same cluster. For this aim, each source node in the network sequentially asks for the credentials of 4 random destinations. We consider each PK request as one iteration. Objects (certificate chains) observed in each iteration are used as the complementary information for clustering in the next iteration.

In each iteration, source node calculates the inconsistency coefficient for each link in the cluster tree using *inconsistent* function in Matlab. In the next step, the cluster tree will be cut in points equal to the inconsistency coefficient values applying *cluster* function, which calculates the inconsistency coefficient value. Each cut-point value leads to different cluster divisions. Figure 7 demonstrates node clustering labels by cutting the dendrogram in 6 different points.

Nodes	Cluster label					
1	1	3	3	3	3	2
2	1	3	3	3	3	2
3	1	3	3	3	3	2
4	1	3	3	3	3	6
5	1	3	3	3	3	2
6	1	3	3	3	3	2
7	1	1	1	1	1	1
…	…	…	…	…	…	…
35	1	2	2	2	2	3
36	1	5	5	5	5	5

Figure 7. Nodes and their corresponding clustering labels, considering different cut-points.

Results are based on the scenario, in which the network consists of 30 nodes including 6 malicious Sybil nodes. Each Sybil node has one corresponding shadow identity (31 to 36 are the shadow IDs of the malicious nodes {4, 6, 9, 11, 21, 28} respectively).

Based on the cluster labels in each columns, the Aggression matrix will be calculated as shown in table 4. In Initialization phase all elements of the matrix are set to zero; representing no aggression between nodes. After cluster labeling phase, the aggression value of nodes that are grouped in the same cluster considered as 0. Nodes who are grouped into different clusters set the aggression value toward each other equal to 1.

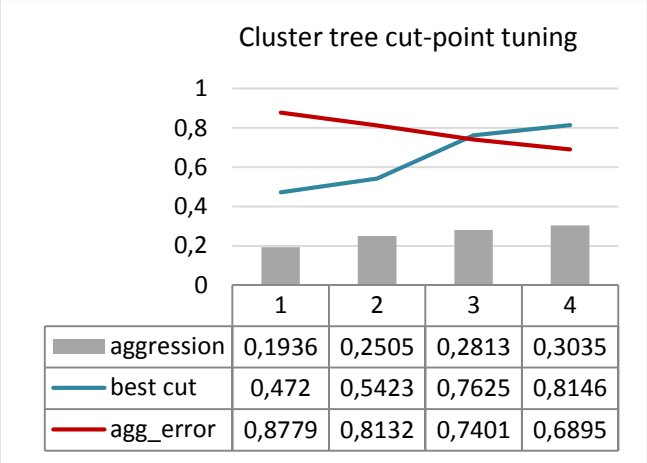
We assume the overall aggression value of all nodes in the network toward a malicious node is 1. In this work this value plays the role

Table 4. Aggression matrix

Nodes	**1**	**2**	**…**	**N**
1	0	Agg_{12}	…	Agg_{1N}
2	Agg_{21}	0	…	Agg_{2N}
…	…	…	0	…
N	Agg_{N1}	Agg_{N2}	…	0
Agg	Agg_{1}	Agg_{2}	…	Agg_{N}

of the discrimination between normal (mate) and Sybil (non-mate) nodes in the network. Therefore, for each group labeling gathered in previous step (each column in figure 7) the overall aggression value of all nodes, including Sybil nodes and their hidden identity, is calculated (last row in table 4). We call this row as aggression vector.

Cutting the cluster tree with different cut-points (inconsistency coefficient), we calculate the aggression vectors for each point. We select the cut-point that results in an aggression vector with the highest aggression values for all malicious Sybil node as the best cluster tree cut-point. Furthermore, we consider the difference between maximum possible aggression, 1, and the average of all calculated aggression values of malicious Sybil nodes. We call this metric as aggression error.



	1	2	3	4
aggression	0,1936	0,2505	0,2813	0,3035
best cut	0,472	0,5423	0,7625	0,8146
agg_error	0,8779	0,8132	0,7401	0,6895

Figure 8. Cluster tree cut-point tuning

Figure 8 represents the cluster tree cutting-point tuning. The values are averaged over 5 different topologies consist of 30 static nodes in the network. In each topology, 5 different source nodes request the PKs of 4 different destinations. The results are also averaged over all malicious numbers in the network.

As it is shown above, generally going through first PK request to the fourth request, the overall aggression values of nodes increase, we reason that, more destination's PK request provides more information about the network environment consisting of malicious (non-mate) nodes. Moreover, the aggression error metric decreases while the number of iterations increases. It is also demonstrated that the best cut and aggression error curves meet, normally when cut-point is between 0.7 and 0.8 for all scenarios.

In figure 9, we investigate the effect of the movement on nodes' aggression values, by increasing the number of Sybil nodes in the network.

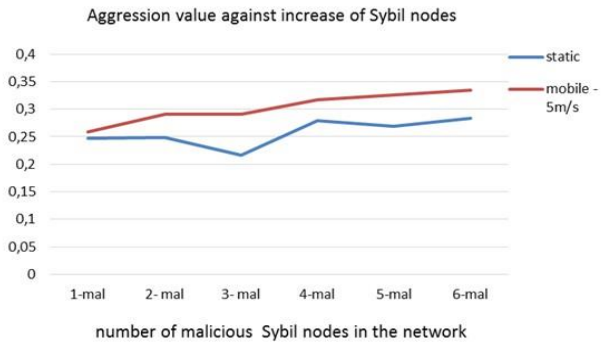


Figure 9. Aggression value against increase of Sybil nodes

It represents that the more malicious nodes in the network the more aggression level, nodes have toward each other. This aggression increasing tendency is observed for both static network and network with mobile nodes.

In addition, it is shown that mobility causes convergence to the higher overall aggression level in the network. We reason that nodes' movement leads to the faster information gathering about the network environment, in compare to the static networks.

Considering 6 Sybil nodes in the network, and 1 shadow identity for each malicious Sybil node, figure 10 demonstrates the overall aggression values toward every node.

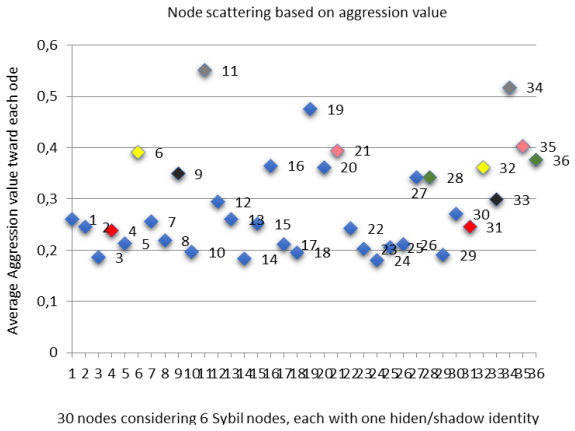


Figure 10. Node scattering based on aggression value

It shows that, most of the normal nodes in the network (nodes with blue color) has the overall aggression value less that 0.3; and most of the malicious nodes has the aggression value above 0.3. Although, we can see some exceptions.

In this work, trough experimental results we try to find the best cluster tree cut-point, by only maximizing the aggression value of the malicious Sybil nodes in the aggression vector. Therefore, in figure 10 some normal nodes such as node 19, 16, 20 and 27 have the same aggression value as the malicious Sybil nodes such as nodes 6, 9, 21 and 28. In the process of finding the best cluster tree cut-point, no parallel controlling for minimizing the aggression value of normal nodes was considered.

Node 4 (with red color) is also a malicious node but located in the group of nodes with aggression value less than 0.3. Some reasons are, either this node in some scenarios selected as the source node, or located in out of the radio rang of other nodes, so that couldn't be visited easily in PK certificate chains. In both case there is no enough information about it to group this node into a correct cluster.

An interesting observation in Figure 10 is that most of the Sybil nodes encountered almost the same aggression value as their corresponding shadow identity: {(6, 32), (9, 33), (11, 34), (21, 35) and (28, 36)}. Even node 4 is almost in the same aggression height as its shadow identity, node 31.

5. CONCLUSION AND OPEN RESEARCH ISSUES

In this work we proposed a Sybil attack discrimination/detection model inspired from the social life of ants. In this regards, the new concept of aggression metric was introduced to discriminate between the normal nodes (mate) and malicious Sybil nodes (non-mate) in the network. In order to enable the source node to discriminate the Sybil attacker(s) from other nodes, we made experimental tests for finding the appropriate cut-point value on the cluster tree. However, the best cut-point value calculated in this work must be tested on a set of observed certificate chains; and be evaluated by the Sybil discrimination results.

The simulation results show that passing through the initiation phase, the average aggression value of each node toward the malicious nodes and their hidden identities are almost in the same level. Therefore, node clustering based on aggression value, could be one of the promising approaches to find the Sybil nodes and their corresponding shadow identities in a self-organized PK authentication mechanism. This work focused on the initialization phase of the Sybil aware authentication mechanism. As the next step we plan to test our proposed model also on the data routing phase.

Furthermore, the results demonstrated that when the number of malicious nodes in the network increases, the average nodes' aggression to the malicious nodes is also increasing. This indicates that nodes understand the potential threat in the environment after the initial phase.

In this work, we investigated the performance of the initial phase. For the next step, we are going to use the clustering results to update the nodes' pheromone and aggression values, which could help the source node to make a decision by choosing the correct public key of the requested destination. Moreover, we plan to investigate our approach in more sophisticated environments, in which the malicious nodes have more than one hidden identities.

6. Acknowledgment

Implementing our proposed authentication protocol in Qualnet, was a challenging task with a lot of effort that was patiently performed by Hang Zhang. She also contributed in the design and data analysis part. Therefore, she is contributing as the co-first author.

We would like to thank Sanaz Karimkhani Asl for providing us information about clustering algorithms.

We would also like to appreciate David Richard Nash in department of Biology, university of Copenhagen, for his valuable suggestions about related biological literatures.

7. REFERENCES

[1] Douceur, J. R. 2002. The sybil attack. Peer-to-peer Systems. Springer Berlin Heidelberg, 251-260.

[2] Newsome, J.,Shi, E., Song, D., and Perrig, A. 2004. The sybil attack in sensor networks: analysis & defenses. In Proceedings of the 3rd international symposium on Information processing in sensor networks (IPSN '04). ACM, New York, NY, USA, 259-268. DOI=10.1145/984622.984660 http://doi.acm.org/10.1145/984622.984660.

[3] Parno, B., and Perrig, A. 2005. Challenges in securing vehicular networks. In Workshop on hot topics in networks (HotNets-IV), 1-6.

[4] Dressler, F., and Akan, O. B. 2010. A survey on bio-inspired networking.Computer Networks, 54(6), 881-900.

[5] Guerrieri, F. J., Nehring, V., Jørgensen, C. G., Nielsen, J., Galizia, C. G., and d'Ettorre, P. 2009. Ants recognize foes and not friends. Proceedings of the Royal Society B: Biological Sciences, rspb-2008.

[6] Aboudagga, N., Refaei, M. T., Eltoweissy, M., DaSilva, L. A., and Quisquater, J. 2005. Authentication protocols for ad hoc networks: taxonomy and research issues. In *Proceedings of the 1st ACM international workshop on Quality of service \& security in wireless and mobile networks* (Q2SWinet '05). ACM, New York, NY, USA, 96-104. DOI=10.1145/1089761.1089777 http://doi.acm.org/10.1145/1089761.1089777

[7] Hashmi, S., and Brooke, J. 2008. Authentication mechanisms for mobile ad-hoc networks and resistance to Sybil attack. In Emerging Security Information, Systems and Technologies, 2008. SECURWARE'08. Second International Conference. IEEE, 120-126.

[8] Hashmi S, and Brooke J. 2010.Towards Sybil Resistant Authentication in Mobile Ad Hoc Networks[C]//Emerging Security Information Systems and Technologies (SECURWARE), 2010 Fourth International Conference on. IEEE, 2010: 17-24.

[9] Capkun, S., Buttyán, L., and Hubaux, J. P. 2003. Self-organized public-key management for mobile ad hoc networks. Mobile Computing, IEEE Transactions on, 2(1), 52-64.

[10] Dahshan, H., and Irvine, J. 2010. A robust self-organized public key management for mobile ad hoc networks. Security and Communication Networks, 3(1), 16-30.

[11] Memarmoshrefi, P., Seibel, R.,Hogrefe, D., 2012. Bio-inspired Self-organized Public Key Authentication Mechanism for Mobile Ad-hoc Networks, Bio-Inspired Models of Network, Information, and Computing Systems. In 5th International ICST Conference, BIONETICS 2010, Boston, USA, 375-386

[12] Abbas, S., Merabti, M., Llewellyn-Jones, D., and Kifayat, K. 2013. Lightweight Sybil Attack Detection in MANETs. Systems Journal, IEEE, 7(2), 236-248.

[13] Memarmoshrefi, P., Zhang, H., and Hogrefe, D. 2013. Investigation of a bio-inspired security mechanism in Mobile Ad hoc Networks. In Wireless and Mobile Computing, Networking and Communications (WiMob), 2013 IEEE 9th International Conference. IEEE. 709-716.

[14] Roulston, T. H., Buczkowski, G., and Silverman, J. 2003. Nestmate discrimination in ants: effect of bioassay on aggressive behavior. Insectes Sociaux, 50(2), 151-159.

[15] Tan, P., Steinbach, M., and Kumar, V. 2005, Introduction to Data Mining. Addison-Wesley. ISBN: 0321321367.

[16] QualNet 5.2.0 Programmer's Guide, 2011, Scalable Network Technologies, Inc.

[17] ALGLIB Reference Manual http://www.alglib.net/translator/man/manual.cpp.html

[18] MATLAB 7.6 and Statistics Toolbox 7.6, the MathWorks, Inc., Natick, Massachusetts, United States.