On the Use of Radio Resource Tests in Wireless ad hoc Networks *

Diogo Mónica João Leitão Luís Rodrigues Carlos Ribeiro
INESC-ID/IST
{diogo.monica, joao.c.leitao, ler, carlos.ribeiro}@ist.utl.pt

Abstract

Sybil identities are a major threat to the dependable operation of wireless ad hoc networks. This paper defines a framework to assess the power and performance of radio resource tests (RRT), a technique that allows the detection of Sybil identities. Several RRTs are analyzed and compared using this framework, including two novel RRTs and an optimization of a RRT previously proposed in the literature. Finally, we show how these tests can be used to detect the presence of Sybil identities in an one-hop population, and conclude that different tests are suitable for different network scenarios.

1. Introduction

The Sybil Attack is a relevant threat to the secure and dependable operation of wireless ad hoc networks. The name for this attack was first coined by Douceur in [6]. It consists in having a malicious node simultaneously assuming multiple identities, commonly called sybil identities. Such a node can easily disrupt the operation of distributed protocols, such as distributed storage, routing, data aggregation, voting, intrusion detection, and resource sharing [11, 7]. A radio resource test (RRT) is a technique that allows detection of Sybil identities and, therefore, is a fundamental building block for developing dependable architectures for wireless ad hoc networks.

RRTs are a particular case of the more general class of arbitrary resource tests. Resource tests operate under the assumption that it is possible to establish a bound to the resources available to a single node. Two non-sybil identities must, therefore, be capable of demonstrating that they own more aggregate resources than those available to a single node. Different kinds of resources can be tested, including computational power, storage capacity, and network bandwidth. Different tests have thus been proposed in the past [10, 11, 4, 1]. RRTs assume that each node has access to a single radio device and builds upon the limitations of these devices. RRTs have the potential to support protocols that do not require pre-configuration, nor pre-shared secrets, improving the scalability of the network.

This paper makes the following contributions: *i)* we propose a framework to assess the power and performance of RRTs; *ii)* we propose a number of novel RRTs; *iii)* we make a comparative analysis of different RRTs; *iv)* we discuss how these tests can be used to test a population of identities, and determine the cost of such combined test.

The remainder of this paper is organized as follows: Section 2 introduces the assumptions made in this paper and describes the radio resource tests that will be studied. Section 3 discusses the power and performance of different RRTs. Section 4 addresses the use of RRTs for testing larger populations of identities. Section 5 provides a discussion of adequate scenarios for the use of different RRTs. Section 6 concludes the paper and establishes some directions for future work.

2. Radio Resource Tests

2.1. Assumptions

A radio resource test is typically based on the following assumptions about radio devices [7]: *A1)* Each node in the network owns at most a single radio device; *A2)* Each device can operate over, at most, K different channels; *A3)* No radio device can simultaneously transmit on two different channels; *A4)* No radio device can listen simultaneously on two different channels; *A5)* A node cannot detect a collision while transmitting.

Each particular test may use only a subset of these assumptions. Nodes with more than two radio devices can be modeled as multiple colluding nodes.

Each RRT is characterized by a set of parameters $RRT(h, c, w)$ as follows. Parameter h is the size of the set $S = \{s_1, s_2, ..., s_h\}$ of distinct identities that can be tested simultaneously, in a single test. Parameter c is the number of *challenger* identities (not in S) that need to actively

*This work was partially supported by FCT under grants PTDC/EIA/65588/2006 and PTDC/EIA/71752/2006, and by LEMe in the context of project WiMesh.

participate in the test. Parameter w is the number of *tester* nodes that can extract information from the test.

2.2. Tests Being Analyzed

Simultaneous Sender Test (SST) Newsome el al. [7] proposed a RRT based on having the entities to be tested transmit simultaneously on different channels. In their original paper, the authors only call their test "radio resource test" but, here, we dubbed it Simultaneous Sender Test (SST) to distinguish it from other tests. As originally proposed, SST is a $RRT(K, 1, 1)$, *i.e.*, a test that allows a single node to test simultaneously as many identities as the number of channels available to the radio devices.

The test operates as follows: The challenger assigns a different channel to each identity being tested. Then, these identities start transmitting simultaneously. According to *A3*, sybil identities will be unable to simultaneously transmit on their assigned channels. The challenger can then listen to a channel at random, to verify if the corresponding identity is actually transmitting. If one of the tested identities does not transmit on the assigned channel, it is assumed to be a sybil identity. Since the challenger can only check one channel at a time, the reply transmissions have to be repeated r times, to achieve the desired probability of detection. The required number of rounds r will be derived at a later section.

Optimized Simultaneous Sender Test (oSST) As the name implies, the Optimized Simultaneous Sender Test (oSST) is an optimization of SST. Although this optimization is not referred in any way in [7], we do not list this test as a novel test, since the optimization is incremental (although powerful, as we will see later).

The oSST is based on the fact that the test proposed in [7] can be used as a $RRT(K, 0, N - K)$, where N is the number of nodes in the one-hop neighborhood of the nodes being tested. In fact: *i)* for a set of k identities ($k < N$) to be tested, it is possible to devise a deterministic channel assignment algorithm, thus avoiding the need for an explicit challenger ($c = 0$), and; *ii)* any node in the one-hop neighborhood of the nodes being tested can be a tester, *i.e.* several nodes can detect, at the same time, the existence (or nonexistence) of sybil identities in the set being tested.

Simultaneous Receiver Test (SRT) Both the SST and the oSST have the disadvantage of being very asymmetrical in resource usage: all the tested identities need to transmit during the test. A set of malicious nodes may, therefore, drain the power resources of the network by issuing successive challenges, possibly with distinct sybil identities.

The Simultaneous Receiver Test (SRT) is a novel $RRT(K, 1, 1)$, that we now propose. As with SST, the challenger (whose need may be avoided) assigns a different channel to each of the K tested identities. However, in the SRT test, tested identities have to listen in those channels. The challenger then sends a message in one of these channels, chosen at random. The corresponding identity is then required to echo this message. If one or more of the identities being tested are sybil, accordingly to *A4*, they will be unable to listen in all channels simultaneously, and there is a probability that the message will not be echoed. As before, the challenger may need to perform multiple rounds, to ensure that sybil identities can be reliably detected.

Forced Collision Test (FCT) All the tests described so far require the radio devices to operate in more than one channel (i.e., $K \geq 2$). We now propose a test that can be performed in settings where radio devices are limited to a single channel.

The test is based on assumption $A5$, a known limitation of radio devices [2]. The Forced Collision Test (FCT) is a $RRT(2, 1, 1)$, where one challenger can test two different identities, s_1 and s_2. The test operates as follows: s_1 is required to transmit a message M to s_2. If s_2 receives M, it should retransmit it. During the transmission of M by s_1, the challenger randomly decides to i) cause a collision on the wireless medium, or ii) listen to the medium to verify compliance of s_1. If s_1 and s_2 are different identities, s_2 will be able to retransmit M if there was no collision, and unable to do so otherwise. If s_1 and s_2 are sybil identities, the malicious node that controls them will have to guess if a collision was generated or not. As in all previous tests, the test must have r rounds, in order to be conclusive.

2.3. Other Tests

In [8], a somewhat different kind of test was proposed. Contrarily to the SST, this test is completely passive, in the sense that it does not require the active participation of nodes. Despite this difference, the main assumptions remain the same. The authors proposed a protocol, named PASID-GD, that identifies sybil identities, by comparing the number of expected and observed collisions. However, this kind of passive approach has a practical problem: it allows the normal participation of sybil identities in the network until they are detected. This means that an attacker can continuously generate new sybil identities to take the place of those that are detected, thereby continuously participating in the network with multiple identities. Conversely, the use of active tests, such as the RRTs, allow us to guarantee that no sybil identities participate in the network before being properly tested.

Additionally, a number of techniques that rely on location information to detect sybil identities have been proposed [3, 5, 8, 9]. Such information can be either in-

ferred using radio signal strength indication [5], or by relying on external components to provide such information *e.g.* GPS [9]. While the existence of external location sources is plausible for vehicular networks, they are not typical for ad hoc networks[1]. On the other hand, radio strength based approaches can be easily attacked, by varying the transmission power, leading to inaccurate detection of sybil identities.

3. Analysis

We now discuss the power and performance of each of the four tests presented earlier: SST, oSST, SRT and FCT. This analysis considers the following metrics: vulnerability to collusion, message cost, ratio of resource consumption between legitimate and sybil nodes, and synchronization requirements.

3.1. Vulnerability to Collusion

Collusion happens when two or more malicious nodes coordinate their efforts to protect one or more sybil identities. For instance, some malicious nodes may vouch for the sybil identities of other malicious nodes being tested, making it impossible to identify such identities as being sybil.

Intuitively, one can circumvent colluding nodes by testing simultaneously more identities than the existing number of colluding nodes. This would ensure that, in these tests, all colluding nodes would have to vouch for one of their own identities, and, thus, all remaining sybil identities would eventually be identified as so.

More precisely, to ensure that a radio resource test $RRT(h, c, w)$ operates correctly in environments with at most m colluding identities[2], we must have $h \geq m$. Due to this fact, SST, oSST, and SRT can tolerate as many as $h - 1$ colluding nodes. Notice that these protocols are limited by parameter h, which depends on the total number of radio channels available to nodes. On sharp contrast, because FCT can only be performed on a pair of nodes, this protocol cannot operate correctly in the presence of colluding nodes. A single pair of colluding nodes can vouch for an arbitrary large number of sybil identities.

3.2. Message Cost

We now discuss the message cost of each test. We also derive the number of rounds (r) required to detect sybil identities with a given target probability. Before discussing r, let us look at the cost of a single round for each test (mt). In SST and oSST, each round requires every tested node to send one message ($mt = h$). In SRT, at most two messages ($mt = 2$) are exchanged (one from the challenger and its echo from the tested node). In FCT, two messages are generated ($mt = 2$), one from one of the tested nodes and another from the other tested node (no forced collision) or from the challenger (forced collision case).

The probability of detecting a given sybil identify in S after r rounds of a $RRT(h, c, w)$ test (p_d) is given by:

$$p_d = 1 - \left(1 - \frac{1}{h}\right)^r$$

Solving in order to r, one can calculate the number of rounds required to attain a specific detection probability:

$$r = \frac{log\,(1 - p_d)}{log\left(1 - \frac{1}{h}\right)}$$

3.3. Resource consumption

We now discuss the onus that a $RRT(h, c, w)$ test imposes on legitimate nodes when a malicious node is involved in the test. The malicious node can be the challenger or the owner of one sybil identity being tested. In this context, we define the *resource consumption cost* as the difference Δ between the number of messages sent by correct nodes and messages sent by the malicious node. Notice that RRTs with higher cost levels are more vulnerable to denial of service (DoS) attacks.

In SST, if the challenger is malicious, it sends a single message to initiate the test and then each of the h correct nodes send a message on the assigned channel. Since there are r transmission rounds, the value of Δ is $rh - 1$. If both the malicious node and its sybil identity are in S, than the sybil will not reply to the challenger query and, therefore, $\Delta = rh + 1 - 3r$. If a malicious node is being tested, but its sybil is not in S, then $\Delta = rh + 1 - 2r$.

In SRT, if the malicious node is the challenger, then Δ has a value of 0 (zero), since the challenger has to send a message for each reply. If both a malicious node and its sybil identity are being tested, $\Delta = \frac{2h-3}{h} \cdot r$. If a malicious node is being tested, but its sybil is not, then $\Delta = \frac{2h-2}{h} \cdot r$.

Finally, for FCT, Δ has a value of r, if the malicious node is the challenger, $-0.5 \cdot r$ if both the malicious node and its sybil are being tested, and $0.5 \cdot r$ if a malicious node is being tested but its sybil is not.

Thus, we have that the SST has the worst cost of all the tests, since it is possible for an attacker to consume a large number of network resources, with a low corresponding effort.

3.4. Synchronization Requirements

All RRTs compared in this paper assume that the participants in the test have exclusive access to the medium for

[1]For instance, it is not typical for laptops to be equipped with GPS receivers.

[2]Notice that, in scenarios without colluding nodes: $m = 1$.

the duration of the test. Otherwise, nonparticipating nodes in the test may generate an unbound number of collisions that, in turn, would make the tests inconclusive.

Also, some tests (such as SST and oSST) require nodes to transmit "simultaneously". However, in practice, nodes are not required to have a perfect synchronization; it is enough to ensure that the time to transmit a message is orders of magnitude larger than the allowed amount of desynchronization among nodes (such that a node cannot leverage on the desynchronization to send a message on both channels).

4. Using the RRTs for Population Control

For any $RRT(h, c, w)$, $h \leq K$. Naturally, the number of identities that need to be tested may be greater than h. Therefore, to test a population $\mathcal{P}$ composed of N identities, one has to execute a given $RRT(h, c, w)$ several times. Thus, the final cost of using a given RRT to test an entire population depends on both the cost of each individual test and also of the number of required tests.

It will assumed that all nodes in the system are in radio range (*i.e.* a single hop scenario), and that a common Time Division Multiple Access (TDMA) scheduler exists, to avoid collision and simplify the scheduling of individual tests.

In order to check if there are any sybil identities in $\mathcal{P}$, each node n must test every group of size h ($h < N$) in $\mathcal{P} \setminus \{n\}$. However, detection of a sybil identity can only occur in groups that include all the colluding malicious nodes ($m < h$) and their sybil identities. There are $\mathcal{G}$ such groups, where:

$$\mathcal{G}(N, h, m) = \binom{N - m - 2}{h - m - 1}.$$

Taking this into consideration, the probability that a challenger node detects a particular sybil identity becomes:

$$p_d^+ = 1 - \left(1 - \frac{1}{h}\right)^{r.\mathcal{G}(N,h,m)}. \quad (1)$$

Notice that $p_d^+ \geq p_d$ for the same r. In order to ensure a consistent view of $\mathcal{P}$ by all non-malicious nodes in $\mathcal{P}$, we require each sybil identity to be detected by every non-malicious node. The overall probability of detection of a particular sybil identity is thus given by:

$$p_d^* = (p_d^+)^{N-m-1}. \quad (2)$$

One could avoid requiring all nodes to perform all the tests, since a node could, upon detecting a sybil identity, simply broadcast a warning. That would make the remaining nodes ignore the sybil identity and avoid further testing. Unfortunately, although this approach allows large improvements on the performance of RRTs, one has to consider that it also creates the opportunity for a simple attack against the group membership of correct participants. If a malicious node broadcasts sybil notifications concerning correct participants they will be expelled from the group, which can be used at a later time by the malicious node to attack the system (*e.g.* to attack some majority based protocol being executed in the system).

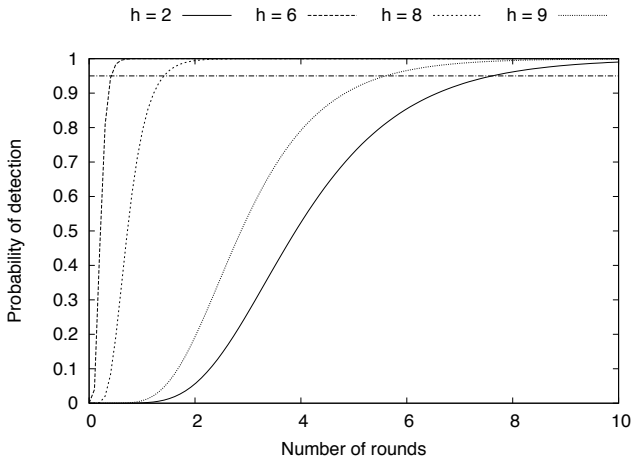

Figure 1. Probability of detection (p_d^*) as a function of the number of rounds (r)

Figure 1 shows the value of p_d^* for a network composed of 10 identities ($N = 10$) for distinct values of r and h. For reference we also represent $P_d = 0.95$ which is the value we will use in following sections[3].

Solving equations (1) and (2) for r allows the specification of the number of rounds required for a desired probability of detection.

$$r = \mathcal{R}(N, h, m, p_d^*) = \left\lceil \frac{log\left(1 - \sqrt[N-m-1]{p_d^*}\right)}{\mathcal{G}(N, h, m).log\left(1 - \frac{1}{h}\right)} \right\rceil$$

4.1. Number of Tests

We now express the total number of tests ($\mathcal{NT}$) as a function of N, h, m and p_d^*.

As previously described, each node will perform r rounds of tests to all possible combinations of h identities of all remaining nodes in the system ($N - 1$). Considering this, one can easily derive function $\mathcal{NT}$ for this protocol as being:

[3] Although we consider $p_d^* \geq 0.95$ as a case study, tests can be configured to any desired target value of p_d^*.

$$\mathcal{NT}(N,h,m,p_d^*) = \mathcal{R}(N,h,m,p_d^*) \cdot \binom{N-1}{h} \cdot N$$

However, considering the optimization of SST (oSST), the total number of tests decreases substantially, since nodes can avoid testing combinations of h identities that have already been monitored when other nodes performed their tests. If, on each test, there are $0 < w \leq N-h$ tester nodes (the challenger node and the passive tester nodes), the function that describes $\mathcal{NT}$ becomes:

$$\begin{aligned}\mathcal{NT}(N,h,m,w,p_d^*) &= \mathcal{R}(N,h,m,p_d^*) \cdot \binom{N-1}{h} \cdot \frac{N}{w} \\ &= \mathcal{R}(N,h,m,p_d^*) \cdot \binom{N}{h} \cdot \frac{N-h}{w}\end{aligned}$$

This equation clearly shows the advantage of oSST, when configured with the maximum allowable value for w (*i.e.* $w = N-h$), in relation to the remaining RRTs (where $w = 1$), for the number of required tests:

$$\mathcal{NT}^{SST/SRT/FCT} = \mathcal{NT}^{oSST} \cdot (N-h)$$

4.2. Total Message Cost

The message cost of a RRT is defined as the total number of messages transmitted to complete the protocol. This metric is closely associated with the energy consumption in the system due to execution of each RRT. The number of messages transmitted by each protocol ($\mathcal{MT}$) can be expressed as the product between the number of tests $\mathcal{NT}$ and the number of messages transmitted on each test (mt):

$$\mathcal{MT}(N,h,m,w,mt,p_d^*) = \mathcal{NT}(N,h,m,w,p_d^*) \cdot mt$$

Comparison Table 1 parameterizes the $\mathcal{MT}$ function for each type of test, given the specificities of their operation. In SST, SRT and FCT the test is only carried by one challenger at a time, $w = 1$. On the other hand, in the optimized version of SST (oSST) every node not being tested is testing the group ($w = N-h$). FCT can only test two nodes at a time ($h = 2$) and, thus, can not handle collusion ($m = 1$).

Figure 2 plots the functions in Table 1 as a function of h (FCT is not plotted because it cannot handle $h \neq 2$), with $N = 20, m = 1$ and $p_d^* = 0.95$. All three plots show the same behavior. The number of transmitted messages is higher for intermediate values of h. Therefore, the number of simultaneously tested identities (h) should be either very low ($h = 2$) or very high ($h = N-1$). However, the choice of h is also dependent on the maximum number of colluding nodes (m) that we want to tolerate, and the number of channels K available ($m \leq h \leq K$). The non-optimized version of SST is always worse than the remaining two, but the optimized version (oSST) is better than SRT for $h < \frac{2N}{3}$ and worse otherwise. Therefore, SRT is better for high collusion scenarios ($m > \frac{2N}{3}$) and oSST is better for scenarios where fewer channels are available.

Figure 3 plots the $\mathcal{MT}$ in table 1 as a function of N with $m = 1$, $p_d^* = 0.95$ and $h = N-1$ for SRT, and $h = 2$ for the others (best case of each of them). FCT is always worse than the other three. As expected, for $h = 2$, the optimized version of SST behaves better than the remaining tests.

Test	Messages transmitted	Test parameters
SST	$\mathcal{MT}(N,h,m,1,h,p_d^*)$	$w = 1, mt = h$
oSST	$\mathcal{MT}(N,h,m,N-h,h,p_d^*)$	$w = N-h$, $mt = h$
SRT	$\mathcal{MT}(N,h,m,1,2,p_d^*)$	$w = 1, mt = 2$
FCT	$\mathcal{MT}(N,2,1,1,2,p_d^*)$	$h = 2$,w = 1, $m = 1, mt = 2$

Table 1. Number of transmissions per test

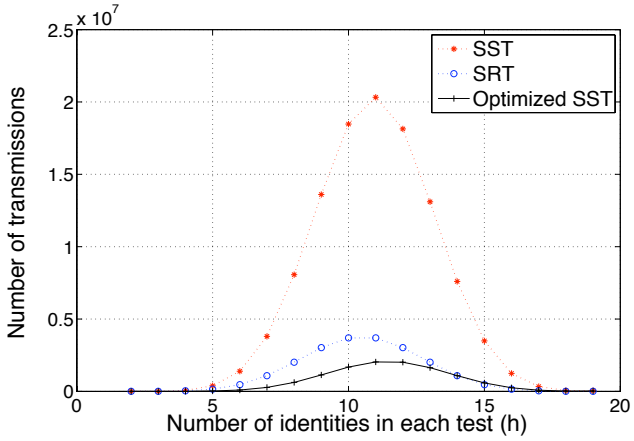

Figure 2. Total number of message transmissions ($\mathcal{MT}$) as a function of the number of simultaneously tested identities (h).

5. Discussion

None of the proposed solutions is better than all the others in every scenario. Table 2 characterizes the scenarios where each solution performs better, when compared with other RRTs. The optimized version of SST (oSST) is most adequate for scenarios with low and medium collusion and where there is no danger of denial of service attacks, because it requires the lowest number of messages of all RRT.

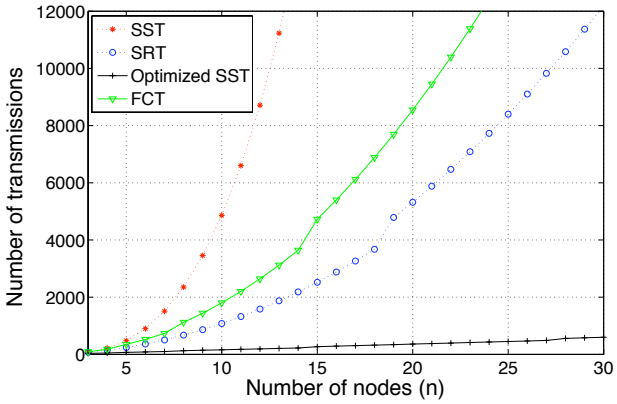

Figure 3. Total number of message transmissions ($\mathcal{MT}$) as a function of the number nodes (N).

Test	Best application context
oSST	Low collusion and no DoS threat
SRT	High collusion and/or DoS threat
FCT	One channel

Table 2. Best application context for each test

SRT is the best one for scenarios with high levels of collusion ($m > \frac{2N}{3}$), or where denial of service attacks need to be taken into account, because it has the lowest resource cost level. Finally, FCT is best suited for scenarios where there is only one channel available, since all the other RRT require the simultaneous use of more than one channel.

6. Conclusions and Future Work

Radio Resource Tests are a viable mechanism for detecting sybil identities in a wireless ad hoc network. In this paper we proposed a framework to compare the power and performance of RRTs. We have also proposed two novel RRTs and an optimization to a RRT previously proposed in the literature. Furthermore, we have analyzed these tests both in isolation and when used to test an one-hop population. We have shown that each radio resource test is best adapted to a different specific scenario, which we described.

As future work, we would like to explore the idea of having more sophisticated algorithms to test an entire population, leveraging on different cooperation algorithms among correct nodes. We also plan to explore the possibilities of extending the knowledge gained with radio resource tests to multi-hop networks.

References

[1] J. Aspnes, C. Jackson, and A. Krishnamurthy. Exposing computationally-challenged Byzantine impostors. Technical Report YALEU/DCS/TR-1332, Yale University Department of Computer Science, July 2005.

[2] R. Bar-Yehuda, O. Goldreich, and A. Itai. Efficient emulation of single-hop radio network with collision detection on multi-hop radio network with no collision detection. *Distrib. Comput.*, 5(2):67–71, 1991.

[3] R. A. Bazzi and G. Konjevod. On the establishment of distinct identities in overlay networks. In *PODC '05: Proceedings of the twenty-fourth annual ACM symposium on Principles of distributed computing*, pages 312–320, New York, NY, USA, 2005. ACM.

[4] M. Castro, P. Druschel, A. Ganesh, A. Rowstron, and D. S. Wallach. Secure routing for structured peer-to-peer overlay networks. *SIGOPS Oper. Syst. Rev.*, 36(SI):299–314, 2002.

[5] M. Demirbas and Y. Song. An rssi-based scheme for sybil attack detection in wireless sensor networks. In *WOWMOM '06: Proceedings of the 2006 International Symposium on on World of Wireless, Mobile and Multimedia Networks*, pages 564–570, Washington, DC, USA, 2006. IEEE Computer Society.

[6] J. R. Douceur and J. S. Donath. The sybil attack. In *Proceedings for the 1st International Workshop on Peer-to-Peer Systems*, pages 251–260, Cambridge, MA, USA, Mar. 2002.

[7] J. Newsome, E. Shi, D. Song, and A. Perrig. The sybil attack in sensor networks: analysis & defenses. In *Information Processing in Sensor Networks, 2004. IPSN 2004. Third International Symposium on*, pages 259–268, 2004.

[8] C. Piro, C. Shields, and B. N. Levine. Detecting the sybil attack in mobile ad hoc networks. *Securecomm and Workshops, 2006*, pages 1–11, 28 2006-Sept. 1 2006.

[9] B. Xiao, B. Yu, and C. Gao. Detection and localization of sybil nodes in vanets. In *DIWANS '06: Proceedings of the 2006 workshop on Dependability issues in wireless ad hoc networks and sensor networks*, pages 1–8, New York, NY, USA, 2006. ACM.

[10] H. Yu, P. Gibbons, M. Kaminsky, and F. Xiao. Sybillimit: A near-optimal social network defense against sybil attacks. In *Security and Privacy, 2008. SP 2008. IEEE Symposium on*, pages 3–17, May 2008.

[11] H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman. Sybilguard: defending against sybil attacks via social networks. *SIGCOMM Comput. Commun. Rev.*, 36(4):267–278, 2006.