Key Exchange for Service Discovery in Secure Content Addressable Sensor Networks

Hans-Joachim Hof, Ingmar Baumgart, and Martina Zitterbart

Institute of Telematics, Universität Karlsruhe (TH), Germany
{hof,baumgart,zit}@tm.uka.de

Abstract. *Secure Content Addressable Network (SCAN)* is an architecture for service discovery in service centric sensor networks that enables dynamic service composition. This paper proposes two new security mechanisms for SCAN: *Single Path Key Exchange (SPX)* and *Multi Path Key Exchange (MPX)*. Both security mechanisms allow two arbitrary nodes of SCAN to exchange a symmetric key for secure communication. We also propose to use replication service information and majority vote to achieve security.
We evaluated the performance and security of Secure Content Addressable Networks with Single Path Key Exchange, Multi Path Key Exchange and replication using a worst case attack model. It has been found, that in a network with 1000 nodes and 5% malicious nodes the probability of a successful lookup operation is still 80%. The results of the simulation indicate, that the overhead and the security level of SCAN with SPX and MPX scale with an increasing number of nodes. The simulation results also show that SCAN is suitable for networks with 100 to 1000 nodes.

1 Introduction

Sensor networks, as we consider them, are resource constrained with respect to memory and computing power of the sensor nodes. Therefore, public key cryptography is not possible because it involves a lot of computing power. We also expect that there is no infrastructure like a public key infrastructure or some powerful server which is available all the time. We focus on scenarios like assisted living, health care, and home automation. In these scenarios, hundreds or thousands of nodes can be in use and the density of nodes in the network is typically high. New services can be dynamically created during the lifetime of the network. To exploit the full power of dynamic service composition, nodes need a way to find available services and to discover the properties of these services (e.g. the address or the position).

Secure Content Addressable Network (SCAN) [1][2][3] is an architecture for secure service discovery in sensor networks that allows for dynamic service composition. In this paper, we propose two new security mechanisms for SCAN: *Single Path Key Exchange (SPX)* and *Multi Path Key Exchange (MPX)*. The feasibility of both mechanisms for service discovery in sensor networks is evaluated by simulation.

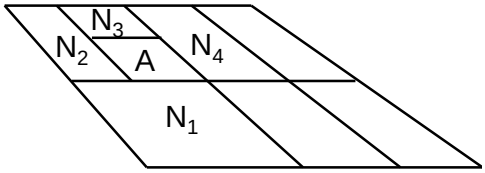


Fig. 1. Example of an unfolded 2-dimensional SCAN space, showing the four neighbor zones (N_1, N_2, N_3, N_4) of a zone A.

2 Secure Content Addressable Network

This section gives an overview of Secure Content Addressable Networks (SCAN). A more detailed description of SCAN can be found in [1], [2], and [3].

Secure Content Addressable Network is based on Content Addressable Network (CAN) [4], which is an overlay network implementing a distributed hash table. SCAN uses a logical virtual d-dimensional coordinate space on a d-torus to store *(service name,service description record)*-pairs. This space is called SCAN space in the rest of this paper. The *service name* is mapped on a coordinate in SCAN space by using a hash function on the service name. The corresponding hash value is interpreted as coordinate in the d-dimensionial coordinate space, e.g. in the case of $d = 2$ the hash value is split into two equal-sized parts x and y and (x, y) is the corresponding coordinate in SCAN space.The *service description record (SDR)* is stored at this coordinate. Service description records hold information about a service, e.g. the network layer address of the service provider. The SCAN space is divided into zones, each owned by one node of SCAN. Hence, if a service description record is stored at a coordinate in SCAN space, the corresponding zone owner stores the service description record. Fig. 1 shows a zone A and its neighbors N_1, N_2, N_3, and N_4 in a 2-dimensional SCAN space.

The secure join operation is used to securely integrate new nodes into a SCAN. During the join operation, the joining node gets assigned a part (zone) of the SCAN space. Each SCAN node maintains a list of network layer addresses of its neighbors in the SCAN space. To avoid that communication between two SCAN neighbors can be attacked, *symmetric keys between SCAN neighbors* are established during the join operation. These symmetric keys can be used to protect the integrity and confidentiality of messages between overlay neighbors hop-by-hop.

SCAN allows for routing in the SCAN space: a SCAN node forwards a message to the SCAN neighbor that is in SCAN space closest to the destination. Nodes that want to retrieve service description records (SDRs) for a specific service compute the hash value of the service name, interpret it as a coordinate in the SCAN space, and send a request message to the calculated coordinates in the SCAN space.

For a d-dimensional SCAN space with N nodes the average routing path length h is:

$$h = \frac{d}{4} N^{\frac{1}{d}} \tag{1}$$

See [4] for details. In section 3 we will show, how the path length h is related to the security of our proposed key exchange protocols.

3 Single Path Key Exchange and Multi Path Key Exchange

When a node stores a (service name, service description record)-pair in SCAN, the message is sent from SCAN node to SCAN node using the routing in SCAN space. Each SCAN node encrypts the service description record (SDR) for its neighbor node. The response is sent in the same way. However, there is no end-to-end encryption between a node A that wants to insert a SDR and a node B that stores that SDR, because A and B have no symmetric key in common. In SCAN, symmetric keys exist only between neighbors, but A and B need not to be neighbors. If A and B have a symmetric key in common (and the network layer address of B is known), it is not necessary to send the service description record using the routing protocol in SCAN space, but it can be sent on network layer. As one hop in the SCAN space can be multiple hops on the network layer, communication on network layer is more efficient than communication in SCAN space. Hence, a key exchange protocol is needed. If a key is exchanged between A and B, updates of service description records are also more efficient. Updates are necessary in SCAN because SCAN requires every service description record to be updated on a regular basis to deal with node failure. If a key exists, the update message can be sent on message layer because integrity and confidentiality of the message can be protected with the key. Otherwise, the update message must be sent in SCAN space.

We propose two key exchange protocols: *Single Path Key Exchange* (SPX) and *Multi Path Key Exchange* (MPX).

3.1 Single Path Key Exchange

If node A uses *Single Path Key Exchange* it creates a symmetric key and sends the key in plain text in the overlay to node B (see Fig. 2). Whenever a node of the overlay forwards a message to one of its neighbors, it uses the corresponding symmetric neighbor key to encrypt the message. This overlay hop-by-hop encryption is possible because symmetric keys between neighbors have been established during the join operation.

For efficiency reasons, the reply of node B is not sent over the overlay but on network layer. Node B sends an encrypted acknowledge. Only after this successful key exchange node A encrypts the SDR and sends it to node B on the

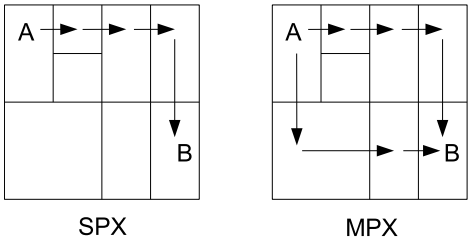


Fig. 2. Single Path Key Exchange (SPX) and Multi Path Key Exchange (MPX)

network layer. Node B uses the exchanged key to decrypt the message and stores the Service Description Record.

Using SPX results in only marginal communication overhead, because only the key is sent over the overlay. The SDR, which is usually larger than a key, is sent directly on the network layer. Hence, this procedure produces less overhead compared to the original SCAN insert operation. Nodes may store the exchanged keys for later updates of the SDRs. As SDRs are stored soft-state, regular updates are necessary. If a key exists, it is no longer necessary to use the overlay for secure communication but the more efficient communication on network layer can be used.

During the key exchange each node on the overlay path between node A and node B can read the symmetric key. Thus these nodes will be able to perform a man-in-the-middle attack on the network layer to tamper with the SDR. This manipulation can not be detected. To accomplish this attack, it is necessary, that the eavesdropper is on the overlay path between node A and node B.

The probability that an overlay path is free of malicous nodes, if the average path length is h and a fraction of m of all nodes in the network are malicious, is $(1-m)^h$.

Thus the probability p of a successful key exchange in a d-dimensional SCAN with N nodes using equation (1) is:

$$p = (1-m)^{\frac{d}{4}N^{\frac{1}{d}}} \tag{2}$$

Single Path Key Exchange can also be used when a node wants to retrieve service descritption records of a service: SPX is executed with the coordinate that is calculated by using a hash function on the service name. The node which stores the service desctiption record then sends back all matching service description records encrypted with the exchanged key on network layer. In this case, SPX is of great use because the list of service description records is much larger than the key, and with SPX, this list need not be transfered in SCAN space but on the more efficient network layer.

3.2 Multi Path Key Exchange

To increase the probability of a successful insert operation in presence of malicious nodes, *Multi Path Key Exchange* uses during the key exchange different paths in the overlay and sends only parts of the key along each overlay path (see also fig. 2). A similar approach has been proposed by [8] for communication on the network layer, but we are, to our best knowledge, the first to apply the idea to a structured overlay network. If node A uses MPX, it creates a symmetric key, splits it into n parts and sends each part in plain text along one of n paths. A very simple approach to split a key into parts is to randomly choose $n-1$ key parts and calculate the last key part by using the XOR-operation ($\oplus$) on the preceding $n-1$ key parts and the key itself:

$$part_n = part_1 \oplus part_2 \oplus ... \oplus part_{n-1} \oplus key$$

We use the fact that in a SCAN with dimension d exist with high probability d nearly optimal distinct paths between two arbitrary nodes because of the structure of the SCAN space. Node B reconstructs the key and sends an encrypted acknowledge message back to node A. In the example above, the key is reconstructed by simply using the XOR-operation on all received keys:

$$key = part_1 \oplus part_2 \oplus ... \oplus part_{n-1} \oplus part_n$$

Node A then encrypts the SDR and sends it to node B on the network layer. To manipulate a SDR without getting noticed, the cooperating attackers must be on each overlay path between node A that uses the insert operation and node B that will store the SDR. More advanced secret sharing schemes [9] can be used so that the key can be reconstructed with only k out of n key parts. These schemes avoid that a single attacker on one overlay path between A and B can hamper the key exchange by inserting fake key parts for a denial-of-service attack. The probability p of a successful key exchange with MPX in a d-dimensional SCAN with N nodes is:

$$p = \sum_{i=k}^{n} \binom{n}{i} ((1-m)^{\frac{d}{4}N^{\frac{1}{d}}})^{i} (1-(1-m)^{\frac{d}{4}N^{\frac{1}{d}}})^{n-i} \tag{3}$$

MPX can also be used to authenticate if a node is at a certain coordinate in the SCAN space: only the node at the destination coordinates of MPX in SCAN space can legitimately receive all the key parts which are sent by MPX. Hence, the knowledge of the exchanged key proves, that the node is really at that coordinate in SCAN space. Thus, if MPX is used for the insert operation, attackers can only claim to be at a fake coordinate in SCAN space if they are on each overlay path and if they cooperate. Hence, a node which uses the insert operation to store a SDR can after MPX be sure with high probability, that it is talking to the node that is expected to store the SDR.

Multi Path Key Exchange can also be used when retrieving service description records. The mechanism is the same as with Single Path Key Exchange (see above).

4 Security by replication of service information

Redundancy can be used to secure the integrity of service description records (SDRs). If a node stores an SDR at different locations of the SCAN space and nodes use multiple of these locations to retrieve service description records, an attacker needs to attack the majority of the SDRs to achieve a high probability of success. The coordinates, at which an SDR is stored, are determined using a hash function. The hash value of the service name is interpreted as a coordinate in SCAN space and the SDR is stored at the node that owns the corresponding zone. One way to store an SDR on multiple nodes is to use multiple hash functions in the computation of the location. If a hash algorithm h is given, several hash functions $(h_1, h_2, h_3, \ldots)$ can be constructed by simply concatenating $(|)$ a number to the string that will be hashed:

$$h_1(value) = h(value|1), h_2(value) = h(value|2), \ldots$$

5 Simulation

To evaluate the feasibility of using an overlay for service discovery in sensor networks and to evaluate the performance of the overlay, we implemented SCAN in the network simulator GloMoSim [10].

5.1 Simulation Settings

The following simulation settings are used for the simulation experiments: the simulation area is 500x500 meters. In this area, 100, 250, and 1000 nodes are placed randomly. The 802.11 MAC layer protocol of GloMoSim is used with a communication range of 158m. These simulation settings are similar to the settings used in other papers. The nodes join the network in constant intervals (120s). Hence, the total simulation time is $number_of_nodes * 120$s.

To simulate the service centric sensor network, each node randomly chooses how many services it will offer (0 to 7 services) and how many services it will use (0 to 10 services) before the node joins the network. A total of 100 different services is present in the network. After joining the network, a node registers all the services that it offers. Later, it searches for the services that it needs. Every 30 minutes of simulation time, a node will re-register its Service Description Records and it will call the lookup operation again for each service it uses. Every second of simulation time, nodes fail with a certain probability. With the same probability, a node searches again for a service that it uses.

The simulation uses static routing with a predefined loss rate. This was the only efficient way to simulate sensor networks with a huge number (> 1000) of nodes in GloMoSim as it turned out that the implementations of routing protocols for GloMoSim do not scale well with the number of nodes. However, this does not affect the conclusion about the simulation results because the underlying routing protocol does not have a high impact on the overlay as long

as the network does not get partitioned. We expect a huge density of sensors, so it is very unlikely that the network gets partitioned. For each combination of parameters several runs with different seeds were done.

5.2 Attack Model

The attack model states "worst case" attackers: it is assumed that all malicious nodes of the network cooperate and that nodes get compromised after being deployed. The probability of a node to get compromised is identical for all nodes of the network. Hence, over time more and more nodes get malicious and the attackers get more powerful. The attackers do not show any suspicious behaviour to their neighbors and they are conforming to the protocol most of the time. So, attackers can not be detected until they start the attack. The simulation marks every message as compromised that passes a malicious overlay node. The number of malicious nodes on the communication paths between two nodes is used to determine success or failure of any operation.

6 Results

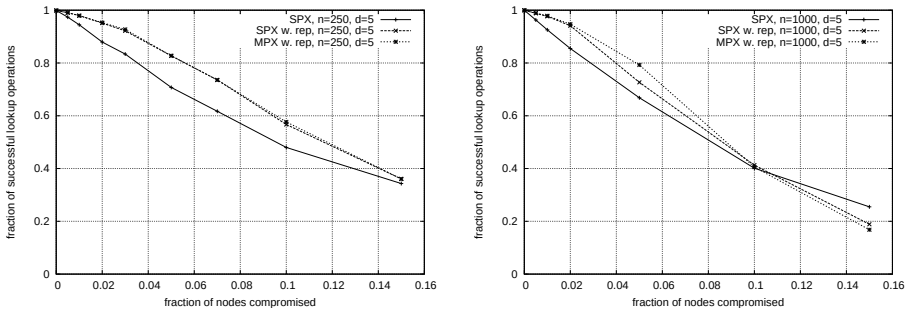
This section presents some simulation results of the GloMoSim implementation of SCAN focusing on security and communication overhead.

6.1 Successful lookup operations

If a single call of the lookup operation returns the Service Description Records of a specific service, the lookup operation is successful. The lookup success is evaluated separately for each node that stores a SDR. If, for example, ten nodes store one Service Description Record each and the lookup operation retrieves only eight of these SDRs, the probability of a successful lookup is 80%. In many scenarios, only one of a number of similar services is really needed, hence the lookup operation would be considered successful by the user, if at least one Service Description Record is retrieved. In such scenarios, the proposed architecture does perform significantly better than presented here. However, we decided to use the more strict definition of lookup success as described above, so the results presented in this paper should be viewed as "worst case" results.

Fig. 3(a) shows the probability of a successful lookup in a network with 250 nodes and a SCAN dimension of 5. For the secure insert operation, we used the methods described in sections 3 and 4: Single Path Key Exchange (SPX), Single Path Key Exchange with replication of SDRs (SPX+Rep), and Multi Path Key Exchange with replication of SDRs (MPX+Rep). For the secure lookup operation, we used SPX without replication of SDRs if no replicates were used by the insert operation; otherwise, we use SPX with replication of SDRs.

SPX with replication of SDRs and MPX with replication of SDRs offer a higher lookup probability than SPX without replication of SDRs. MPX with replication of SDRs has nearly the same probability for a successful lookup than



(a) $n = 250$ nodes (b) $n = 1000$ nodes

Fig. 3. Successful lookup operations in a network with (a) 250 and (b) 1000 nodes, different key exchange methods (SPX, MPX), and with and without replication of Service Description Records.

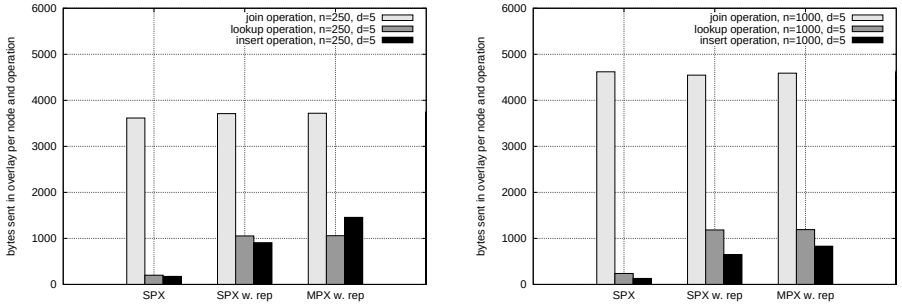
SPX. For example, if 5% of all nodes (=13 nodes) are malicious, we could still retrieve 83% of all Service Description Records. Similar probabilities were found in a network with 1000 nodes: with 5% malicious nodes (=50 nodes) it is still possible to retrieve about 80% of all Service Description Records (see Fig. 3(b)). In the simulation with 1000 nodes, MPX with replication of Service Description Records produces a better possibility of a successful lookup, unless the fraction of malicious nodes is higher than 10%. In this case the majority of involved lookup paths is malicious, thus the majority vote fails. Because MPX without replication performs similarly to SPX with replication, we omit the MPX results in the plots.

Fig. 3 shows that the use of replicates of Service Description Records significantly enhances lookup probability whereas the use of MPX+Rep has an impact only in the simulation with 1000 nodes. The reason for this is the small number distinct paths in small networks.

The simulation results concerning the probability of a successful lookup show that it is possible to retrieve a reasonable number of SDRs even if a moderate fraction of nodes is malicious. In sensor networks, it is often only needed to find only one out of many identical services. Here, SCANs are an ideal solution. The results also show, that SPX with replication of the SDRs should be used by the secure insert operation and the secure lookup operation. MPX should be used only in networks with many nodes. However, MPX offers authentication of a node's coordination in SCAN. SPX does not offer this feature.

6.2 Overhead

In the following, we concentrate on the communication overhead on the overlay layer. Fig. 4 shows this overhead per operation (join, lookup, and insert) and node.



(a) $n = 250$ nodes (b) $n = 1000$ nodes

Fig. 4. Overhead on overlay layer of the join operation, lookup operation, and insert operation with a SCAN dimension of $d = 5$ in a network with (a) 250 and (b) 1000 nodes.

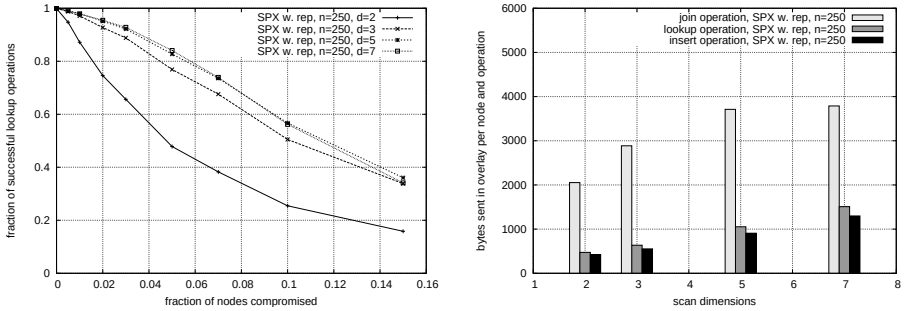
It is clear that the use of replicates multiplies the communication overhead of the lookup operation and of the insert operation because the operations exchange keys for all destinations. MPX with replication produces about a third more traffic than SPX with replication. If we compare Fig. 4(a) and Fig. 4(b) we see that the overall communication overhead only slightly increases with a higher number of nodes.

The simulation results concerning the communication overhead of SCANs show, that the join operation is costly whereas the insert and lookup operation have moderate costs. Although the join operation looks expensive compared to the lookup and insert operation, the join communication costs are incurred only once in the lifetime of a sensor node. In contrast each node performs several hundred lookup and insert operations. Consequently the join operation poses only a small fraction of the total communication overhead.

Because SCAN is an overlay, one hop in SCAN space typically involves multiple hops on the network layer. In our simulation, one overlay hop involved on average 2-3 underlay hops. Thus, to get the total communication overhead on network layer the costs shown in Fig. 4 have to be multiplied by this factor. The resulting costs for insert and lookup operations look very promising for networks with up to 1000 nodes.

6.3 Influence of system parameter

The dimension d of the SCAN space is a parameter of SCAN. It can be chosen freely. However, if we increase d, we also increase the memory usage of every node, because more dimensions result in more neighbors and thus larger neighbor tables must be stored. The neighbor tables store information about the neighbors, e.g. the network layer address and a symmetric key for each neighbor. In SCAN, each node has in average $2d$ neighbors. The advantage of an increased d is, that the average path length between two arbitrary nodes in the overlay



(a) Successful lookup operations. (b) Overhead.

Fig. 5. Influence of SCAN dimension d on successful lookup operations (a) and overhead (b).

decreases, because each node has a higher connection degree (more neighbors). Thus, the probability to run across a malicious node on this path is reduced. As one hop in the overlay usually corresponds to multiple hops on network layer, communication overhead can also be significantly reduced by a higher dimension d. This makes a trade-off between memory and communication overhead possible with SCAN.

We studied how a change of d affects the probability of a successful lookup due to shorter overlay paths and a larger number of overlay neighbors. Fig. 5(a) shows how d affects successful lookup operations in a network with 250 nodes and SPX with replication of Service Description Records used by the insert operation and the lookup operation. We use d replicates of the SDRs. A dimension of $d = 5$ seems to be the ideal choice for the given scenarios (100 to 1000 nodes) as a dimension d of seven does not increase the probability of a successful lookup, because there are not enough nodes in the network to further increase the number of SCAN neighbors. However, this statement may not hold for other numbers of nodes. Fig. 5(b) shows how costly an increase of d is in the same scenario. The increase in replicates causes an increase in communication overhead.

7 Related Work

Several protocols and architectures for service discovery exist. Popular architectures and protocols in infrastructure based networks include e.g. the Service Location Protocol [11] or the Secure Service Provision Protocol [15].

Architectures that use an infrastructure of any kind (e.g., a central server or a public key infrastructure) are not suitable for sensor networks as we see them (see Introduction). The security concepts of service discovery protocols like the Secure Service Provision Protocol, the Secure Service Discovery Protocol [13], or the Secure Service Discovery Protocol [14] are based either on public key

cryptography or preshared secrets (passwords). Both concepts are not suitable for the sensor networks we consider: public key cryptography is at the moment computationally too expensive and preshared secrets are difficult to setup and do not scale well.

There are several methods for key exchange: Diffie-Hellman [16] is computationally too complex for the sensor nodes that we consider. Promising key exchange methods for sensor networks are random-key predistribution protocols, e.g., [17]. Random-key predistribution protocols assign each sensor a random subset of keys out of a very large reservoir of keys. If two nodes want to communicate and they have a key in common, they can use this key. Otherwise, neighbors are used to construct a key. This idea is extended in [8] by using multiple redundant paths to increase the security of the exchanged keys.

8 Conclusion

This paper presented two new security mechanisms for Secure Content Addressable Networks: Single Path Key Exchange (SPX) and Multi Path Key Exchange (MPX). Both mechanisms allow for secure insert and lookup of Service Description Records and both mechanisms allow for more efficient subsequent updates of Service Description Records. MPX also allows for authentication of the node which stores the Service Description Records.

The paper also presented a simulation of Secure Content Addressable Networks with SPX and MPX. The results show that Secure Content Addressable Networks with SPX and MPX provide a reasonable level of security for service centric sensor networks. If, for example, in a network with 1000 nodes 5% of all nodes are malicious, 80% of all lookups are still successful. The results indicate, that Secure Content Addressable Networks with SPX and MPX scale with an increasing number of nodes concerning security level and overhead. The achieved security level can easily be adapted by carefully choosing the dimension d of Secure Content Addressable Networks and the number of replicates at the cost of an increased communication overhead. The simulation results show that Secure Content Addressable Networks are suitable for networks with 100 to 1000 nodes.

References

1. H.-J. Hof, E.-O. Blass, T. Fuhrmann, and M. Zitterbart, "Design of a secure distributed service directory for wireless sensor networks," First European Workshop on Wireless Sensor Networks, Berlin, Germany, Jan. 2004.
2. H.-J. Hof, E.-O. Blass, and M. Zitterbart, "Secure overlay for service centric wireless sensor networks," First European Workshop on Security in Ad-Hoc and Sensor Networks (ESAS 2004), Heidelberg, Germany, Aug. 2004.
3. H.-J. Hof and M. Zitterbart, "SCAN: A secure service directory for service-centric wireless sensor networks," *Computer Communications*, July 2005.
4. S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker, "A scalable content-addressable network," ACM SIGCOMM 2001, San Diego, California, USA, Aug. 2001.

5. J. R. Douceur, "The sybil attack," in *IPTPS '01: Revised Papers from the First International Workshop on Peer-to-Peer Systems*. London, UK: Springer-Verlag, 2002.
6. F. Stajano and R. J. Anderson, "The resurrecting duckling: Security issues for ad-hoc wireless networks," in *Proceedings of the 7th International Workshop on Security Protocols*. London, UK: Springer-Verlag, 2000, pp. 172–194.
7. D. Balfanz, D. K. Smetters, P. Stewart, and H. C. Wong, "Talking to strangers: Authentication in ad-hoc wireless networks," Symposium on Network and Distributed Systems Security (NDSS'02), San Diego, California, USA, Feb. 2002.
8. H. Chan, A. Perrig, and D. Song, "Random key predistribution schemes for sensor networks," 2003 IEEE Symposium on Security and Privacy, Oakland, California, USA, May 2003.
9. A. Shamir, "How to share a secret," *Commun. ACM*, vol. 22, no. 11, 1979.
10. X. Zeng, R. Bagrodia, and M. Gerla, "Glomosim: A library for parallel simulation of large-scale wireless networks," Workshop on Parallel and Distributed Simulation, Banff, Alberta, Canada, 1998.
11. E. Guttman, C. Perkins, J. Veizades, and M. Day, "Service Location Protocol, Version 2 ," RFC 2608 (Proposed Standard), June 1999, updated by RFC 3224. [Online]. Available: http://www.ietf.org/rfc/rfc2608.txt
12. S. Czerwinski, B. Zhao, T. Hodes, A. D. Joseph, and R. H. Katz, "A secure service discovery service," ACM/IEEE International Conference on Mobile Computing and Networks (Mobicom 1999), Seattle, Washington, USA, Aug. 1999.
13. F. Almenárez and C. Campo, "Spdp: A secure service discovery protocol for ad-hoc networks," 9th Open European Summer School and IFIP Workshop on Next Generation Networks, Balatonfured, Hungary, Sept. 2003.
14. Y. Yuan and A. William, "A secure service discovery protocol for manet," 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2003), Beijing, China, Sept. 2003.
15. R. Handorean and G.-C. Roman, "Secure service provision in ad hoc networks," First International Conference on Service-Oriented Computing, Trento, Italy, Dec. 2003.
16. W. Diffie and M. E. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.
17. L. Eschenauer and V. D. Gligor, "A key-management scheme for distributed sensor networks," Proceedings of the 9th ACM conference on Computer and communications security, Washington, DC, USA, 2002.