A Low-cost Anonymous Routing Protocol in MANETs

Lianyu Zhao and Haiying Shen
Department of Computer Science and Computer Engineering
University of Arkansas, Fayetteville, AR 72701
Email: {lxz014, hshen}@uark.edu

***Abstract*—With the wide use of mobile devices in mobile ad hoc networks, maintaining anonymity is becoming an increasingly important issue. Existing routing algorithms either rely on hop-by-hop encryption or local broadcasting for anonymous routing, which lead to high overhead. We propose a low-cost anonymous routing algorithm, which can provide both source/destination and routing anonymity. It dynamically divides the network into hierarchical zones to obscure routing path and randomly chooses nodes as intermediate relay nodes. Therefore it splits routing path to multiple steps that contains no specific routing pattern. Furthermore, we present analysis of the ability of our algorithm to withstand certain attacks. Experiment results show that our algorithm trades some routing efficiency for anonymity, but gains better delivery rate than GPSR under short transmission range.**

I. Introduction

As the fast development of mobile ad hoc networks (MANETs), anonymity in MANETs becomes imperative. Anonymous routing hides the identifiers of data providers, requesters or routing path. Current anonymous routing methods generally can be classified into three categories: hop-by-hop encryption which uses asymmetric key or symmetric key to ensure anonymity, but leads to high computing time; local broadcasting, which is also performed at each hop to hide the routing path or source/destination, it consumes much extra hops; anonymity zone, which is similar to local broadcasting, but it is performed in destination to maintain the anonymity of destination.

In this paper, we propose a low-cost anonymous routing protocol for MANETs, which provides source node, destination node as well as routing anonymity. Compared to other existing anonymous routing approaches that use hop-by-hop encryption, the proposed method costs less computing energy and time because of the greatly reduced encryption/decryption needs. In addition, the proposed protocol reduces the cost due to broadcasting. The approach uses geographic routing algorithm in every step of our routing process. Moreover, the approach dynamically generates hierarchical zone and randomly chooses a node within a zone as a relay node to provide the anonymity. The routing protocol provides near-optimal routing efficiency while offers anonymity protection to the data providers, requesters and routing path.

The remainder of this paper is organized as below. In Section II, we describe related anonymous routing approaches in MANETs. In Section III, we present the design of the proposed routing protocol. The analysis of the protocol is presented in Section IV. Performance of the protocol is evaluated in Section V. Conclusion and future work are given in Section VI.

II. Related Work

Anonymous routing algorithms in MANETs has been studied in recently years. In the following, we discuss the anonymous routing algorithms in two main categories of routing algorithms in MANETs: virtual topology based routing and geographic (Position) based routing.

Virtual topology based routing. In this category, there are some approaches that utilize hop-by-hop encryption to provide anonymity [1]–[4]. ANODR [1] is the first routing protocol that requires no identification of nodes. It incorporates a symmetric encryption called trapdoor boomerang onion (TBO) instead of high-cost public key encryption, as well as local broadcast to provide anonymity. In MASK [2], neighborhood authentication is used in routing path discovery. It performs routing tasks by utilizing identifiers and keys established during the authentication phase. SEAD [3] is a secure routing protocol based on DSDV [5]. It utilizes inexpensive cryptography in the protocol to obtain both efficiency and resistance to attacks. Discount-ANODR [4] is built using the same techniques in ANODR. It compromises slightly on anonymity guarantee, thus is able to maintain lower computation and communication complexities.

There are some approaches incorporate local broadcasting [6], [7]. Ariadne [6] is based on DSR protocol. It is an on-demand and dynamic routing algorithm. Rather than directly applying cryptography to an existing protocol to achieve security, it edited DSR [8] protocol message to meet the needs of efficiency, thus is applicable in a variety of routing protocols. Aad *et al.* [7] combines onion routing with multicast to thwart attackers.

MAPCP [9], unlike previous methods, is a middleware lies between application layer and network layer. It uses controlled and probabilistic broadcasting to provide anonymity while avoids the use of step-by-step encryption. In addition, it utilizes multiple path in routing to provide a higher degree of anonymity.

Geographic (Position) based routing. Though geographic based routing avoids the overhead of virtual topology maintenance, its exposure of location information is an obstacle to achieving anonymity. Local broadcasting is used in

some geographic based protocols [10], [11] for anonymity. In AO2P [10], node position instead of node identification is used for routing. However, AO2P still reveals destination information while protecting the communication anonymity. R-AO2P [10] is further proposed in order to improve destination privacy without significantly downgrading the performance. ASR [11] relies on both hop-by-hop encryption/decryption and local broadcasting to ensure both security and anonymity.

Hop-by-hop encryption is usually used to preserve anonymity in geographic routing. AODPR [12] encrypts the position of destination and uses the encrypted position in routing, thus can effectively control the leak of position information to nodes that do not belong to the network. Zhi et al [13] proposed a secure routing algorithm that uses GPSR-like greedy forwarding and anonymous location service without compromising the efficiency of geographic routing. It decouples location information and identity to provide location privacy. However, it does not provide routing anonymity. Secure vehicular ad-hoc networks (VANET) service in [14] is also based on GPSR routing algorithm. It authenticates the locations of anonymous nodes in order to provide location authentication and location privacy. It is claimed that the service maintains high delivery rate even when a big proportion of nodes are malicious.

Some routing protocols use geographic zone to provide privacy. Mix zones [15] does not reveal the positions of mobile users in order to keep users' movement from being traced by attackers. Each user has a pseudonym and his/her real identity is not traceable by applications, whenever a user enters a zone, its pseudonym changes. ZAP [16] uses fuzzy position in routing to prevent malicious nodes from obtaining the real position of a node. It uses a concept called anonymity zone, in which there are a number of nodes to obscure the destination. Though the zone concept is also a key feature in our proposed protocol, zone in the protocol is quite different from the anonymity zone in ZAP. Firstly, our zone is hierarchically divided while ZAP's zone does not have hierarchy. Moreover, the intention to use zone in our method is to obscure the routing path and the identifier of the destination while ZAP's goal is only to hide the destination.

III. Low-cost Anonymous Routing Protocol

A. Neighbor Maintenance

In a MANET, each node periodically sends "hello" message to its neighbors about its position. To maintain one's own anonymity, each node only sends the message with a pseudonym and its current location. Each node maintains a routing table which keep its neighbors' pseudonyms associated to their locations. In order to protect its own identifier, each node changes its pseudonym periodically. If a node does not receive updated information from one of its neighbors after a certain period of time, it simply discards the information of the neighbor in its routing table. Due to the changing pseudonyms, a node cannot correlate tuples of the same node in its routing table. Therefore, a node always searches from the most recently recorded neighbors to find a proper node for routing. Two nodes may have the same pseudonyms if they merely hash their positions for the pseudonyms. In order to avoid pseudonym collision, previous work [10] uses time and position for hashing. However, the minor error of locating system might still result in two nodes at the same location at the same time. In order to completely overcome collision, we use consistent hash function such as MD5 [17] to hash the concatenation of a node's MAC address, time, and position.

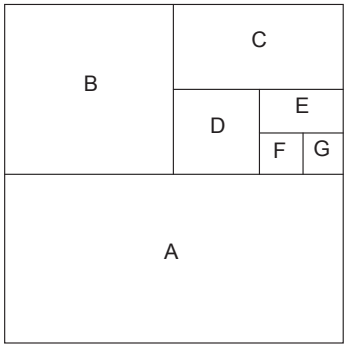

Figure 1. Zone generation.

B. Routing Algorithm

In node communication, a source node sends a request message (RREQ) to a destination and the destination responds with data message (DATA). We temporarily assume a node knows the location of destination, and will discuss the location service in III-C without the assumption.

Figure 2 shows the complete process of the routing algorithm. A message is routed from the message *source* to the destination through a number of *random-forwarders*. The message is routed from the source to a random-forwarder, between random-forwarders and from a random-forwarder to the destination through a number of *relay-nodes* using GPSR-like [18] greedy forwarding.

Before we explain the details of the routing algorithm, let us introduce the concept of zones in a MANET. Without the loss of generality, we assume the network area is a rectangle, in which nodes are randomly disseminated. The area can be divided in a hierarchical manner. Figure 1 shows hierarchically divided zones. The entire network area is firstly partitioned into zones A and B, then zone B is further divided into B and C, and so on. A zone generated after the entire field has been divided n times is called an n^{th} partitioned zone. For instance, zone A is a 1^{st} partitioned zone, and zone C is a 2^{nd} partitioned zone.

In a nutshell, in the routing protocol, a message source (source and random-forwarder) dynamically divides its zone until itself and message receiver (random-forwarder and destination) are in different zones, and sends the message to a randomly chosen random-forwarder in the other zone. The last random-forwarder broadcasts the message to all the nodes in the destination's zone. The protocol aims to achieve k-anonymity [19], where k is a pre-defined integer. That is, a

message is broadcasted to k nodes in order to hide the identifier of the destination. We call the zone having k nodes where the destination resides in as *destination zone*. For example, the shaded zone in Figure 2 is destination zone. Based on k and node density, the source can calculate the upper-left and bottom-right coordinates of the destination zone, and the number of divisions to generate the zone, denoted by n_{max}.

When a source sends a message, it divides the entire field to two zones. It then checks whether itself and destination are in the same zone. In this case, the source further divides the zone they both reside in. This process is repeated until the source and the destination are not in the same zone or the number of divisions reaches n_{max}. In the former case, the source randomly chooses a location in the other zone. It relies on GPSR-like [18] greedy forwarding to send the message to a random-forwarder near the randomly chosen location through a number of relay-node. Upon receiving the message, like the message source, the random-forwarder repeats the same process. In the latter case, the message arrives at the destination zone. The receiver broadcasts the message to all nodes in the zone.

A RREQ message contains a symmetric key which will be used for encrypting the data sent from the destination to the source. In order to protect this key private, the source encrypts it using the destination's public key. We use d_ul to represent the upper-left coordination, and use d_br to represent the bottom-right coordination of the destination zone. Specifically, a source sends RREQ in the form of

$$< RREQ, s\_pd, d\_ul, d\_br, n, rf, n_{max}, E_{K_{pub}}(K_{sym}) >,$$

where s_pd is the pseudonym of source, n is the number of partitions, rf is the next random-forwarder's coordinate, K_{pub} denotes the destination's public key, and $E_{K_{pub}}(K_{sym})$ is the encrypted result of the symmetric key denoted by K_{sym}.

After a source randomly chooses a location, using GPSR-like [18] greedy forwarding, it looks up in its routing table for the relay-node that can most greatly reduce the distance to the chosen location, and sends this message to the neighboring node. Then, the relay-node conducts the same operation until reaching a node, which cannot find a neighbor closer to the location. In this step, the routing proposal does not depend on the right-hand perimeter method in GPSR, which route a message around the location to find the nearest node. This is because the routing protocol does not need to find the node precisely closest to the chosen location due to its randomization feature.

For example, in Figure 2, source node S first divides the entire zone to zone A and zone B, and randomly chooses a location in zone B. S then sends RREQ to RF_1 near the location as the next random-forwarder. RF_1 divides zone B to zone B and C, and randomly chooses location in zone C. The RREQ from RF_1 is routed to RF_2 which is close to the location. Using the same process, the RREQ is routed from RF_2 to RF_3, and then to RF_4. When RF_4 finds that the number of zone partition reaches n_{max}, it locally broadcasts the RREQ to the zone it resides.

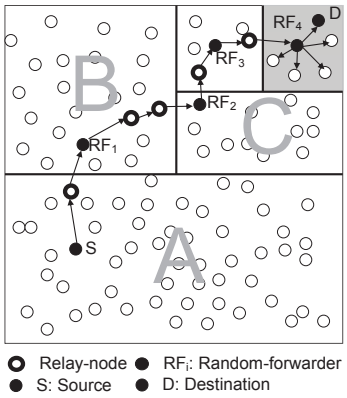

Figure 2. Routing among zones

Data transmission from the destination to the source is similar to RREQ transmission except the data is encrypted by K_{sym} sent along with RREQ rather than K_{pub}.

In the routing protocol, each random-forwarder acts as a temporary destination, and it has no relationship to the final destination. Thus, it is difficult for an intruder to find routing nodes and the path. Also, because this random relay choosing policy, the statistical pattern of transmission could not be observed.

C. Location Service

In our approach, Location service is necessary when destination's position is not available, because it is difficult to know the location merely depending on the destination node's ID. The provided position contains the same boundary description of destination as in the RREQ message. Also this query result must be encrypted using a key only known to nodes within the network. This rule can avoid malicious attempts to correlate a node ID given by an attacker to its location. We incorporate a scheme similar to DISPOSER [20] which is used in [10], [12], it is a distributed position service in which the whole region is divided into grids. In our approach, a number of trusted nodes play the role of position servers. Every node is mapped to a server, and follows a hash function known to every node within the network. Position servers could handle malicious location request by filtering repetitive requests that contains no actual valid connection proof.

The location servers in our approach only provide a whole zone's location which is calculated from the true location. Therefore, the proposed protocol does not need to prevent nodes from position abuse by continuously sending position requests of a target node, because the sever sends imprecise location information which is of no use to a malicious node. In dynamic environment of MANETs, nodes have to periodically report their changes of locations and pseudonyms. In our method, the interval is determined according to each node's current moving speed, the faster a node moves, the more frequent it reports to the server. In addition, a node needs to report its updated location information together with its updated pseudonym to the servers it is mapped to.

IV. Property Discussion

In this section, we discuss the property of the proposed routing protocol to deal with certain attacks that could be issued by a malicious individual or party. Our discussion is based on the categories of attacks studied in [21].

A. Anonymity

a) Routing Anonymity.: Routing anonymity is to ensure that any node cannot identify any part of the routing path. Our routing protocol dynamically keeps splitting zones into smaller ones in order to enable a message to approach the destination. It also maintains the randomization feature of the routing path by randomly choosing random-forwarders for routing. Therefore, the path of data transmission is not fixed. Malicious attackers that try to monitor data transmission cannot find nodes responsible for routing, because every node in a zone has the chance to route data. Therefore, even when two nodes always transmit data, attackers still cannot find the routing path.

b) Source Anonymity.: Source anonymity is to hide source node from any other node within the network. In our approach, every source uses pseudonym as its identity, which is a hashed value and this value changes as time and its position change periodically. The length of the period that a node's pseudonym stays the same is related to anonymity, because the longer this pseudonym remains the same, the higher possibility that the node may be recognized. In addition, the source anonymity is ensured because the source does not embed its precise position in a message, but only the zone where source resides. Therefore, if any node in the network intercepts this message, it cannot tell the position of the source.

c) Destination Anonymity.: Destination anonymity is to ensure that destination is not known to any other nodes. From the packet formation listed above, we know the destination is not encrypted. Rather, it is a vague location which is specified as an zone. Since there are k nodes in this area, the routing protocol with broadcasting at the last step achieves k-anonymity of the destination.

B. Contextual Attacks

The proposed routing protocol is able to deal with certain contextual attacks that are particularly effective in MANETs.

In communication pattern attack, by observing the communication patterns of nodes, an attacker may collect their communication profile to identify the sources and destinations. When a group of nodes are sending packages, another group may stay silent, this synchronization pattern will become clearer as the attacker keeps monitoring and thus it can recognize the two groups as sources and destinations. In our approach, when routing a message, the selection of random-forwarders is random, though the random-forwarders can only be selected in one specific zone, this effect can be still reduced greatly: even for two communicating nodes, zones' division pattern is random, thus the routing path is different every time. In addition, the fact that many nodes communicate in the network can further blur communication pattern. Moreover, because the random-forwarders can exist in any place within the network and the relay-nodes are randomly chosen as well, every node has the opportunity to transmit messages and there will be no observable communication pattern. From this point of view, packet counting attack could also be avoided, because the random chosen routing path ensures that every node has the same opportunity to receive and route other nodes' messages.

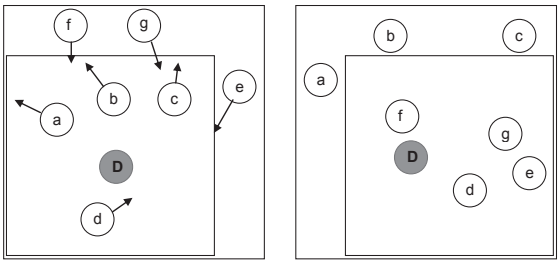

Figure 3. Formation of intersection.

Intersection attack is to extract information from repeated observations of nodal communication. It is especially effective in MANETs because under mobile environment, intersection forms frequently as node moves in and out a zone. Figure 3 shows how an intersection forms in our approach. Figure 3(a) is the status after a message is delivered to an entire zone using broadcasting according to our routing algorithm. We see a, b, c, d, and D (destination node) are in the same zone. Figure 3(b) is the next time that a message is sent between the same pair of nodes. This time nodes d, e, f, g and D are in the zone. Notice that the intersection of the in-zone nodes in both figures are d, and destination D, it means the destination is partially exposed. The longer an attacker watches the whole process, the smaller set of suspicious nodes can be determined to contain the destination.

To mitigate this observation of increasingly small intersection. Wu *et al.* [16] seek to dynamically enlarge the range of anonymous zone. In the proposed protocol, the smallest split zone can be treated as the anonymous zone. This strategy could reduce the possibility that the attacker finds the destination, although it inevitably increased the communication overhead. In our algorithm, we can enlarge the minimum zone size or increase the TTL of broadcast messages in order that more nodes can receive these messages to make intersection attack harder and maintain the anonymity level. Because under such circumstance, the attacker may not be able to observe a useful intersection due to two reasons: Firstly, increasing the size of destination zone can increase the number of nodes with a zone, thus it will take much longer for the attacker to observe a small enough intersection. Also, the attacker has more probability to encounter a situation that destination node moves outside its original zone even before this attacker could identify it using intersection attack. Therefore, the destination's anonymity can be better protected.

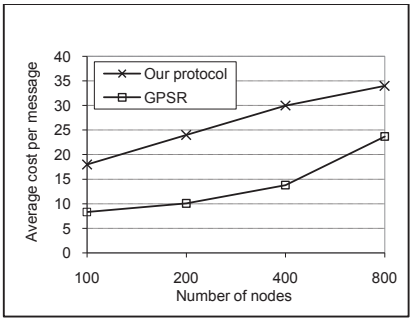

Figure 4. Performance of average hops

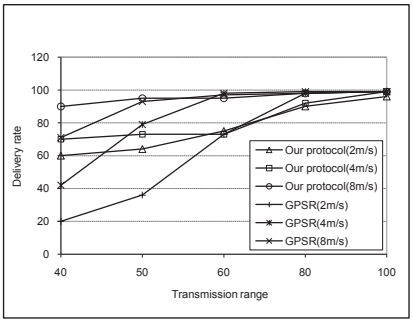

Figure 5. Performance of delivery rate

V. Performance Evaluation

Our experiment is based on an event-driven simulator [22]. We evaluated the proposed protocol compared to the GPSR [18] protocol. In GPSR, a packet is always forwarded to the node nearest to the destination; when such a node does not exist, it uses perimeter forwarding to find the next hop which is the closest to the destination. Messages were randomly generated at the speed of 10 queries per second. The number of nodes was set to 400 in a field of 1000m$\times$1000m. The times of zone partition is set to 6 in all tests. We use following metrics for performance evaluation:

(1) Average cost per message. It is measured as the cumulated routing hop counts divided by the number of messages sent. This metric shows the efficiency of routing algorithms.

(2) Delivery rate. It is measured by the fraction of messages that are successfully delivered to destination node. This metric shows the robustness of a routing protocol to adapt to mobile network environment.

In the first experiment, we simulate different network scale by varying the number of nodes, where node moving speed is 2m/s and transmission range is 100m.

Figure 4 shows the average number of hops versus the node number within network, we see that our approach always spends some more hops than GPSR, because firstly, the routing path of our approach is divided to several steps and each step has a random temporary destination; also the local broadcast consumes several hops. This extra cost is inevitable and acceptable because we do not need hop-by-hop encryption or step by step local broadcasting, our protocol merely broadcast in the destination zone to ensure the anonymity of destination, which costs little.

Figure 5 shows delivery rate as a function of transmission range. We alter the transmission range between 40m-100m, and conduct the tests under node moving speed of 2m/s-8m/s. We observe that when transmission range is 80-100m, the delivery rates of both approaches at different moving speeds are over 90%. When the transmission range is between 50-60m, the delivery rate of GPSR downgrades significantly while our protocol can still maintain a relative steady delivery rate. Especially when nodes in GPSR move at 2m/s, its delivery rate drops to less than 40%. This is because when transmission range is about 50m, the connections between nodes are very sparse. If nodes move at 2m/s, new connections can not be established easily for nodes to delivery messages. This phenomenon is more obvious when the transmission range is 40m, under which our protocol maintains more steady delivery rate and outperforms GPSR under all speeds. The reason is the incorporation of local broadcasting in the last step, which can deliver the message even when the receiver has moved out of the zone a little. On the contrary, nodes in GPSR know the position of destination, but when the destination has moved far from the original position, the delivery may fail. This result shows that our protocoal can maintain a comparable or even better delivery rate than GPSR while maintains complete anonymity of source, destination and routing path.

VI. Conclusion and Future Work

In this paper, we propose a low-cost anonymity routing protocol that provides efficient routing algorithm using nodes' position information in MANETs without heavy encryption/decryption or local broadcast cost. Because both source and destination only embed in messages the position of the zone they resides instead of location of themselves, their anonymity can be protected. Moreover, the use of hierarchical zones and randomly chosen intermediate random-forwarders can ensure an anonymous and random routing path. From the analysis of security and experiments, we prove that our approach can handle various attacks efficiently while maintains good performance.

Future works lies in more thorough simulation, and making this protocol more sophisticated and robust. In addition, current method needs a proactive mechanism to better solve intersection attacks.

Acknowledgment

This research was supported in part by U.S. NSF grants CNS-0834592 and CNS-0832109.

References

[1] J. Kong, X. Hong, and M. Gerla, "ANODR: Anonymous on demand routing protocol with untraceable routes for mobile ad-hoc networks," in *Proc. of MobiHoc2003*, 2003, pp. 291–302.

[2] Z. Y, W.Liu, and W.Luo, "Anonymous Communications in Mobile Ad Hoc Networks," in *Proc. of INFOCOM2005*, 2005.

[3] Y.-C. Hu, D. B. Johnson, and A. Perrig, "SEAD: Secure Efficient Distance Vector Routing for Mobile Wireless Ad Hoc Networks," in *Proc. of WMCSA '02*, 2002, p. 3.
[4] L. Yang, M. Jakobsson, and S. Wetzel, "Discount Anonymous On Demand Routing for Mobile Ad hoc Networks," *SECURECOMM*, vol. 6, p. 2006.
[5] C. E. Perkins, "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers," 1994, pp. 234–244.
[6] Y.-C. Hu, A. Perrig, and D. B. Johnson, "Ariadne: a secure on-demand routing protocol for ad hoc networks," *Wirel. Netw.*, vol. 11, no. 1-2, pp. 21–38, 2005.
[7] I. Aad, C. Castelluccia, and J. Hubaux, "Packet coding for strong anonymity in ad hoc networks," in *Proc. of IEEE Securecomm*, 2006.
[8] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks," *IEEE Mobile Computing*, 1996.
[9] C.-C. Chou, D. S. Wei, C.-C. J. Kuo, and K. Naik, "An efficient anonymous communication protocol for peer-to-peer applications over mobile ad-hoc networks," in *Selected Areas in Communications, IEEE Journal on*, 2007, pp. 192–203.
[10] X. Wu, "AO2P: Ad Hoc On-Demand Position-Based Private Routing Protocol," *IEEE Transactions on Mobile Computing*, vol. 4, no. 4, pp. 335–348, 2005, fellow-Bhargava,, Bharat.
[11] B. Zhu, Z. Wan, M. S. Kankanhalli, F. Bao, and R. H. Deng, "Anonymous Secure Routing in Mobile Ad-Hoc Networks," in *Proc. of LCN '04*, 2004, pp. 102–108.
[12] S. M. M. Rahman, M. Mambo, A. Inomata, and E. Okamoto, "An Anonymous On-Demand Position-Based Routing in Mobile Ad Hoc Networks," in *Proc. of SAINT '06*, 2006, pp. 300–306.
[13] Z. Zhi and Y. K. Choong, "Anonymizing Geographic Ad Hoc Routing for Preserving Location Privacy," in *Proc. of ICDCSW '05*, 2005, pp. 646–651.
[14] P. V., D. Yao, and I. L., "Securing location aware services over VANET using geographical secure path routing," in *Proc. of ICVES*, 2008.
[15] A. R. Beresford and F. Stajano, "Mix zones: User privacy in location-aware services," in *Proc. of PERCOMW04*, 2004, pp. 127–131.
[16] X. Wu, J. Liu, X. Hong, and E. Bertino, "Anonymous Geo-Forwarding in MANETs through Location Cloaking," *IEEE TPDS*, vol. 19, no. 10, pp. 1297–1309, 2008.
[17] H. Dobbertin, "Cryptanalysis of MD5 Compress," in *In Rump Session of EuroCrypt 96*. Annoucement, 1996, p. 68442.
[18] S. Ratnasamy, B. Karp, S. Shenker, D. Estrin, R. Govindan, L. Yin, and F. Yu, "Data-centric storage in sensornets with GHT, a geographic hash table," *Mob. Netw. Appl.*, vol. 8, no. 4, pp. 427–442, 2003.
[19] L. Sweeney, "k-anonymity: a model for protecting privacy," *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, vol. 10, no. 5, pp. 557–570, 2002.
[20] X. Wu, "DISPOSER: distributed secure position service in mobile ad hoc networks: Research Articles," *Wirel. Commun. Mob. Comput.*, vol. 6, no. 3, pp. 357–373, 2006.
[21] J. francois Raymond, "Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems," in *Proc. of International Workshop on Design Issues in Anonymity and Unobservability*, 2001, pp. 10–29.
[22] "www.netlab.tkk.fi/tutkimus/dtn/theone/."