Themed Section : Engineering and Technology
© 2018 IJSRSET | Volume 5 | Issue 3 | Print ISSN: 2395-1990 | Online ISSN: 2394-4099

Identification of Malicious and Liar Nodes by Query Processing In Manets

R.Gangadevi[1], T.Ravichandran[2]

[1]PG Scholar, Department of ECE, Annamalai University, Tamil Nadu, India

[2]Associate Professor, Department of ECE, Annamalai University, Tamil Nadu, India

ABSTRACT

Mobile Ad-hoc Networks (MANETs) is a collection of mobile nodes sharing a wireless channel without any centralized control (or) established communication backbone. MANETs require fundamental changes to network routing protocols. These are characterized by the mobility of nodes, which can move in any direction and at any speed that may lead to arbitrary topology and frequent partition in the network. In the existing method, the malicious nodes are identified based on the grouping. In MANETs, if a normal node becomes an attacker, the malicious node tries to disrupt the operations of the system. The main drawback in the above existing method that it doesn't address the critical Liar Nodes (LNs) responsible for the network in security. In this proposed work, malicious node are identified based on the top-k-query processing methods and it is also designed to locate and identify LNs by the False Notification Attacks (FNA) using the Dynamic Source Routing (DSR) reactive routing protocol and to prevent the liar nodes from performing False Notification Attack (FNA). So that the message can be sent with more security by using the Rivest Shamir Adleman (RSA) Algorithm and Secure Hashing Algorithm 1(SHA1) methods and also with less power, thereby increasing the battery life. To analyse the following parameters such as Throughput and packet loss were simulated through Network Simulator-2 (NS-2). In the proposed work, the output produced promising results to identify and eliminate the LNs, and thus reducing the number of unwanted nodes thereby improving security due to unwanted nodes using majority of the network performance and metrices.

Keywords: Ad-Hoc Networks, Attacker, Liar Nodes, Malicious Nodes, Query Processing

I. INTRODUCTION

Mobile Ad-hoc Networks (MANETs) is a collection of mobile nodes sharing a wireless channel without any centralized control (or) established communication backbone Recently, there has been an increasing interest in MANET, which is constructed by only mobile nodes. Since such self-distributed networks do not require pre-existing base stations, they are expected to apply to various situations such as military affairs and rescue work in disaster sites. In MANETs, since each node has poor resources (i.e., the communication bandwidth and the battery life of mobile nodes are limited), it is effective to retrieve only the necessary data items using top-k query, in which data items are arranged according to a particular attribute score, and the query- issuing node acquires the data items with k highest scores in the network (the global top- k result).On the other hand, in MANETs, if a normal node becomes malicious owing to an attack from outside the network, the malicious node tries to disrupt the operations of the system. In this case, the user whose network holds the malicious node will typically continue to operate

the system normally, unaware of the threat, while the malicious node may execute a variety of attacks (e.g. Denial of Service (DoS) attack such as blackhole attack).Basically, malicious nodes attempt to disrupt query- issuing node's acquisition of the global top-k result for a long period, without being detected. However, DoS attacks in MANETs have been actively studied for long years, and as a result, using existing techniques, such attacks can be exposed by the query-issuing node or intermediate nodes. Here, a unique characteristic of top- k –query processing is that the query-issuing node does not know the unique top-k result beforehand. Therefore, even if a malicious node replaces high- score data items with its own low-score ones, when relaying the data items, it is difficult for the query-issuing to detect the attack, and it may believe that all the received data items with k highest scores are the unique top-k result. In this paper, we introduce a new type of attack called data replacement attack (DRA), in which a malicious node replaces the received data items (which we call the local top-k result) with unnecessary yet proper data items (e.g., its own low-score data items). Since DRAs are a strong attack, and more difficult to detect than other traditional types of attack, some specific mechanism for defending against DRAs are required. This method also identifies the liar nodes to prevent it from the False Notification Attack (FNA).

II. RELATED WORK

W.T.Balkey et al. [1] suggested the advantages of best top-k queries in the context of distributed peer-to-peer information infrastructures and show how to extend the limited query processing in current peer-to-peer networks by allowing the distributed processing of top-k queries, while maintaining a minimum of data traffic.

Kejun Liu et al. [2] focus on the nodes which drop the information to send forward considered as a malicious nodes. Specifically, nodes may participate in the route discovery and maintenance process but refuse to forward data packets. To detect such misbehavior and more efficient detection process, the 2ACK technique is analysed. The main idea of the 2ACK scheme is to send two-hop acknowledgement packets in the opposite direction of the routing path.

Minji Wu et al. [3] exploits the semantics of top-k query and developed an energy-efficient monitoring approach called FILA. The basic idea is to install a filter at each sensor node to suppress unnecessary sensor updates. Filter setting and query re-evaluation upon updates are two fundamental issues to the correctness and efficiency of the FILA approach. They develop a query re-evaluation algorithm that is capable of handling concurrent sensor updates and also present optimization techniques to reduce the probing cost.

R.Zhang et al. [4] suggested three schemes whereby the network owner can verify the authenticity and completeness of fine-grained top-k query results in tired sensor networks, which is the first work of its kind. This schemes are built upon symmetric cryptographic primitives and force satisfied master nodes to get both authentic and complete query results to avoid being caught.

P.Dewan et al. [5] developed reputation systems for peer-peer networks to protect the network without using any central component, and thereby harnessing the full benefits of the peer-peer network. The reputations of the peers are used to determine whether a peer is a malicious peer (or) good peer. Once detected, the malicious nodes are obstracized from the network as the good nodes do not perform any transactions with the malicious peers. This technique allows secure exchange of reputation information between the two peers participating in a transactions.

Jing Shi et al. [6] developed a two-tier sensor network with resource-rich nodes at the upper tier and resource-poor sensor nodes at the lower tier. Source node collect data from sensor nodes and answer the queries from the network owner. The reliance on

master nodes for data storage and query processing raises serious concern about both data authentication and query-result correctness.

Yao-Tung Tsou et al. [7] implies each sensor having many sensing capabilities periodically route the multidimensional sensed data to the storage node, which responds to the queries. Unfortunately, node compromises pose the great challenge of securing the data collection. The sensed data could be leaked to or could be manipulated by the compromised nodes.

Y.Sasaki et al. [8] proposed a two-phase query processing method for top-k query processing. In this method, the query-issuing node collects the information of scores of data items held by each node in the first phase. Based on the received information, it determines the threshold of the score, i.e., the estimated k-th highest score. Then, in the second round, the query-issuing node transmits a query attached with the threshold, and each node that received the query sends back only its own data items whose scores are equal to or larger than the threshold. This can further reduce the traffic and also keep high accuracy of the query result.

Baljeet Malhotra et al. [9] discussed the importance of underyling structure and proposed the use of Dominating Set Tree (DST) for processing the top-k queries. Leveraging on the properties of a DST they proposed a new algorithm, EXTOK and proved its correctness. Simulation, using real and synthetic data sets, revealed the effectiveness and superior efficiency of the combination EXTOK-DST for processing the top-k queries in WSN.

Rui Zhang et al. [10] took multidimensional range queries as an example to investigate secure cooperative data storage and query processing in UTSNs. They present a suite of novel schemes that can ensure data confidentiality against master nodes and also enable the network owner to verify with very high probability the authenticity and completeness of any query result by inspecting the spatial and temporal relationships among the returned data and the performance evaluations confirm the high efficacy and efficiency of the proposed schemes.

III. PROPOSED METHOD

In our new suggested top- k query processing method, the query issuing node first floods a query over the entire network, and each node receiving the query stores information on all possible routes to the query- issuing node. Then, each receiving node replies with data items with the k highest scores to two neighbor nodes. Each node includes, in its reply message, information on the reply message forwarding routes which consist of pairs of sender node and next node IDs. Based on this attached information, the query- issuing node can detect an attack occurring along a reply message route. In MANETs, since the network topology dynamically changes due to the mobility of nodes, radio link disconnections can occur between nodes. Therefore, if a node detects a radio link disconnection along one of its two reply routes, it sends back the data items to a different node, to which those data items have not yet been sent, to ensure that they are sent back along two different routes.

In our proposed malicious node identification method, a query- issuing node that detects a DRA narrows down the malicious node candidates based on the received reply messages. Then, the query-issuing node determines whether a given reply message sent back by a malicious node candidate includes replaced data items or not, by sending inquiries to nodes receiving reply messages from this candidate.

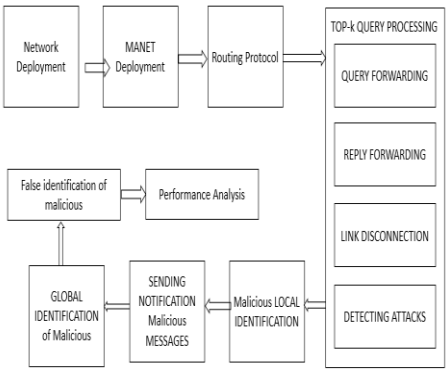

Figure 1. Architecture Diagram

In this way, the query- issuing node can detect the malicious node. Here, each node tends to identify the neighboring malicious nodes, but hardly identify the malicious nodes which are far from it. Therefore, nodes share the information on identified malicious nodes to detect the malicious nodes quickly. Specifically, after each node identifies malicious nodes, it floods the information on the identified malicious nodes within the network. When each node has received a certain number of queries, it performs malicious node identification procedures based on the received information. Specifically, it divides nodes into relevant groups based on similarities of the information on malicious nodes detected by those nodes, and then identifies malicious nodes based on the results of malicious node identifications by these groups.

Liar nodes which provides the false information can be identified by the False Notifcation Attacks (FNA).

IV. TOP- K QUERY PROCESSING

A. Query Forwarding

First, the query- issuing node floods a query over the entire network. The query consists of the node identifier of the query- issuing node (Query-issuing node ID), the query identifier of the query (Query ID), the number of requested data items (k), the query condition, and a list of the node identifiers of nodes on the path along which the query message is to be transmitted (Query path). Specifically, the query- issuing node, M_p, specifies the query condition and the number of requested data items, k. Then, M_p transmits a query message whose Query path includes its identifier, M_p, to its neighbor nodes. A node, M_q, which receives the query, transmits it. Hop count denotes the number of hops to the query- issuing node, based on the number of nodes included in the Query path. Then, M_q sets a waiting time for reply (RD) according to the following equation:

$$RD = (hop_{max} - hop_{cnt}) \cdot T_{wait} \quad (1)$$

where hop_{cnt} denotes the number of hops to the query- issuing node, hop_{max} denotes the maximum number of hops (calculated based on the area size of the network and the radio range of nodes), and T_{wait} is a positive constant. In this equation, as hop_{cnt} increases, RD decreases. When M_q receives the query later again, it stores the ID of the query sender node as its neighbor node, as well as, the Query path and the number of hops.

B. Reply Forwarding

When RD has transfered, each node sends back a reply message, which includes its own node identifier (Sender node ID), the identifier of the next node along the reply route (Dest node ID), a list of the data items (including their scores) and the node identifiers of the nodes possessing them (Data list), and a list summarizing the reply message routes, i.e., a list of the pairs of sender and next node identifiers (Forwarding Route). Node M_r sends a reply message when its RD has passed. Here, REP denotes a reply message and REP. FR denotes the forwarding route list consisting of (Sender node ID, Destination node ID), which denotes the list of sender and next node identifier pairs, and R denotes the maximum number of reply messages to be re-sent. M_r selects the next node from its neighboring nodes, which has the least hop count and least overlap between its Query path and the parent node's Query path.

C. Link Disconnection

In MANETs, the network topology changes dynamically due to the movement of nodes. When a radio link disconnection to the parent node or next node occurs, a replying node, M_r, cannot send a reply message, resulting in reduced accuracy of the query result. Therefore, if a node sends a reply message R times but does not receive an ACK from the parent or next node, the sending node detects a radio link disconnection; at which point the node sends the reply message to another neighbor node among those whose routes to the query- issuing node include the least overlap between Query path in the replay

message and their own Query path. If the sending node has no neighbor nodes which satisfy this condition, it sends the reply message to a neighbor selected in the same way as in selecting the next node in "Reply Forwarding" process among nodes which have not been sent the reply message.

D. Detecting Attacks

After the query- issuing node, Mp, obtain all the reply messages, it detects a DRA. Top- K result denotes the data items with the k highest scores, acquired by the query- issuing node, REP. Data and reply forwarding respectively denote the data list and forwarding route included in the reply message, REP, and Send route denotes the set of node identifiers along the route from the node possessing a given data item to the query issuing node (the query- issuing node can know the Send route from the forwarding route information). If the nodes which have data items in the top-k result are included in Send route (or REP.FR), but the data items in the top- k result are not included in REP. Data, the query- issuing node detects a DRA and initiates the malicious node identification process and also the liar nodes from the False Notification Attacks (FNA). If the query-issuing node does not detect a DRA, it completes the top-k query processing.

E. Global Identification

In our method, each node individually identifies malicious nodes and liar nodes using the shared information by the two steps; node grouping and malicious node identification. Each node divides nodes in the network into some groups based on the information in the notification messages received by the nodes.

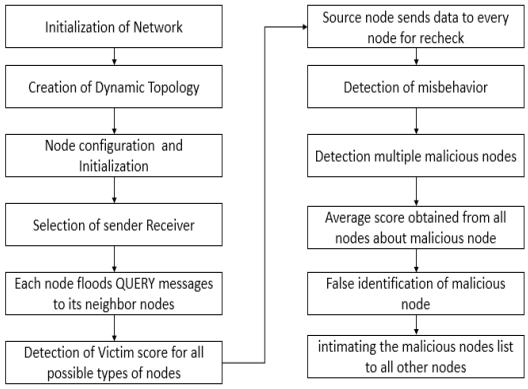

Figure 2. Flow Diagram

After the node grouping, each node conclusively determines malicious nodes and liar nodes based on the information about malicious nodes and lair nodes identified by nodes in each group. Here, there are three types of groups, i.e., a group composed of (i) only normal nodes, (ii) only malicious nodes, and (iii) both liar and malicious nodes. The nodes identified as malicious and liar nodes by all nodes in a group of (i) or (iii) are surely malicious nodes and liar nodes. Here, since liar nodes are generally minorities in the entire network, majority based judgment (and pruning) works well for malicious node and liar nodes identification. Therefore, in our method, nodes are confirmed to be malicious and liar when they are determined to the number of groups equal to or larger than a certain threshold. Malicious and Liar nodes can be identified by the following steps:

Step 1:

In our proposed system, the MANET nodes are initialized with basic configuration of node.

Step 2:

While concentrating 'Ad-hoc' network we have created dynamic topology with the nodes.

Step 3:

Nodes are initialized as per the topology as well as working behavior.

Step 4:

User can select the sender and receiver where the secured communication is important.

Step 5:

Each node floods QUERY messages to its neighbour nodes and the source node.

Step 6:

Top-k Query gives Victim score for all possible types of nodes with respect of source node.

Step 7, 8 & 9:

Source node sends data to every node for recheck the misbehaving function in the network.

Step 10, 11 & 12:

To detect falsification of malicious node, average score obtained from all nodes about malicious node list and it detects the good node from that malicious node list. Intimating the malicious nodes list to all other nodes.

V. DYNAMIC SOURCE ROUTING

Dynamic Source Routing (DSR) is a on demand reactive protocol based on the source route approach. In DSR, the protocol is based on the link state algorithm in which source initiates route discovery on demand basis. The sender determines the route from source to destination and it includes the address of intermediate nodes to the route record in the packet. DSR was designed for multi- hop networks for small diameters.

VI. ALGORITHM

A. Rivest Shamir Adleman Algorithm (RSA)

Ron Rivest, Adi Shamir, and Leonard Adleman popularized an asymmetric algorithm. Asymmetric algorithm means it uses two different keys (i.e) public and private key. Public key is given to everyone and the private key is kept private. Most importantly, RSA contrivance a public-key cryptosystem, as well as digital signatures. RSA is motivated by the published works of Diffie and Hellman.

The process of exchange the session key:

- A uses RSA asymmetric algorithm to generate their own public key (n, e) and the private key (n, d) and sends the information to B that contains the public key (n, d) and ID of A
- B gets the session key k and uses the public key encrypt the message to A, $m^e \bmod n$.
- A uses his private key to decrypt the $C^d \bmod n$, then can get K. In this way, A and B can communicate with symmetric encryption algorithm which was symmetrric and the session key K.

The RSA security depends on the computational difficulty of factoring large integers. Obviously the longer a number is the harder is to factor, and so the better the security of RSA. Strength of encryption is directly tied to key size, and key length is doubled delivers an exponential increase in strength, although it does debase performance. RSA keys are commonly 1024 or 2048 bits long, but experts believe that 1024 bit keys could be broken in the future. The main advantages in using extremely of long length is the computational overhead involved in encryption/ decryption. The main drawback is only if a new factoring technique arises that requires keys of such lengths to be used that important key length increases much faster than the average speed increase in computers utilizing the RSA algorithm.

B. Secure Hashing Algorithm

A hashing algorithm is a mathematical function that compress data to a fixed size. Hashes are convenient for situations where computers may want to identify, compare, or otherwise run calculations against files and strings of data. One of the key properties of hashing algorithms is determinism. Hashing algorithms are used in all sorts of ways – they are used for password storing, in computer vision, in databases.

SHA-0:

A retronym applied to the original adaptation of the 160-bit hash function published in 1993 under the name “SHA”.

SHA-1 :

A 160-bit hash function which is same as that of the Message Digest (MD) 5 algorithm. This was

described by the National security agency to be part of the digital signature.

SHA-2:

A family of two identical hash functions with different block sizes known as sha-256 and sha-512. They differ in word size.

SHA-3:

A hash function formerly called keycap, chosen in 2012 after a public competition among non NSA-designers.

VII. SIMULATION RESULTS

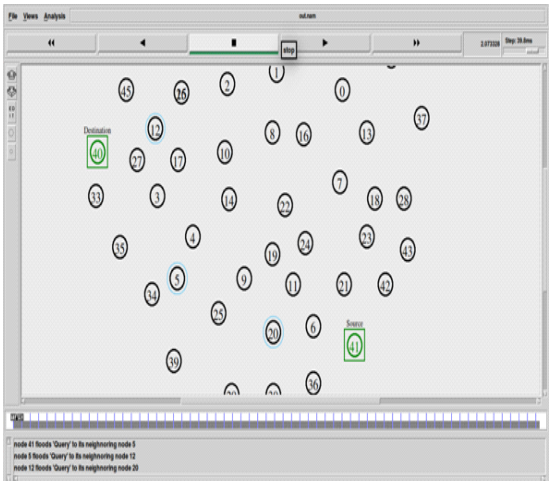

Figure 3. Selection of Source and Destination

Here the source is selected as node 0 and destination is selected as node 41 to transmit the above given message.

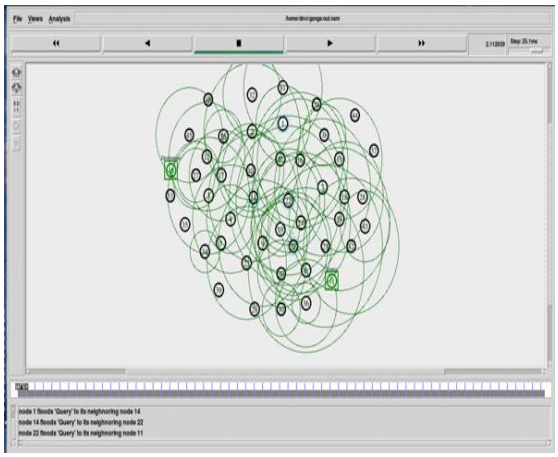

Figure 4. Node Floods Query To Neighbouring Nodes

Processing action with each and every nodes are described in the figure 4

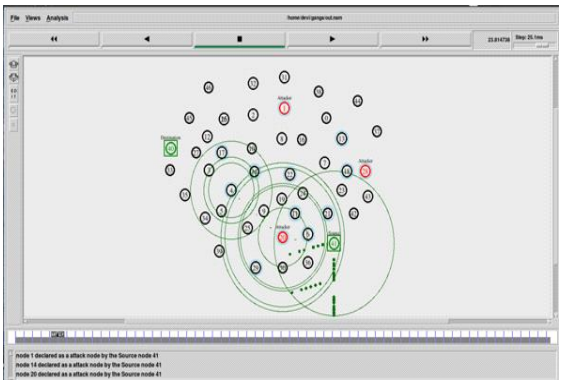

Figure 5. Identification of Malicious Nodes

Here the malicious nodes are identified one after the other using the victim score values.

Figure 6. Identification Of Liar Nodes

After the identification of malicious nodes, based on the message authentication code liar nodes are identified. Here the malicious and liar nodes are identified one after the other using two different techniques.

A. Parameter Comparison Throughput

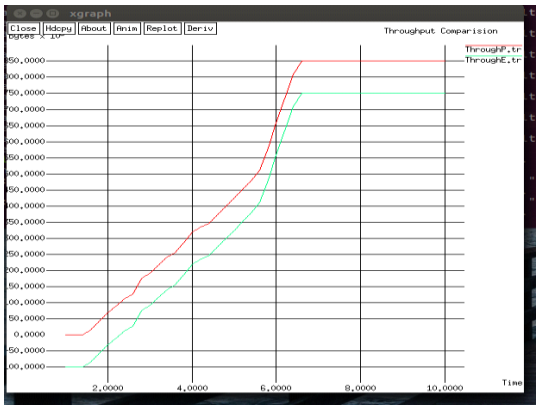

Figure 7. Throughput Comparison

Throughput is the total number of packets delivered over the total simulation time. It is found to be more in SHA algorithm than in the RSA algorithm.

B. Packet Loss

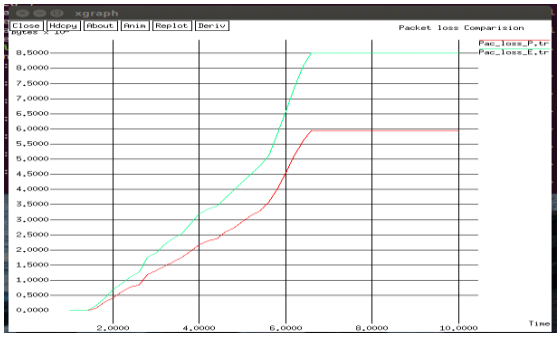

Figure 8. Packet Loss Comparison

When one or more packet of data travelling across a network fail to reach their destination. Packet loss is more in RSA algorithm than in the SHA algorithm

VIII. CONCLUSION

Malicious and liar nodes will result in the lack of security factor. So malicious and liar nodes are identified one after the other by two different techniques. In removing these nodes will improve the security factor. Then the parameters such as Throughput, packet delivery ratio, Energy consumption, packet loss were compared with RSA and SHA algorithms. SHA algorithm found to produce the better performance and results.

IX. REFERENCES

[1] W.T. Balke, W. Nejdl, W. Siberski, and U. Thaden, "Progressive distributed top-k retrieval in peer-to-peer networks," in Proc. ICDE, Apr. 2005, pp. 174–185.

[2] Kejun. Liu, J. Deng, P. K. Varshney, and K. Balakrishnan, "An acknowledgment-based approach for the detection of routing misbehavior in MANETs," IEEE Trans. Mobile Comput., vol. 6, no. 5, pp. 536–550, May 2007.

[3] Minji. Wu, J. Xu, X. Tang, and W. C. Lee, "Top-k monitoring in wireless sensor networks," IEEE Trans. Knowl. Data Eng., vol. 19, no. 7, pp. 962–976, Jul. 2007.

[4] R. Zhang, J. Shi, Y. Liu, and Y. Zhang, "Verifiable fine-grained top-k queries in tiered sensor networks," in Proc. INFOCOM, Mar. 2010, pp. 1–9.

[5] P. Dewanand, P. Dasgupta, "P2P reputation management using distributed identities and decentralized recommendation chains, "IEEE Trans Knowl. Data Eng., vol. 22, no. 7, pp. 1000–1013, Jul. 2010.

[6] J. Shi, R. Zhang, and Y. Zhang, "A Spatiotemporal Approach for Secure Range Queries in Tiered Sensor Networks," in Proc. INFOCOM, Jan.2011, pp. 945–953.

[7] C.M.Yu, Y.T.Tsou, C.S.Lu and S.Y.Kuo, "Practical and secure multi dimensional query frame work in tiered sensor networks," IEEE Trans. Inf. Forensics Security, vol.6, no.2, pp.241-245, Jun 2011.

[8] Y. Sasaki, T. Hara, and S. Nishio, "Two-phase top-k query processing in mobile ad hoc networks," in Proc. NBiS, Sep. 2011, pp. 42–49.

[9] B.Malhotra, M.A. Nascimento, and I. Nikolaidis, "Exact top-k queries in wireless sensor networks," IEEE Trans .Knowl. Data Eng., vol.23, no.10, pp. 1513–1525, Oct. 2011.

[10] Rui Zhang, Jing Shi, Yanchao Zhang and Jinyuan Sun, "Secure Cooperative Data Storage and Query Processing in Unattended Tiered Sensor Networks" , in IEEE Journal on selected

areas in communications, Vol. 30, No. 2, February 2012.

[11] Y. Yi, R. Li, F. Chen, A. X. Liu, and Y. Lin, "A digital watermarking approach to secure and precise range query processing in sensor networks," in Proc. INFOCOM, Apr. 2013, pp. 1950–1958.

[12] D. Amagata, Y. Sasaki, T. Hara, and S. Nishio, "A robust routing method for top-k queries in mobile ad hoc networks," in Proc. MDM, Jun. 2013, pp. 251–256.

[13] C.M. Yu, G.K. Ni, I.Y. Chen, E. Gelenbe, and S.Y. Kuo, "Top-k query result completeness verification in tiered sensor networks," IEEE Trans. Inf. Forensics Security, vol. 9, no. 1, pp. 109–124, Jan. 2014.

[14] T. Tsuda, Y. Komai, Y. Sasaki, T. Hara, and S. Nishio, "Top-k query processing and malicious node identification against data replacement attack in MANETs," in Proc. MDM, Jul. 2014, pp. 279–288.

[15] J. Shi, R. Zhang, and Y. Zhang, "Secure top –k – query processing in unattended tiered sensor networks," in Proc. INFOCOM, Nov 2014, pp. 945–953.

[16] Y. Zhang, G. Wang, Q. Hu, Z. Li, and J. Tian, "Design and performance study of a topology-hiding multipath routing protocol for mobile ad hoc networks," in Proc. INFOCOM, Mar. 2012, pp. 10–18.

[17] S. Chen, Y. Zhang, Q. Liu, and J. Feng, "Dealing with dishonest recommendation: The trials in reputation management court," Ad Hoc Network., vol. 10, no. 8, pp. 1603–1618, Nov. 2012.

[18] Z. Li and H. Shen, "A hierarchical account-aided reputation management system for large-scale MANETs," in Proc. INFOCOM, Apr. 2011, pp. 909–917.

[19] X. Liu, J. Xu, and W. C. Lee, "A cross pruning framework for top-k data collection in wireless sensor networks," in Proc. MDM, May 2010, pp. 157–166.

[20] M. Yiu, Y. Lin, and K. Mouratidis, "Efficient verification of shortest path search via authenticated hints," in Proc. IEEE ICDE, Long Beach, CA, USA, Mar. 2010, pp. 237–248.

[21] D. Ma, C. Soriente, and G. Tsudik, "New adversary and new threats: Security in unattended sensor networks," IEEE Netw., vol. 23, no. 2, pp. 43–48, Mar. 2009.

[22] M. Ye, X. Liu, W.-C. Lee, and D. L. Lee, "Probabilistic top-k query processing in distributed sensor networks," in Proc. IEEE ICDE, Long Beach, CA, USA, Mar. 2010, pp. 585–588.

[23] Q. Wang, K. Ren, W. Lou, and Y. Zhang, "Dependable and secure sensor data storage with dynamic integrity assurance," in Proc. IEEE INFOCOM, Rio de Janeiro, Brazil, Apr. 2009, pp. 954–962.

[24] R. Lu, X. Lin, H. Zhu, and X. Shen, "TESP2: Timed efficient source privacy preservation scheme for wireless sensor networks," in Proc. IEEE ICC, May 2010, pp. 1–6.