Abstract - A self-configuring Mobile Ad-Hoc Network (MANET) is a network of connected mobile nodes by wireless links. Primarily the recent research work has focused on the routing protocols. An essential technique to achieve power efficiency is a data fusion technique in sensor nodes. Here we propose an efficient Secured Quality of Service (QoS)-Aware Data Fusion technique for distributed Wireless Sensor Networks is proposed. Detection of misbehavior of a node which defers the data packet for an extra period of time is a key feature of secure data fusion. The proposed scheme is been investigated by using simulation results efficiently detects the malicious nodes. Expected throughput is a new metric introduced, for the comparison of throughput in multi-hop networks, and to show how the use of explicit link failure notification (ELFN) techniques can significantly improve TCP performance.

Keywords --- performance evaluation, data fusion; malicious nodes; power efficiency; QoS

1. INTRODUCTION

Without losing any important data, Data fusion is a process of aggregating certain combination of packets. Data fusion is performed in a distributed fashion based on the available local information. It can be performed at the intermediate nodes as well as at the end nodes. Data can be fused at the intermediate nodes to satisfy end-to-end delay constraint whereas at the end nodes it is performed to balance the delay and buffer overflow.

Recent development in the manufacturing of electronic components such as microprocessors, memory chips and development in the field of wireless networking have led to the development of WSNs. The collection of large number of low cost sensor nodes having sensors, processors, memory chips and wireless transceivers are deployed in a network. Sensor network is used in many applications such as military, agriculture, detection of forest fire, nuclear attacks and health.

Security in WSNs has six challenges.
1. Wireless nature of communication
2. Resource limitation in sensor nodes
3. Deployment of sensor nodes
4. Lack of fixed infrastructure
5. Unknown network topology prior to deployment
6. High risk of physical attacks to unattended sensors.

Motivation: Data fusion can be used to decrease the traffic load and energy consumption thus increasing the network lifetime. Data fusion is performed by specifying the delay constraint within which data is collected at the intermediate nodes and fusion is performed. This in turn increases the packet drop. Some nodes increase the delay constraint maliciously and further increase the packet drop. So, there is a necessity to propose a solution to determine the malicious nodes efficiently and thus reduce the packet drop.

Contribution: In the proposed secured data fusion technique, data fusion is performed with reliability. Some malicious nodes are introduced in the network which defers the packet for extra period than the actual specified time. Node’s with malicious nature increases the packet drop. To overcome this problem, neighboring nodes of a particular node monitors the node and if it finds the stated misbehavior then that node is excluded from the network. Hence, the network is secured with reduced packet drop. It should be noted that packet drop in secured data fusion technique is nearly equal to the data fusion technique that is not secured. Thus, the proposed scheme efficiently provides the security to the network.

Agenda of this paper is done: In Section 2, research works in the data fusion techniques and security issues are discussed. Problem definition is stated in Section 3 and Data Fusion Technique is explained in Section 4. Algorithm to provide a security to Data Fusion Technique called SQDF is proposed in Section 5. Implementation details are explained in Section 6. and Conclusions in Section 7.

2. RELATED WORK

Barton et al., [1] explain the traffic patterns in Wireless Sensor Networks as many-to-one or one-to-many communication. Performance can be characterized by the rate at which data can be fused at the collector center. The fusion rate of \(0\left(\log n/n\right)\) is optimal and it can be achieved using time-reversal communication.

J. Ghayathri et al.,[2] have analysed and simulated results show that SEF can effectively detect false reports even when the attacker has obtained the security keys from a number of compromised nodes, as long as those keys belong to a small number of the key pool partitions.

Tara et al., [3] explain about the intermittent...
connectivity network. It is a network in which connected path between source and destination exists very rarely because of limited transmission range. A node generates and stores the data and on reaching the communication range of another node it replicates the data. Multiple copies of the packet decrease the time to offload data to the destination, but increases energy and storage used in the system. Resource-delay tradeoff and capacity of intermittent connectivity with QoS restrictions such as communication bandwidth is quantified.

Hong et al., [4] recommend a protocol called Reliable Data Aggregation which associates packet's reliability in data transmission with the amount of information it contains and gives higher reliability to the packet which has more information. The reliable data aggregation can jointly optimize both information reliability and energy efficiency in sensor networks with data fusion.

Bhaskar et al., [5] evaluate the traditional end-to-end routing scheme and data-centric routing scheme. Data centric is a mechanism that performs in-network aggregation of data needed for energy efficient information flow. The impact of source destination placement and communication network density on the energy costs and delay associated with data fusion is presented. It is shown that data-centric routing offers significant performance gains across a wide range of operational scenarios over traditional end-to-end routing scheme. Information about the location of control center is continuously propagated throughout the sensor field to keep all sensor nodes updated with data reports. This leads to both excessive drain of sensor's limited battery power and increased collisions in wireless transmissions.

Fan et al., [6] portray Two-Tier Data Dissemination that provides scalable and efficient data delivery to multiple mobile sinks. Node compromise leads to severe security threats in Wireless Sensor Networks and the security protection breaks down when threshold is exceeds.

Hao et al., [7] converse an idea to overcome the threshold limitation and achieve resiliency against node compromise. Location-based approach is proposed in which the secret keys are bound to geographic locations and each node stores few keys based on its own location. The location-binding property constraints the scope for which individual keys can be used, thus limiting the damages caused by a collection of compromised nodes.

Jing et al., [8] considers routing security. This paper deals with the attacks against sink holes and hello floods in sensor networks. Multipath routing to multiple control centers is analyzed to provide tolerance against individual control center attacks and disguise the location of control center from eavesdroppers, explains relocation of control center in case of damage and enhances resiliency of the network.

Maarten Ditze et al., [9] in attendance the result of a study on the effects of data fusion for multi-target in WSNs. Normally WSNs has limited bandwidth and nodes with limited computing power and limited battery life. The main aim is to accurately track multiple targets crossing an area observed by Wireless Sensor Networks, while limiting the amount of network traffic. Various computing power aware data aggregation strategies are presented which helps in reducing energy consumption and tracking accuracy.

R. Lath et al.,[10] proposed method finds the packet loss and its position effectively. Using this method the packet dropping is identified easily. Sleep scheduling for event monitoring is used for energy minimization

Przydatek et al., [11] projected a novel frame work for secure information aggregation in large sensor networks. Certain nodes in a network, called aggregators, help in aggregating information requested by query, which reduces the communication overhead. By constructing efficient random sampling mechanisms and interactive proofs, it is possible for a user to verify the answer given by the aggregator which is the good approximation of a true value when aggregator and a fraction of the sensor nodes are corrupted.

Lageweg et al., [12] contrast the fusion strategies for multi-target tracking. The effect of noise (considered as false contacts) on the accuracy is verified. Central track algorithm is proposed to associate the individual measurements with tracker and estimates the target's position and velocity. The tracker can ideally separate true targets from false contacts.

Intanagonwiwat et al., [13] anticipated a Directed Diffusion. The basic idea of this protocol is to construct data fusion tree which collects the data at the control center where it is rooted. When data is delivered from any node to control center, aggregation occurs at the control center without interacting with the node to eliminate redundancy and to reduce transmission energy. They have mainly concentrated on energy aspect of data aggregation and have considered reliability.

Younis et al., [14] future the autonomous WSNs as a service platform whose mission is to provide dependable information to satisfy QoS requirements.

Carlos et al., [15] have through a survey on Wireless Sensor Networks, their technologies, standards and applications. Many routing, power management and data dissemination protocols with energy awareness as an important design issue are discussed.
James et al., [16] measure the quality of spatial resolution. In a network all the sensors participate equally in the network at the same time it conserves the energy and maintains the desired spatial resolution. The parameters such as mean and variance of the QoS are taken into consideration to control the network performance. Jie Gao et al., [18] formulated the problem of performing the data aggregation for sparse nodes. When the sensors are deployed to detect relatively rare events, each node which participates in the fusion must be queried. Instead of blindly querying all the nodes in the network, it is feasible to discover the interesting nodes to get statistical summaries. The key idea is the capability for two nodes that wish to communicate at the same time to discover each other at a cost that is proportional to the network distance.

Hartley et al., [19] discussed the problem of inferring per node loss rates from passive end-to-end measurements in Wireless Sensor Networks. They have shown how to adapt network interference, so that loss rates in WSNs are inferred. This includes per node loss rates and considering the unique characteristics of Wireless Sensor Networks. The problem of Maximum-Likelihood Estimation is solved using Expectation-Maximization. The result of inference procedure can be used to streamline the data collection process.

3. PROBLEM DEFINITION

AIM: The goal of the proposed concept is to perform data fusion process at the intermediate nodes to decrease the traffic load, reduce energy consumption and obtain reliability.

The following conditions are assumed for the WSNs.

1. Packet origination rate at all the nodes follows an exponential distribution.
2. All nodes defer the packet for same period of time before data fusion process.
3. Routing paths are predetermined and the routes from individual node to the control center is identified by the edges between various nodes and is modified only when a malicious node is detected in the network.
4. Control Center cannot be a malicious node.
5. The transmitting range of each node is assumed to be 50 meters.

4. DATA FUSION TECHNIQUES

It can be performed at couple of levels in a network. Such as:

a. Fused data is transmitted from end nodes to the control center through intermediate nodes.
b. Collecting the data packets at the end nodes.

At the end nodes, multiple samples are collected and fused together into a single packet for energy efficiency purpose and when an intermediate node receives a packet, forwards the packet or performs local processing. To achieve QoS constraints it is necessary that intermediate nodes perform data fusion process before forwarding a packet.

To perform data fusion at intermediate nodes, each node defers for a particular time for collecting set of packets over a period of time. Those packets can be used for data fusion process in order to decrease network load and energy efficiency. Under such scenarios, malicious nodes defer for extra period of time than expected. Due to such malicious nature of a node, the following situations arises;

1. Threat to lose some important information because of extra deferred time.
2. Decrease in the successful packet delivery to the control center from the sensor nodes.
3. End-to-end delay increases which sternly decreases the network lifetime.

Identify such malicious nodes and exclude them from the network topology and there is a necessity that the simulation has no packets should be forwarded to such malicious nodes and descendents of those nodes should be connected to the nearest well-behaving node for the continued network operation.

5. ALGORITHM

A malicious node considered is called delay hole. Delay hole defers the packet for extra time than the specified time. In a network there may be one or more malicious nodes, which reduce the network throughput. Every node in a network defers the packet for a certain period of time for data fusion process. As a result of this, most of the packets will be dropped because their exists an end - to- end delay constraint which increases the packet drop. Increase in number of such malicious node lead to network failure.

To achieve this, an algorithm for Lock based Secure QoS-Aware Data Fusion process with the routing scheme is proposed in this paper. Algorithm effectively detects the malicious node, advertises this information to the network and new routing information are announced.

In the first phase each node broadcasts the hello messages in the network to determine its neighbors and also to know the average transmission time required to transfer the packet between any two nodes.

In the second phase, when a packet is received by a node, it defers the packet for certain time which is required for data fusion process and sends the packet to its parent node. This packet is received by the source node which checks the behavior of a node. A node defers the packet for extra time over many transmissions, then the node is considered to be a malicious node. This observation is
communicated throughout the network in the third phase by broadcasting report message which finally reaches the control center.

Control center receives the report from different nodes. Upon receiving this report, control center broadcasts the decision message to the network and also communicates the new routing information.

**Algorithm :** Lock based Secure QoS-Aware Data Fusion (SQDF)

<table>
<thead>
<tr>
<th>First Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>begin</td>
</tr>
<tr>
<td>Broadcast Hello Message</td>
</tr>
<tr>
<td>Determine average transmission time</td>
</tr>
<tr>
<td>Send Data Packet</td>
</tr>
<tr>
<td>end</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Second Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>begin</td>
</tr>
<tr>
<td>Receive Data Packet</td>
</tr>
<tr>
<td>if (malicious node)</td>
</tr>
<tr>
<td>Reset the timer for (defer time + extra time)</td>
</tr>
<tr>
<td>else</td>
</tr>
<tr>
<td>Reset the timer for defer time</td>
</tr>
<tr>
<td>Fuses data till timer expires</td>
</tr>
<tr>
<td>if (timer expires)</td>
</tr>
<tr>
<td>Send Data Packet</td>
</tr>
<tr>
<td>end</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Third Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>begin</td>
</tr>
<tr>
<td>Check parent node’s behavior</td>
</tr>
<tr>
<td>if (parent node is malicious)</td>
</tr>
<tr>
<td>Broadcast Report Message</td>
</tr>
<tr>
<td>else</td>
</tr>
<tr>
<td>Send Data Packet</td>
</tr>
<tr>
<td>end</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fourth Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>begin</td>
</tr>
<tr>
<td>Receive Report Message</td>
</tr>
<tr>
<td>Find new routing path excluding malicious node</td>
</tr>
<tr>
<td>Broadcast Decision Message</td>
</tr>
<tr>
<td>end</td>
</tr>
</tbody>
</table>

All the neighboring nodes receive this message and sends an acknowledgment by including the current time stamp in the message. Source node receives this message and finds the transmission time required to transfer the packet from source node to the node which has sent an acknowledgment. Initially each node called source node broadcasts the hello message which includes the time stamp at which the packet has been sent. After transferring n such packet between any two nodes, average transmission time is calculated.

Average transmission time is determined by considering n such data transfers. The motivation for considering n such data transfer is that, a node may not find out the exact transmission time because of channel error, interference, traffic load, etc., and legitimate nodes may be accused as malicious for deferring the packet for extra time. Average transmission time(T) when the n packets are sent by, say, the first node is, 

\[ T_{1i} = \frac{(t_{11} + t_{12} + t_{13} + \ldots + t_{1n})}{n} \]

Second node is:

\[ T_{2i} = \frac{(t_{21} + t_{22} + t_{23} + \ldots + t_{2n})}{n} \]

In general for N nodes’ surrounding a particular node is, 

\[ T_{ni} = \frac{\sum_{j=1}^{n} t_{nj}}{n} \]

After calculating the average transmission time, during data transfer when a node receives the data packet from the leaf node, it defers the packet for t time, assigns the current time and forwards the packet to its parent node. Since, the data transfer takes place through omni directional antenna, source node receives this packet and checks whether the intermediate node is malicious or not. In order to find out the defer time, source node considers the packet origination time and the time at which the source node receives it from its parent. The difference gives the dt ime.

A node is said to be behaving according to the protocol specification if the defer time dt ime is

\[ d_{\text{time}} \leq 2(T) + t \]

If dt ime is greater than the time required for transmission and defer time, then that node is considered as a malicious node. A threshold is maintained which indicates the maximum number of packets which can be deferred for time greater than dt ime. If the threshold exceeds i.e., dt ime threshold \( \geq d_{\text{time}} \), then a node is accused as a delay hole and the source node broadcasts the report message into the network which finally reaches the control center.

WSNs should be robust, since such networks can be efficiently used in real time application.

In the final phase, control center receives the report message about a particular node as delay hole. If majority of nodes send the report that a node is delay hole then control center declares the node as malicious node and broadcasts this decision message to the network. Malicious node is determined when all the neighboring nodes decides that it is a delay hole by observing the transmission time taken for the transmitting n packet by m nodes. Thus, a node is malicious when,
Control center is a central authority which has the routing information of all the nodes which has to exclude the malicious node from the network and specify the new routing scheme. Control center verifies its routing information to determine the descendents of the malicious node and finds the nearest non-malicious node. It directs the descendents of a malicious node to get connected to the non-malicious node. This process continues and if there are any such malicious nodes in the network, they are excluded from the network. The proposed algorithm efficiently detects the misbehaving node at the same time successful packet delivery probability can be increased.

For instance, consider a topology depicted in Figure 1. A represents the control sender which is responsible for taking major decisions such as whether the node is behaving properly and broadcasting this decision in the network. Sensor nodes are responsible for performing data fusion process and transmitting such fused data to the control center.

The aggregation of data requires certain waiting time by all the intermediate nodes which increases the end-to-end delay. The restriction on end-to-end

control center receives the reports from node X and node Y and it comes to the conclusion that node M is misbehaving. Now the control center broadcasts the decision saying that node M is misbehaving and it also instructs node B and node C to get connected to node X and node Y respectively and not with node M. Thus, a malicious node M is successfully excluded from the network as shown in Figure 3.

6. IMPLEMENTATION

Implementing QoS- Secured Data Fusion involves three different types of messages other than the data that has to be sent. Hello message is used to evaluate the average time required to send a data from a node to its parent. Every node transmit a hello message periodically to update the node's information at that particular instance. This message is transmitted to only one hop neighbors and average time is calculated.

A source node detects an intermediate node behaving maliciously, it broadcasts a report message to the control center informing about the malicious node. A report message consists of malicious node's information with the higher time-to-live value. Decision message is sent by the control center to indicate the source nodes about its new routing path to the control center.

The aggregation of data requires certain waiting time by all the intermediate nodes which increases the end-to-end delay. The restriction on end-to-end
delay called delay constraint D is used to increase the efficiency of the network. The value of D emphasize on every data packet to be received within a fixed time interval. The packet is dropped when the difference between the time of receiving packet by control center and its origination time is greater than the delay constraint D.

The origination of a report message by a source node depends on the defer time by its parent. When a parent node sends data after aggregation, the source node checks the time parent has waited for aggregation. If this time is greater than the normal defer time, then the parent node is said to be misbehaving. Because of a traffic load in a network, the transmission time may vary. Hence, a threshold of twice the transmission time is considered to detect a malicious node. Since the control center has a routing information of all the nodes, it takes a decision of a new routing path for a node from which it receives a report message. A new routing information is broadcasted to a node through decision message. Exclusion of malicious nodes is implemented by nodes not sending a data packets to misbehaving nodes.

7. CONCLUSIONS

The proposed scheme is evaluated using fixed sensor networks. This approach can be combined with appropriate routing technique and thus can be used in dynamic network environment. Secured data fusion technique reduces the packet drop which is caused by the malicious nodes at the same time reduces the traffic load and increase the network life time. The proposed approach includes the fusion of data at the intermediate nodes simultaneously providing security to the network. This paper proposes an efficient lock based QoS-Aware Data Fusion for Wireless Sensor Networks. This makes the network secure and reliable. Hence, it can be efficiently used in real time applications such as in military applications where sensor networks may be exposed to hostile environment.

REFERENCES

[16] James Kay and Jeff Frolik, "Quality of Service Analysis and Control for Wireless Sensor Networks", in

