PROTOCAL FOR SECURE DATA AGREEGATION TECHINIQUES USE FOR WIRELESS SENSOR NETWORKS IN THE PRESENCE OF COLLUSION ATTACKS

K. Suganya* & A.Anitha**

* PG Scholar, Department of Master of Computer Applications, Dhanalakshmi Srinivasan Engineering College, Perambalur, Tamilnadu

** Assistant Professor, Department of Master of Computer Applications, Dhanalakshmi Srinivasan Engineering College, Perambalur, Tamilnadu

Abstract:

Due to limited computational power and energy resources, aggregation of data from multiple sensor nodes done at the aggregating node is usually accomplished by simple methods such as averaging. However such aggregation is known to be highly vulnerable to node compromising attacks. Since WSN are usually unattended and without tamper resistant hardware, they are highly susceptible to such attacks. Thus, ascertaining trustworthiness of data and reputation of sensor nodes is crucial for WSN. As the performance of very low power processors dramatically improves, future aggregator nodes will be capable of performing more sophisticated data aggregation algorithms, thus making WSN less vulnerable. Iterative filtering algorithms hold great promise for such a purpose. Such algorithms simultaneously aggregate data from multiple sources and provide trust assessment of these sources, usually in a form of corresponding weight factors assigned to data provided by each source. In this paper we demonstrate that several existing iterative filtering algorithms, while significantly more robust against collusion attacks than the simple averaging methods, are nevertheless susceptive to a novel sophisticated collusion attack we introduce. To address this security issue, we propose an improvement for iterative filtering techniques by providing an initial approximation for such algorithms which makes them not only collusion robust, but also more accurate and faster converging.

Introduction:

As wireless network increases in leaps and bounds because of the inherent benefits, the numbers of challenges also increase proportionally. The existing challenging issues like energy, throughput, link detection, resources, overheads, security etc. have to be overcome as newer problems crop up. This is due to the absence of topology and lack of continuous power supply to the devices operating the networks. All wireless WSN Networks basically are designed for situations having no preinstalled infrastructure and these networks impose severe limitations which are the challenges to be addressed on the participating nodes. These include energy constraints, low processing power, and the need for operating in challenging environments.

The main causes were due to the frequently changing topology as a result of mobility, obstacles in the path of transmission resulting in or leading to node failures which are common. Understanding the MANET WSN is very vital in that it is an autonomous collection of mobile nodes that communicate over relatively constrained bandwidth wireless links with the nodes moving dynamically in different routes, the network topology is mobile which may change unpredictably and rapidly over a period of time. Such mobile networks are decentralized, where all network activity like discovering the (neighborhood) topology and also delivering messages at the same time must be executed by the nodes themselves. In short the routing functionality is incorporated into the mobile nodes themselves.

Thus summing up, the major challenges in ad hoc routing pertaining include: (1) Fewer Overheads and (2) good Route (path) selection, and (3) control over energy. (4) Reliable link (neighborhood) detection, and (5) accurate information about the topology. Many novel solutions and approaches have been presented earlier, which have one drawback or another. However, simple solutions have been either eluded or ignored. Existing solutions consume more overheads and when the focus shifts to fewer resources usage there is a drop in throughput or security lags.

Modules Description:

WSN Setup and Configuration:

In this module Wireless Sensor nodes are setup with information collection and dissemination to the sink or master node. The WSN are placed in a remote locations with a sink connected to the network. According to the number of cluster heads, the nodes are randomly placed in a network. As events occur randomly the WSNS transmit the datas to the sink node or master node. Each node is assumed to be calculating the energy independently. The data transmission takes places. Whenever the particular node is used for data transmission, an energy level should be reduced.

The WSNs which act as relays also lose energy when relaying the datas of the WSN's. Thus each node is acting independently when event occurs and transmits energy according to differing energy levels. The multi-channel contributes in enhancing the network performance and extending its lifetime duration when the network density is high. As future works, we plan to evaluate the performance of our algorithm on a largest network topology. A theoretical study should be done to determine the upper working limit of such large networks. A dynamic distributed algorithm which dynamically assigns one frequency channel per cluster when the sinks are assumed randomly deployed into a monitoring region. Finally, the sinks' mobility could be considered according to application requirements.

Collusion Attack:

The adversary can launches a sinkhole attack: he compromises a few nodes, uses the cryptographic information obtained from the compromised nodes to produce replicas, and then inserts the replicas into the network. The compromised nodes and replicas are fully controlled by the adversary and can communicate with each other at any time. Also, same as previous protocols we assume nodes controlled by the adversary still follow the replica clone-detection protocol, since the adversary always wants to keep him unnoticed to others. The adversary will try to protect its replicas. Assume that the nodes are stationary, at least during the execution of replica-detection protocol. Each node has a private key K−1 and can use the private key to sign its location claim. Other nodes are also able to verify the signature. Now several libraries for sensor networks are available. We also assume the communications between any two nodes are protected

Detection Module – If Algorithm:

This is because if any replicas are detected, besides starting a revoke process to revoke the replicas, the network may start a sweeping process to sweep the compromised nodes out and may draw during the execution of clone replica detection protocol, the adversary can select a limited number of nodes to disable (i.e., compromise or jam) for protecting his replicas. He is able to do that because the time taken by the execution of protocol may be long enough (e.g., the delay caused by synchronization error, processing delay in each hops, and sleep schedule of the network). Also, since jamming a node is more quickly than compromising it, the adversary can jam a node first and compromise it later. A general assumption that is the adversary can disable a small number of nodes only. Each node broadcasts a signed location claim used witness's node to detect the clone.

Each of the node's neighbors probabilistically forwards the claim to some randomly selected nodes. If a sensor node meets another sensor node at an earlier time and sends a random number to at that time. Then when they meet again, can ascertain whether this is the node met before by requesting the random number. The effectiveness relies on the simple challenge and response framework, which obviously holds. Nevertheless, the performance varies according to different network settings. Thus, this section is devoted to validating the effectiveness through a simulation. Within a period of time with length properly chosen according to the offline step, the number of encounters with the genuine node and the number of encounters with the replicas can be distinguished well if the threshold is set in a way indicated. We discuss how the parameters, such as communication range and node velocity, affect the detection.

The extremely Efficient Detection linear model and Efficient Distributed Detection closed form model, for byzantine attack detection in sensor networks. The idea behind is motivated by the observation that , if a sensor node meets another sensor node at an earlier time and sends a random number to at that time, then, when and meet again, can ascertain whether this is the node met before by requesting the random number. Note that, in XED, we assume that the nodes cannot collude with each other but this assumption will be removed in the next solution. In addition, all of the exchanged messages should be signed unless specifically noted. Specifically, the scheme is composed of two steps: an offline step and an online step. The former is executed before sensor deployment while the latter is executed by each node after deployment.

Detection:

Each node a broadcasts signed locations claim to its neighbors. The claim has such a format <1><2> where 1 and 2 is A's location (e.g., location (x,y) in 2D) and _ is the concatenation. When hearing the claim, each neighbor verifies the signature and checks the plausibility of la (e.g., the distance between two neighbors cannot be bigger than the transmission range). Then with probability p, each neighbor randomly selects g nodes (or g locations4) and uses geographic routing to forward the claim to the g nodes (or nodes closest to the chosen g locations). Each chosen node that receives the claim of a, first verifies the signature. Then it stores the claim and becomes a witness node of a. Also, it will start a t-step random walk in the network (t is a system parameter, and we will analyze its value in by sending the location claim together with a counter of walked steps (sc) initiated to 1, to a random neighbor. The neighbor will also become a witness node of a. It adds counter sc by one and continues to forward the message to a random neighbor, unless counter sc reaches t. A node finds a collision (two different location claims with a same node ID), the node will broadcast the two conflicting claims as evidence to revoke the replicas. Each node receiving the two claims independently verifies the signatures. If the two signatures are valid, it terminates the links with replicas.

It is easy to see that the physical node of the starting node has the biggest walked times in a t-step random walk. If they walked times of this node is still less than 2, then they walked times of all the other visited physical nodes are also less than 2, and the number of selected physical nodes by a random walk must be no less than t/2. When analyzing the walked times of the starting node, we find it is hard to use general concepts in random walk, such as hitting time and commute time.

Revocation:

Then we describe the process of revocation. When receiving a location claim, a node will first find the entries which have the same node ID as the claim in its trace table. Then if any entry is found, the node will compute the digest of the claim using equation 1 and compare the digest with the digest in the entry. When the two digests are different, the node detects a clone attack. If the node stored the location claim of the entry, it will flood the network with the two location claims to revoke replicas. Otherwise it will flood a HELPREV request with only one location claim. Any node receiving the HELPREV message will check locally that if it stored a location claim conflicting with the received one. If such a location claim is found, it will flood the stored location claim into the network as evidence. In such revocation process an algorithm for duplicate message suppression can be employed. This is because when receiving a location claim, a witness node will compare the claim's node ID in its trace table at first. Thus even if two nodes' location claims passing the witness node have the same claim Digest, given that the two nodes have different node IDs, they will not be falsely detected as a clone attack.

The proposed simplified, linear q-out-of-m scheme that can be easily applied to large size networks. The basic idea is to find the optimal scheme parameters at relatively small network sizes through exhaustive search, and then obtain the fusion parameters for large network size by exploiting the approximately linear relationship between the scheme parameters and the network size.

It is observed that the proposed linear approach can achieve satisfying accuracy with low false alarm rate. However, there are chances of violating the problem constraint. To enforce the miss detection constraint and improve the data fusion accuracy. It is further proposed to use the linear approximation as the initial point for the optimal exhaustive search algorithm. With this enhanced linear approach, near-optimal solutions can be obtained with much lower computational complexity compared with that of the pure exhaustive search approach. This will enhance the efficiency of the WSN detection rate by upto 90%. Also in an effort to search for an easier and more flexible distributed data fusion solutions that can easily adapt to unpredictable environmental changes and cognitive behavior of malicious nodes, we plan derive a closed-form solution for the q-out-of-m fusion scheme based on the central limit theorem.

Existing System:

In recent years, there has been an increasing amount of literature on IF algorithms for trust and reputation systems. The performance of IF algorithms in the presence of different types of faults and simple false data injection attacks has been studied where it was applied to compressive sensing data in WSNs.

In the past literature it was found that these algorithms exhibit better robustness compared to the simple averaging techniques; however, the past research did not take into account more sophisticated collusion attack scenarios. If the attackers have a high level of knowledge about the aggregation algorithm and its parameters, they can conduct sophisticated attacks on WSNs by exploiting false data injection through a number of compromised nodes.

Disadvantages of Existing System:

Although the existing IF algorithms consider simple cheating behaviour by adversaries, none of them take into account sophisticated malicious scenarios such as collusion attacks. Although the existing IF algorithms consider simple cheating behaviour by adversaries, none of them take into account sophisticated malicious scenarios such as collusion attacks.

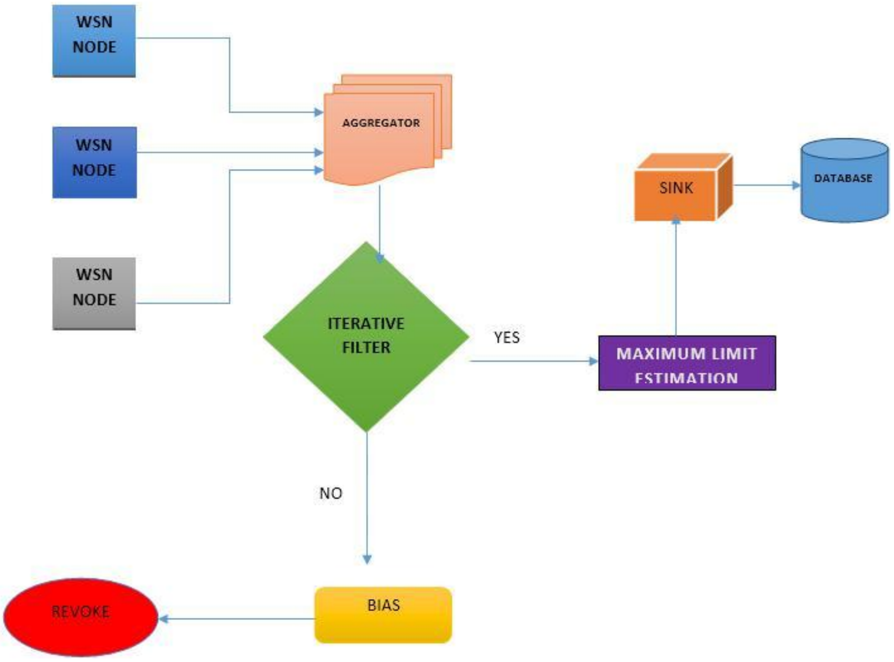
Proposed System:

This paper presents a new sophisticated collusion attack scenario against a number of existing IF algorithms based on the false data injection. In such an attack scenario, colluders attempt to skew the aggregate value by forcing such IF algorithms to converge to skewed values provided by one of the attackers. In this paper, we propose a solution for vulnerability by providing an initial trust estimate which is based on a robust estimation of errors of individual sensors. Identification of a new sophisticated collusion attack against IF based reputation systems which reveals a severe vulnerability of IF algorithms. A novel method for estimation of sensors' errors which is effective in a wide range of sensor faults and not susceptible to the described attack. Design of an efficient and robust aggregation method inspired by the MLE, which utilises an estimate of the noise parameters obtained using contribution above. Enhanced IF schemes able to protect against sophisticated collusion attacks by providing an initial estimate of trustworthiness of sensors using inputs from contributions.

Advantages of Proposed System:

We provide a thorough empirical evaluation effectiveness and efficiency of our proposed aggregation method. The results show that our method provides both higher accuracy and better collusion resistance than the existing methods. To the best of our knowledge, no existing work addresses on false data injection for a number of simple attack scenarios, in the case of a collusion attack by compromised nodes in a manner which employs high level knowledge about data aggregation algorithm used.

System Architecture:



System Maintenance:

Wireless sensor networks hold great promise as an enabling technology for a variety of applications. Data collection and event detection are two such classes of applications that are broadly representative and which have received considerable attention in the literature. While wireless multi-hop data collection has achieved operational lifetimes on the order of a year, we are unaware of lifetimes exceeding a few days or weeks for wireless multi-hop event detection sensor networks. This project is that sensor networks for event detection are constrained by two factors which do not similarly affect data collection sensor networks. The first factor is that no appropriate sensing, signal conditioning, and signal processing architecture has been broadly implemented to support event detection in distributed systems that are simultaneously energy, space, time, and message complexity-constrained. The second factor is that middleware for services such as time synchronization, localization, and routing are predominantly and unnecessarily proactive.

A comparison of data collection and event detection will serve to illustrate the subtle but important differences between these applications. Fundamentally, data collection is a signal reconstruction problem in which the objective is to centrally reconstruct observations of distributed phenomena with high spatial and temporal fidelity. Performance metrics for such applications include the accuracy and precision of the signal reconstruction, the correlation between the observed signal and the underlying physical phenomena, and the lifetime of the sensor network.

Physical phenomena such as light, temperature, humidity, and barometric pressure change at very low frequencies and can be sampled faithfully at periods of a minute or more. System performance can be adjusted by introducing compression and aggregation, or by varying the duty-cycle, sampling and communication rates, allowing sensor lifetimes to approach a year or more. In contrast with data collection, sensor network applications for event detection must continuously observe noise for the rare presence of a burst of high-frequency signal.

Screen Shots:

Output Screens:

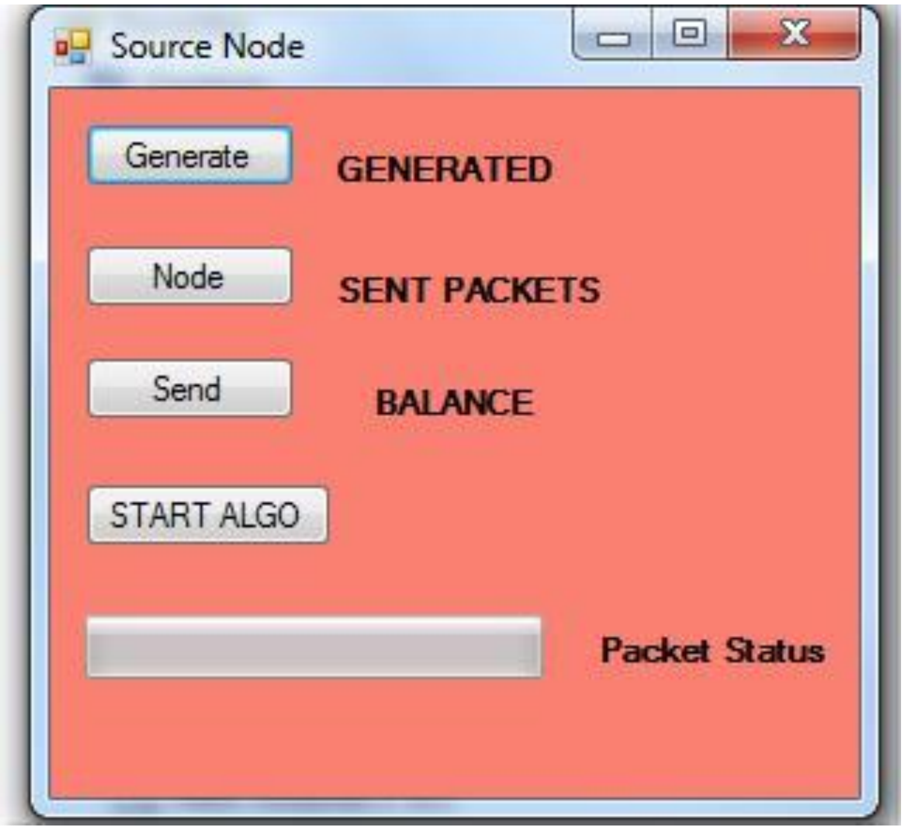
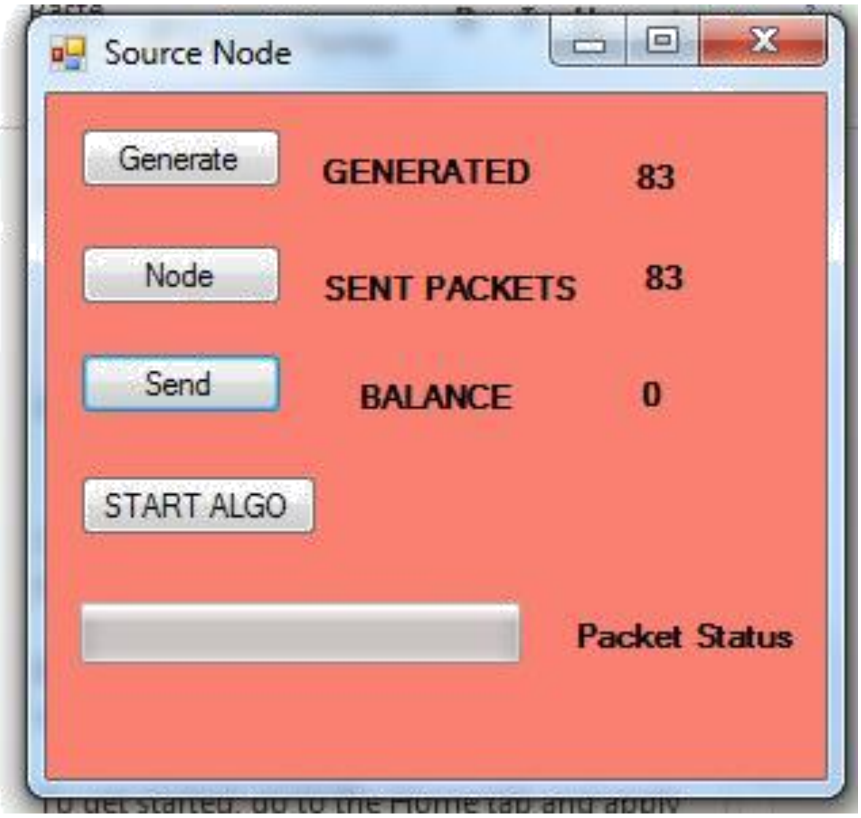


Figure 2: Home Page for node



Source Node

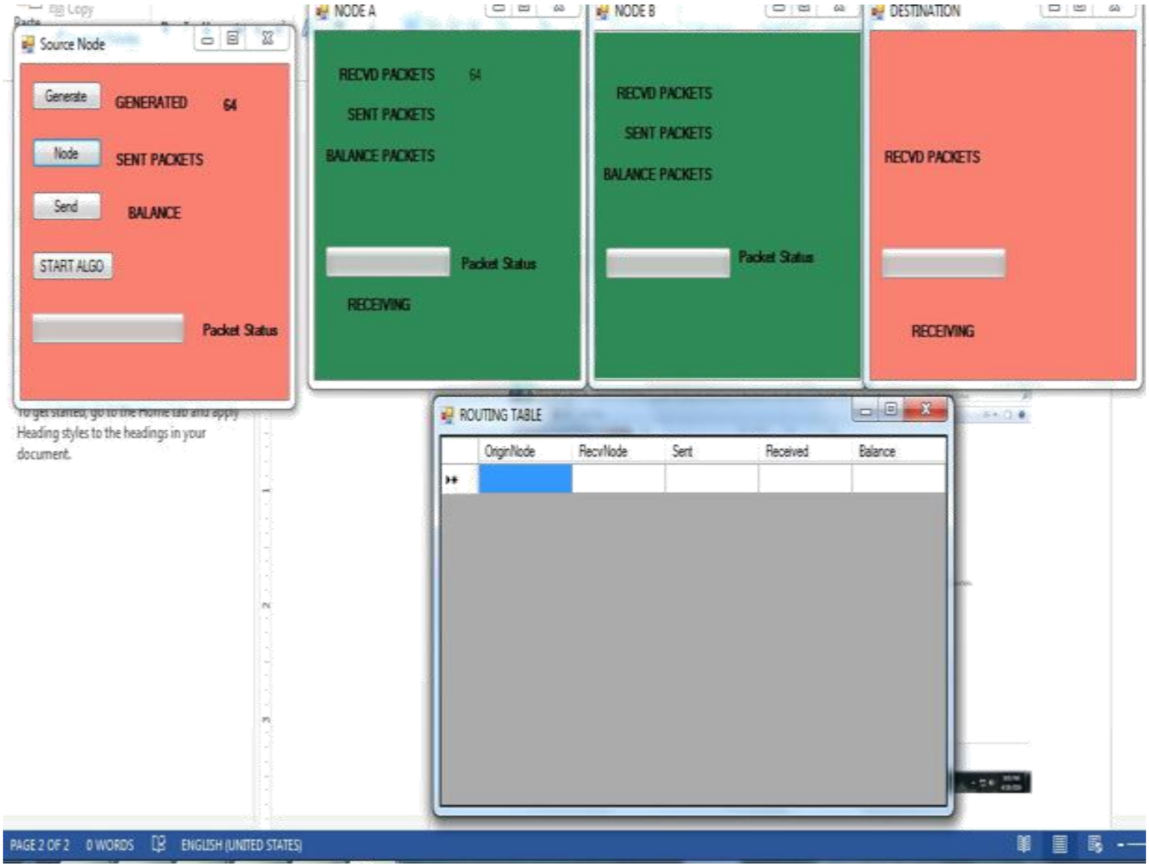
GENERATED 83

SENT PACKETS 83

BALANCE 0

Packet Status

Figure 3: Node Status



Source Node

GENERATED 64

SENT PACKETS

BALANCE

Packet Status

NODE A

RECVD PACKETS 64

SENT PACKETS

BALANCE PACKETS

Packet Status

RECEIVING

NODE B

RECVD PACKETS

SENT PACKETS

BALANCE PACKETS

Packet Status

DESTINATION

RECVD PACKETS

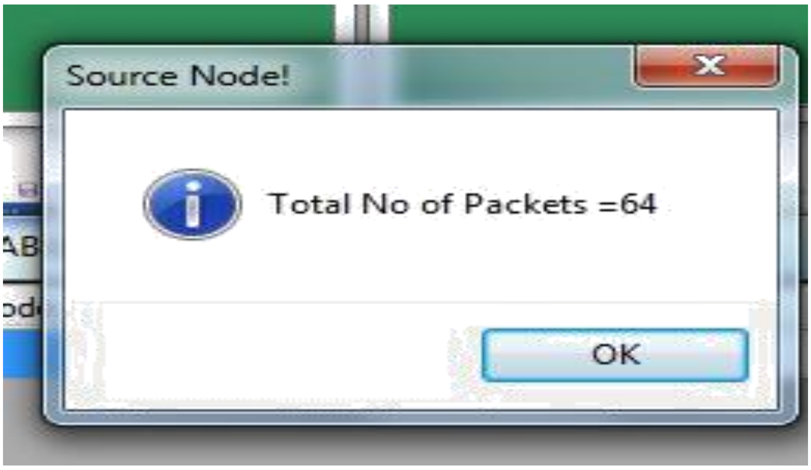
RECEIVING

ROUTING TABLE

OriginNode	RecvNode	Sent	Received	Balance

Generate
Node
Send
START ALGO
To get started, go to the Home tab and apply Heading styles to the headings in your document.
PAGE 2 OF 2
0 WORDS
ENGLISH (UNITED STATES)

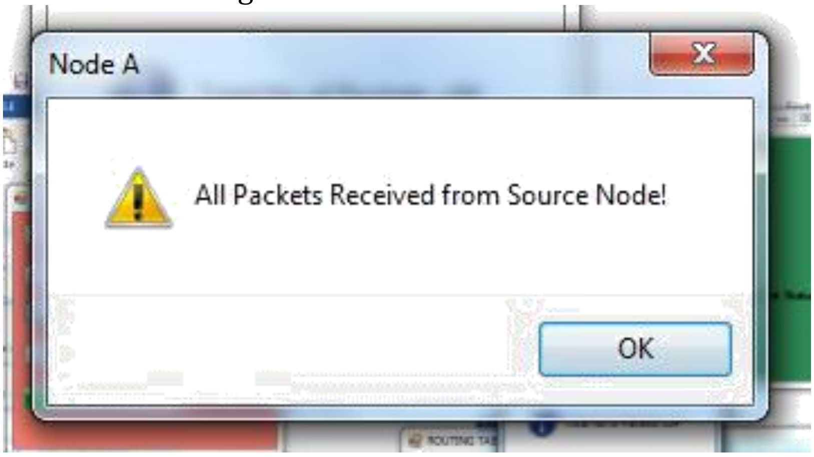
Figure 4: Multiple Nodes



Source Node!

Total No of Packets =64

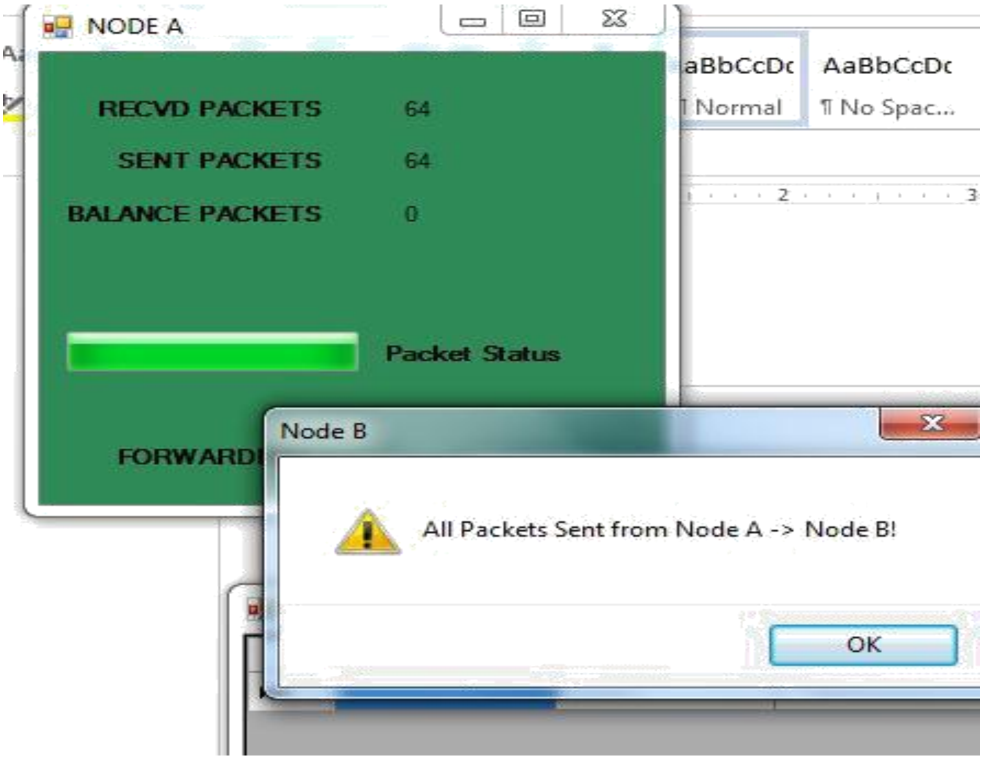
Figure 5: Total Packets of Root



Node A

All Packets Received from Source Node!

Figure 6: Received From Source Node



NODE A

RECVD PACKETS 64

SENT PACKETS 64

BALANCE PACKETS 0

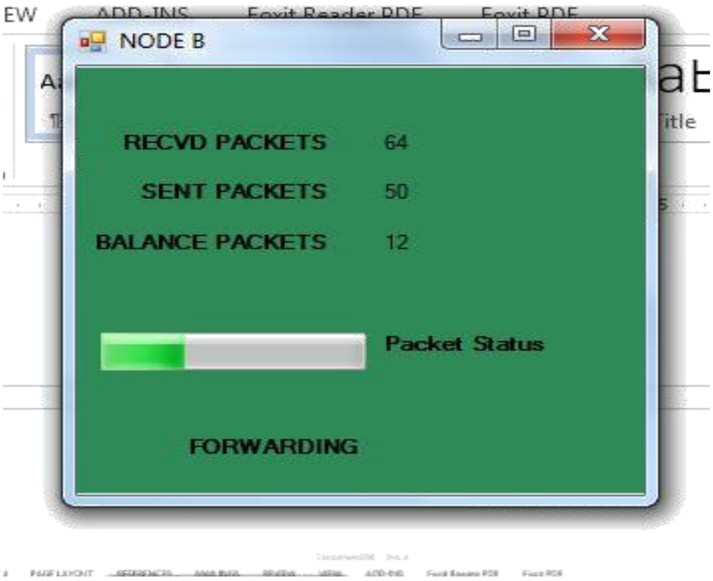
Packet Status

FORWARDI

Node B

All Packets Sent from Node A -> Node B!

Figure 7: Total Number of Send and Received Packets

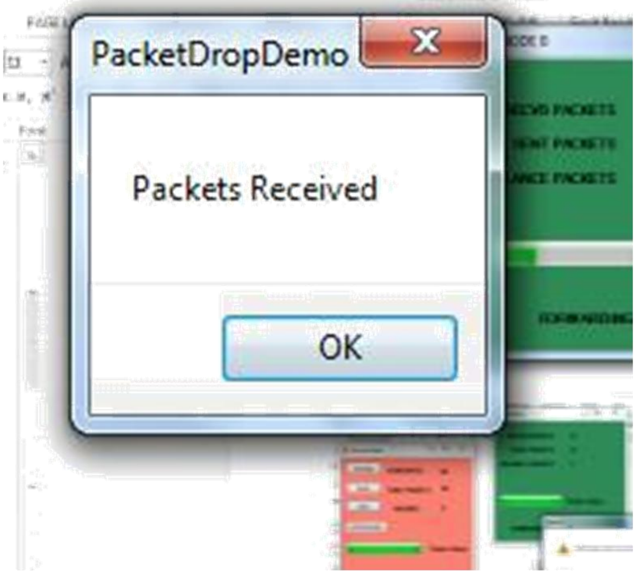


RECVD PACKETS 64

SENT PACKETS 50

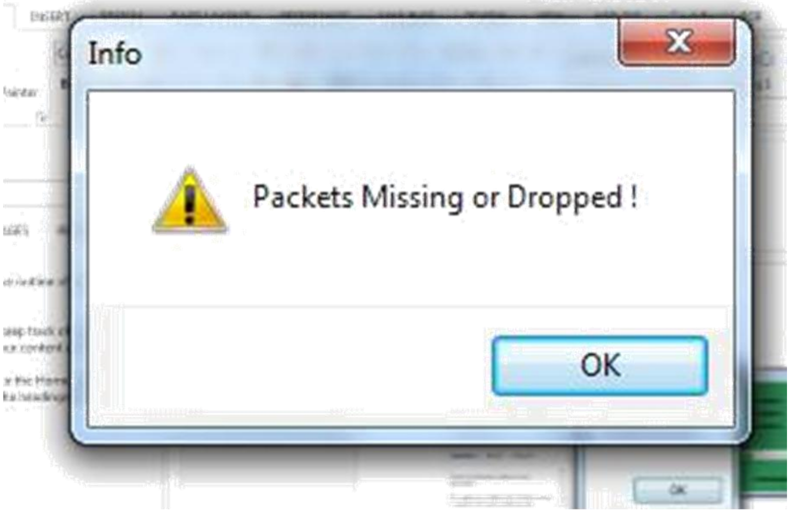
BALANCE PACKETS 12

Figure 8: Total Number of Packets Status



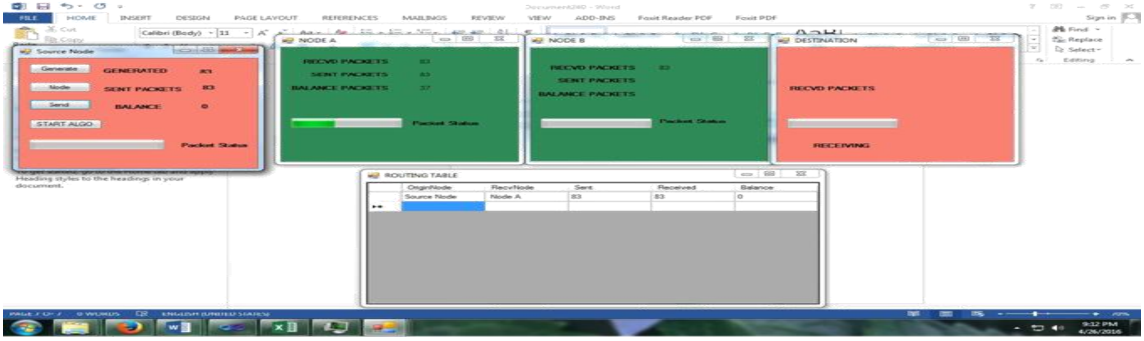
Packets Received

Figure 9: Packets Received



Packets Missing or Dropped !

Figure 10: Packets Missing or Dropped



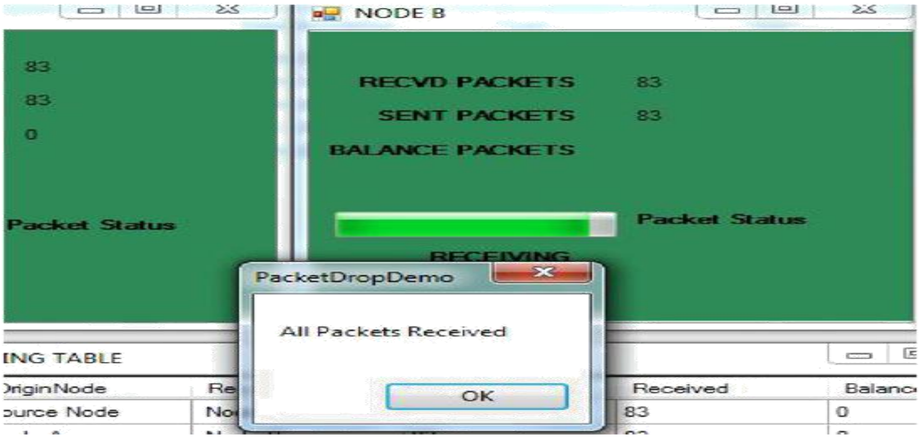
GENERATED 83
SENT PACKETS 83
BALANCE 0

RECVD PACKETS 83
SENT PACKETS 83
BALANCE PACKETS 37

RECVD PACKETS 83
SENT PACKETS
BALANCE PACKETS

RECVD PACKETS

OriginNode	RecvNode	Sent	Received	Balance
Source Node	Node A	83	83	0



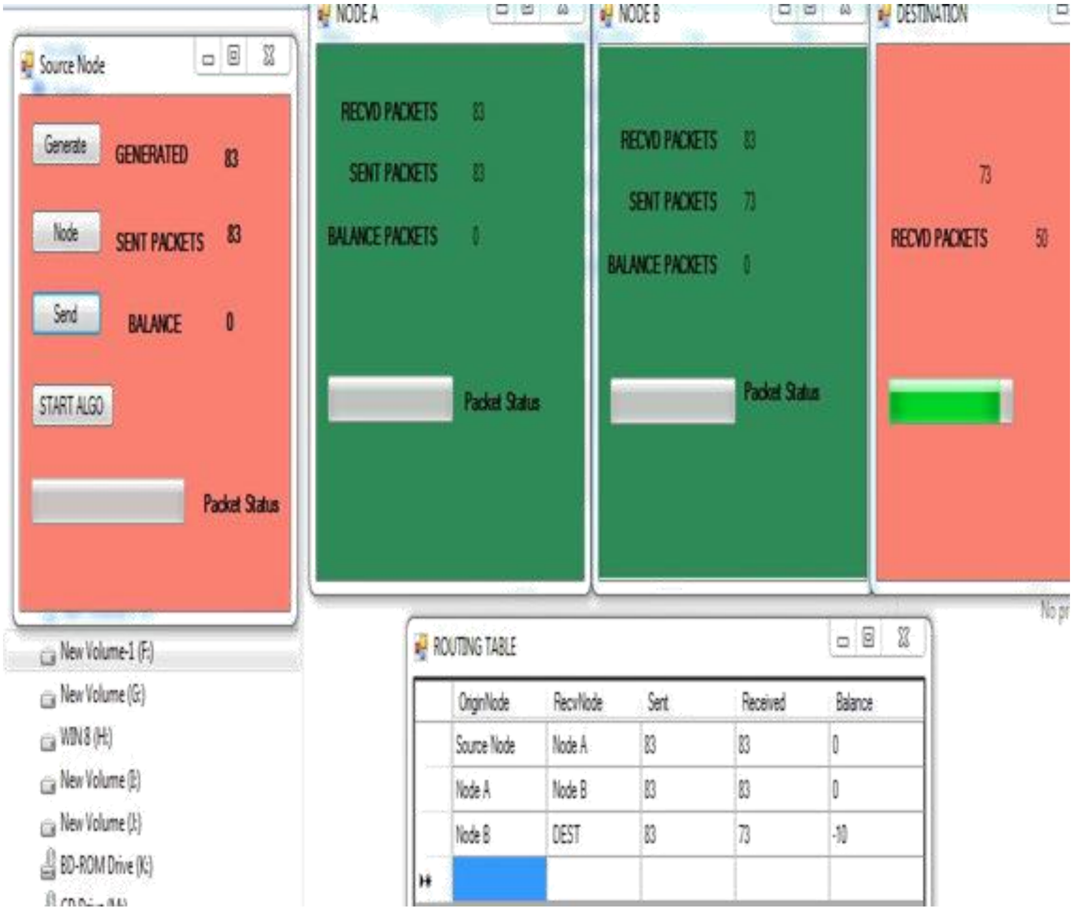
83
83
0

RECVD PACKETS 83
SENT PACKETS 83
BALANCE PACKETS

All Packets Received

)riginNode	Re	Received	Balanc
ɔurce Node	No	83	0

Figure 11: All Packets Received By Node B



GENERATED 83
SENT PACKETS 83
BALANCE 0

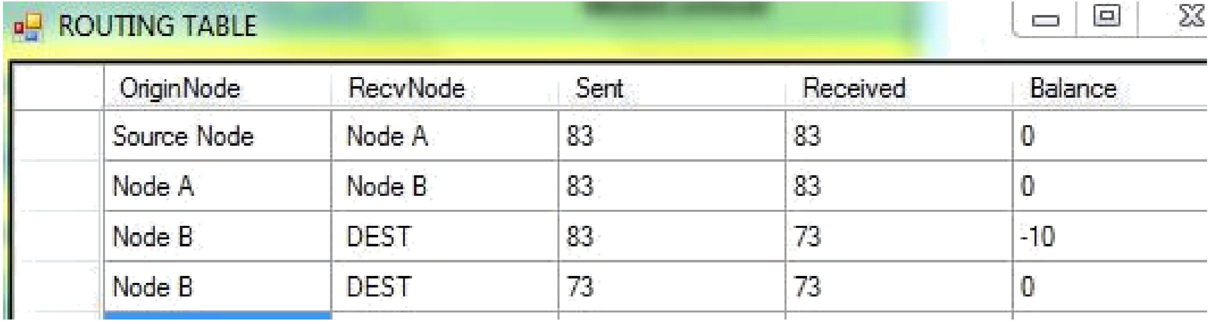
RECVD PACKETS 83
SENT PACKETS 83
BALANCE PACKETS 0

RECVD PACKETS 83
SENT PACKETS 73
BALANCE PACKETS 0

73
RECVD PACKETS 50

OriginNode	RecvNode	Sent	Received	Balance
Source Node	Node A	83	83	0
Node A	Node B	83	83	0
Node B	DEST	83	73	-10

Figure 12: Status of Node A and B



	OriginNode	RecvNode	Sent	Received	Balance
	Source Node	Node A	83	83	0
	Node A	Node B	83	83	0
	Node B	DEST	83	73	-10
	Node B	DEST	73	73	0

Figure 13: Routing Table

Conclusion and Future Enhancement:

Thus the proposed model is superior to the existing WSN collusion routing protocols considering the energy awareness model, link detection, trust value, selfish node detection, and throughput along with low overheads. Thus the model proves decisively in case of models with low cost and high values solving certain inherent problems in WSN ad hoc routing in the process. The model provides a significant amount of overhead reduction while being simple to implement and integrate. In future this may be adapted to Wireless Sensor networks, opportunistic routing networks and sensor networks as well.

References:

1. J. Bao, Y. Zheng, M.F. Mokbel, "Location-based and Preference Aware Recommendation using Sparse Geo-Social Networking Data," In Proceeding of 20th International Conference on Advances in Geographic Information Systems, ACM New York, pp.199-208, 2012.
2. M. Ribeiro, A. Lacerda, A. Veloso, and N. Ziviani, "ParetoEfficient Hybridization for Multi-objective Recommender Systems," In Proceeding of 6th ACM Conference on Recommender Systems, pp. 19-26, 2012.
3. K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A Fast and Elitist Multi-objective Genetic Algorithm: NSGA-II," IEEE Transaction on Evolutionary Computations, vol. 6, no. 2, pp. 182- 197, 2002.
4. H. Nasiri, M. Maghfoori, "Multiobjective Weighted Sum Approach Model reduction by Routh-Pade approximation using Harmony Search," Turkish Journal of Electrical Engineering and Computer Science, vol. 21, no. 2, pp. 2283-2293, 2013.
5. J. Abimbola, "A Non-linear Weights Selection in Weighted Sum Information, vol. 27, no. 3, 2012. [19] Paired t test. Wiley Encyclopedia of Clinical Trials, 2008