5
Consensus Aware Sociopsychological Model for WSN

Sociopsychological trust model is an efficient model which uses three factors namely, ability, benevolence and integrity for developing the trust of sensor nodes. The performance metrics such as detection and recovery time, scalability, reliability, space and time complexity presented the model as efficient in comparison to other models. However, the limitation of sociopsychological model was the efficiency in detection of colluding fraudulent nodes i.e. the model failed when the number of colluding fraudulent sensor nodes outnumbered the number of benevolent nodes in the network. In this chapter, a solution is provided to the above problem which in turn also proved to be a solution for colluding data nodes attack in WSN.

The consensus aware sociopsychological model uses three factors namely, ability, benevolence and integrity for computing trust of sensor nodes. Additionally, it uses the consensus approach for dealing with the colluding node problem. Consensus is mostly used in problem where an agreement is required. The consensus aware psychological model proved efficient even if the number of colluding fraudulent sensor nodes outnumbers the number of benevolent nodes in the network. Likewise, the model also provides a solution for colluding data nodes attack in WSN. The organization of the chapter is as follows: Section 5.1 illustrates basic building blocks for the proposed work. In Section 5.2, the consensus aware sociopsychological model is described for detection of fraudulent nodes in the network. The new model uses consensus approach to increase the detection rate of colluding fraudulent nodes close to *100%*. Section 5.3 presents the performance analysis of the proposed work. Section 5.4 concludes the discussion and suggests some possible future directions.

5.1 BUILDING BLOCKS FOR CONSENSUS AWARE SOCIOPSYCHOLOGICAL MODEL

Trust is an essential requirement in decision making of any network where uncertainty is a factor. Due to limited computational power and energy resources in WSN, methodologies used for computing trust should focus on these limitations. The proposed work is lightweight and proves to be an efficient model for detection of fraudulent nodes in the network. The model is efficient even in the presence of higher number of fraudulent nodes in comparison to the benevolent nodes in the network. Further, it also addresses the issue of colluding data nodes [Rezvani et al, 2014; Kerchove and Dooren, 2010; Rezvani et al, 2013].

The section explains sociopsychological concepts for building trust and agreement. Trust is the glue that holds the agreement/people together. Trust is critical for negotiations which enable us in constructing an agreement. It enhances the enforcement of deals and helps in determining the creditability of others.

5.1.1 Sociopsychological Account: Trust Model

As discussed in Chapter 4, an in-depth analysis of sociopsychological account unveils that trust is a function of three factors: ability (A), benevolence (B) and integrity (I) where:
Ability denotes the capacity of a person in performing a given task.
Benevolence is the stable disposition which capacitates the trustee to do good for the trustor.
Integrity is the virtue to confirm consistency and coherence of positive outputs the person had delivered in the past.

The optimum trust between individuals appears when there exists a perfect congregation of ability, benevolence and integrity.

5.1.2 Building Trust in a Group

Cooperative and collaborative learning are increasingly influential areas of research and practice in psychology and education. Cooperative learning refers to the learning that occurs when a group works together to accomplish shared learning goals [Hogan, 2013]. The group members should have a higher level of dispositional trust amongst each other. In social psychology, the behavior of people in a group is judged by attaching meaning to the information they provide, with the help of attribution theory. Attribution theory is a core element of social-psychological thinking [Malle, 2003]. It deals in giving explanation to the behavior of individuals. Attribution theory deals with how the social perceiver uses information to arrive at causal explanations for the events. It examines what information is gathered and how it is combined to form a causal judgement [Fiske and Taylor, 2013]. Kelley's covariation model is a model which explains the attribution theory [Kelley, 1967]. The term covariation can be defined in a situation where a person receives information from multiple observations, at different times and situations. The person can perceive the covariation of an observed effect and its causes. It has three basic rules in reaching to the judgement:

(1) *Consensus*: it gives the extent to which other people behave in the same way in a similar situation.
(2) *Distinctiveness*: it gives the extent to which the person behaves in the same way in different situations.
(3) *Consistency*: it gives the extent to which the person gives consistent results in similar situations.

Among the above factors, consensus and consistency can play an important role in building trust in WSN since they focus on reaching to a judgment in similar situations. Consistency gives the extent of coherent results produced by the person in similar situations. Thus, consistency can be computed with the help of integrity factor which is used for modelling trust. Computation of consensus is explained in the following section.

5.1.3 Consensus Algorithm

Consensus means arriving at a common agreement. The consensus algorithm refers to the rule that defines how to reach to an agreement, such as the position, velocity, orientation etc., among a group of autonomous agents in a dynamical system. Consensus algorithms are used in multi agent systems where agents are connected in a distributed manner with each other. Since the sensor nodes are autonomous, having the capability of making independent decisions, agents can be perceived as the wireless sensor nodes [Wooldridge, 2009]. At any given time t node will have some state information x. State of node I is defined at a given time t as $x_i(t) \epsilon R$ where $t \epsilon R$. State information can be anything like the position, velocity, temperature, orientation etc. Here state information can be said as the data acquired by the node. In order to obtain the consensus, dynamics of the network needs to be understood. Laplacian matrix helps in understanding the dynamics of the network.

WSN can be considered as a dynamical system which has a state vector to describe the states/measurements of all the nodes. The topology of WSN can be represented in the form of a tree or a graph as shown in Figure 5.2. It can be considered as a weighted tree $T_\omega(S, \omega)$, where $S = \{s_1, s_2, s_3 \ldots s_n\}$ be the set of vertices, $\omega = \{\omega_1, \omega_2 \ldots \omega_n\}$ be the corresponding importance/weights of the vertices. For weighted graphs, the adjacency matrix A can be defined as follows,

$$a\omega_{ij} = \begin{cases} \omega_j & if\ i\ and\ j\ are\ adjacent \\ 0 & else \end{cases} \quad (5.1)$$

The adjacency matrix tells the connection cost between any two adjacent vertices, which depends on the vertex weights. The Laplacian matrix is defined as follows [Merris, 1994]:

$$L\omega_{ij} = \begin{cases} \sum_{j \neq i} a\omega_{ij} & if\ i = j \\ -aw_j & otherwise \end{cases} \quad (5.2)$$

Laplacian matrix represents the structure of a system and by using it in consensus algorithms gives insight knowledge into the dynamics of a system. Some of the properties related to Laplacian matrix are [Fiedler 1973]:

(1) Row sum of the matrix is equal to zero.
(2) The Laplacian matrix L_ω has eigenvalues $0 = \lambda_1 \geq \lambda_2 \geq \ldots.. \geq \lambda_n$ which states at least one eigenvalue is equal to zero i.e. $\lambda_1 = 0$.
(3) The right eigenvector $[v_r]$ corresponding to λ_1 is $\alpha.S_{1\times N}$ and the left eigenvector $[v_l]$ corresponding to λ_1 is $\beta.S_{1\times N}$ where $S_{1\times N}$ is a vector of all 1.
(4) The weighted Laplacian Matrix L_ω satisfies $[v_l]\ L_\omega = L_\omega\ [v_r] = 0$ which leads to:
$$U_N = \lim_{t\to\infty} e^{-L_\omega t} = [v_r][v_l]^T \quad (5.3)$$
where U_N is a $N \times N$ square matrix, each of its element is equal to $1/\sum_{i=1}^{N} \omega_i$.
(5) For a connected network, the second smallest eigenvalue of the matrix L_ω is $\lambda_2 > 0$.

Now, the rate of change of information of the agent/node which is solely dependent on the neighbouring state and its weight can be defined as:

$$\dot{x}(t) = \sum_{j=1, j\neq i}^{N} a\omega_{ij}(x_j(t) - x_i(t)) \quad (5.4)$$

By solving the above consensus algorithm for all the nodes in the system the following equation is obtained:

$$\dot{X}(t) = -L_\omega X(0) \quad (5.5)$$

where L_ω is the weighted Laplacian matrix, which is already defined in Eq. (5.2). The weighted Laplacian matrix for a connected network has row sum equal to 0 which has the lowest eigenvalue $\lambda_{\omega 1} = 0$.

Now consider a weighted multi-node sensor network $T(S, \omega)$, where S is the set of nodes and ω is the respective trust/weights of nodes in the network. By applying the above linear consensus algorithm for such weighted multi-node WSN, following weighted average consensus result is obtained [Porfiri and Stilwell, 2007]:

$$\alpha = \frac{\sum_{i=1}^{N} \omega_i x_i}{\sum_{j=1}^{N} \omega_j} \quad (5.6)$$

The linear consensus algorithm for the weighted multi-node WSN results to a weighted Laplacian matrix L_ω. Solving the above linear consensus algorithm the following result is obtained:

$$X(t) = e^{-L_\omega t} X(0) \quad (5.7)$$

From properties 3 and 4 of Laplacian matrix the solution of the above linear consensus algorithm is obtained as [Olfati-Saber et al, 2007]:

$$X(t) = [v_{\omega r}]\left[v_{\omega l}^T\right] X(0) \quad (5.8)$$

Additionally, a weighted system achieves consensus only if it is connected and all agents have non zero weights. Consensus aware sociopsychological trust model is proposed in the next section.

5.2 CONSENSUS AWARE SOCIOPSYCHOLOGICAL MODEL

Taking inspiration from the sociopsychological account, the thesis introduces a novel process for computing trust of sensor nodes. Sociopsychological norms use factors like Ability *(A)*, Benevolence *(B)* and Integrity *(I)* for computing trust of a person as discussed in Chapter 4 with *Ability* calculated in a similar manner too. To detect maximum number of fraudulent nodes in the network, the *Benevolence* and the *Integrity* factors are computed with different criteria.

Ability of WSN is set to 1 if the sensor node is able to send the packets.

Benevolence of a node is calculated with the help of modified error function. In order to check the goodness of the sensor, the metric is derived as the difference between the actual and the perceived readings (error). For calculation of perceived readings the average of previous consensus output is used. Precisely, weights ω are treated as benevolence factor *(B)* (described in Section 5.1.2). It implies that the current readings which the sensor node sends are to be compared with the average of previous consensus output. More the error between the two, lesser is the benevolence factor. Similarly, lesser the error more is the benevolence factor of the sensor node.

Integrity of a node is calculated by checking the 'quality' of the sensor node. Here, the quality of the sensor node is evaluated in the light of the number of positive outcomes given by the sensor nodes from the initial time instant. In chapter 4, integrity was calculated with a history of 10 kept in mind. However, integrity in consensus aware sociopsychological model is computed by considering the positive outcomes from the start.

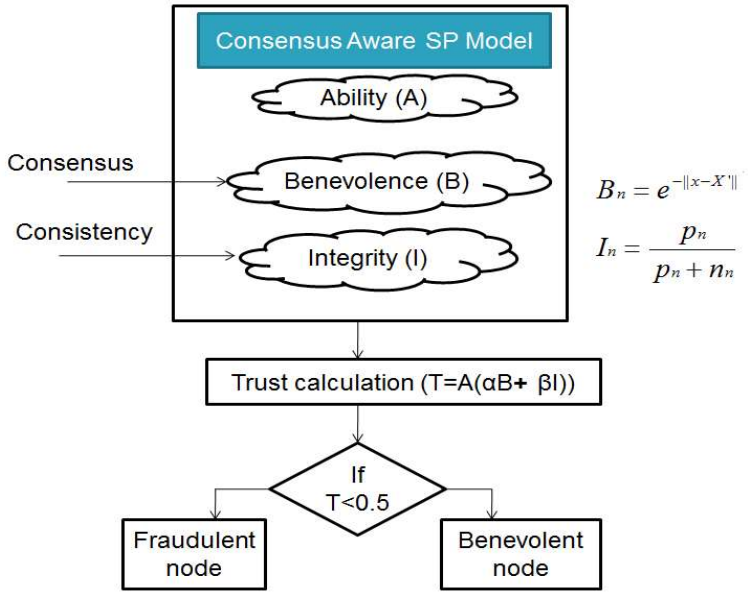


Figure 5.1: Flowchart for Consensus Aware Sociopsychological Trust Model

Trust *(T)* is then defined as the function of A, B and I, having weights α, β given to factors B and I respectively:

$$T_n = A_n(\alpha B_n + \beta I_n) \tag{5.9}$$

where $\alpha + \beta = 1$

$$A_n = \begin{cases} 1 & functional \\ 0 & otherwise \end{cases} \tag{5.10}$$

$$B_n = e^{-(||E_n||)} \tag{5.11}$$

$$I_n = \frac{p_n}{p_n + n_n} \tag{5.12}$$

where,

$B_n = \{\omega_n\}$ = weight of node

E = x_n - X_n'

x_n = current reading from a sensor node n

X_n'= the average of previous consensus output i.e. $\frac{\sum_{i=1}^{n-1} X_i}{n-1}$ and X_i is calculated using Eq. (5.6).

p = number of positive responses = number of times error is below the threshold in previous time steps i.e., $|B| < \delta$ (margin of error) as calculated in Chapter 4.

n = number of negative responses

The flowchart for consensus aware sociopsychological model is illustrated in Figure 5.1. It uses the concept of consensus and consistency in understanding the behavior of sensor nodes. While, consistency and integrity are closely associated with each other, consensus is used in computation of benevolence factor.

Algorithm 5.1 describes the consensus aware sociopsychological model. It is a type of centralised algorithm which runs at the host controller.

The justification of using weighted consensus approach is: A graph is said to be disconnected when there exists a vertex v and a vertex u which are not connected. In the weighted multi agent system, the connection between the nodes depends on its weights (Eq. (5.6)). Due to high deviation in E for fraudulent node, node's benevolence factor tends to zero which makes the weights of the nodes equal to zero. As a result, some elements of the adjacency matrix will be zero. Hence, the nodes with zero weights get disconnected from the system which in turn makes the consensus unachievable. To overcome this issue, weighted average of consensus result is used as shown in Eq. (5.6).

ALGORITHM 5.1 CONSENSUS AWARE SOCIOPSYCHOLOGICAL MODEL

Input:

n: Number of sensor nodes

A, B, I, T: Ability, Benevolence, Integrity, Trust respectively

X': Average of previous consensus

Sensor measurements ($x^{(1)}, x^{(2)}.....x^{(n)}$), where $x^{(i)} \epsilon R^{(n)}$

Output: Trust ratings for the nodes.

```
1  while till convergence do
2      for i← 1 to n do
3          if measurement received then
4              A_i = 1
5          Else
6              A_i = 0
7          B_i = e^{-(||x_i - x'_i||)}
8          if B_i < 0.001 then
9                  p_i = p_i+1;
10         else
11                 n_i = n_i+1;
12         I_i = p_i / (p_i+n_i)
13         T = 0.5(B_i + I_i)
14         if T < 0.5 then
15                 Node is fraudulent
16         Else
17                 Node is benevolent
```

5.3 EXPERIMENTS AND RESULTS

This section presents the performance analysis of the proposed consensus aware sociopsychological trust model, for different scenarios. Studies have been conducted to show the effectiveness of the model for different sized networks, impact of introduction of fraudulent nodes at the early/late stages on trust measurement, convergence of consensus in the presence/absence of fraudulent nodes. The performance studies are conducted on both simulations and experimental setup on a 16-node test bed. In addition, a comparative analysis is carried out between the proposed model and the machine learning based models for the performance metrics such as robustness, space and time complexity and attacks.

5.3.1 Simulation Results

Extensive simulations are carried out in LabVIEW for both smaller and larger network sizes. Ability of the sensor node is set to 1 since there are no communication faults in the network.

Studies over a Smaller Network: Simulations are carried out for a network size of 11 nodes where 6 of them behave in a fraudulent manner and send faulty information in the range [35-40] which is different from the benevolent range [20-30]. The topology of the network is as shown in Figure 5.2. S1 is the gateway node, S2 and S3 are the relay nodes and S4 to S11 are the end nodes. Nodes S6 to S11 start sending fraudulent information for the time ranging from 11 to 50 sec.

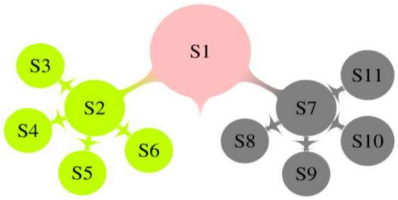


Figure 5.2: Topology of a Smaller Network: S1 is the gateway node, S2 and S7 are the relay nodes, S3 to S6 and S8 to S11 are the end nodes where S6 to S11 send fraudulent information

Table 5.1 shows a trace example of consensus aware sociopsychological model for three time instances. The sensor readings in the first three rows of the table indicate the measurements acquired at time instances 10, 11 and 12. For each iteration, the lower part of the table illustrates the computation of benevolence, integrity and trust ratings of the sensor nodes at the respective time instances. At time $t = 11$, 6 of the sensor nodes start sending fraudulent measurements. Due to this behavior, benevolence of these sensor nodes reduces drastically which in turn reduces the trust ratings of these nodes. As the fraudulent behavior affects the overall network, the reduction of trust ratings of benevolent nodes can be seen as well. Finally, the consensus output shows that despite having incorrect readings from fraudulent nodes, the gateway node is able to estimate genuine readings (around 25). This is because of the benevolence factor. The weights of the consensus are changed according to the benevolence factor which in turn helps in reaching the desired output. Consensus is achieved before $t = 11$ because when there is a fraudulent node in the system, reduction in its weight (importance) occurs, as a result the edge cost between the fraudulent node and any other node reduces. In other words change in the weight restricts the fraudulent node to share information with other connected node. A system is said to reach consensus only when all of the nodes in the system agrees to a common state (information). However, in this scenario, restriction happens by changing the weight of the fraudulent node. When all the nodes are benevolent, no reduction in

the weights of the nodes occurs; as a result the cost on all the edges is good enough, for every node to share information. So from the Eq. (5.7), the system reaches consensus.

Figures 5.3(a), (b) and (c) describes the corresponding benevolence, integrity and trust factor for the network. Nodes S6-S11 starts behaving fraudulently in the time period 11-50 sec which in turn reduces the benevolence factor. The band in the upper region is the benevolence of benevolent nodes and the fall depicts the benevolence of fraudulent nodes. Integrity factor also reduces since it is inversely proportional to the total number of responses which is the sum of positive and negative responses. Therefore, due to fraudulent data, the number of negative responses increases which in turn reduces the integrity factor.

Table 5.1: A trace Example of Consensus Aware Sociopsychological Model

Sensor readings												
Time	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	
10	22.2716	22.0654	21.6977	21.5970	25.7176	29.5207	29.3219	29.5807	20.0007	24.6329	23.6185	
11	25.9292	23.1773	23.0581	29.9275	27.2357	**36.5619**	**39.4781**	**39.3839**	**37.4326**	**38.3034**	**39.8594**	
12	28.3523	22.5027	23.5298	20.2951	29.9945	**38.4106**	**37.3479**	**36.6300**	**37.6309**	**35.4652**	**36.6292**	
Time	Sensor benevolence (B)											
10	0.2875	0.0191	0.0132	0.0119	0.7354	0.0303	0.0370	0.0286	0.0024	0.2486	0.0901	
11	0.9908	0.0632	0.0561	0.0185	0.2733	**2.43e-5**	**1.31e-5**	**1.44e-6**	**1.01e-5**	**4.26e-6**	**8.99e-7**	
12	0.1448	0.0199	0.0592	0.0022	0.0280	**6.20e-6**	**1.79e-5**	**3.68e-5**	**1.35e-4**	**1.18e-4**	**3.68e-5**	
Time	Sensor integrity (I)											
10	1	1	1	1	1	1	1	1	1	1	1	
11	1	1	1	1	1	**0.9091**	**0.9091**	**0.9091**	**0.9091**	**0.9091**	**0.9091**	
12	1	1	1	1	1	**0.8333**	**0.8333**	**0.8333**	**0.8333**	**0.8333**	**0.8333**	
Time	Trust ratings											**Consensus**
10	0.6437	0.5095	0.5066	0.5060	0.8677	0.5152	0.5185	0.5143	0.5012	0.6243	0.5451	25.1594
11	0.9934	0.5316	0.5281	0.5093	0.6366	**0.4546**	**0.4545**	**0.4545**	**0.4545**	**0.4545**	**0.4545**	25.9980
12	0.5734	0.5099	0.5296	0.5011	0.5140	**0.4167**	**0.4167**	**0.4167**	**0.4167**	**0.4167**	**0.4167**	25.0896

At time $t = 50$ sec, the nodes start giving benevolent readings which in turn increases the benevolence and integrity factors. The dashed line is the integrity of benevolent nodes and dark line is the integrity of fraudulent nodes. This type of attack, where sensor nodes change their behavior alternatively is known as ON-OFF attack [Lopez et al, 2010]. Figure 5.3(c) shows the corresponding change in the trust ratings when fraudulent nodes are introduced in the network. While there is no change in the trust factor of benevolent nodes, a significant drop in the trust ratings of the fraudulent nodes is observed.

Figure 5.4 gives the variation of consensus output. Without the benevolence factor, the consensus output is not within the normal range. However, when benevolence factor is taken into account, consensus result is within the valid range *[25-30].*

Figure 5.5(a) and (b) illustrate the result of the study to identify the impact of early/late introduction of fraudulent nodes on the rate of change in integrity factor. Nodes behave fraudulent for a period of 40 sec. Figure 5.5(a) illustrates the rate of change of integrity factor of fraudulent nodes at different time instances. Early introduction of fraudulent nodes has a high drop in integrity factor when compared to that of the late introduction of fraudulent nodes. This makes sense in sociopsychological account, if the person's trust is high for a longer time period; his/her integrity factor reduces slowly. Figure 5.5(b) presents the rate of change of increase in integrity factor of fraudulent nodes when they start giving valid measurements again. The

integrity factor reduces to 0.1 for fraudulence behavior shown at time $t = 10$, whereas it is 0.55 at time $t = 50$. The rate of increase in integrity is high for the node which gives benevolent measurements for longer time.

Figures 5.6(a) and (b) explain how the consensus is achieved in absence/presence of fraudulent nodes in the networks. When there are no fraudulent nodes in the network, the consensus converges quickly and is achieved within 50 iterations. However, in the presence of fraudulent nodes in the network, the consensus converge at $t \rightarrow \infty$. The second eigenvalue of the weighted Laplacian matrix $L\omega$ when there is no fraudulent node in the system is greater than the second eigenvalue of the weighted Laplacian matrix $L\omega$ when there exists a fraudulent node as shown in Figure 5.6(c). Smaller the eigenvalues, more is the connectivity in the system which helps in converging to consensus. Similarly, higher the eigenvalues, lesser the connectivity in the system which takes more converging time. Therefore, the weighted consensus algorithm takes a longer time to achieve consensus in comparison to the system with all the nodes benevolent.

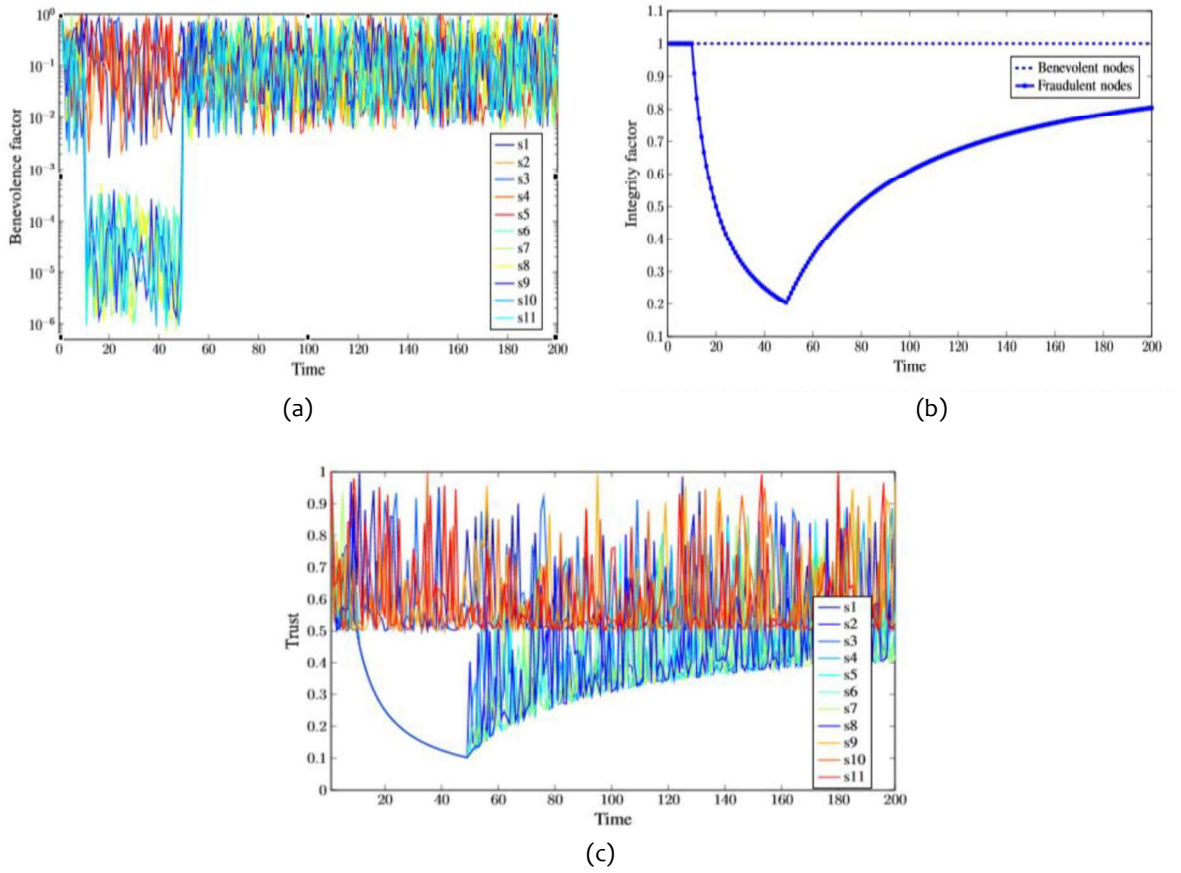


Figure 5.3: Results for a Network Size of 11 Nodes of which 6 are Fraudulent Nodes: From t = 10 to 50 sec, nodes S6 to S11 behave fraudulently, the corresponding change in benevolence factor, integrity factor and trust ratings can be observed in (a), (b) and (c)

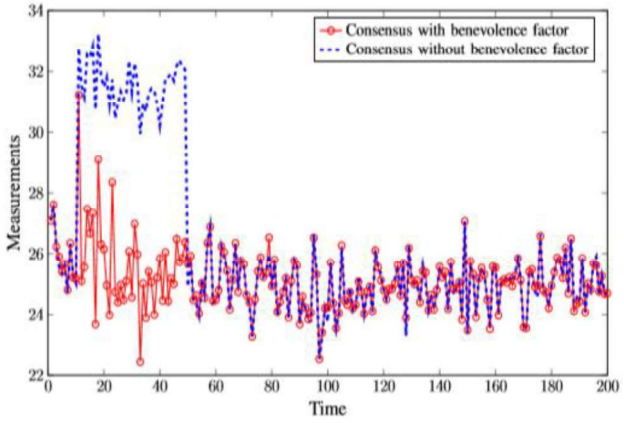
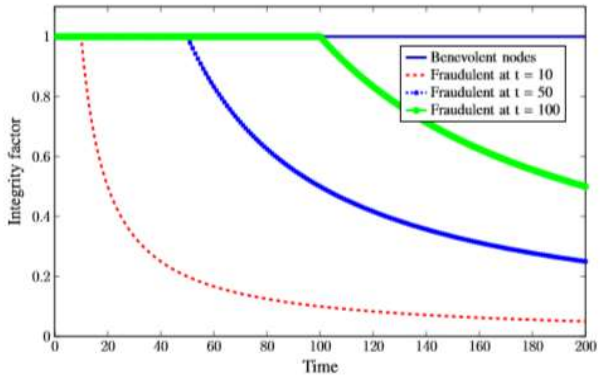
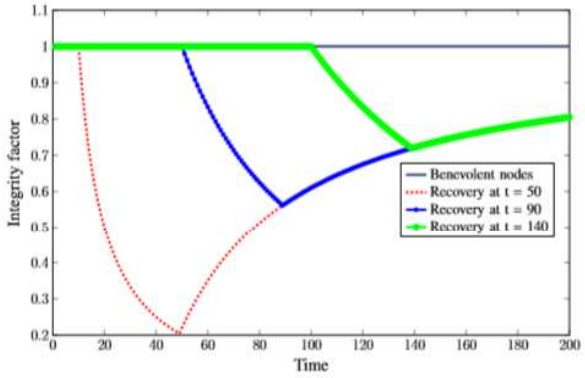


Figure 5.4: Consensus Output with and without Benevolence Factor: It describes the importance of consensus in computing benevolence factor



(a)



(b)

Figure 5.5: The Impact of Early/Late Introduction of Fraudulent Nodes on the Rate of Change in Integrity Factor: (a) Integrity change for nodes behaving fraudulent at t = 10, 50 and 100; (b) Integrity change for nodes behaving benevolent at t = 40, 90 and 140

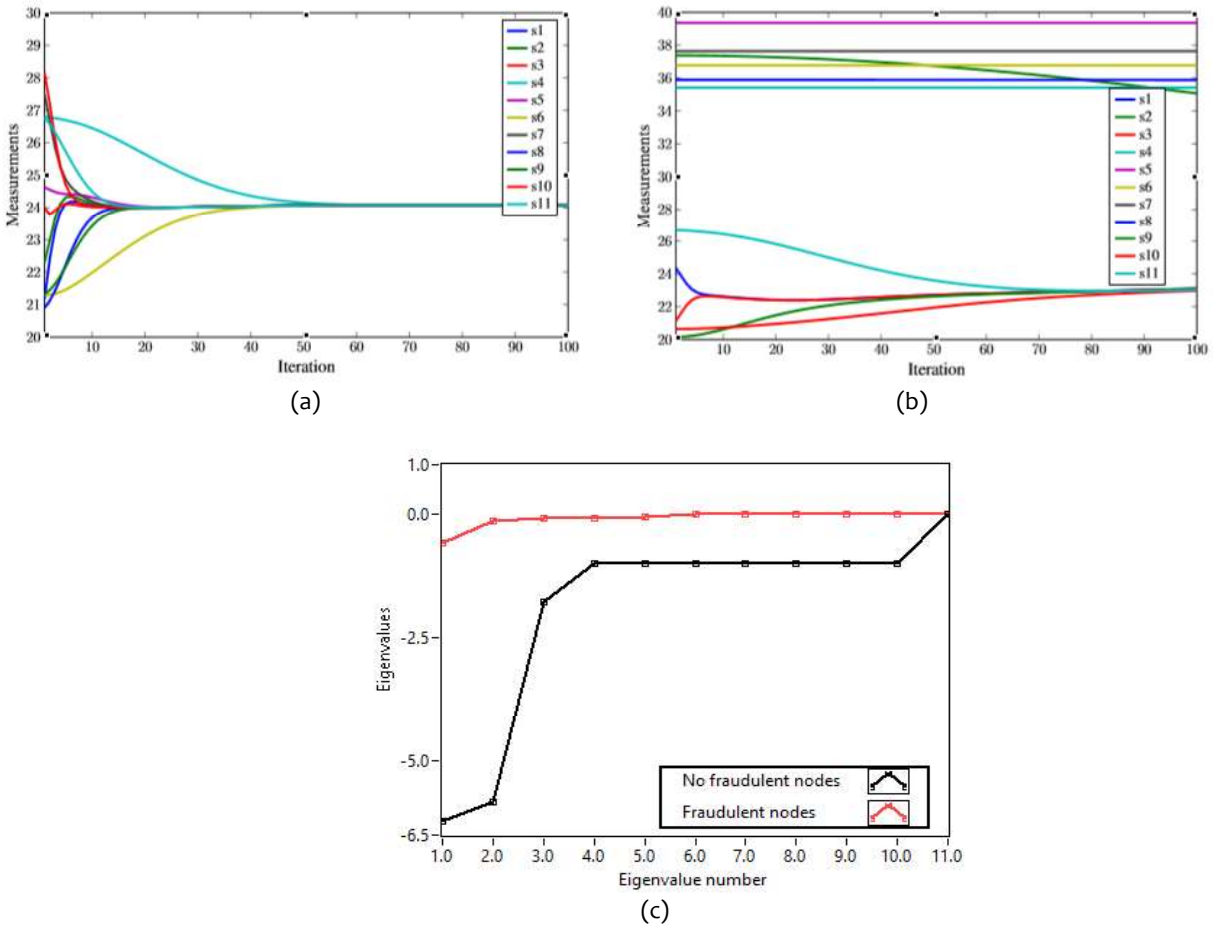


Figure 5.6: Convergence of Consensus: (a) Absence of fraudulent nodes; (b) Presence of fraudulent nodes; (c) Eigen values of the Laplacian matrix in the absence and presence of fraudulent nodes

Studies over a Larger Network: To study the performance of the proposed model on a larger network, similar studies were carried out over a network size of 341 nodes. The network is densely populated as seen in Figure 5.7. The number of fraudulent nodes is kept as 171 and these nodes start sending fraudulent information in the time period t = 10 to 20 sec.

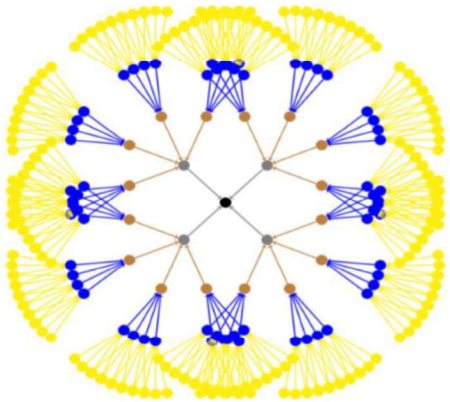
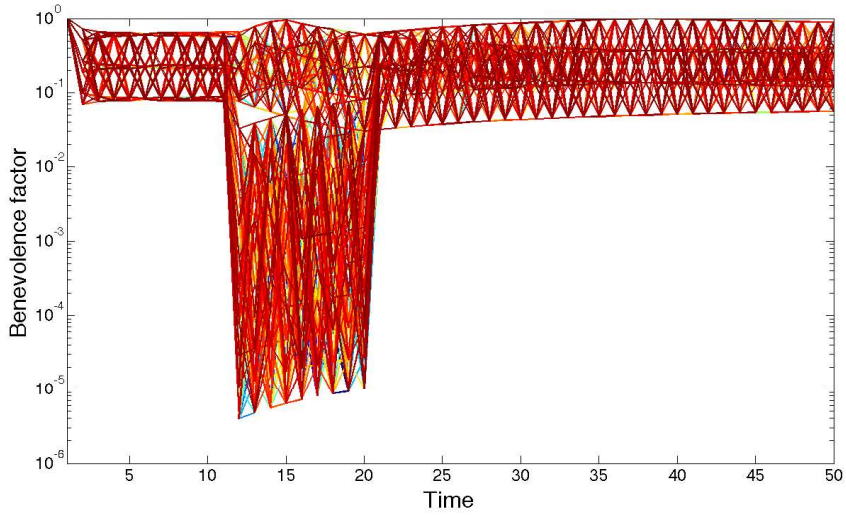
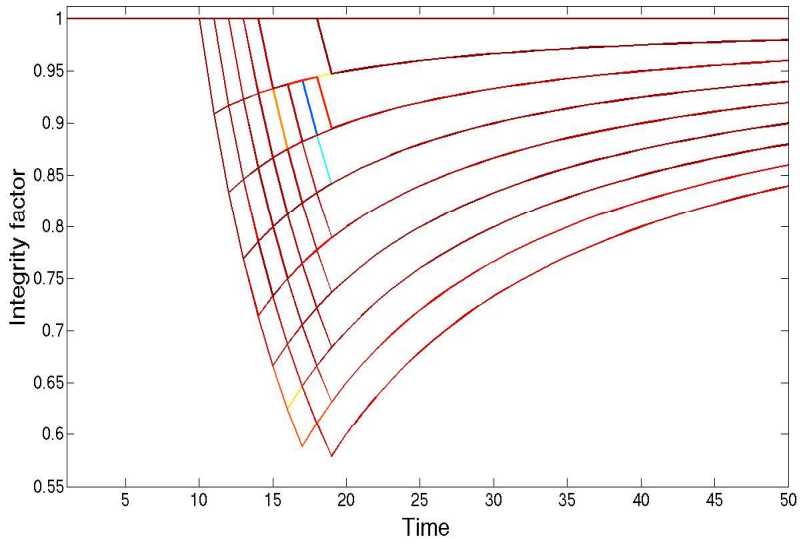


Figure 5.7: 341 Network Topology: 171 nodes start sending fraudulent information at t=10 sec

The change in variation of benevolence and integrity for the network is shown in Figures 5.8(a) and (b) respectively. The variations of all the parameters are self explanatory from the figures and similar to the trends seen in the smaller network. When fraudulent nodes are present in the network, the network gets affected which in turn reduces the trust ratings of the nodes as seen in Figure 5.9(a) and (b). Figure 5.9(b) shows the distribution of trust ratings of the nodes. 180 nodes have trust ratings around 0.1 and about 55 nodes have trust ratings greater than 0.5 which is set as the threshold of benevolent nodes. Figure 5.10 shows the corresponding consensus output using a benevolence factor. The consensus result by taking benevolence factor into account is within the valid range [25-30].

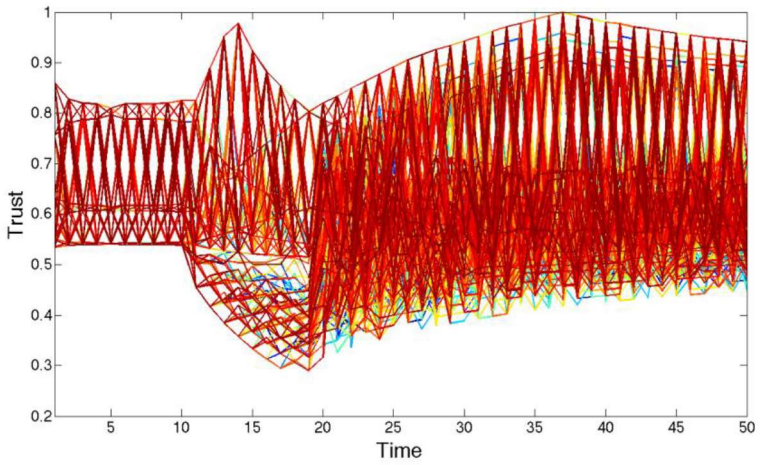


(a)

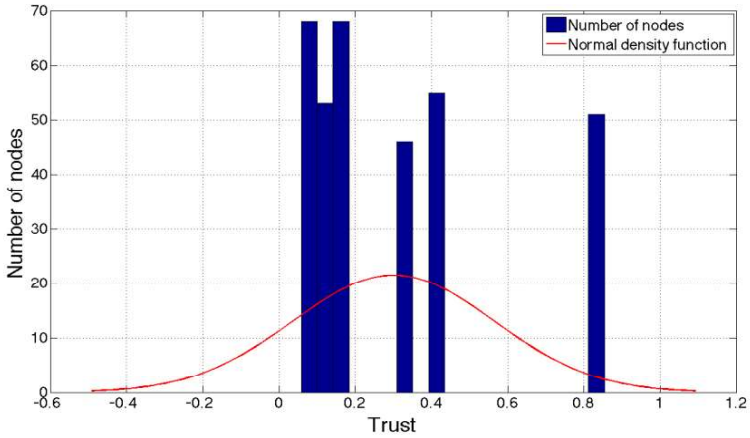


(b)

Figure 5.8: Results for a Network Size of 341 Nodes: From t = 10 to 20 sec, 171 nodes start sending fraudulent information: (a) Benevolence factor dropped at t = 10 sec and increases at t = 50 sec for the change in the behavior of fraudulent nodes; (b) Change in integrity factor



(a)



(b)

Figure 5.9: Change in the Trust ratings for 341 Network Size: (a) Shows the trust ratings trend for the network; (b) Shows the actual trust ratings for the whole network

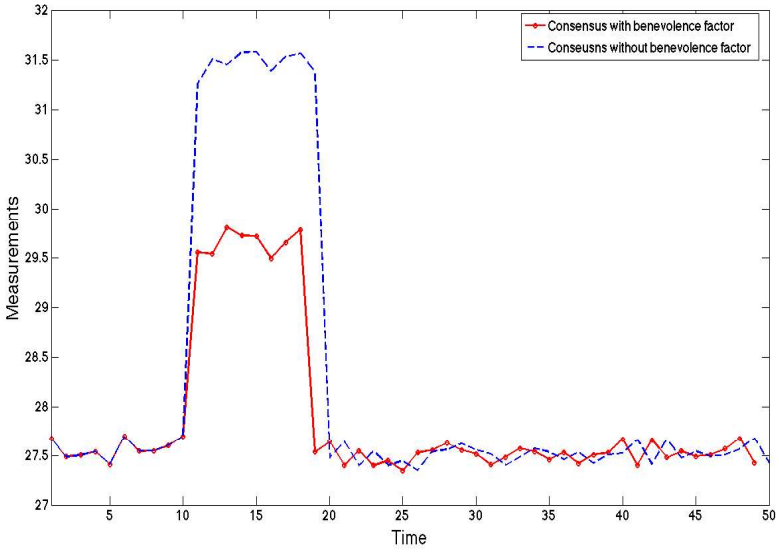


Figure 5.10: Consensus with and without Benevolence Factor: With benevolence factor, consensus results comes out in the normal acceptable range

5.3.2 Experimental Evaluation

Testbed setup: A wireless sensor network was setup in Indian Institute of Technology Jodhpur (IIT Jodhpur) field of terrain size *38m×12m* with NI Gateway 9792 and NI sensor nodes WSN-3202 as shown in Figure 5.11(a). There are 4 sensor nodes deployed in the field. Each node has 4 channels. Each sensor is tied to soil moisture sensor probe VH400 as shown in Figure 5.11(b). VH400 from Vegetronix is selected due to its moderate cost and higher sensitivity. Thus, four soil moisture sensors are attached to each sensor node thereby making a setup of 16 sensor nodes in the field. The nodes can be operated in two configurations end-node mode and relay mode. The nodes deployed in this setup run in end-node mode. In this mode, the nodes send the sensor data received at the analog channels to the gateway via a relay node. After sending the data, the node goes into sleep in order to conserve battery power. In the relay mode, the node does not sense through any channel. It just acts as an agent in multi-hopping of data packets. It never goes to sleep and hence it is recommended to keep it running on mains power supply. The modes of the node can be configured using NI LabView, NI Real Time Module and WSN 1.5 driver software. Each node is powered from 4 AA alkaline batteries. In this mode, one terminal of the node acts as a voltage source supplying 12V and current upto 20mA. The communication protocol used in this wireless sensor network is ZigBee which operates at 2.4 GHz and is a license-free band. The sampling rate of the nodes is configured as 1 minute. The nodes are enclosed in an NI WSN-3294 Outdoor IP Enclosure in order to protect them from wear and tear arising due to harsh weather conditions and the relative ease of clamping and mounting the nodes.

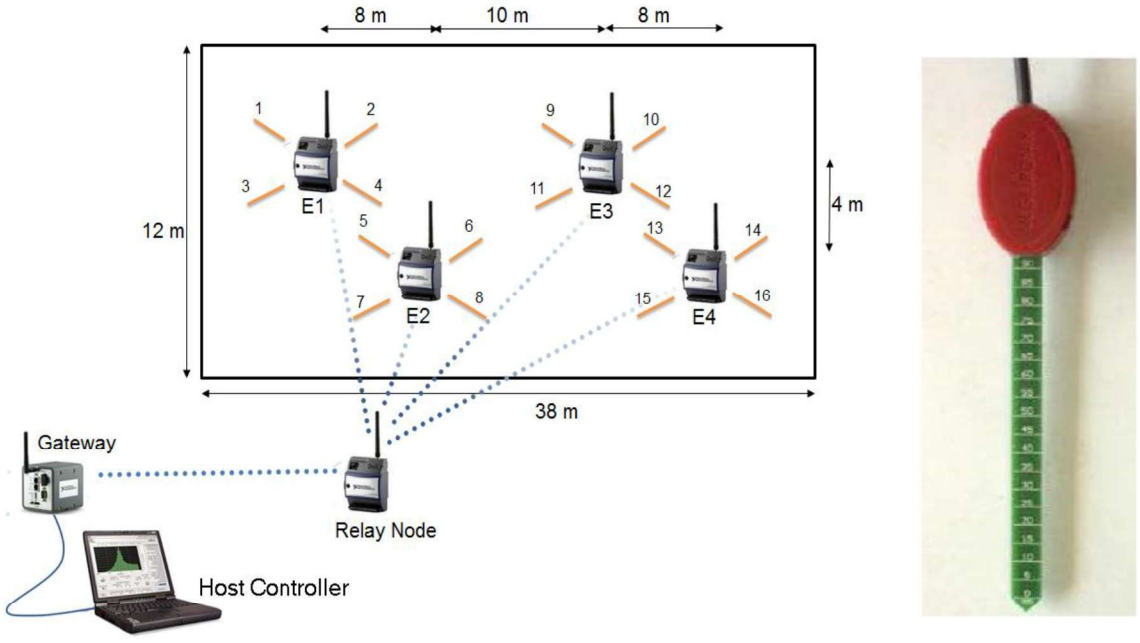
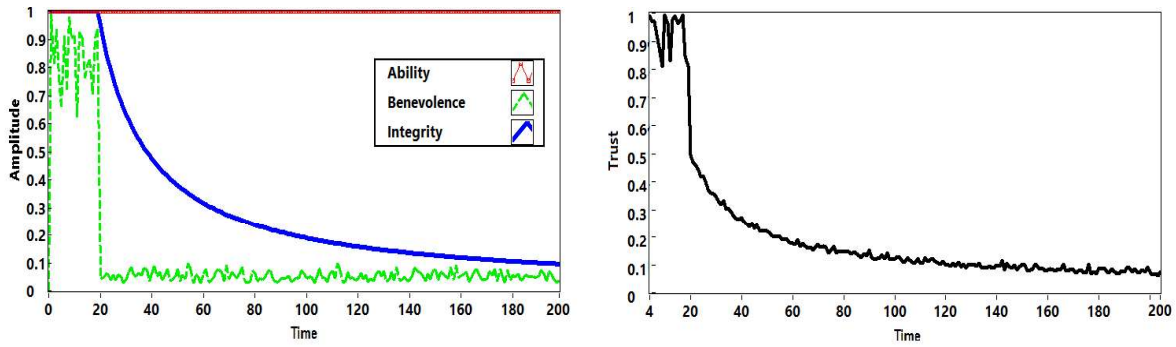


Figure 5.11: Field setup at IIT Jodhpur: (a) E1-E4 are end nodes each connected with four sensors making a 16 sensor nodes setup. Relay node is used to collect the data from the end nodes and send to the host controller; (b) Soil moisture sensor probe VH400

Experimental Setup: Different studies are conducted for different types of faults as described in Chapter 2. Faults were introduced in the network via various means. While few nodes were

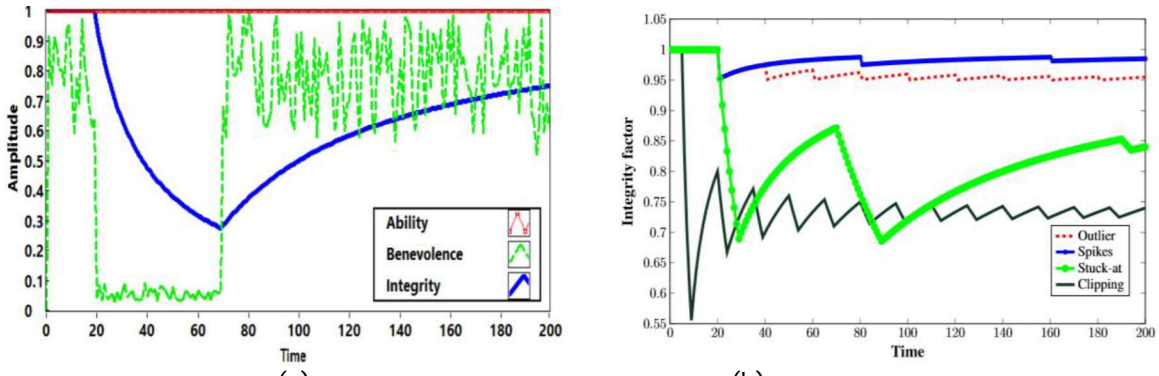
instigated with loose connections and low batteries, few were made prone to environmental perturbations and environment out of range. Few of the nodes were not calibrated properly; and a few of them were introduced with noise. The pattern of various faults can be seen from Figure 2.3.

Initially the setup was tested by introducing noise in the network. Among 16 sensor nodes, 9 nodes had loose connections due to which excessive noise was observed after 20 minutes. The variation of three factors viz. A, B and I of sociopsychological trust model is as shown in Figure 5.12(a). B and I decrease as soon as the fraudulent sensor nodes send fraudulent information due to battery voltage issues. Figure 5.12 (b) shows the trust ratings of the fraudulent nodes.



(a) (b)

Figure 5.12: Changes in the Three Factors on a Common Time Frame: (a) Sensor node is assumed to send packets continuously, thereby making ability a constant function with value 1. Benevolence and integrity factors change in the case of a fraudulent node. Initially the node is benevolent; however at t = 20 minutes, it starts behaving fraudulently, thereby making benevolence and integrity factors fall; (b) Corresponding change in the trust ratings of fraudulent node



(a) (b)

Figure 5.13: Illustration of Various Faults: (a) Illustration of ON-OFF attack: node changing the behavior alternatively; (b) Shows the integrity pattern for different faults

The consensus aware sociopsychological model effectively worked for ON-OFF attack as shown in Figure 5.13(a). Initially the nodes had loose connections. Later, corrections were made

and the corresponding change can be reflected in the figure. It was observed that the ability, benevolence and integrity factors were proportional to the change of state of node.

Later it was tested for other faults. It was observed that for the attacks such as noise, calibration, connection or hardware failure and low battery failure, the integrity pattern was smooth. This makes sense, as the sensor nodes gives fraudulent information continuously. However, in case of faults such as outliers, spikes, stuck-at, clipping the integrity pattern was not smooth as shown in Figure 5.13(b). The pattern is not smooth because of the occurrence of the fault on different time instances.

5.4 COMPARATIVE ANALYSIS

A comparative analysis between the different models is tabulated in Table 5.2 on the following metrics:

Second-hand information: In case of machine learning model, the sensor nodes learn the pattern of measurements via their own data. Therefore, it does not require any neighbouring node information for detection of the fraudulent nodes. However, in case of sociopsychological method, the neighbouring node information is required for computation of trust.

Space complexity: Space complexity is defined as the amount of memory required in computation of the algorithm. In case of machine learning model complexity is $O(s^2)$, where s is the number of samples of each sensor node. In case of sociopsychological model, the model is less complex with complexity of $O(n)$, where n is the number of nodes. In consensus aware sociopsychological model, the complexity comes out to be $O(n^2)$, since Laplacian and adjacency matrix are used for calculating the consensus of the system.

Time complexity: Time complexity is defined as the amount of time required for the execution of the program. The time complexity of machine learning model is $O(s^3)$, which gets reduced to $O(n)$ in case of simple sociopsychological model. In consensus aware sociopsychological model, the complexity comes out to be $O(n^2)$ since the computation of weighted Laplacian matrix is performed before finding the consensus of the system. If the matrix is structured and sparse the complexity is same $O(n^2)$ since matrix and vector multiplication requires n^2 time [Buluc et al, 2009].

Attacks: The Machine learning model cannot detect outliers, spikes, clipping and on-off attack as this model follows online learning where the quality of learning depends on the behavior of the majority samples. If continuous fraudulent measurements come in, the learning will be skewed towards the fraudulent set. Hence, there would be possibility of treating fraudulent measurement as benevolent measurement and vice versa. Sociopsychological model suffered from colluding data node attack. The consensus aware sociopsychological can detect all types of attack because of the integrity component. If continuous fraudulent measurements come in, node integrity factor will reduce. Similarly, integrity factor increases if the node starts behaving benevolently again. Furthermore, it can be seen from the results that high integrity variation makes the fraudulent node less reputable and the highest trust level would never be achieved again.

Robustness: Robustness deals with the capability of working effectively in the presence of high number of fraudulent nodes. For the machine learning model, though the efficiency in detection of fraudulent nodes is high, it depends on the quality of the training set. This model pre-assumes that the data set used during the training time is good which makes the system learn efficiently. However, choosing such a training set is an ideal case. Sociopsychological model, an effective yet powerful model fails when the number of colluding fraudulent nodes exceeds the number of benevolent nodes. The detection percentage comes out to be *49.5%* on a set of 200

nodes. However, consensus aware sociopsychological model has a high detection percentage of colluding fraudulent nodes because of the addition of the consensus component. Consensus aware sociopsychological model can detect upto *n-1* fraudulent nodes provided one node is sending good data. Thus, the accuracy comes out to be *99.5%*.

Table 5.2: Comparative Analysis

Model	**Machine Learning Model**	**Sociopsychological Model**	**Consensus Aware Sociopsychological Model**
Second-hand information	No	Yes	Yes
Space complexity	$O(ns^2)$	$O(n)$	$O(n^2)$
Time complexity	$O(s^3)$	$O(n)$	$O(n^2)$
Attacks	Do not work against spikes, outliers and on-off attacks	Do not work against colluding data node attack	Effective against all data attacks
Robustness (Colluding nodes)	*99.5%*	*49.5%*	*99.5%*

Note: n is the number of nodes, s is the number of samples

It can be inferred from the above evaluation that the machine learning complexity is high and detection is based on the quality of the training set. Simple sociopsychological model is a lightweight model which takes in societal norms for the detection of fraudulent nodes. However, the detection percentage is less. Consensus aware sociopsychological model is efficient in such ways that even if there is one benevolent node in the entire network it can detect all the fraudulent nodes present in the network.

5.5 SUMMARY

One of the major challenges in wireless sensor network is security. Dynamic and adaptive nature of WSN makes it prone to various types of attacks. A trust model is required to find fraudulent, selfish and compromised nodes by evaluating trustworthiness of the sensors from the network. It is difficult to find genuine and correct measurements when the number of colluding fraudulent nodes is more than the number of benevolent nodes in the network. Providing this management would notably increase the security in such a sentient environment.

A consensus based sociopsychological model for establishing trust among WSN nodes is proposed in this chapter. In a WSN, as discussed in Chapter 4, ability, benevolence and integrity of sensor nodes make the network trustworthy. The three factors play an important role in determining the trust of the sensor node. Here, the calculation of benevolence factor is based on a consensus based approach. In general, consensus means to come to an agreement. Inclusion of consensus factor increased the detection percentage of fraudulent nodes in the network. A set of

experiments over static WSNs were carried out. Experimental results validate the results of the sociopsychological model. The suggested model is compared against the models discussed in the thesis. It shall be interesting to extend the present study to dynamic WSN where the readings change due to environmental events.

…