TBSD: A Defend Against Sybil Attack in Wireless Sensor Networks

Rupinder Singh[†], Jatinder Singh[‡], Ravinder Singh[‡]

[†]Research Scholar, IKG PTU, Kapurthala, Punjab, India. E-mail: rupi_singh76@yahoo.com
[‡]IKG PTU, Kapurthala, Punjab, India. E-Mail: bal_jatinder@rediffmail.com

Summary
Security has become a most important issue for several significant applications provided by wireless sensor networks (WSNs). The intrinsically susceptible features of WSNs employ them susceptible to a diversity of attacks. This paper has centered on how to protect from a principally destructive type of attack called Sybil attack. A Sybil node using only one physical device may produce an random number of extra node identities and can be used to interrupt standard performance of the WSNs, like multi-hop routing which is utilised to discover numerous disjoint paths among source and destination. Recently, there has been a increasing attention in leveraging WSNs to mitigate Sybil attacks. Digital certificates are a way used to show individuality, however, it is not feasible in sensor networks. This paper has proposed a Trust Based Sybil Detection (TBSD) technique to detect Sybil nodes in WSNs. The TBSD scheme is based on manipulative trust values of adjacent sensor nodes and the nodes with the trust values less than a threshold value are detected as Sybil node. The feasibility of TBSD method is demonstrated systematically, while experimental results of TBSD in exposing Sybil attacks is expansively assessed equally mathematically and numerically. The acquire consequence show that the TBSD attains significant attack detection rate than existing techniques.

Keywords:
Wireless sensor networks, Malicious, Sybil, Attack, Trust based system.

1. Introduction

WSNs is defined as a self-configured and infrastructure-less wireless networks which is used to monitor environment or physical conditions, such as temperature, sound, wind direction, humidity, pressure, illumination intensity, speed, chemical concentrations, vibration intensity, sound intensity, pollutant levels, power-line voltage, etc. WSNs considerately send the information collected from the sensors to a centre position or sink [6]. This information is processed for more processing and to take different decisions. WSNs have limited capacity of processing speed, communication bandwidth, and storage. The WSNs due to limitations are inherently resource constrained and are vulnerable to various attacks. The inbuilt complexity of the applied security algorithms also adds to the difficulty of providing security to WSNs [5]. The proposed security techniques for WSNs in the history supposed that almost all sensor nodes are cooperative as well as trustworthy, but the same is not true for most of the case for various sensor network adavantages presently. A large number of attacks
has been feasible in WSN which contains tampering, jamming, hello flood, exhausting, wormhole, collision, sinkhole, Sybil, flooding, denial-of-service, cloning etc. [6].
Sybil attack in WSNs is the important attacks in this malevolent sensor node intentionally and illegally presents many forge or false identities to other sensor nodes. This is done by either creating new (fake) identities or by stealing legal identities from others sensor nodes. A variety of countermeasures against Sybil attack have been proposed in the literature that we discussed in our previous work [1]. Each of the countermeasures has its own limitation and need improvement for producing more efficient one. In this paper we first discuss the Sybil attack, trust based system and related works. In the next section of the paper, we describe our TBSD (Trust Based Sybil Detection) technique for countermeasure against Sybil attack in wireless sensor networks. The proposed scheme is based on calculating trust values of adjacent nodes and the nodes with the trust values less than threshold value are detected as malicious (Sybil) nodes. The proposed technique is designed and implemented in NS-2 tool.

2. Sybil Attack, Trust Based Systems, and Related Works

A variety of attacks are possibly in WSNs and Sybil is one of them in which a malicious node illegitimately takes multiple identities. Sybil attack can result in badly affecting the routing in the sensor networks. A large number of network security schemes are available for the protection of WSNs from Sybil attack. In this section of the paper, we discuss Sybil attack and trust based system along with proposed countermeasures.

2.1 Sybil attack

In WSNs, each node is recognized as one entity and just one single abstract idea is presented of an identity.

Manuscript received November 5, 2016
Manuscript revised November 20, 2016

Therefore, in WSNs nodes are susceptible to any scheme that allows identities to be falsified or forged. An attack that results in such a malicious activity is called the Sybil attack. So, a single node in Sybil attack intentionally and illegitimately produces numerous false or forge identities to sensor nodes in the WSN. This is done by either stealing legal identities of other nodes or creating new (fake) identities [1]. A Sybil node in the network is a disobedient nodes extra identity. As a result, a single entity of the network may get a selected number of times (depending on number of identities) in order to contribute in network operations that basically relies on redundancy, thereby in this way it can control the outcome of the operation in order to defeat the redundancy mechanisms. Sybil attack can be activated while broadcasting without the use of any central authority. This central authority of the network may help in the identification of the identities of sensor nodes [2]. Sybil attacker can have different identities; this is done by sending messages with multiple identifiers. Such a malicious sensor node replicates its multiple copies in order to damage the network. One of important observation done about the Sybil attack is that it violates one-to-one mapping between entities and identity in WSN. Figure 1 provides a scenario of Sybil attack [3].
For detecting the Sybil attack it is very necessary to understand the ways in which the network is attacked. The attack can be divided into following three ways:

1) Direct and Indirect Communication
In direct Sybil attack, the legal nodes communicate openly with the Sybil nodes in the network, whereas in indirect attack, this communication is done with the help of malicious nodes [5].

2) Fabricated and stolen identities
In this type of Sybil attack, a malicious node constructs a new identity for itself. This new identity is based on the identities of the legitimate nodes. The process when these malicious nodes communicate with their next neighboring nodes, they make use of any one of fake identities. This result in confusion in the network and it may collapse the entire network. In stolen identities case, the attacker first identifies legitimate existing identities and stole it. This type of Sybil attack may go unidentified in the network in the case of destroying of the node whose identity has been stolen. Node identity replication is done in the case when the same identities are used for a number of times in the same places in the sensor network [6].

3) Simultaneous and Non-simultaneous attack
In the simultaneous type of Sybil attack, all the Sybil identities participate simultaneously in the sensor network. Due to one identity appearing at a time, cycling through the identities will make it to appear simultaneously. In non-simultaneous Sybil attack, the number of identities that are used by assailant is equal to the quantity of physical devices that are present, where each of the devices presents dissimilar identities at different times [5].

A variety of techniques have been proposed in the literature [1] for tackling with Sybil attack. These include Message authentication and passing method, TDOA method, Random password comparison method, Neighbourhood RSS based approach, SYBILSECURE technique, Genetic

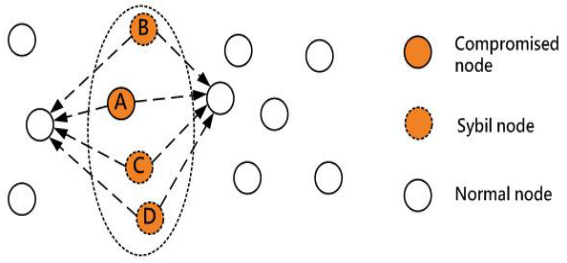

Figure 1: Sybil attack

algorithm, Genetic algorithm two-hop messages approach, P2DAP approach, Compare and match approach, Energy and hop based detection, Threshold elgamal key management scheme, Optimized secure routing protocol, RSSI-based scheme, RSSI channel-based detection, TBSD ranging-based information etc. We will make use of trust based system for the detection of malicious node as described in the next part of this section of the paper.

2.2 Trust based concept

Wang et al. [2] proposed the concept of trust computation based IDS for mobile ad hoc networks (MANETs) based on the trust variations along with a chain of evidences. In this trust based systems, the evaluation of the node is approved out regularly. A trust evaluation and reputation interchangeability based IDS mechanism is presented by Ebinger et al. [3]. In this work, the mixture of trust, reputation, and confidence along with trustworthiness is used as an improvement for the intrusion detection. Various trust management mechanisms are proposed in [4][5][6] for WSNs. The main work of these proposed works comprises security of the systems along with the reliability of information. In [7], a trust based IDS is proposed for cluster based WSNs. In the above proposed work, Cluster head (CH) is used to perform the trust computation and assessment of sensor nodes in the cluster. Honesty, supportiveness, and energy consumptions are the assessment metrics used for detection of malicious activity. The base station is used to evaluate the trust level of CH. In [9][10][11] mean node detection is carried out based on the neighbor node calculation.
Trust is a term that is used for the dependability of an entity. It is a possibility of an person node A that have

individual node B to execute a specified task at a particular time. The idea of reputation (that is collecting data concerning the status of a consecutive sensor node) is linked to the trustworthiness. Trust depends upon the ratings of consecutive nodes in the WSN. If the ratings of the consecutive node in the network are over the threshold (accepted value) then the node for further transfer of data will be trusted. Trusting on self detecting misbehavior of nodes in the network is risky. That is why collaboration among neighboring sensor nodes is necessary. The information transfer situation is shown in figure 2. Node A via node D manages trust of node D for future transfer of data. When node A forward information to D, node D receives the data as well as acknowledges this to node A. Node D may or may not transfers data to the subsequently succeeding node in the sensor network. If the node A someway knows that the node D successfully forwarded data, then node A is going to assume that node D is trusted one. After repetitive transfers of data, if the trust value reaches lower than the threshold value, then node A is going to compare trust value of its neighboring node B along with node C that are used for transferring information through the node D. If sensor nodes D is trusted with nodes B and C, then node A is going to set up a new route for data transfer by avoiding node D.

3. Assumptions and Attack Design

Here in the existing work, the proposed work considers an IEEE 802.15.4. TBSD-depends on WSN comprising of M sensor nodes. These nodes are frequently allocated in an

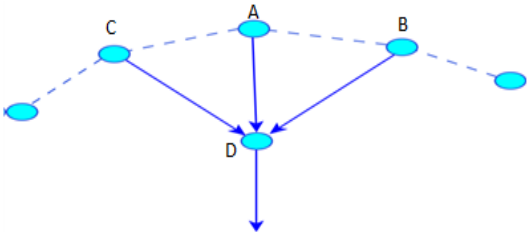

Figure 2: Trust scenario

operation region of E quadratic metric values. Inside the model, we suppose to the nodes are all still as well as uninformed of their positions. Furthermore, supposed about the nodes which transmit through each other with the wireless radio channel which transmitting in an all the ways form which covering a round region of radius R. Whenever a node transfers a message, then specific transmission is acquired from the nodes that are inside the transmission variety specified after this is either ‘‘neighboring nodes’’ or basically ‘‘neighbors’’. Also, think about that no node should be totally confidential as no presently spread confidence model present. Several of reasonable nodes are tamper as well as rebuild for an adversary’s reason to be able to release a Sybil attack from the WSN.

As a supporter can totally get more nodes, so think about an adversary which would not be outnumber genuine sensor nodes by duplicating capture nodes as well as producing other ones in adequately various sections of the network. The beginning of a Sybil attack could be shown by utilizing various parameters; (a) communication, (b) uniqueness, and (c) contribution. Correspondence is worried through the Sybil nodes that acquainted with the true blue ones within the system. The various possible methods for transmitting: the direct, where Sybil transmits directly with straightforwardly nodes, and the indirect, as the nodes may not ready for discussing specifically through the Sybil nodes, however rather communicate through the Sybil nodes. Their character parameters represent the technique through this Sybil node be able to obtain its personality. The feasible techniques are of two type’s i.e. the stolen and the created personalities. In the primary strategy, a Sybil sensor node have to take personality of a genuine to goodness sensor node with mimicking the last mentioned. Another strategy includes the creation of fabrication of arbitrary new identity. At last, the contribution measurement is worried through the interest of the Sybil sensor nodes within correspondence among the genuine nodes of the system. This type of nodes may take an interest at the same time or non-at the same time. In the synchronous investment, the vindictive nodes take part with every one of his personalities on the double, though in the non-concurrent mode, the noxious nodes show countless over a timeframe. As per the above classes, in this research, the immediate, concurrent Sybil assault with together stolen and manufactured characters is measured. The compromised node is known as the malicious node, while the rest of the nodes inside the system are eluded to as honest to goodness nodes. The assault model accepts that the malevolent nodes manufacture different, new characters, one for every substance which makes it. This type of nodes is alluded to as Sybil nodes. The primary task of the malevolent node is to trap the true blue nodes into trusting which have neighbours. Therefore the Sybil sensor nodes are not consistent; so the existence may genuinely disturb various network protocols or significantly cause to be not curable

4. Evaluated Description

As predicting that, TBSD comes under the classification with the guideline depends on ADSs. In standard depending on recognition, the abnormality locator utilizes fixed principles to characterize information focuses as peculiarities as well as normalities. While checking the system, these standards are chosen fittingly and connected

to the observed information. On the off chance that the tenets characterizing an irregularity are fulfilled, a peculiarity is announced. Inside TBSD, the fundamental finder takes after four stages towards identifying Sybil assaults in WSNs. In the initial step, the neighbour revelation stage happens. Neighbour disclosure comprises of the trading of extending empowered hi bundles (likewise alluded to as reference points) among the neighbour nodes. This is utilized for extending evaluation are measured pairs, with the main distinction with the estimation of a particular piece in the PHY header (PHR) i.e. physical header known as the "running bit", i.e. locate with the help of transferring PHY for edges proposed for running.

Inside TBSD, the hidden locator takes after four stages towards distinguishing Sybil assaults in WSNs. In the initial stage, the neighbour disclosure stage happens. Neighbor disclosure comprises of the trading of extending enabled hello packets (additionally alluded to as reference points) among the neighbor nodes. This is utilized for running evaluation are standard parcels, through the main distinction with the estimation of a particular bit inside PHY header (PHR) known as the " ranging bit", i.e. locate by the transferring PHY for edges planned i.e. going.

In the second step, every node develops a table containing the privately figured extending evaluation, i.e., the separation $\mathrm{j}_{\mathrm{up}}^{\mathrm{a}}$ from each neighboring hub it identified. Note that $\mathrm{j}_{\mathrm{up}}^{\mathrm{a}}$symbolizes the evaluated separation between hub $h_k$and hub $h_k$as measured by hub h_u. However, the separation estimation is not blunder free. Extending mistakes, which we mean from this point forward by e metric units, exist either because of the remote way of the running correspondence and the flaws of the fundamental PHY, and/or because of a malevolent hub playing out a separation diminishing or expanding assault.

Therefore, by $\mathrm{j}_{\mathrm{up}}^{\mathrm{a}}$ we represent the real separation between node $h_u$and node h_k. Clearly, it holds that de $\mathrm{j}_{\mathrm{up}}^{\mathrm{a}} - \frac{\mathrm{a}}{2} = \mathrm{j}_{\mathrm{up}}^{\mathrm{b}} \leq \mathrm{j}_{\mathrm{up}}^{\mathrm{a}} + \frac{\mathrm{a}}{2}$ at normal for every node h_u, h_k where $u, k, \in \mathrm{z}$. In the third step, each and every hub in the system freely plays out various separation coordinating checks. This implies hub h_u thinks about the going estimations of each conceivable pair of nodes h_k and h_l having in its neighbourhood directory in this $k, l \neq \mathrm{a}, 1 \leq k, l \leq \mathrm{z}$

$$\text{if} \begin{cases} \left|\mathrm{j}_{\mathrm{up}-}^{\mathrm{a}} \mathrm{j}_{\mathrm{ul}}^{\mathrm{a}}\right| < a, \\ \mathrm{j}_{\mathrm{p}-}^{\mathrm{u}} \mathrm{j}_{\mathrm{ul}}^{\mathrm{a}}\right| \geq \mathrm{a}, \end{cases} \qquad (1)$$

The above rule expresses that on the off chance node h_u observes two other, particular nodes, meant by h_k and h_l, have a distinction in separation not as much as e quadratic metric values, after that the node playing out the separation check having a Sybil assault is dynamic and continues with the way toward boycotting hubs h_k and h_l. As evident, this supposition could produce a false (positive) alarm1 in the event that the two separation coordinating hubs, h_k and h_l , are honest to goodness sensor nodes. Thus, the execution consequently, thus immaterialness of the planned Sybil assault identification calculation profoundly relies on upon the false alert likelihood. To improved opinion for our examination performance, in the resulting area, so build up the expository system which precisely processes this likelihood and permits developmental assessment to happen. Now, express the third stage of the calculation i.e. repeating stage, implying the separation checks are obtained intermittently.

With the time duration every node obtained the TBSD extending depends on Sybil assault identification calculation relies on upon the recurrence with that every node transmit a message, that communication will be acquired from the nodes that are inside the sender's transmission area specified also think about ''neighboring nodes'' or normally ''neighbours''. Also, suppose the node should be completely confidential as no presently spread confidence model present. Several of legitimate nodes are tampered with as well as rebuild for an adversary's reason to be able to release a Sybil attack from the WSN node enters the neighbour disclosure stage searching for new neighbours in its region. Every period a node hunts down presently or new neighbours, it re-runs the separation checks. This fourth step is important to guarantee that separation checks are dependably progressive between the recently included neighbours and each various present sensor node in the area list. As indicated by the situation expressed before, while an honest to goodness node finds a separation match between no less than two unmistakable sensor nodes, it increases an alert attempting to repudiate the Sybil sensor nodes. In renouncing Sybil nodes, the true blue sensor node that is the TBSD-fit locator, transfers a caution message to the sink node (SN) empowering the system executives to obtain countermeasures. It likewise boycotts the sensor nodes maintaining a strategic distance from any future increments of them in its neighbourhood list. On the other hand, if no separation matching exists, the node proceeds with its typical operation. And the process, will transfers and receives system bundles among its neighbours satisfying the detecting undertakings doled out to it.

As the proposed TBSD going depends on Sybil assault identification calculation has completely disseminated, implying the information accumulation, observing as well as identification procedures are implemented on various areas in the system. Such a design clearly suggests, to the point that every one of the nodes of the system are fit for

running the proposed inconsistency depends on location calculation. Also, no participation or data contribution is required between the nodes so as to renounce a noxious node. Henceforth, no correspondence overhead is caused for recognition purposes. At long last, in recognizing abnormalities, our methodology works with restricted review information, to be specific the nodes going evaluations. This implies every node works as an autonomous oddity based location framework (ADS), and in that capacity, it is in charge of identifying assaults just for itself. Calculation 1 outlines the diverse periods of the fundamental TBSD running based Sybil assault recognition calculation.

Algorithm: TBSD varying-depends on Sybil attack detection
Take the sensor node black list (NBL)
Take a clock for development of the neighborhood finding stage
for each period the neighborhood discovery-relevant clock expires do
for each node i in the network, i=2: M do
stage 1: Replace a varying-open the beacon through each neighbor sensor node
Stage 2: Build a table obtaining the varying evaluations for each detected neighbor node $k, = l, 2, Z$
Step 3: execute distance entry
for every feasible couple of nodes in the neighbor list do
if $\left|j_{up-}^{a} j_{ul}^{a} < a, \right| \leq k, l, \leq z, k, l, \neq u$ then
increase an alarm
withdraw nodes k=l by placing in the NBL
else
go on with regular process
end
end

4.1 Problem Formulation

Let's suppose it has a set as well as finite quantity of z sensor nodes is consistently spread in a sensor field. The sensor field includes a location of E quadratic metric values. For ease, we suppose that a sensor node covers a small region on the field. Therefore, it's feasible the particular every sensor node has situated on peak of a new sensor node. Every sensor node has been considered to protect a round transmission area of radius r (in parameter values) which represents the transmission area form of a sensor node in the nature of disk. A separate ring formation is inferred (the sensor node is at the middle of the disk). The disk is splits into concentric rings which contains the similar size, e. Recall, that e is the standard varying evaluation fault well-informed with sensor node (in metric values). With no failure of overview, let's think a separation of the network nodes, i.e., nodes h_u;h_k; h_l with node ni being the detector node, and nodes $h_k, h_l \neq h_k$ the nodes under study by node ni used for Sybil attack. The target set to analyse the possibility of more nodes are situated inside similar region enclosed with round ring. The area when the nodes h_k; h_l might exist, creating a false (negative) alarm, makes a round ring. At last identify that possibility coexistence region possibility, as well as we calculate it in the next parts. The primary regulation creates a false alarm is:dij and dik

$$\left|j_{up-}^{a} j_{ul}^{a}\right| < a, 1 \leq u, p, l \leq z, k, l \neq u \qquad (2)$$

After that, we utilize the idea of arithmetical probabilities to be able as to build the study.

4.2 Fake alarm: Possiby allocation of a fake alarm

The possibility of wrongly increasing an alarm should be produced through computing the possibility of at least one node varying at least two nodes in the similar round ring region inside their transmission radius, r. Indeed, every node increases an alarm while detecting at least a couple of nodes lying in the similar round ring. Exactly the similar result will be acquired as soon as further two nodes or more are found in the similar round ring. Actually, it is the rationale behind the at least declaration.

So, what we should seek is the possibility j^u of at least one node, h^u , varying at least two other nodes, h^k; h^l, in the similar round ring region, known that the transmission radius of node h^u is R, as well as every of the total M nodes are consistently spread in a sensor area of E quadratic metric values. The possibility of an at least event could derived by not including all other possible occurrences. Therefore, the required possibility is computed as follows:

$$j^u = 1 - o^u(t) - o^u(1) - o^u(2), l(2) - \cdots - o^u(z - 1) l^u(z - 1) \qquad (3)$$

$o^u(0)$: Possibility that nodeh^u has no neighbors.
$o^u(1)$: Possibility that node h^u has exactly 1 neighbor.
$o^u(2)$: Possibility that node h^u has exactly 2 neighbors.

- $o^u(z-1)$: possibility that node ni has $z-1$ neighbors, i.e., that each other sensor node as neighbour.
- $l^u(2)$: possibility to facilitate that there is no couple of two existing neighbours located in the similar round ring inside node h_u′s range._ li
- $l^u z - 1$ possibility to facilitate that there is no couple of $z-1$ neighbors located in the similar round ring inside node h_u′s range.

The computation of the Eq. (2) enables the purpose of the possible allocations $q^u$and $l^u(y)$. We contribute the these three subparts for that's reason.

4.2.1 Possibility of allocation of sensor node's neighbors

In that part, it decides the possibility density function (pdf) of a single node h_u which contain accurately y neighbors. By taking into consideration that z total nodes are consistently spread in a sensor field of area E, the possibility$o^u(y) \le y \le z-1$, is derived as follows:

$$o^u(y) = sv(Y-y) = \binom{z-1}{y}\alpha^y(1-\alpha)^{z=(y+1)} \qquad (4)$$

Where $\alpha \le 1$ represents the arithmetical possibility of node h_k to be inside the communication radius R of node h_u, where $h_u - h_k$. This possibility is shown by:

$$\alpha = \frac{\text{Area of favorable region}}{\text{Area of total region}} = \frac{\pi r^2}{f}, \pi r^2 \le f \qquad (5)$$

4.2.2 Coexistence region possible allocation

In essence, the coexistence region possibility, denoted with m^u , shows that the possibility that a single node h_u ranges at least two other nodes $h_k, h_l \ne h_u$ in the similar round ring. Then firstly obtain the possibility of a single node ni to discover the accurately two other nodes $h_k, h_u \ne h_k$ in the similar round ring of width e. suppose that node h_u is located on a sensor region E. Specified that nodes $h_k h_l$ nk are neighbours of node h_u , node ni's possibility of detecting either node h_k or node h_k inside his coverage region is single. As node h_k is able to placed with other round ring with possibility one, the possibility of node h_l to be detected in the similar round region as node h_k is known with the arithmetical possibilites:

$$m^u \frac{\text{Area of favorable region}}{\text{Area of total region}} = \frac{\text{Average coexistence area}}{\text{Circle area}} \qquad (6)$$

Let's calculate the arithmetical possibility wi, then firstly require to decide the regular coexistence region.

4.2.3 The approximate distance calculated with node $h_u$when varying one more node h_k (inside his coverage disk region) might reduce inside three zones, namely (a) the inner zone, (b) the middle zone (c) the outer zone. The inner zone is defined as $t < j^a_{uk} \le r - \frac{a}{2}$, the middle zone is defined as $\frac{a}{2} < j^a_{uk} r \frac{a}{2,}$ and the outer zone is defined as $r\frac{a}{2} < j^a_{uk} \le r$

The motivation behind is to show the node h_u detects node h_k to be de j^g_{uk} metric value distant from his place. As extended distance is not as much of as $\frac{a}{2}$, all feasible real positions of the node k form an inner zone, actually a circle, having radius de j^a_{uk} where $t\frac{a}{2} < j^a_{uk} \le r$. As soon as de j^a_{uk} increases, the middle zone is arrived where all feasible actual positions of the node h_k at present form round rings. Every round ring is surrounded with the boundary of two concentric circles of various radii; assuming that be $j^a_{uk} - \frac{a}{2} \le j^a_{uk} j^b_{uk} + \frac{a}{2,}$ the primary or inner circle has radius de $j^a_{uk} + \frac{a}{2}$ v and the second or outer circle has radius de $j^a_{uk} - \frac{a}{2}$. H ence, once the external connectivity limits of node ni are arrived, the feasible positions of the node nj form the outermost ring.

Now this node ni is conscious of his higher transmission threshold R, therefore they set an higher boundary to the feasible locations of node nj. Therefore, the outermost ring types the outer zone where. $r\frac{g}{2} < j^g_{uk} j^g_{uk}$

Lemma 4.3 The standard value of every coexistence region is $\overline{br} = \frac{5\pi rg}{2}$

proof. By equating Eq. (5) and 4.2, now verify the coexistence region possibility:

wi ¼ Average coexistence area

$$m^u = \frac{\text{Average coexistence area}}{\text{Area}} = \frac{\overline{br}}{\pi r^2}\frac{5\pi rg/2}{\pi r^2} = \frac{5g}{2r} \qquad (7)$$

Given that the coexistence region possibility is known, the pdf of li ðxÞ could be simply known.

Lemma 4.4 The possibility of having no feasible pair of nodes out of x total nodes in the similar coexistence region is li

$$l^u(y) = jv(Y-y) = (1-m^u)^{\frac{u(y+1)}{2}}, \quad \text{where} \quad 2 \le (y) \le z \text{ and} z \ge 0 \qquad (8)$$

Proof. In computing the sum of quantity of node pairs having x total nodes, the recurrence is excluded, as the events of detecting the nodes nj and nk exist in the same area, as well as the event of detecting whether nodes nk and nj coexist in the same area, are the same assumed that node ni has x neighbours, the number of all feasible pair of combinations not including recurrence is $y(y-1)/2$ such as, assume that node ni has three neighbours,p, l, m, then, $y-3$ nd all feasible pairs of combinations are j $p \leftrightarrow l\, p \leftrightarrow m, l \leftrightarrow m,$ that is $(y-1)/2 = 6/2 = 3$.

4.5 Generalized coexistence area possibility allocation

By equating Eq. (3) and Lemma 4.3, so decide the generalized coexistence region possibility $t^u(y)$. The latter possibility express the possibility that for node ni, which has accurately x neighbors, no feasible pair of x out

of M neighbors exist that lie in the similar round ring of width e inside node ni's communication area of radius R. The probability is $t^u(y)$ given by:

$$t^u(y)l^u(y) = \binom{z-1}{y} \alpha^y (1-\alpha)^{z=(y+1)} (1-m)^{u(y+1)/2} \quad (9)$$

By replacing Eq. (7) to Eq. (2), the probability Pi can now be formulated
as follows:

$$j^u(y) = o^u - o^u(2)l^u(2) - o^u(3)l^u(3) - \cdots \quad (10)$$

$$-o^u(z-1) = 1 - 1(\alpha-1)^{(z-1)} - (z-1)\alpha(\alpha-1)^{(z-2)} \quad (11)$$

$$-t^u(2) \ldots = o^u(z-1) = 1 - (\alpha-1)^{(z-1)} - (z-1)\alpha(\alpha-1)^{(z-2)} \quad (12)$$

$$- \cdots - \sum_{m-2}^{z-1} \binom{z-1}{m} \alpha^m (1-\alpha)^{z=(y+1)} (1-m^u)^{m=(m-1)/2} \quad (13)$$

4.6 False alarm probability

The shown investigation keeps simply that single node performs varying. Specifically Eq.(8) produces the possibility that a single node, node ni, causes a wrong (positive) alarm. Though, in a sensor network of M nodes, any node may cause a wrong alarm. Therefore, we look for the possibility of at least one node causing a wrong alarm in a sensor network comprising of M nodes. The possibility that denote by p is required as shown:

Lemma 4.7 The wrong alarm possibility in a sensor network of M nodes spread in the region E, and every node having a transmission radius R, as well as varying experiences an standard fault of e, is

$$s = 1 - (1 - s^2)^z \quad (14)$$

Proof. So Look for the possibility of at least one node causing a wrong alarm. The amount $1 - s^2$ produces the possibility that a one sensor node not produce a wrong alarm. Therefore, the total $(1 - s^2)^z$ M gives the possibility in which all nodes in the network do not produce a wrong alarm. In that the possibility $1 - (1 - s^2)^z$
Proof: We look for the possibility of at least one node tr expresses the event that at least one node triggers a false alarm.

5. Simulation-Based Implementation and Experimental Results

The implementation of TBSD was systematically tested in a simulation environment for wireless sensor network simulated in NS-2, through the simulation parameters utilized has been defined in Table 1.

Table 1: Simulation parameters

Parameter	Value
Simulator used	NS 2.35
Area (meter)	800X800
No. of nodes	100 to 500
Routing protocol	DSDV
Channel type	Wireless
Packet size	512 byte
Mobility model	Two ray ground Propagation model

The first step in the simulation is to organize WSN by explaining the network sender node and reciever nodes. We first create a sensor network with 38 nodes and then trigger Sybil attack as shown in figure 3. The sensor network is divided into different clusters. Each cluster has its own CH (Cluster Head) and is represented with a different colour as shown in the figure 3. The neighboring nodes of the reciever node will reply back to sender node including the path reply packets. The sender node choose best route through sender to reciever with the help of sequence number as well as hop count. It is supposed in the simulation that a single Sybil node existed in the sensor network including all nodes containing same IEEE 802.11b hardware. The sensor nodes with the exemption of the Sybil node are randomly distributed inside the network area. Node 7 is the malicious (Sybil) node in the network as its represent a different identity. After the implementation of TBSD technique this malicious node is isolated from the sensor network as shown in figure 4. In the next part of this section of paper, we provide various graphs to show the effectiveness of the proposed technique.

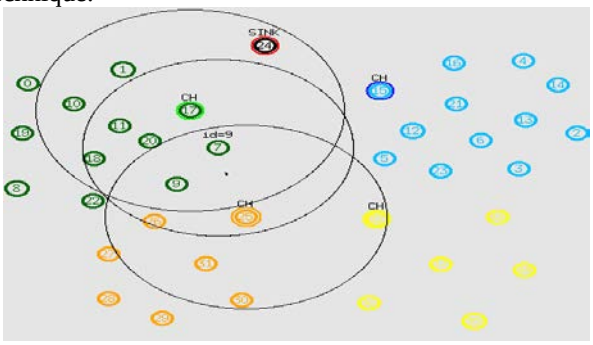

Figure 3: Sybil attack scenario in WSN

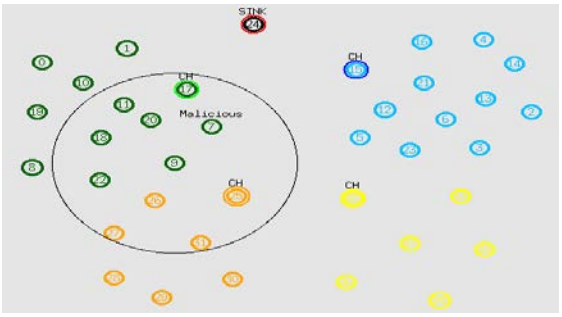

Figure 4: Isolation of malicious (Sybil) node

The formula for the TBSD performs since the cause holds back for your destination for a post acknowledgement going without running shoes soon after any eleventh packet. In the event origin will get the particular acknowledgement coming from desired destination, then there's simply no bad behavior inside WSN as well as approach remains since normal. However, if the desired destination is unable to know the results boxes for a while period of time, after that recognition method will start it is functionality. The particular proven journey is going to be analyzed in order to discover as well as identify position connected with detrimental nodes (if any) from your WSN. Ideas submit an application the particular planned solution connected with TBSD strategy to uncover almost any probable Sybil for the duration of the path development process. The cluster head make use of trust value matrix to locate and isolate the Sybil nodes in the sensor networks.

In order to assess the efficiency and competence of proposed technique i.e. TBSD and other some well known sybil detection techniques, NS-2.3 based simulation is done for WSNs coding organizations and run sybil detection techniques.

5.1 Experimental Set-up

The existing and proposed sybil detection techniques are implemented on a Linux workstation (2.4 GHz Intel i5 processor with 10 GB RAM and 102 GB memory). The simulation is done several time, by considering different set of sensor nodes every time. And sybil attack is applied on the given simulation by considering random nodes. The evaluated data is organized and build up for training and testing purpose. Then MATLAB 2013 a tool is used to evaluate the performance of the proposed technique.

5.2 Performance Measures

The primary metrics in this paper considers Accuracy (A_{cc}), F1 score (F_1) and Matthews correlation coefficient (M_{cc}). These metrics can be defined as follows.

a) Accuracy (A_{cc}): The A_{cc} represents the effectiveness of the given sybil detection techniques. It states how much effective the detection rate is, which is calculated as:-

$$A_{cc} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (15)$$

Here TP: represents the accurate prediction of those in which sybil attacks are detected successfully, whereas FP represents in which non-sybil nodes are detected as attackers. TN indicates those in which non sybil nodes are evaluated successfully, whereas FN represents in which sybil nodes are detected as genuine nodes.

b) F1 score (F_1): The F_1 can be demonstrated as a weighted mean of the precision and recall, where an F_1 attains its effective value at 1 and worst score at 0.

$$F_1 = \frac{2*TP}{2*TP+FP+FN} \qquad (16)$$

c) Matthews correlation coefficient (M_{cc}): M_{cc} represent the degree of correlation between the actual sybil nodes and predicted sybil nodes. M_{cc} lies between -1 to 1, close to 1 is effective in the sybil detection techniques.

$$M_{cc} = \frac{TP*TN-FP*FN}{\sqrt{(TP+FP)*(TP+FN)*(TN+FP)*(TN+FN)}} \qquad (17)$$

5.3 Experimental results

The proposed and the existing well-known sybil detection techniques are applied on the designed simulation, 15 times. The mean values of the simulation are taken for evaluating, the best technique. However the nodes are varied between 50 to 500 only, but the proposed and existing work is not limited to this set only.

Table 1 clearly demonstrate that the proposed technique has optimistic sybil detection rate than existing techniques. The mean A_{cc} of the best known sybil detection technique in literature i.e. [21] is 0.8946. Whereas in the case of the proposed technique it is 0.8962. Therefore proposed technique has minimum improvement in terms of A_{cc} is 0.0141 i.e. 1.41 %. Thus proposed technique is more effective than most of the existing techniques.

Table 2 represents that the proposed technique has optimistic sybil detection rate than existing techniques. The mean F_1 of the best known sybil detection technique in literature i.e. (16) [21] is 0.7234. Whereas in the case of the proposed technique it is 0.7451. Therefore proposed technique has minimum improvement in terms of A_{cc} is 0.0317 i.e. 3.17 %. Thus proposed technique is more effective than most of the existing techniques.

Table 3 prove that the proposed technique has optimistic sybil detection rate than existing techniques. The mean Mcc of the best known sybil detection technique in

literature i.e. (16) [21] is 0.89126. Whereas in the case of the proposed technique it is 0.89446. Therefore proposed technique has minimum improvement in terms of A_{cc} is 0.018 i.e. 1.8 %. Thus proposed technique is more effective than most of the existing techniques.

6. Conclusion and Future WorK

This paper, has proposed a novel technique called TBSD, for improving the detection rate of Sybil attack in WSNs. The sensor network is divided into clusters with each having a CH. In this technique trust values are assigned to each node after gathering information about the location and adjacent nodes of each sensor node. The node which change identification has different adjacent nodes each time, this information will reduce the trust value of the node

Table 1: Accuracy (A_{cc}) analysis

Nodes	[1]	[2]	[3]	[4]	[5]	[6]	[20]	[21]	TBSD
50	0.64	0.700	0.720	0.75	0.890	0.892	0.891	0.894	0.895
100	0.65	0.660	0.710	0.73	0.760	0.891	0.893	0.895	0.896
150	0.66	0.689	0.620	0.75	0.789	0.892	0.895	0.894	0.897
200	0.66	0.720	0.740	0.76	0.892	0.894	0.896	0.897	0.898
250	0.66	0.710	0.720	0.74	0.891	0.892	0.894	0.894	0.897
300	0.64	0.720	0.700	0.72	0.892	0.890	0.892	0.893	0.896
350	0.66	0.689	0.689	0.74	0.770	0.789	0.894	0.896	0.899
400	0.65	0.689	0.700	0.72	0.789	0.890	0.892	0.895	0.896
450	0.64	0.720	0.689	0.73	0.892	0.789	0.893	0.893	0.895
500	0.66	0.660	0.740	0.76	0.760	0.894	0.896	0.895	0.897

Table 2: F1 score (F_1) analysis

Nodes	[1]	[2]	[3]	[4]	[5]	[6]	[20]	[21]	TBSD
50	0.6248	0.6896	0.6134	0.6389	0.673	0.7400	0.6891	0.7160	0.7260
100	0.6326	0.6689	0.6046	0.6221	0.6568	0.6891	0.7089	0.7260	0.7430
150	0.6446	0.6763	0.6264	0.6389	0.6643	0.7100	0.7260	0.7160	0.7520
200	0.6326	0.6134	0.6307	0.6568	0.7200	0.7160	0.7430	0.7520	0.7610
250	0.6171	0.6046	0.6134	0.6307	0.6891	0.7300	0.7160	0.7160	0.7520
300	0.6446	0.6134	0.6896	0.6134	0.7200	0.6730	0.7100	0.7089	0.7350
350	0.6248	0.6763	0.6763	0.6307	0.6650	0.6640	0.7160	0.7350	0.7610
400	0.6136	0.6763	0.6896	0.6134	0.6640	0.6730	0.7200	0.7260	0.7430
450	0.6326	0.6134	0.6763	0.6221	0.7420	0.6640	0.7089	0.7089	0.7260
500	0.6136	0.6689	0.6307	0.6568	0.6561	0.7160	0.7430	0.7260	0.7520

Table 3: Matthews correlation coefficient (M_{cc}) analysis

Nodes	[1]	[2]	[3]	[4]	[5]	[6]	[20]	[21]	TBSD
50	0.6893	0.6670	0.6760	0.7160	0.7640	0.7731	0.763	0.8902	0.8912
100	0.7898	0.6380	0.6660	0.6890	0.7350	0.7632	0.7892	0.8912	0.8931
150	0.7520	0.6570	0.6891	0.7160	0.7540	0.7738	0.8912	0.8902	0.8970
200	0.7260	0.6760	0.7060	0.7350	0.7730	0.8902	0.8931	0.8970	0.8950
250	0.7520	0.6660	0.6760	0.7060	0.7630	0.7737	0.8902	0.8902	0.8970
300	0.7120	0.6760	0.6670	0.6760	0.7730	0.7647	0.7731	0.7892	0.8921
350	0.7260	0.6570	0.6570	0.7060	0.7440	0.7548	0.8902	0.8921	0.8950
400	0.7898	0.6570	0.6670	0.6760	0.7540	0.7645	0.7732	0.8912	0.8931
450	0.6893	0.6760	0.6570	0.6896	0.7730	0.7545	0.7892	0.7892	0.8912
500	0.7520	0.6389	0.7060	0.7350	0.7350	0.8902	0.8931	0.8912	0.8970

Each node in the cluster calculates trust value of neighbor nodes and sends it to the CH in the form of a massage for further processing. The node which has average trust value less than a predefined threshold is detected as the Sybil nodes and is isolated from the sensor network. The simulation result of the proposed technique has been compared with some well known sybil attack detection techniques. The comparisons have clearly indecates that

the proposed technique outperforms over the available techniques. However this work has not considered the effect of mobility of sensor nodes, therefore in near future this work will be extended for the mobile sensor nodes.

Acknowledgement

Authors are highly thankful to the Department of RIC, IKG
Punjab Technical University, Kapurthala, Punjab, India for providing opportunity to conduct this research work.

Conflict of Interest
The authors declare no conflict of interest

References

[1] Thiago Bruno M. de Sales, Angelo Perkusich, Leandro Melo de Sales, Hyggo Oliveira de Almeida, Gustavo Soares, Marcello de Sales, ASAP-V: A privacy-preserving authentication and sybil detection protocol for VANETs, Information Sciences, Volume 372, 1 December 2016, Pages 208-224
[2] Riccardo Pecori, S-Kademlia: A trust and reputation method to mitigate a Sybil attack in Kademlia, Computer Networks, Volume 94, 15 January 2016, Pages 205-218
[3] Panagiotis Sarigiannidis, Eirini Karapistoli, Anastasios A. Economides, Detecting Sybil attacks in wireless sensor networks using UWB ranging-based information, Expert Systems with Applications, Volume 42, Issue 21, 30 November 2015, Pages 7560-7572
[4] Bo Yu, Cheng-Zhong Xu, Bin Xiao, Detecting Sybil attacks in VANETs, Journal of Parallel and Distributed Computing, Volume 73, Issue 6, June 2013, Pages 746-756
[5] Kuo-Feng Ssu, Wei-Tong Wang, Wen-Chung Chang, Detecting Sybil attacks in Wireless Sensor Networks using neighboring information, Computer Networks, Volume 53, Issue 18, 24 December 2009, Pages 3042-3056
[6] M. Conti, R. Di Pietro, A. Spognardi, Clone wars: Distributed detection of clone attacks in mobile WSNs, Journal of Computer and System Sciences, Volume 80, Issue 3, May 2014, Pages 654-669
[7] R. A. Shaikh, H. Jameel, B. J. Auriol, H. Lee, S. Lee, and Y. J. Song, "Group-based trust management scheme for clustered wireless sensor networks," IEEE Transactions on Parallel and Distributed Systems, Vol. 20, No. 11, PP. 1698-1712, Nov. 2009.
[8] Fenye Bao, Ing Ray Chen, Moon Jeong Chang, and Jin Hee Cho, "Trust-Based Intrusion Detection in Wireless Sensor Networks," IEEE International Conference on Communications, 2011.
[9] Stetsko, L. Folkman, and V. Matyayas, "Neighbor-Based Intrusion Detection for Wireless Sensor Networks," in International Conference on Wireless and Mobile Communications Los Alamitos, CA, USA, pp. 420-425, 2010.
[10] Liu, X. Cheng, and D. Chen, "Insider attacker detection in wireless sensor networks," in Proceedings of IEEE INFOCOM, pp. 1937-1945, 2007.
[11] Li, J. He, and Y. Fu, "A group-based intrusion detection scheme in wireless sensor networks," in Proceedings of GPS - Workshops, pp. 286-291, IEEE, 2008.
[12] Renyong Wu, Xue Deng, Rongxing Lu, and Xuemin Shen, "Trust-based anomaly detection in wireless sensor networks," IEEE International Conference on Communications in China (ICCC), 2012.
[13] S. Zheng and J. Baras, "Trust-assisted anomaly detection and localization in wireless sensor networks," in Proc. IEEE Conf. on Sensor, Mesh and Ad Hoc Comm. and Network (SECON), pp. 386394, 2011.
[14] Aditi Paul, Somnath Sinha, and Sarit Pal, "An Efficient Method to Detect Sybil Attack using Trust based Model," Proc. of Int. Conf. on Advances in Computer Science, AETACS, Elsevier, 2013.
[15] Reza Rafeh and Mozhgan Khodadadi, " Detecting Sybil Nodes in Wireless Sensor Networks using Two-hop Messages," Indian Journal of Science and Technology, Vol. 7(9), 1359–1368, September 2014.
[16] Weichao Wang, Di Pu, and Alex Wyglinski, "Detecting Sybil Nodes in Wireless Networks with Physical Layer Network Coding," IEEE IIFIP International Conference on Dependable Systems & Networks (DSN), 2010.
[17] Tong Zhou, Romit Roy Choudhury, Peng Ning, and Krishnendu Chakrabarty, "P2DAP – Sybil Attacks Detection in Vehicular Ad Hoc Networks," IEEE Journal On Selected Areas in Communications, Vol. 29, No. 3, March 2011.
[18] Philip W. L. Fong, "Preventing Sybil Attacks by Privilege Attenuation: A Design Principle for Social Network Systems," IEEE Symposium on Security and Privacy, 2011.
[19] Sohail Abbas, Madjid Merabti, David Llewellyn-Jones, and Kashif Kifayat, "Lightweight Sybil Attack Detection in MANETs," IEEE Systems Journal, Vol. 7, No. 2, June 2013.
[20] Guojun Wang, Felix Musau, Song Guo, and Muhammad Bashir Abdullahi, "Neighbor Similarity Trust against Sybil Attack in P2P E-Commerce," IEEE Transactions o Parallel and Distributed Systems, December 2013.
[21] Neil Zhenqiang Gong, Mario Frank, and Prateek Mittal, "SybilBelief: A Semi-Supervised Learning Approach for Structure-Based Sybil Detection," IEEE Transactions on Information Forensics and Security, Vol. 9, No. 6, June 2014.
[22] Lin Cai and Roberto Rojas-Cessa, "Containing Sybil Attacks on Trust Management Schemes for Peer-to-Peer Networks," IEEE ICC Communication and Information Systems Security Symposium, 2014.
[23] Jatinder Singh, Dr. Savita Gupta, and Dr. Lakhwinder Kaur, "A Cross-Layer Based Intrusion Detection Technique for Wireless Networks," The International Arab Journal of Information Technology, Vol. 9, No. 3, May 2012.