J Med Syst (2016) 40: 161
DOI 10.1007/s10916-016-0515-4

SYSTEMS-LEVEL QUALITY IMPROVEMENT

Distributed Denial of Service Attack Source Detection Using Efficient Traceback Technique (ETT) in Cloud-Assisted Healthcare Environment

Rabia Latif[1] · **Haider Abbas**[1,2] · **Seemab Latif**[1] · **Ashraf Masood**[1]

Received: 29 September 2015 / Accepted: 2 May 2016 / Published online: 17 May 2016

Abstract Security and privacy are the first and foremost concerns that should be given special attention when dealing with Wireless Body Area Networks (WBANs). As WBAN sensors operate in an unattended environment and carry critical patient health information, Distributed Denial of Service (DDoS) attack is one of the major attacks in WBAN environment that not only exhausts the available resources but also influence the reliability of information being transmitted. This research work is an extension of our previous work in which a machine learning based attack detection algorithm is proposed to detect DDoS attack in WBAN environment. However, in order to avoid complexity, no consideration was given to the traceback mechanism. During traceback, the challenge lies in reconstructing the attack path leading to identify the attack source. Among existing traceback techniques, Probabilistic Packet Marking (PPM) approach is the most commonly used technique in conventional IP- based networks. However, since marking probability assignment has significant effect on both the convergence time and performance of a scheme, it is not directly applicable in WBAN environment due to high convergence time and overhead on intermediate nodes. Therefore, in this paper we have proposed a new scheme called Efficient Traceback Technique (ETT) based on Dynamic Probability Packet Marking (DPPM) approach and uses MAC header in place of IP header. Instead of using fixed marking probability, the proposed scheme uses variable marking probability based on the number of hops travelled by a packet to reach the target node. Finally, path reconstruction algorithms are proposed to traceback an attacker. Evaluation and simulation results indicate that the proposed solution outperforms fixed PPM in terms of convergence time and computational overhead on nodes.

Keywords Traceback · Probabilistic Packet Marking (PPM) · Cloud- assisted WBAN · Distributed Denial of Service (DDoS) Attack

This article is part of the Topical Collection on *Systems-Level Quality Improvement*

✉ Haider Abbas
hsiddiqui@ksu.edu.sa; haiderabbas-mcs@nust.edu.pk

[1] National University of Sciences and Technology, Islamabad, Pakistan

[2] King Saud University, Riyadh, Saudi Arabia

Introduction

With the increasing popularity of cloud- assisted WBAN for critical health applications, the demand for securing these networks is also increasing. One of the major threats to these networks is Distributed Denial of Service (DDoS) attacks that not only exhaust the network capacity but also prevent these networks to perform their desired tasks [1, 2].

In DDoS attack, the key issue lies in detecting an attack and invoking the appropriate traceback mechanism. Several techniques are available in literature for detecting DDoS attack in sensor networks, but very limited amount of work is found on traceback mechanism [3]. This research work is an extension of our previous work in which a machine learning based attack detection algorithm is proposed to detect DDoS attack in cloud-assisted healthcare environment [3–6]. However, in order to avoid complexity, no consideration was given to the traceback mechanism.

Figure 1 shows the cloud-assisted healthcare architecture being considered for this research [5]. The green dotted circle shows the entities that are the victims of DDoS attacks. The arrows shown in yellow are the attack path reconstructed by the victim node in order to identify an attacker and further block it.

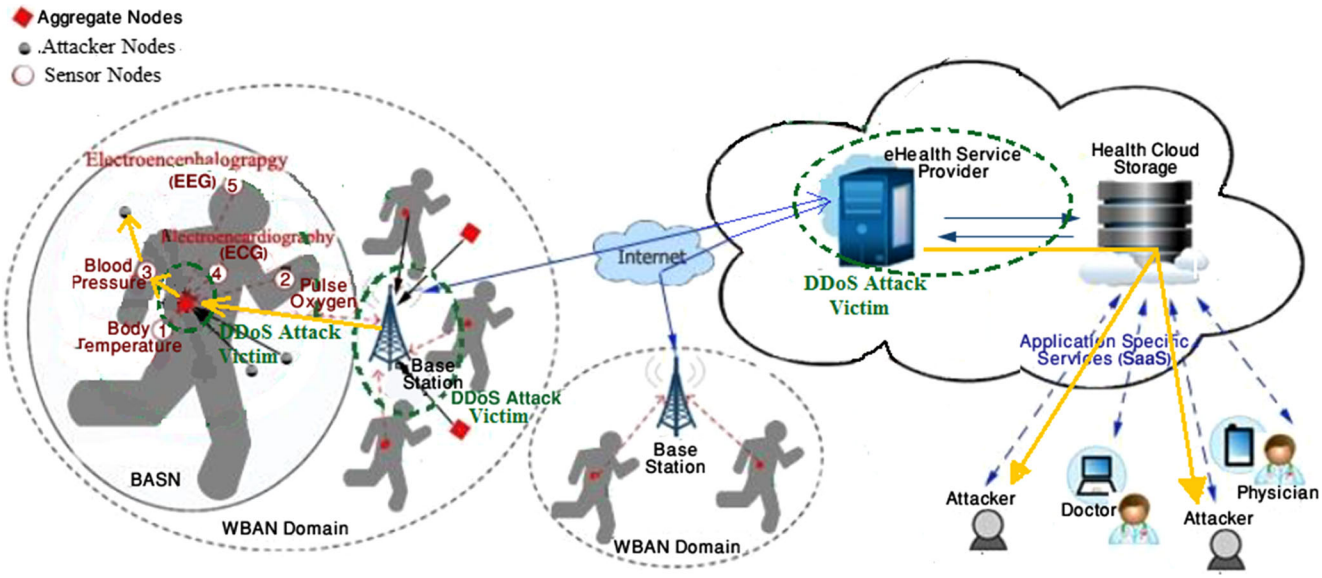

Fig. 1 Cloud-Assisted Healthcare Architecture

Although the given architecture presents the complete cloud-assisted healthcare environment, however for this research, the key focus is to reconstruct the attack path and identify the attack resource within WBAN domain.

Traceback requires reconstructing the attack path and identifying the source of DDoS attack [7]. Traceback techniques proposed for conventional IP- based networks [8–11] are not directly applicable on resource constrained Wireless Body Area Network (WBAN) environment due to additional overhead requirements and high convergence time. Similarly, several traceback techniques are also available for Mobile Ad-hoc Network (MANET) [12] and Wireless Sensor Network (WSN) [13] that overcomes the limitation of overhead but at the cost of additional processing and storage requirements [14].

Analysis shows that none of the available solutions are appropriate for traceback of DDoS attack in cloud-assisted WBAN environment. Among the available techniques, Fishbone Traceback (FBT) [15] is specifically proposed for hierarchical WSN. It is based on edge sampling approach [10] and appears to be more appropriate than other techniques because it is lightweight and easily implemented in WSN. FBT uses marking probability distribution function that assigns fixed marking probability to all the nodes in order to minimize the convergence time but, concurrently, it increases the overhead on nodes.

In this research, we propose a new traceback technique called Efficient Traceback Technique (ETT), to be deployed specifically in resource constrained WBAN environment. The proposed technique assigns the dynamic marking probability to each node based on the number of hops the packet travelled once it originates from the source. The number of hops can be calculated as the distance travelled by the packets from the source. Finally, a path reconstruction algorithm is proposed to traceback the attacker.

Results and comparison shows that the proposed technique has less convergence time as compared to fixed Probability Packet Marking (PPM) approach. Similarly, the proposed technique results in less computational overhead on nodes as compared to other available schemes.

The paper is organized as follows:

Section "Literature Review", details the existing traceback techniques in both standard based IP networks and mobile Ad-hoc networks. Section "Preliminaries" gives an introduction to PPM and explains the problems related to choosing marking probability. Section "Proposed Traceback Technique" describes the proposed technique. The proposed packet marking technique is described in section "Finding the Traveling Distance" and proposed traceback algorithms are presented in section "Working Example". In section "Performance Evaluation", results of simulations and comparative performance evaluation is given. Finally, the paper is concluded in section "Conclusion".

Literature Review

Traceback Techniques for Standard IP- Based Networks

There are few major techniques which exist in the literature that deals with traceback problem in standard IP-based networks.

Bellovin [8] introduces the concept of Internet Control Message Protocol (ICMP) traceback technique which utilizes ICMP packets that contains the information about the preceding and the succeeding routers and sends this information to the destination and the origin of the original packet. Using this additional ICMP packet, the target node easily reconstructs the attack path. However, this technique is not appropriate for resource constraint WBAN network because it requires the

WBAN network to make use of full TCP/IP protocol stack. Also maintaining extra ICMP packet throughout the transmission and traceback requires additional memory and computational resources.

Snoeren et al., [9] proposed the hash-based IP traceback that creates audit stream for network traffic and can track the source of a single packet given by the network recently. These techniques require adequate amount of memory and storage space to record and transfer these network audit trails. The implementation of hash-based traceback is not practical in WBAN. These techniques are only good for traceback in conventional IP based networks where storage space is sufficient for logging traffic data.

Savage et al., [10] proposed the Probabilistic Packet Marking (PPM) technique, in which each router not only forward the packet but also mark individual packets with a low marking probability. This mark is a unique identifier analogous to that specific router. As compared to other techniques, PPM has small implementation and management overhead due to the probabilistic nature of algorithm. However, the computational overhead and the convergence time is high, which is the time taken by victim node to reconstruct the attack path by collecting at least one marked packet from each intermediate router. This results in limiting the usefulness of PPM for fast traceback in WBAN environment.

Andrey and Nirwan [11] proposed a Deterministic Packet Marking (DPM) technique, which like PPM also requires each router to mark individual packets. Moreover, the DPM approach requires all the internet routers to be updated for every packet marking, which in turn requires a huge amount of spare bits in IP packets. Therefore, the scalability of DPM is very limited. Also it requires a huge amount of storage space for packet logging of routers. For this reason, DPM is not a good solution for traceback in WBAN [11].

All of the traceback techniques discussed above are for conventional IP-based networks. However, these techniques are not appropriate when deployed in resource constrained WBAN environment because they require extensive computation and implementation resources.

Traceback Techniques for Mobile Ad-hoc Networks

A number of traceback approaches exists in literature that are proposed specifically for MANETs. These are discussed as follows:

Jin et al. [12], proposed traceback technique based on node sampling in which a complete network is split into various zones where each node is familiar with its zone ID to which it belongs. Upon the arrival of packet, each node first writes its zone ID into the packet with a certain probability and then passes it. Upon the detection of DDoS attack, the victim node reconstructs the complete path by gathering adequate number of these marked packets. Analysis shows that the reconstruction process of this technique is less accurate to efficiently traceback the source of an attack.

Things et al. [16], proposed a scheme for MANET named ICMP traceback with Cumulative Path (CP). This scheme conceals the complete information of attack path in ICMP traceback CP message. Nevertheless, this scheme needs to overload some fields of the IP header and thus, needs a heavy protocol stack which is unavailable in resource constraint WBAN environment.

Bo Chao et al. [15] proposed a traceback scheme specifically for hierarchical WSN environment. The proposed scheme is based on two layer labeling technique and a Marking Probability Distribution Function (MPDF) that assigns a fixed marking probability assignment to each node for ease. Using fixed marking probability requires a large amount of packets for reconstructing an attack path which results in high convergence time. As the packets are overwritten with same marking probability (by all routers) results in unfairness marking. The evaluation of proposed scheme is done by comparing the results qualitatively rather than quantitatively.

Preliminaries

In sensor network environment, one of the key features is that the source node itself inserts its source address in the Medium Access Control (MAC) header before it sends any packet. This allows a number of anonymous attacks on sensor networks [10].

A number of approaches are available to traceback the source of an attack and packet marking is one of them. In packet marking approach, each node places some path information in every passing packet until it reaches the victim. The victim node reconstructs the attack path by collecting a certain number of packets along the network path.

Among packet marking approaches, PPM is considered as the most well-known solution for traceback of DDoS attack because PPM has small implementation and management overhead due to the probabilistic nature of the algorithm [17].

Probabilistic Packet Marking (PPM)

A PPM based traceback can be classified into packet marking and path reconstruction phases. In packet marking phase, each originating packet is marked with some probability τ as it passesintermediate nodes along the attack path. In reconstruction phase, a victim node uses the recorded path information in the packet to reconstruct the attack path and locating the source of an attack. For recording path information, node sampling, node append and edge sampling are widely used techniques [10].

Key Issues in Selecting Probability

In DDoS attack, traceback mechanism is carried out between an attacker and the victim. Attackers hide their identity using

spoofing and restrict the number of attack packets. However, the victim needs to choose appropriate marking scheme to locate the attacker. For efficient PPM mechanism, the key issue lies in selecting a suitable marking probability τ for easy and accurate traceback in WBAN environment [17].

At-Least-One-Marking per Sensor Node

According to the graphical network topology shown in Fig. 2, let A be the attack path such that A = {a, n_1, n_2… n_N, v}, where 'a' represents the attacker, v denotes a victim of DDoS attack and n_i (i = 1, 2… N) represent N sensor nodes (including aggregate node) along the attack path.

Suppose node n_i has a marking probability τ_i. The residual probability φ_i is defined as the probability that an attack packet has lastly been marked by node n_i and not by any other node further down the attack path. From the perspective of victim v, φ_i helps the victim v to know that the node n_i is on the attack path after inspecting this incoming packet. Residual probability φ_i is represented as:

$$\varphi_i = \begin{cases} \prod_{j=1}^{n} (1-\tau_j) & i = 0 \\ \tau_i \prod_{j=i+1}^{n} (1-\tau_j) & 1 \le i < N \\ \tau_i & i = n \end{cases} \quad (1)$$

Consider all nodes have the fixed marking probability then $\tau_1 = \tau_2 = \ldots = \tau_n \equiv \tau$. From Eq. 1, we have

$$\varphi_i = \tau (1-\tau)^{N-1} \quad for\ 1 \le i \le N \quad (2)$$

From Eq. 2, it is concluded that the residual probability φ_i for node n_i is geometrically smaller, i.e. the node is closer to the attacker. It is given as:

$$\varphi_1 < \varphi_2 < \ \ldots \ < \varphi_N \quad (3)$$

From Eq. 3, it is concluded that the node n_1 has minimum possibility whereas node n_N has the maximum possibility to send its marking information to the victim node v. It is not possible for victim v to figure out that node n_i is on attack path until v receives a packet that contains a marking left by node n_i. Therefore, the victim must receive at-least-one-marking from each node n_i on the attack path for the successful reconstruction of attack path.

Let P be the attack measure from an attacker a to the victim v. To fulfill the need of at-least-one-marking per node n_i, an efficient PPM-based traceback must meet the following criteria:

$$P\varphi_1 = P\tau\ (1-\tau)^{N-1} \ge 1 \quad (4)$$

In Fig. 3, a graph is plotted that shows the possible values of residual probability φ_1 for node n_1 with respect to marking probability τ and number of nodes N using Eq. 2. It is evident from Fig. 1 that for different number of N, the peak value occurs at τ = 1/N e.g., for $N = 25$, the peak value occurs at 1/25 for which $\varphi_1 = 0.0277$. As the value of N (total number of nodes between a and v) varies and is unknown to victim, therefore it is difficult to decide the ideal marking probability a priori.

One possible solution is to select a small τ, again doing this allows the attacker to lessen the attack volume so that a limited range of τ are available for successful attack.

Spoofing

In spoofing, the attacker besides spoofing source address may also spoof the packets marking field by falsifying data in order to conceal his/her identity or attack path. This whole process is termed as spoofed marking attack [17].

From the victims perspective if a packet remains unmarked along the path i.e., the packet remains unmarked by any intermediate node n_i, the false data in the marking field left by an attacker may lead to inaccurate path reconstruction. The probability that the packet remains unmarked is expressed as:

$$\varphi_0 = (1-\tau)^N \quad (5)$$

Taking φ_0 along y-axis, a graph is plotted with respect to τ and N using Eq. 5 as shown in Fig. 4. The graph shows that φ_0

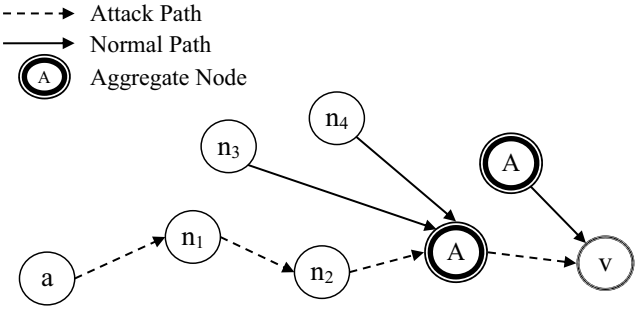

Fig. 2 Graphical Network Topology

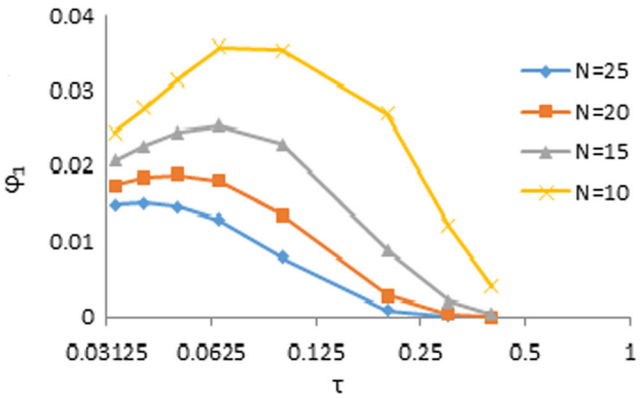

Fig. 3 Residual Probability (φ_i) for node (n_1)

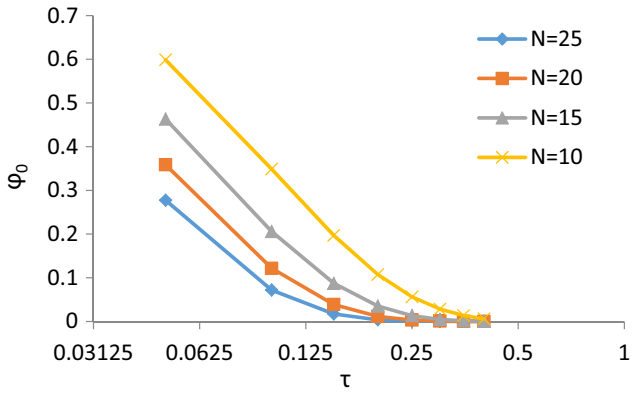

Fig. 4 Unmarked Probability (φ_0)

is inversely proportional to τ with different number of nodes N, which means that φ_0 is a decreasing function of τ. It is concluded that the possibility of packet remain unmarked is decreasing with increase in marking probability τ, so as the marking probability starts increasing along the attack path, the chances that the packet gets spoofed is decreased.

Uncertainty

Packets whose marking fields are spoofed with false data also cause uncertainty in traceback. The concept of uncertainty was introduced by Park and Lee [18] and explained with the help of Fig. 5. Back to the previous assumption in which an attack path is defined as A = {a, n_1, n_2… n_N, v}. As shown in Fig. 5. An attacker initiates an attack by spoofing the marking field with the false data (l_1, n_1), where l_1 is the legitimate node which is spoofed. Before reaching the victim node v, if the spoofed packet remains unmarked by other nodes along the path, it is considered as legitimate packet originating from l_1. A similar scenario is assumed for other nodes l_2, l_3,…, l_K, where K is the uncertainty factor and defined as a total number of fake sources of an attack besides the actual attacker a. Hence, the total number of false sources of an attack identified by a traceback technique is (K + 1).

As discussed before, the node closer to an attacker has least residual probability φ_i as compared to other nodes. The attacker takes advantage of this scenario by keeping all the spoofed packets unmarked and send them to victim v showing

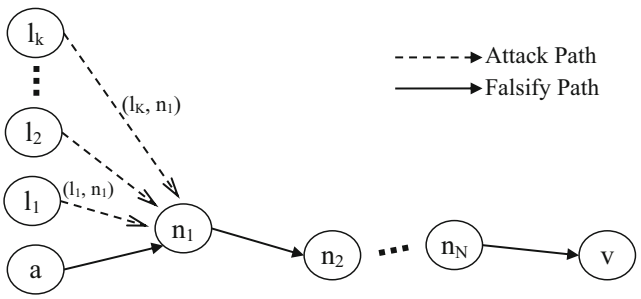

Fig. 5 Falsify Paths

them as these were marked by node n_1. This scenario is represented as:

$$K\varphi_1 = \varphi_0 \tag{6}$$

Solving Eq. (6) by putting values of φ_1 and φ_0, we obtained

$$K = \frac{1}{\tau} - 1 \tag{7}$$

From Eq. (7), it is observed that marking probability τ is inversely proportional to uncertainty. As in original PPM approach, the marking probability τ is fixed. Increase in fixed marking probability τ results in the decrease of uncertainty factor K.

Proposed Traceback Technique

The existing PPM approaches proposed for sensor networks uses fixed marking probability τ_i, for packet marking which results in high convergence time, additional overhead and uncertainty as discussed in section "Key Issues in Selecting Probability". The root cause of this variance is the assignment of uneven probability φ_i to sensor nodes n_i along the attack path.

Liu et al., [17] introduces the concept of a Dynamic Probability Packet Marking (DPPM) approach, in which the marking probability is assigned to each node based on the distance travelled by the packet. DPPM uses Time-to-Live (TTL) field in IP-header to determine the travelling distance of each packet passing by the router.

As in sensor networks, we are dealing with MAC protocol, determining the travelling distance is a key issue. Using a TCP protocol in WSN itself increases the overhead due to three-way handshake [13].

In the following section, we will present a new traceback technique specifically for resource constrained WBAN. The proposed technique is based on DPPM and uses MAC header instead of IP header.

The proposed technique has following features:

- It assigns a uniform probability φ_i to all the nodes n_i along the attack path with the aim to reduce the overall convergence time.
- It reduces the overhead on all the nodes by assigning the variable marking probability in descending order as the packet travels along the attack path towards the victim node.
- It ensures that each packet should mark at least once in order to remove the uncertainty caused by spoofed marking.

The proposed technique works as follows:

Let d denotes the travelling distance of a packet such that $(1 \leq d \leq i)$, where i is the total number of nodes along the attack path. Each node n_i marks the packet r with the marking probability which can be calculated as the distance travelled by packet r from its source until reach that particular node. It can be expressed as:

$$\tau_i = \frac{1}{d} \tag{8}$$

Taking into account the working of proposed technique, the key issue lies in how to find the travelling distance of each packet r from its source? In the following section, we will answer this question. To the best of our knowledge, it is the first attempt to deploy DPPM in WSN environment.

Finding the Traveling Distance

Before finding the traveling distance of each packet from its source, first we look into the WBAN network topology shown in Fig. 6. The network topology can be either multi-hop or single-hop. Figure 6a shows the multi-hop WBAN topology in which sensor nodes transmit their data to an aggregate node via intermediate nodes. Figure 6b shows the single-hop topology in which each sensor node directly sends its data to an aggregate node and further to Base Station (BS) via intermediate aggregate nodes.

To find the traveling distance of each packet from its source, a small number of bytes are reserved in the data payload of MAC Protocol Data Unit (MPDU) and labeled as DPPM label. Figure 7 shows the MPDU with DPPM label. The labeling mechanism brings a very less change in IEEE 802.15.4 MAC header. In each packet, only 12 bytes are reserved to carry DPPM label for multi-hop WBAN and 10 bytes for single-hop WBAN. As the label uses data payload of MPDU which is variable in length, therefore, it is acceptable to carry this amount of data to perform traceback operation in WBAN environment.

The length of DPPM label depends upon the topology employed for WBAN. Next, we will discuss labeling in detail for both multi-hop and single-hop WBAN topology.

Multi-hop WBAN Topology

For multi-hop WBAN topology, 12 bytes are reserved in MAC data payload and labeled as P(s) = (Source, End, Initial, Head, Tail, Distance) as shown in Fig. 8. Each P(s) represent a packet field marked at each sensor node *s* along the path. (*Source*, *End*) is associated with regular sensor node, where (*Initial*, *Head*, *Tail*) is associated with aggregate nodes which helps in path reconstruction and *Distance* is used to find the distance travelled by each packet from its origin.

The detail of each field is given as follows:

Source: Source is the originating sensor node ID of an edge connecting two sensor nodes e.g., in Fig. 9 A is the source node sending packet to node B. When the attack packet first originates, the source node write its node ID to this field of the packet P(s).

End: It is a node which receives a packet from a source i.e. node at the edge that receives the packet e.g. in Fig. 9, B is End. Upon receiving the packet, the end node first checks the following conditions before writing its ID in the field:

$$\begin{aligned} &Source \quad field \quad ! \quad = \quad EMPTY \\ &Distance \quad field \quad = \quad = \quad 0 \\ &End\ node\ and\ Source\ node \in Same\ Cluster \end{aligned}$$

When the above conditions met, the node writes its ID into the end filed of packet P(s). At this point the distance field becomes 1.

Distance: It is defined as the traveling distance from the source to the victim. This field is incremented by each intermediate node as the packet travels along the attack path.

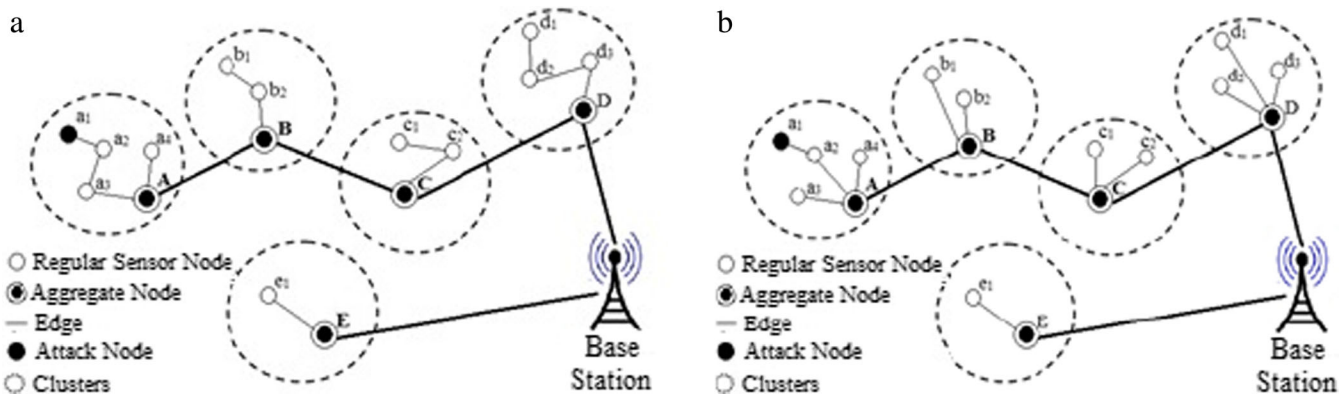

Fig. 6 **a** Multi- Hop WBAN Topology. **b** Single- Hop WBAN Topology

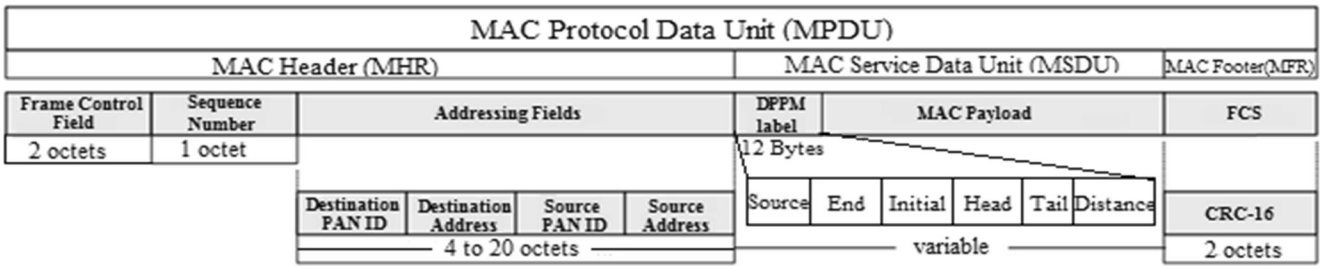

Fig. 7 IEEE 802.15.4 with DPPM label for Multi-hop WBAN

Initial: This field of a packet is written by an aggregate node of the cluster where source node is present and remains same along the path until the packet reaches the victim. The attack node also lies in the same cluster as the aggregate node.

Head: This field is written by aggregate node of current cluster and contains the head of an edge for aggregate nodes. This field is updated by every downstream aggregate node upon the arrival of packet.

Tail: Upon receiving the packet, the aggregate node updates this field with the tail of an edge for aggregate node. This field is also written by aggregate nodes only.

Working Example

A detailed working example for finding the traveling distance of a packet is given in this section. A multi-hop WBAN network topology is shown in Fig. 10a. It consists of four clusters, where each cluster have regular sensor nodes and one aggregate node that acts as a cluster head. Each sensor node either sends its data directly to an aggregate node or via other regular sensor nodes. Similarly, each aggregate node forwards its data to Base Station (BS) either directly or via intermediate aggregate nodes. Suppose attacker a1 launch DDoS attack towards BS by sending out spikes of packets. Figure 10b shows the sequence of packets traveling along the path towards BS. Every node updates each field of a packet P(s) in order to find the distance and reconstruct the path successfully. It is explained as follows:

- Sensor node a_2 writes its ID in the source field of packet P(s). After reaching node a_2, the DPPM label became (a_2, 0,0,0,0,0)
- At node a_3, the DPPM label is updated as $P(a_3) = (a_2,a_3,0,0,0,1)$. At this stage, distance field is incremented by 1.
- Upon reaching at aggregate node A, the DPPM label is updated and becomes $P(A) = (a_2,a_3,A,A,0,2)$.
- When aggregate node B receives the packet, it updates the packet by putting its ID in the tail field as $P(B) = (a_2,a_3,A,A,B, 3)$. The value of initial and head remains the same and distance is incremented by 1.
- Similarly, aggregate node C and D successively update the packet.

Finally, the packet reaches the base station with DPPM label (a_2,a_3,A,C,D,5).

Single-hop WBAN Topology

For single-hop WBAN topology, only 10 bytes are reserved in MAC data payload and labeled as P(s) = (Source, Initial, Head, Tail, Distance). The *End* field is redundant and thus is eliminated. The rest of the marking procedure for single-hop topology is same as discussed in section "Multi-hop WBAN Topology".

Uniform Residual Probability

As discussed in section "Proposed Traceback Technique", the key feature of proposed technique is to maintain a uniform residual probability φ_i. To attain this, each node chooses its marking probability $\tau_i = 1/d$ where d = (1, 2… N) and defined as a traveling distance of a packet from its source until it reaches the victim.

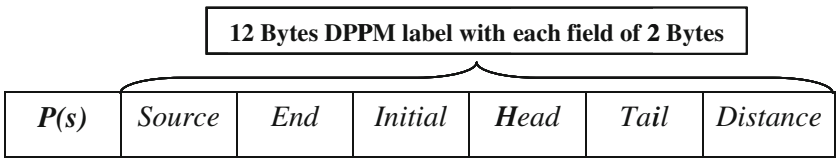

Fig. 8 DPPM label

Fig. 9 Sensor Nodes Connecting with an Edge

For N sensor nodes, the residual probability is given as:

$$\varphi_N = \frac{1}{N} \tag{9}$$

Similarly, for other nodes the residual probability φ_i is calculated by solving Eq. 1:

$$\begin{aligned} \varphi_i &= \tau_i \prod_{j=i+1}^{n} (1-\tau_j) \quad 1 \le i < N \\ \varphi_i &= \frac{1}{N} \qquad for \;\; 1 \le i < N \end{aligned} \tag{10}$$

From Eqs. (9) and (10), it is concluded that each node n_i along the attack path has maintained a uniform residual probability φ_i to mark each packet before it reaches the victim. This shows that each packet has been marked legitimately and no packet has been left unmarked by any node which results in no uncertainty at all. It is further evaluated in section "Performance Evaluation".

DDoS Attacker Traceback

After successful packet marking, the next step is the path reconstruction and identification of an attacker. Based on the collected marked packets, victim v executes the attack path reconstruction process. The proposed technique divides the reconstruction process into two procedures: (1) Aggregate nodes path reconstruction, and (2) Sensor node path reconstruction within the cluster.

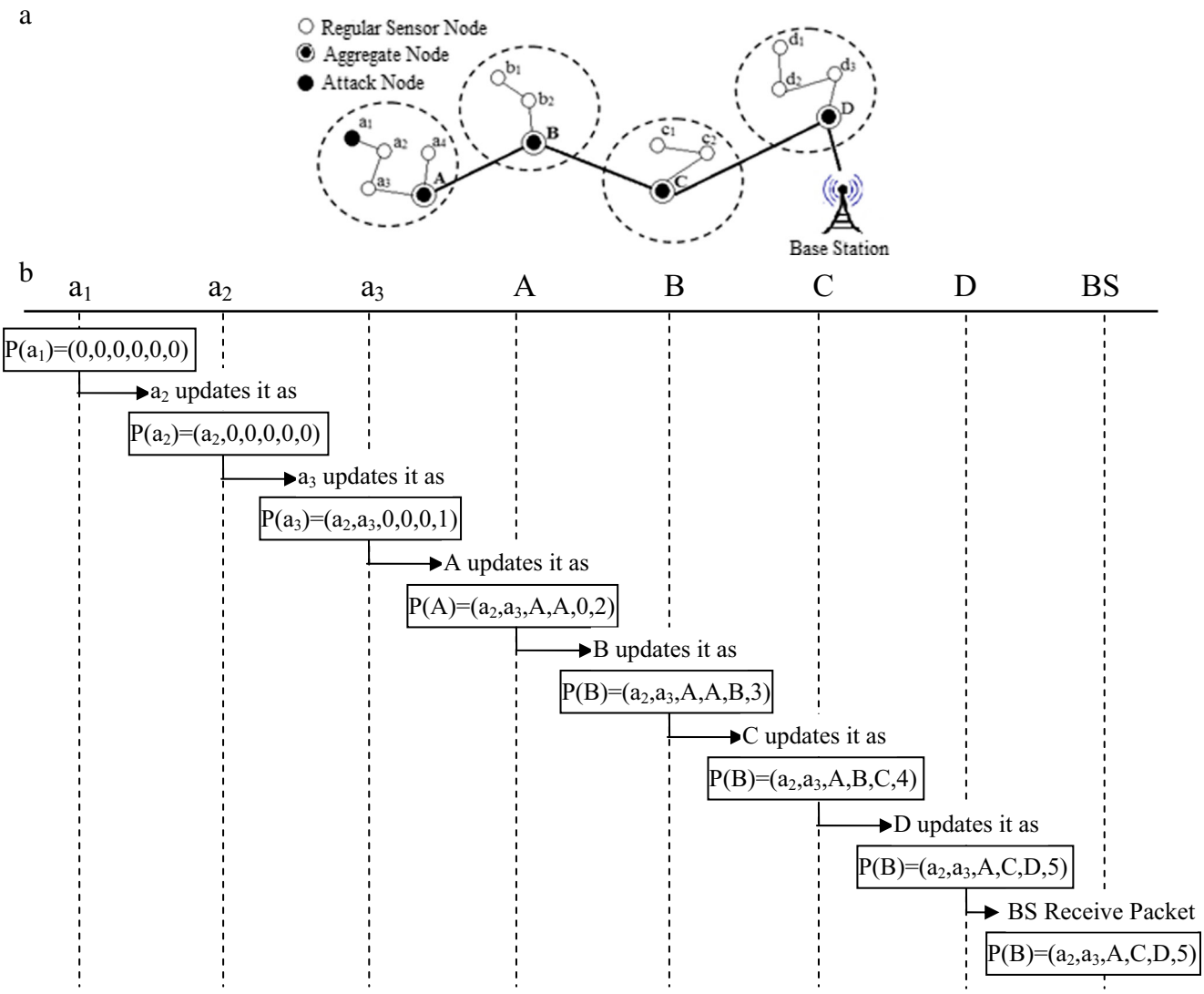

Fig. 10 **a**: Multi-Hop WBAN Topology. **b**: Sequence of Packet Traveling Along the Path

A. **Procedure for Aggregate Node Path Reconstruction**

This procedure reconstructs the path from victim to the aggregate node of the cluster that contains the attacker and the source node. The procedure for aggregate node path reconstruction is shown in Algorithm 1.

Algorithm 1: Aggregate Node Path Reconstruction at Victim

```
Require: S: Set of attack packets at victim v Packet x; y
   Stack S1
   String path
1: BEGIN Procedure PathReconstructionForSensorNode( ).
2: Group the packets in set S based on Initial field
3: for (each Group G1 in S) do
4:     x = FindLeaf (G1)       //Function Call
5:     S1.push(x.Head)
6:     y = FindParent (x,Head,G1)       //Function Call
7:     while y ≠ 0 do
8:        S1.push(x.Head)
9:        x = y.
10:       y = FindParent(x,Head,G1)       //Function Call
11:   end while
12: end for
13: path = AggregateNodePathReconstruction(S1)     //Function Call
14: END Procedure
15:
16: BEGIN Procedure Packet FindLeaf(Group G1)   //Function Definition
17: for (each packet j in G1) do
18:    if j:Tail == 0 then
19:       RETURN path
20:    end if
21: end for
22: END Procedure
23:
24: BEGIN Procedure Packet FindParent(Packet k,G1)  //Function Definition
25: for (each packet j in G1) do
26:    if j:Tail == k then
27:       RETURN j.
28:    end if
29: end for
30: END Procedure
31:
32: BEGIN Procedure String AggregateNodePathReconstruction(Stack S1) //Definition
Require: String path
33: while S1:IsEmpty( ) ≠ 0 do
34:      path += S1.pop( )
35: end while
36: RETURN path
37: END Procedure
```

B. **Procedure for Sensor Node Path Reconstruction**:

This procedure performs the path reconstruction from aggregate node to the source node from where the attack originates. The procedure for sensor node path reconstruction is given in Algorithm 2.

Algorithm 2: Sensor Node Path Reconstruction at Victim

```
Require:
   Packet j; k
   Stack S2
   String path
1: BEGIN Procedure PathReconstructionAtSensorNode().
2: Find Aggregate Node Packet AggPacket at Aggregate Node A.
3: for (every packet i in A) do
4:       if (i.Initial = = A) then
5:          AggPacket = i.
6:       end if
7: end for
8: Find Parent of Aggregate Node A
9: for (every packet i in A) do
10:     if (i.Source = =AggPacket.Source)&&(i.End = = AggPacket.End)&&(i.Initial = 0) then
11:        j = i.
12:     end if
13: end for
14: S2.push(j.End)
15: k = FindParent(j.Source)     //It will return the packet which has End value same as the
                                   input Parameter.
16: while k ≠ 0 do
17:      S2.push(k.End)
18:      j = k.
19:      k = FindParent(j.Source)      //Function Call
20: end while
21: S2.push(k.Source) //k.Source is the Intruder Node
22: path = CompromisedNodePathReconstruction(Stack S2) //Function Call
23: END Procedure
24:
25: BEGIN Procedure String CompromisedNodePathReconstruction(Stack S2) //Definition
Require: String path
26: while S2:IsEmpty( ) ≠ 0 do
27:      path += S2.pop()
28: end while
29: RETURN path
30: END Procedure
```

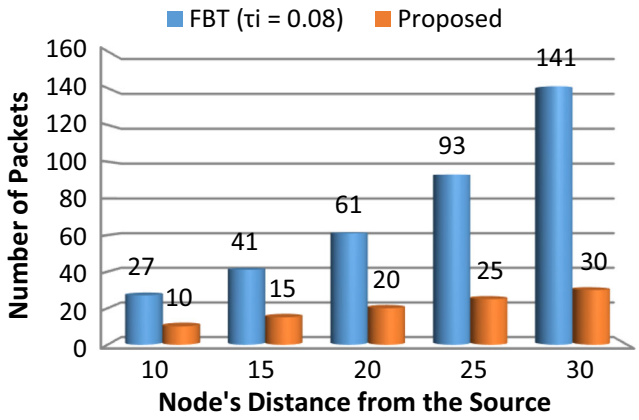

Fig. 11 Number of packets required by proposed technique and FBT ($\tau_i = 0.08$)

Performance Evaluation

In this section, the performance of proposed traceback technique is evaluated. The network simulator NS-2 was used to evaluate the performance in terms of following metrics including: convergence time, overhead and uncertainty. The results are compared with existing traceback techniques for both multi-hop and single-hop sensor network. The results show that the proposed technique is better than FBT [15] that used fixed marking probability.

Convergence Time

Convergence time is measured a number of packets needed for a successful attack path reconstruction [11]. It depends on the uniform residual probability φ_i.

From Eq. (2), the most prominent aspect of the traceback convergence time for PPM [15] is given as: $CT_{FBT}\ \tau(1\text{-}\tau)^{N-1} \geq 1$. Thus keeping τ and N fixed for FBT [12], we get:

$$CT_{FBT} \geq \frac{1}{\tau(1-\tau)^{N-1}} \tag{11}$$

As we learned from Eqs. (9) and (10) that $\varphi_i = 1/N$, therefore for proposed technique the traceback convergence time is given as:

$$CT_{ETT} \geq N \tag{12}$$

Figure 11 shows the number of packets required by proposed traceback technique and FBT to reconstruct the attack path. For FBT, we assume the fixed marking probability of 0.08. The graph clearly indicates that the proposed packet marking technique has less convergence time. For FBT, the convergence time is exponential to the length of attack path which means that the convergence time increases with the increase in path length.

Table 1 compares the numerical values of CT_{FBT} and CT_{ETT} for different number of node's distance from the source. It is evident from the table that the proposed CT_{ETT} requires less amount of packets for attack path reconstruction which means that it has less convergence time as compared to CT_{FBT} with different marking probabilities.

Uncertainty

For PPM, the maximum uncertainty is given as: $(m = (1/\tau)\text{-}1)$ in [7] which shows that PPM locates few possible attackers under spoofed marking attack. Figure 12 shows the uncertainty values of PPM for different marking probabilities τ_i. As the value of τ increases, the uncertainty factor decreases. Again, choosing a large value of τ is not a good solution.

As discussed in section "Working Example", each node n_i along the attack path has maintained a uniform residual probability φ_i to mark each packet before it reaches a victim. Concluding this shows that each packet has been marked legitimately and no packet has been left unmarked ($\varphi_i = 0$) by any node which results in no uncertainty at all which means ($m = 0$) for the proposed technique. This indicated that proposed ETT allows locating actual attacker under DDoS attack.

Table 1 Number of packets required for path reconstruction

No. of Nodes	FBT ($\tau_i = 0.02$)	FBT ($\tau_i = 0.04$)	FBT ($\tau_i = 0.06$)	FBT ($\tau_i = 0.08$)	FBT ($\tau_i = 0.10$)	FBT ($\tau_i = 0.20$)	FBT ($\tau_i = 0.30$)	FBT ($\tau_i = 0.35$)	Proposed Technique
10	59	38	29	28	24	36	79	129	10
15	64	42	37	39	43	125	489	1209	15
20	73	55	56	63	74	337	2835	10,432	20
25	83	66	75	94	125	1102	17,386	87,983	25
30	89	85	99	152	302	3182	101,625	765,292	30

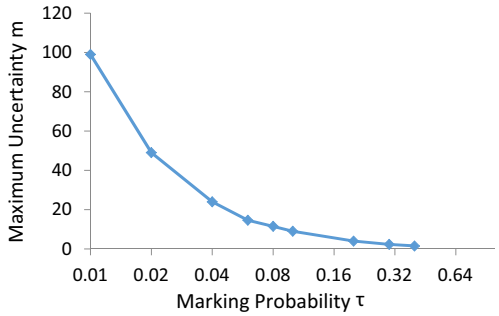

Fig. 12 Uncertainty values for PPM with Different Marking Probabilities

Overhead on Nodes

A key issue of WBAN is its resource scarcity. Therefore, any traceback technique should ensure less overhead cost on WBAN nodes. In this section, we estimate and compare the overhead on nodes under FBT and proposed technique.

The proposed technique has to determine the traveling distance of each node from its origin and therefore, it is expected that its overhead cost is more for marking packets as compared to FBT that uses fixed marking probability. Despite that, this assumption is not correct, because each node only inspects the packet and increment the distance field by one for each incoming packet. Hence, the cost of proposed technique turns out to be very less than FBT.

For simplicity, first we calculate the overhead on individual nodes and then the total overhead on all the nodes along the path has been computed.

For FBT, a fixed marking probability τ_i is assigned to every node for packet marking. If there are n numbers of packets in a DDoS attack, the overhead on every individual node is calculated as:

$$OH_{FBT} = n\tau_i \tag{13}$$

For the proposed technique, every node chooses a marking probability of 1/d (for d = 1,2,…,N) to mark packets. In this case, the overhead on every node turns out to be:

$$OH_{ETT} = \frac{n}{d} \tag{14}$$

Figure 13 gives a comparison of individual nodes overhead for both FBT and the proposed technique, where number of packets are $n = 10{,}000$, total number of nodes are $N = 15$ and marking probability for FBT is assume to be $\tau_i = 0.3$.

It is evident from the graph that under FBT, all nodes have same overhead. On the contrary, under ETT, only first two nodes undergo high overhead after that the overhead drops rapidly as the path length increases.

Similarly, the total overhead for FBT and the proposed ETT depends upon N which defines as the total number of nodes on the reconstruction path.

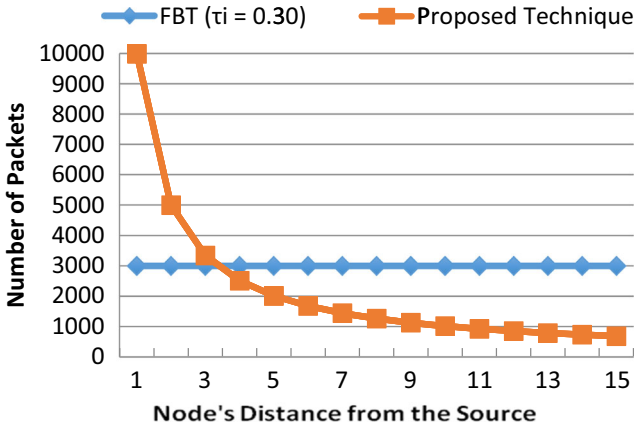

Fig. 13 A Comparison of Overhead on Individual Nodes

Recalling Eqs. (13) and (14), the total overhead under FBT is calculated as:

$$TOH_{FBT} = n\tau_i N \tag{15}$$

For proposed ETT, total overhead is calculated by summing all N terms and is represented as:

$$TOH_{ETT} = n\left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{N}\right) = nH_N \tag{16}$$

Where H_N is the Nth harmonic number. Table 2 shows the comparison of total overhead on nodes under FBT and the proposed technique.

Conclusion

In resource constrained cloud-assisted WBAN, identifying the source of distributed denial of service attack and reconstructing an attack path are the key challenges due to the resource constrained nature of these networks. Traceback techniques proposed for standard IP-based networks are not appropriate for sensor networks due to additional overhead requirements and high convergence time. Similarly existing techniques proposed for mobile ad-hoc networks requires additional processing and storage requirements.

In this paper, an efficient traceback technique is proposed that can be deployed in cloud-assisted WBAN environments.

Table 2 Total overhead on nodes

Number of Nodes	FBT ($\tau_i = 0.20$)	FBT ($\tau_i = 0.30$)	FBT ($\tau_i = 0.35$)	Proposed Technique
10	2	3	3.5	2.93
15	3	4.5	5.25	3.32
20	4	6	7	3.6
25	5	7.5	8.75	3.82
30	6	9	10.5	4

The proposed technique assigns the dynamic marking probability to each node based on the number of hops the packet travelled once it originates from the source. The number of hops can be calculated as the distance travelled by the packets from the source. Finally, a path reconstruction algorithm is proposed that efficiently traceback the attacker.

The performance of the proposed DDoS attack traceback technique is evaluated and compared for the variation in results. The results acquired from simulation experiments were analyzed and compared in terms of convergence time, overhead on individual nodes, total overheads on all nodes and uncertainty in marking packets. The results comparison shows that the proposed technique outperforms existing techniques in all respects.

The paper has few limitations, the number of bytes assigned for DPPM label depends upon the topology of network being deployed. Another limitation is that the proposed scheme uses WBAN with MAC header only. It can be also be deployed and evaluated for IPv6 header.

Acknowledgments The authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for its funding of this research through the Research Group Project no. RG-1435-048. The authors would also like to thank the National University of Sciences and Technology, Pakistan for its support during the research.

References

1. Latif, R., Abbas, H., and Assar, S., Distributed denial of service (DDoS) attack in cloud-assisted wireless body area networks: a systematic literature review. *J. Med. Syst. (Springer)* 38(128):1–10, 2014.
2. Khan, F. A., Ali, A., Abbas, H., and Haider, N., A cloud-based healthcare framework for security and patients' data privacy using wireless body area networks. *Procedia Comput. Sci.* 34:511–517, 2014.
3. Latif, R., Abbas, H., Latif, S., Masood, A. EVFDT: an enhanced very fast decision tree algorithm for detecting distributed denial of service attack in cloud-assisted wireless body area network. *Mob. Inf. Syst*. 2015, Article ID 260594:1–13, 2015.
4. Latif, R., Abbas, H., Latif, S, Masood, A. Performance Evaluation of Enhanced Very Fast Decision Tree (EVFDT) Mechanism for distributed denial of service attack detection in healthcare systems. healthcare on smart and mobile devices. *Ann. Telecommun.*:1–11, 2015.
5. Latif, R. Distributed denial of service (DDoS) Attack detection and prevention mechanisms for cloud- assisted Wireless Body Area Networks (WBANs). Doctoral Thesis, National University of Sciences and Technology, NUST, Pakistan.
6. Irum, S., Ali, A., Khan, F. A., Abbas, H. A hybrid security mechanism for intra-WBAN and inter-WBAN communications. *Int. J. Distrib. Sens. Netw*. 2013, Article ID 842608:1–11, 2013.
7. Waqar, A., Raza, A., Abbas, H., and Khurram Khan, M., A framework for preservation of cloud users' data privacy using dynamic reconstruction of metadata. *J. Netw. Comput. Appl.* 36(1):235–248, 2013. doi:10.1016/j.jnca.2012.09.001.
8. Bellovin, S.M. ICMP Traceback Messages. Internet Draft: draft-ietf-itrace-04.txt, expires. 2003.
9. Snoeren, A. C., Partridge, C., Sanchez, L.A., Jones, C. E. Hash-Based IP Traceback. In: Proceeding in ACM. SIGCOMM, pp 3–14, 2001.
10. Savage, S., Wetherall, D., Karlin, A., Anderson, T. Practical network support for IP traceback. In: Proceeding in ACM SIGCOMM, pp 295–306, 2000.
11. Andrey, B., Nirwan, A. IP Traceback with deterministic packet marking. *IEEE Commun. Lett*. 7(4), 2003.
12. Jin, X., Zhang, Y., Pan, Y., Zhou, Y., and ZSBT, A novel algorithm for tracing DoS attacker in MANETs. *EURASIP J. Wireless Commun. Netw.* 2006:9, 2006.
13. Sy, D., Bao, L. CAPTRA: coordinated packet traceback. In Proceedings of the 5th International Conference on Information Processing in Sensor Networks (IPSN), pp 152–159, 2006.
14. Abbas, H., Magnusson, C., Yngstrom, L., and Hemani, A., Addressing dynamic issues in information security management. *Info. Manag. Comp. Secur.* 19(1):5–24, 2011. doi:10.1108/09685221111115836.
15. Bo-Chao, C., Huan, C., and Guo-Tan, L., FBT: an efficient traceback scheme in hierarchical wireless sensor network. *Secur Commun. Netw.* 2(2):133–144, 2009.
16. Thing, V. L. L., Lee, H. C. J., Sloman, M., Zhou, J. Enhanced ICMP traceback with cumulative path. In proceedings of 61st IEEE Vehicular Technology Conference. (VTC 2005), Vol. 4, Sweden, pp. 2415–2419, 2005.
17. Liu, J., Lee, Z., and Chung, Y., Dynamic probabilistic packet marking for efficient IP traceback. *Comput. Netw.: Int. J. Comput. Telecommun. Netw.* 51(3):866–882, 2007.
18. Park, K., Lee, H. On the On the Effectiveness of Probabilistic Packet Marking for IP Traceback Under Denial of Service Attack. In Proceedings of 2001 I.E. INFOCOM Conference, 2001.