© 2016 IJSRSET | Volume 2 | Issue 2 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

Assured Scheme for Investigating Provenance Falsification & Packet loss Attacks in Wireless Sensor Networks

Shylaja B N, Devaraja S

Rajiv Gandhi Institute of Technology, Bangalore, Karnataka, India

ABSTRACT

Numerous application domains deploy large scale sensor networks and the data they sense are used in critical infrastructure for decision making. The data packet travel from source node along intermediate nodes to destination where aggregation is done for the original message. So while the data packets are travelling through the specified network there may be chances of malicious adversary introducing additional fake nodes to the existing network to track the information or it may compromise the network to get information. Therefore, assuring high data trustworthiness is crucial for correct decision-making. Data provenance represents a key factor in evaluating the trustworthiness of sensor data. The challenging requirements of provenance management for sensor networks are low energy and bandwidth consumption, efficient storage and secure transmission. In this paper, a novel assured scheme is proposed to securely transmit provenance for sensor data. The proposed technique relies on in-packet Bloom filters to encode provenance. Introducing the efficient mechanisms for provenance verification at each node will guarantee us in preserving the provenance falsification by investigation. Another feature concentrated here is packet drop attacks staged by malicious data forwarding nodes. Evaluation of the results prove effectiveness and efficiency of assured scheme for investigating provenance falsification & packet loss attacks in wireless sensor networks.

Keywords : Sensor network, Security, Provenance, Bloom Filtering

I. INTRODUCTION

1.1 Data provenance at sensor network

Sensor networks are used in various areas like cyber physical infrastructure systems, environmental weather monitoring, power grids, etc. Data are originated from a huge number of sensor node sources and they are processed at intermediate hops in networks. These data's finally going to a base station (BS) which performs decision-making about where to go next. The uniformity of data sources creates assurance of the trustworthiness of data. This type of trustworthy information is considered in the decision making process at the base station. The data trustworthiness is assured by data provenance scheme. This is an effective method since it summarizes the history of ownership on the data and the list of actions performed on that information. The big advantage of this provenance scheme is detecting packet loss attacks organized by malicious/compromised sensor nodes. The major disadvantage of this scheme is the use of untrustworthy data at the nodes may create the catastrophic failures (e.g., SCADA systems). Although provenance modeling, collection, and querying have been used extensively in workflows [1] and curated databases [2], provenance at sensor networks has not been fully addressed.

1.2 In packet Boom Filter (iBF)

This is a distributed mechanism in order to encode provenance at the nodes and it will work as centralized algorithm to decode it at the BS. The technical core of this survey is the notion of (iBF) [3]. In this packet consists of a unique sequence number, data value, and an iBF which contains the provenance. The focus of this scheme is a securely transmitting provenance with the data to the BS. In this aggregation framework, securing the data values is an important factor,. The secure provenance technique can be used to obtain a complete

solution that provides security for data, provenance and data-provenance binding. Data-Provenance Binding, so the attacker cannot successfully drop or alter the legitimate/valid data while containing the provenance with the data, or swapping the provenance of two packets

1.3 Detecting Packet Drop Attacks

Provenance encoding could be used for a packet acknowledgement. By using this sensor can transmit more meta-data. For any individual data packet, the provenance record generated by a node will now consist of the node ID and an acknowledgement in the form of a sequence number of the lastly seen (processed/forwarded) packet belonging to that data flow. If the intermediate packet could be drop by the attacker means some nodes on the path do not receive that packet. Hence, during the next round of packet transmission the mismatch between the acknowledgements should be generated from different nodes on the path. This factor could be to detect the packet drop attack and to localize the malicious node.

1.4 Data recovery in case of base station failure

This is an added feature to above which is technique to recover the data when a base station fails due to any power loss at base station or any other physical failures. Base station is only system which has all the data to be aggregated so if that goes wrong there will large amount of data loss which may take long time to recover so we are concentrating on fast recovery of from base station failures.

II. METHODS AND MATERIAL

A. Objectives

- **Confidentiality.** Only authorized parties (e.g., the BS) can process and check the integrity of provenance.
- **Integrity.** An adversary, acting alone or colluding with others, cannot add or remove non-colluding nodes
- **Freshness.** An adversary cannot replay captured data and provenance without being detected by the BS.
- It is also important to provide Data-Provenance Binding,
- Recovery of data when base station fails. more precisely,
- Providing security for all nodes.
- To detect origin forgery in between nodes.
- To detect loss of packet/packet drop.
- Give more Secure Scheme for data transmission.

B. Literature Survey

In 2006 K. Muniswamy-Reddy et al[2], propose “Provenance Aware Storage systems,” .This survey states that in a multi-hop sensor network by using the data provenance scheme the BS can trace the source and forwarding path of an individual data packet. For each packet Provenance must be recorded but there is an important challenge arises due to the heavy storage, energy and bandwidth conditions of sensor nodes. So, it is necessary to provide a light-weight provenance scheme with low overhead.

Disadvantages

- Sensors often operate in a UN trusted environment, so there may chance of attacks.
- The necessary to address security requirements such as confidentiality, integrity and freshness of provenance should be increased.

[4]In 2005 R. Hasan et al proposes “threat model for wireless sensor networks”. The assumption about the BS is it should be a trusted one, but if any other arbitrary node may be attacked means the also be changed to malicious. An attacker can eavesdrop and perform traffic analysis anywhere on the path. In addition to this he/she is able to organize a few malicious nodes, as well as compromise/attack a few legitimate nodes by capturing them and physically overwriting their memory. If an attacker compromises a node means it can extract all key materials, data, and codes stored on that node. The adversary can drop, inject or alter packets on the links which are under the control of attacker. Also the attacker can create the denial of service attacks such as the complete removal of provenance. If a data packet does not contain no provenance records means it considered as highly suspicious data and hence generate an alarm/signal at the BS about this malicious packet arrival. To overcome this type of detection the attacker

attempts to misrepresent the data provenance [5] In 2012 S. Roy et al propose "Secure Data Aggregation in Wireless Sensor Networks," .This work deals with attacks against the synopsis diffusion. This aggregation work presents a lightweight verification algorithm to make verification at the BS. The several synopses generated should be verified independently by the verification protocol at three phases. The phases are query dissemination phase, aggregation phase and the verification phase. In the first phase called query dissemination phase, the BS broadcasts the aggregation name to compute a random seed. In second phase called the aggregation phase, each node computes a sub aggregate value based on the local value and the synopses of its children. The node also randomly selects a set of MACs .From the selected MACs check whether it should be the received ones from its children. Finally, in the third phase called verification phase, the BS computes the final synopses using the messages from its child nodes and verifies the received MACs.

Disadvantages

- Employs separate transmission channels for data and provenance [6] but the provenance only requires a single channel for both.
- Furthermore, traditional provenance security solutions use intensively cryptography and digital signatures [4], and they employ append-based data structures to store provenance, leading to prohibitive cost and time.

[7] In 2008 A. Ramachandran et al proposed "Packets with Provenance" .This scheme catches provenance for network packets in form of per packet tags. The captured information stores a history of all nodes and processes that packet and manipulates those packets. However, this scheme assures a trusted environment which is not practical in sensor networks.

[8]In 2010 W. Zhou et.al proposes "Querying and Maintenance of Network Provenance at Internet- Scale" which describes the history and sub part of the network state. This result came from the execution of a distributed protocol. The disadvantage of this system is also does not address security concerns and is specific to some network use cases.

[9] In 2011W. Zhou, et.al, proposes a "Secure Network Provenance," .This extends network provenance up to the adversarial environments. Even though all of these systems are general purpose network provenance systems but they are not optimized for the resource constrained sensor networks.

[10] In 2010 A. Syalim et al propose a "Preserving Integrity and Confidentiality of a Directed Acyclic Graph Model of Provenance," .The chain model of provenance ensure integrity(no one can change the data other than the original user) and confidentiality(no one can see the data other than original user)through encryption, checksum and incremental chained signature mechanism. Syalim et al. extend this method by applying digital signatures. This signature applied to a DAG model of provenance.

Disadvantages

- These generic solutions are not aware of the sensor network specific assumptions, constraints, etc.
- Since provenance tends to grow very fast, transmission of a large amount of provenanceinformation along with data will incur significant bandwidth overhead, hence low efficiency andscalability.

[11] In 2006 N. Vijayakumar et al proposes "Towards Low Overhead Provenance Tracking in Near Real-Time Stream Filtering,". This system is an application specific system for near-real time provenance collection in data streams. Nevertheless, this system traces the source of a stream long after the process has completed.

[12] In 2010 Chong et al proposes" Self-Identifying Sensor Data". This scheme embeds the provenance of data source within the data. While it reflects the issues related to the confidentiality, Integrity and efficiency but it is not considered as a security mechanism. Also it does not deal with malicious attacks. However practical issues like scalability, data degradation have not been well addressed. In networking applications Bloom Filters are commonly used. In Packet Bloom Filters have only recently gained more attention being utilized in applications such as credential based data path security [13], IP trace back [14], source routing and multicast [15], [16], etc. The basic idea in these works is to encode

the link identifiers constituent to the packet routing path into an In Packet Bloom Filter.

Disadvantages

- The encryption of the whole path is performed by the data source and the intermediate routers check their membership in the In Packet Bloom Filter and forward the packet further based on the decision. This approach is infeasible for sensor networks where the paths may change due to dynamic nature.
- An intermediate router only checks its own membership which may create several integrity attacks such as all-one attack, random bit flips, etc.

III. RESULTS AND DISCUSSION

A. Proposed System

The goal is to design a provenance encoding and decoding mechanism which satisfies security and performance needs. It proposes a provenance encoding strategy in that each node on the path of a data packet securely embeds provenance information within a Bloom filter (BF) should be transmitted along with the data. While receiving the packet the Base Station extracts and verifies the provenance information. The extension of the provenance encoding scheme allows the BS to detect packet drop attack organized by a malicious node. The features are

- Formulate the problem of secure provenance transmission in sensor networks, and identify the challenges specific to this context.
- Design an effective technique for provenance decoding and verification at the base station.
- Extend the secure provenance encoding scheme and devise a mechanism that detects packet drop attacks staged by malicious forwarding sensor nodes.
- Perform a detailed security analysis and performance evaluation of the proposed provenance encoding scheme and packet loss detection mechanism.

Advantages of Proposed System

- The fast message authentication code (MAC) schemes and Bloom filters are fixed-size data structures that efficiently represent provenance.
- Bloom filters make efficient usage of bandwidth, and they yield low error rates
- Claim for Confidentiality: - iBF is computationally infeasible to an attacker to gain data about the sensor nodes included in the provenance.
- Claim for Integrity: - An attacker, acting as single user or colluding with others in the group cannot successfully add or legitimate nodes to the data generated by the compromised/already attack happened nodes.
- An attacker or a set of cooperative attackers cannot selectively add or remove nodes from the provenance of data generated by legitimate nodes.

Corresponding Author: Ms. M. Tharani, Sasuri Academy of Engineering, Coimbatore, Tamilnadu, India.
1058

- A malicious aggregator cannot selectively drop a child node from the provenance.
- Claim for Freshness Provenance replay attacks are detected by the provenance scheme.

B. System Architecture

Investigating the problem of secure and efficient provenance transmission and processing for sensor networks, and using provenance to detect packet loss attacks staged by malicious sensor nodes. The goal is to design a provenance encoding and decoding mechanism that satisfies such security and performance needs. We propose a provenance encoding strategy whereby each node on the path of a data packet securely embeds provenance information within a Bloom filter (BF) that is transmitted along with the data. Upon receiving the packet, the BS extracts and verifies the provenance information. We also devise an extension of the provenance encoding scheme that allows the BS to detect if a packet drop attack was staged by a malicious node system architectures are depicted in figures 1 and figure2.

Sequence diagram

The flows of packets from source to destination via intermediate nodes are depicted in figure 3.

C. Implementation

i. **Secure Provenance** Encoding

We secure provenance technique can be used in conjunction with such work to obtain a complete solution that provides security for data provenance and data-provenance binding. We propose a distributed mechanism to encode provenance at the nodes and a centralized algorithm to decode it at the BS. The technical core of our proposal is the notion of in-packet Bloom filter (iBF). Each packet consists of a unique sequence number, data value, and an iBF which holds the provenance. We emphasize that our focus is on securely transmitting provenance to the Base station. We secure provenance technique can be used in conjunction with such work to obtain a complete solution that provides security for data provenance and data provenance binding.

ii. Provenance Encoding

The Figure 4 shows that to produce the final result, the contributor C5 uses the outputs of contributors C1 and C2 while contributor of C6 uses the output of contributors C3 and C4. Contributor C7 uses the output of C5 and C6 which later used by C8 and C9. C10 is the final process is executed by that processes the outputs of C8 and C9. After each process is executed and the provenance of the process we had created/generated, the provenance is stored in the provenance database. All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

iii. Provenance Decoding

When a Base station receives a data packet .Base station know what the data packet should be checks. Afterwards, upon receiving a packet, it is sufficient for the BS to verify its knowledge of provenance with that encoded in the packet.

Algorithm-1 Provenance Verification:

Input: Received packet with sequence seq and iBFibf.
Set of hash functions H, Data path P = < n l 1 , ..., n 1 , ..., n p >
BF c ⟵ 0 // Initialize Bloom Filter
for each n i ∈ P do
vid i = generateVID (n i , seq)
insert vid i into BF c using hash functions in H
endfor
if (BF c = ibf) then
return true // Provenance is verified
endif
return false

Algorithm-2 Provenance Collection:

Input: Received packet with sequence seq and iBFibf. N
Set of nodes (N) in the network, Set of hash functions H
1. Initialize
Set of Possible Nodes S ⟵ ∅
Bloom Filter BF c ⟵ 0 // To represent S
2. Determine possible nodes in the path and build the representative BF for each node n i ∈ N do
vid i = generateVID (n i , seq)
if (vid i is in ibf) then
S ⟵ S ∪ n i
insert vid i into BF c using hash functions in H
endif
endfor
3. Verify BF c with the received iBF
if (BF c = ibf) then
return S // Provenance has been determined correctly
else
return NULL // Indicates an in-transit attack
endif

IV. CONCLUSION

In this paper we addressed the problem of securely transmitting provenance for sensor networks, and proposed an assured scheme for provenance encoding and decoding scheme based on Bloom filters. The scheme ensures confidentiality, integrity and freshness of provenance. We extended the scheme to in-corporate data-provenance binding, and to include packet sequence information that supports detection of packet loss attacks. Another extension added is to recover the

data loss in case of base station failure. Experimental and analytical evaluation results prove that the proposed scheme is effective and scalable. In future work, we plan to implement a real system prototype and to recover the packets lost during attacks.

V. REFERENCES

[1] I. Foster, J. Vockler, M. Wilde, and Y. Zhao, "Chimera: A Virtual Data System for Representing, Querying, and Automating Data Derivation," Proc. Conf. Scientific and Statistical Database Management, pp. 37-46, 2002.

[2] K. Muniswamy-Reddy, D. Holland, U. Braun, and M. Seltzer, "Provenance-Aware Storage systems," Proc. USENIX Ann. Technical Conf., pp. 4-4, 2006.

[3] C. Rothenberg, C.Macapuna,M.Magalhaes, F. Verdi, and A. Wiesmaier, "In-Packet Bloom Filters: Design and Networking Applications," Computer Networks, vol. 55, no. 6, pp. 1364-1378, 2011.

[4] R. Hasan, R. Sion, and M. Winslett, "The Case of the Fake Picasso: Preventing History Forgery with Secure Provenance," Proc. Seventh Conf. File and Storage Technologies (FAST), pp. 1-14, 2009.

[5] S. Roy, M. Conti, S. Setia, and S. Jajodia, "Secure Data Aggregation in Wireless Sensor Networks," IEEE Trans. Information Forensics and Security, vol. 7, no. 3, pp. 1040-1052, June 2012.

[6] Y. Simmhan, B. Plale, and D. Gannon, "A Survey of Data Provenance in E-Science," ACMSIGMODRecord, vol. 34, pp. 31- 36, 2005.

[7] A. Ramachandran, K. Bhandankar, M. Tariq, and N. Feamster, "Packets with Provenance," Technical Report GT-CS-08-02, Georgia Tech, 2008.

[8] W. Zhou, M. Sherr, T. Tao, X. Li, B. Loo, and Y. Mao, "Efficient Querying and Maintenance of Network Provenance at Internet- Scale," Proc. ACM SIGMOD Int'l Conf. Management of Data, pp. 615-626, 2010.

[9] W. Zhou, Q. Fei, A. Narayan, A. Haeberlen, B. Loo, and M. Sherr, "Secure Network Provenance," Proc. ACMSOSP, pp. 295-310, 2011.

[10] A. Syalim, T. Nishide, and K. Sakurai, "Preserving Integrity and Confidentiality of a Directed Acyclic Graph Model of Provenance," Proc. Working Conf. Data and Applications Security and Privacy, pp. 311-318, 2010.

[11] N. Vijayakumar and B. Plale, "Towards Low Overhead Provenance Tracking in Near Real-Time Stream Filtering," Proc. Int'l Conf. Provenance and Annotation of Data (IPAW), pp. 46-54, 2006.

[12] S. Chong, C. Skalka, and J.A. Vaughan, "Self-Identifying Sensor Data," Proc. Ninth ACM/IEEE Int'l Conf. Information Processing in Sensor Networks (IPSN), pp. 82-93, 2010. [13] T. Wolf, "Data Path Credentials for High-Performance Capabilities- Based Networks," Proc.

[13] ACM/IEEE Symp.Architectures for Networking and Comm. Systems, pp. 129-130, 2008.

[14] R. Laufer, P. Velloso, D. Cunha, I. Moraes, M. Bicudo, M. Moreira, and O.Duarte, "Towards Stateless Single-Packet IP Traceback," Proc. 32nd IEEE Conf. LocalComputer Networks (LCN), pp. 548- 555, 2007.

[15] P. Jokela, A. Zahemszky, C. Esteve, S. Arianfar, and P. Nikander, "LIPSIN: Line Speed Publish/Subscribe Inter-Networking," Proc. ACM SIGCOMM Conf. Data Comm., pp. 195-206, 2009.

[16] A. Ghani and P. Nikander, "Secure In-Packet Bloom Filter Forwarding on the Netfpga," Proc. European NetFPGA Developers Workshop, 2010.

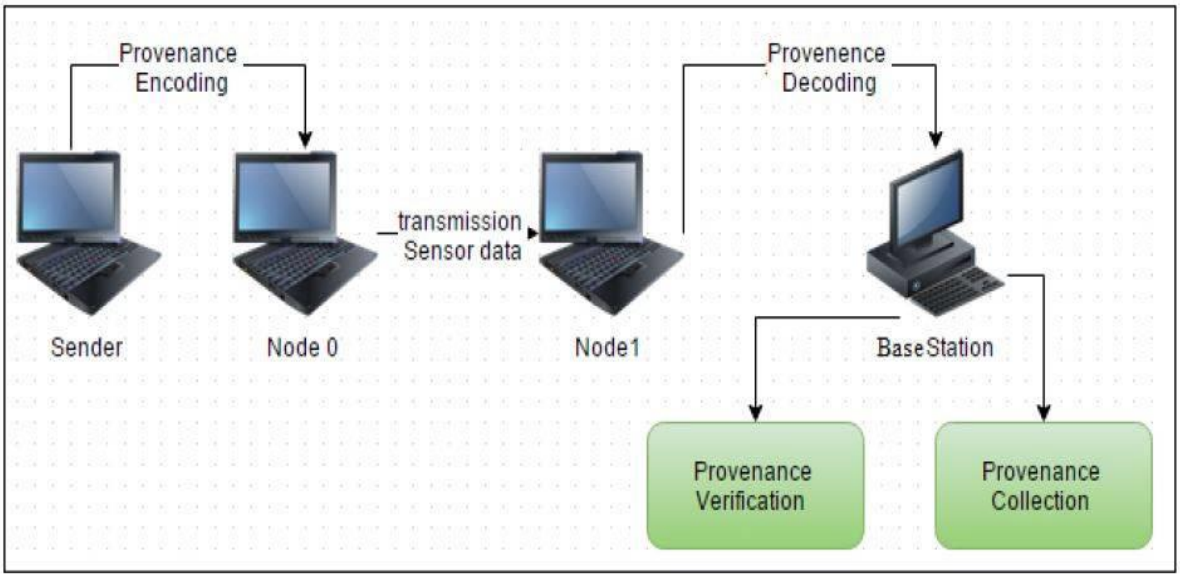

Figure 1: System Architecture

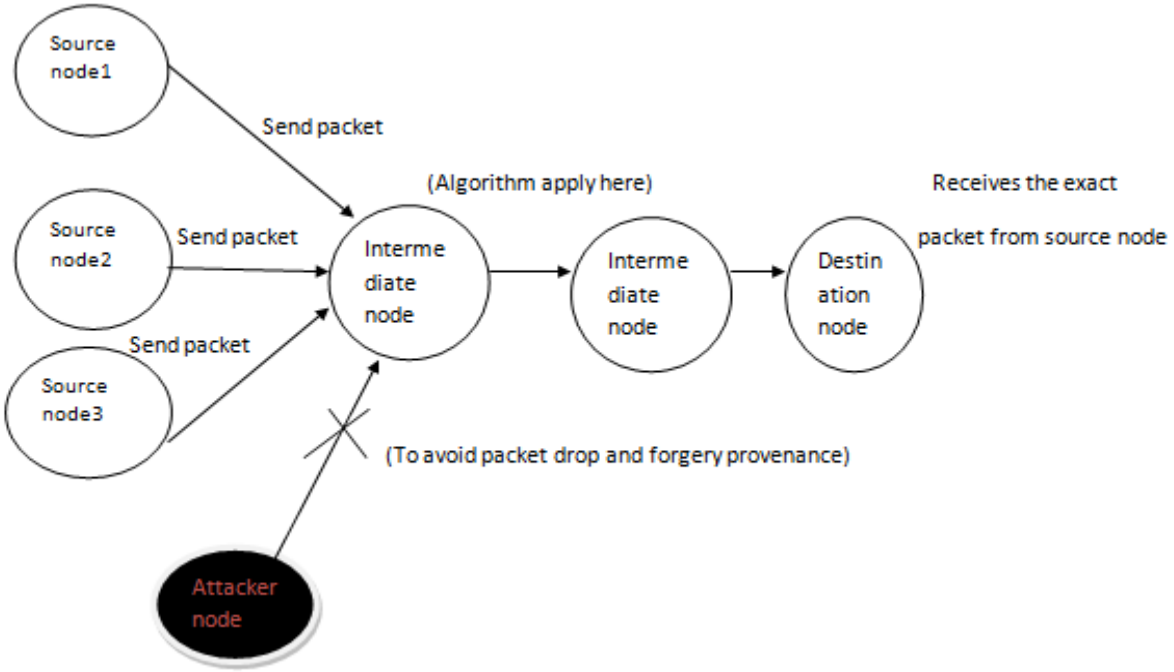

Figure 2 : Internal Architecture Diagram

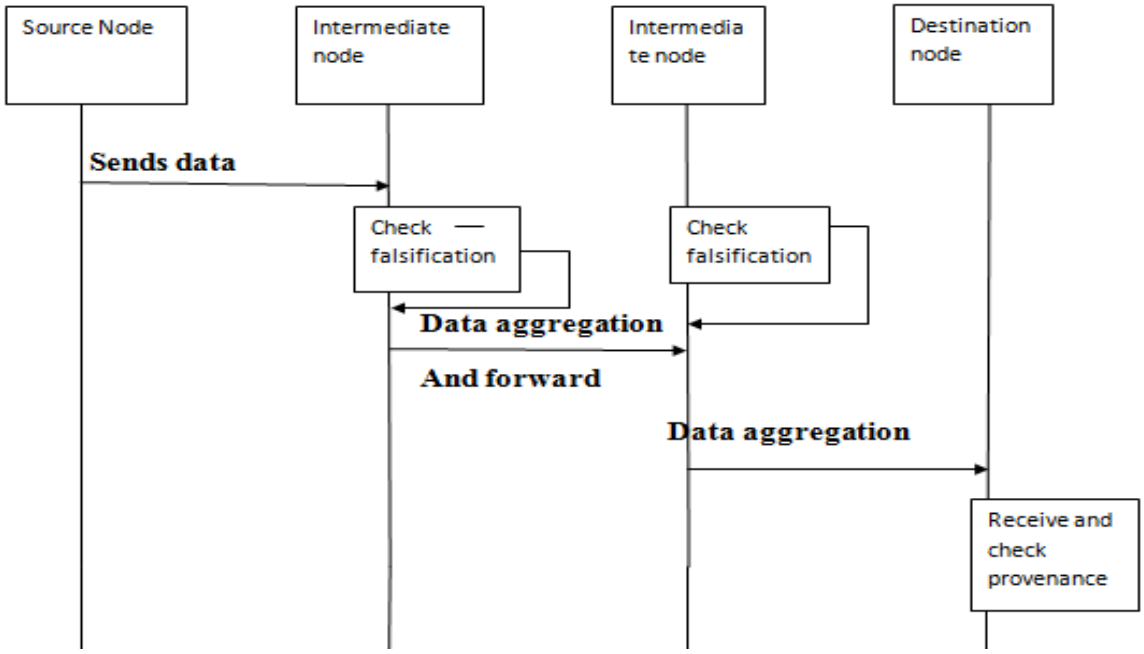

Figure 3 : Sequence Diagram

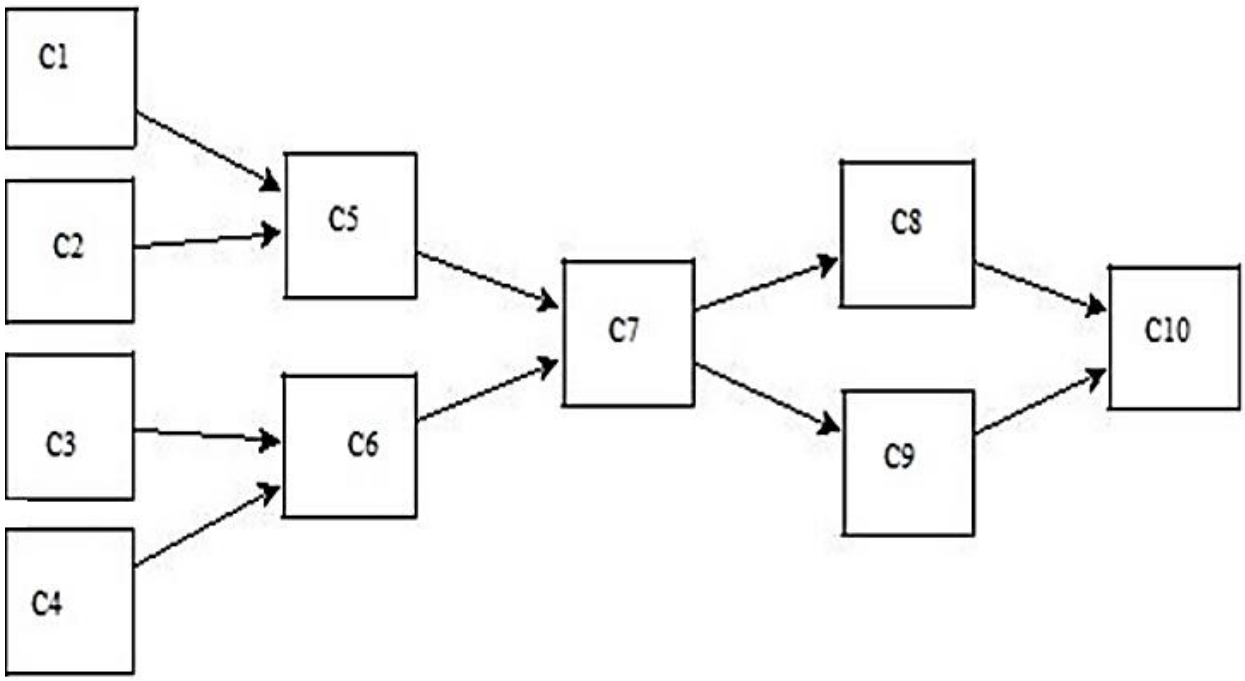

Figure 4 : Provenance Graph