Evaluation of Sybil Attack Detection Approaches in the Internet of Things Content Dissemination

Danilo Evangelista*, **Farouk Mezghani**†, **Michele Nogueira***, **Aldri Santos***
*Federal University of Paraná - UFPR, Curitiba, Brazil
†Université de Toulouse, INP/ENSEEIHT, IRIT, Toulouse, France
*{dfrevangelista, michele, aldri}@inf.ufpr.br, †farouk.mezghani@enseeiht.fr

***Abstract*—The Internet of Things (IoT) comprises a diversity of heterogeneous objects that collects data in order to disseminate information to applications. The IoT data dissemination service can be tampered by several types of attackers. Among these, the Sybil attack emerged as the most critical since it operates in the data confidentiality. Although there are approaches against Sybil attack in several services, they disregard the presence of heterogeneous devices and have complex solutions. This paper presents a study highlighting strengths and weaknesses of Sybil attack detection approaches when applied in the IoT content dissemination. An evaluation of the LSD solution was made to assess its effectiveness and efficiency in a IoT network.**

***Index Terms*—Security and privacy in the Internet of Things, Security in networks, Sybil Detection Techniques**

I. Introduction

The Internet of Things (IoT) consists of a hybrid, open and heterogeneous network which integrates objects, such as lamps, refrigerators, clothes, and computing devices [1]. It provides interaction among objects and also with humans in smart home, industrial environments, and other ones. IoT allows many content dissemination services to be established. Hence, IoT demands cooperation between objects that collaborate by forwarding contents in real time to get services as temperature measurement, object location, and vital signs monitoring [2].

The content dissemination in IoT networks is subject to issues such as link losses, eavesdropping and mobility [3]. The IoT communication must also deal with the diversity of computing resources, once objects and devices have different memory capacity, processing and battery. In this way, attackers can explore such issues to disrupt communication. In a data dissemination, an attacker can drop network packets and send only the ones wanted, as well as personifying the identity of other network users.

Among the malicious actions performed by an attacker to harm the dissemination of data, it is highlighted the identity personification. This action taken by Sybil attack acts on identity manipulation of network devices [4], [5]. This attacker aims to achieve benefits, such as the usage of unauthorized resources, and getting and publishing private information on network users. In IoT, such attack affects the confidentiality and privacy of users, gathering personal information, such as vital data, key of a house or store.

The solutions against Sybil attack can be classified in three perspectives: network features [6], cryptography [7], and relationship between neighbors [8]. Among the most common networking features used by the approaches to identify a Sybil attack is the received signal strength (RSS) and the received signal strength indication (RSSI). However, the mobility of devices can reduce the effectiveness of detection approaches leading to a high rate of false positives. Moreover, cryptographic techniques and relationship between nearby devices require additional communication and processing. In cryptography, the generation of a secure asymmetric key like RSA demand high processing cost. While symmetric keys need of the key change between two nodes to ensure non repudiation, becoming a bottleneck. On the other hand, the relationship between devices approach requires constant updating about legitimate users and attackers, and this can cause overhead on the network. Thus, those approaches imply a *trade-off* between security and performance taking into account resource constraints, scalability and network overhead.

This paper presents a study highlighting strengths and weaknesses of Sybil attack detection approaches when applied in the IoT content dissemination. The goal of this study is identify vulnerabilities related to security and performance in the usage of such approaches in IoT network. The study considers Sybil attacks resulting from stolen and fabricated identities, where for each type of identity an attacker can exhibit churn behavior or multiple identities in a perspective of collusion. Among the existing detection approaches, the Lightweight Sybil Attack (LSD) system, that employs network features, was evaluated in a IoT environment similar to smart home. The LSD effectiveness was measured taking into account four metrics security and one for performance.

The paper is organized as follows: Section II presents related work on effectiveness evaluations on the services of dissemination of content and authentication when subjected to attacks on wireless networks. Section III details the characteristics of the approaches and Sybil attack detection. Section IV evaluates the LSD mechanism under Sybil attacks in an IoT network.

II. Related work

The dissemination of content and the authentication of users are among the most important services on a network, being target of Sybil attacks. Abbas et al. [6] proposed a framework against Sybil attacks on dissemination services in Manets. This framework adopts RSS and RSSI network characteristics to detect the presence of an attacker. However, Both characteristics are affected by eletromagnetic interference damaging the detection. Raghu et al. [9] presented a mechanism employing RSS and hypothesis testing on the dissemination service. Nevertheless, hypothesis testing with RSS can generate false positives due to mobility and interferences. Thus, those adverse effects should be considered by employing network characteristics on a sybil attack detection system on IoT. Silva et al. [10] evaluated the cryptography-based authentication service for MANETs, called PGP-LIKE, against Sybil and blackhole attacks. However, this work employed only performance metrics. Lin [7] proposed an authentication service, called LSR, that applies cryptography based on symmetric keys and a certification authority (CA) to detect the Sybil attack in Vanets. However, the cryptography applied is costly for IoT due to key updates, as well as the usage of CA constraints the network scalability.

In existing IoT evaluating studies, the communication mechanisms taking into account only their efficiency. Blazquez [11], for example, performed an evaluation of the OpenID and Sensei protocols for multimedia data management. Wang et al. [12] evaluated the efficiency of a public key-based scheme on attribute(EBA) for IoT . The KP-EBA and CP-ABE schemes were evaluated on devices with limited resources. However, only performance metrics were considered in the evaluation. One of the major concern on IoT is the security of data traffic on the network. So, an evaluation of attack detection techniques and flaws in the IoT becomes essential.

III. Sybil attack detection techniques

This section describes the operation of Sybil attack detection techniques and discusses their effectiveness in IoT. To support the understanding of these techniques, it is initially defined an IoT network model and the data dissemination model adopted in the work. Next, types of Sybil attacks and behaviors are explained. Finally, it is showed strengths and weaknesses of each detection technique when applied to the IoT.

A. *Network and Sybil Attack Model*

The IoT network model consists on environments composed of objects (*things*) and computing devices (nodes) interconnected. These environments correspond to a residence, a hospital or even an automobile industry. Within the environment, the objects interact with each other and with computing devices, transmitting data collected to a given access point in order to send these data to an application. Further, the users interacting with applications, as showed in Figure 1. The nodes are fixed or mobile, and may have different resource limits. In addition, due to the mobility of nodes the network becomes denser or sparse due to association and dissociation of its components. In the figure, there are interaction between service providers, objects and devices in order to disseminate content to applications and its users. An energy company, for example, collects power information of a residence, and could report to residents the real-time power consumption, allowing control about the energy consumption .

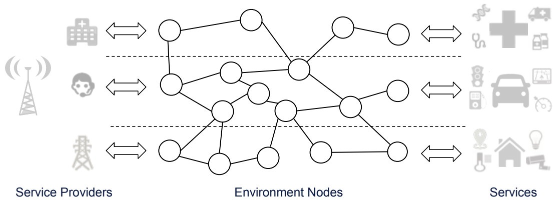


Fig. 1: Network Model

The Sybil attack (SA) manipulates stolen and fabricated identities. In an identity manipulation, an attacker fabricates or exploits the vulnerabilities of wireless networks through packet interception and through the promiscuous mode to get legitimate identities. Then, the attacker chooses one or more identities and requests an association from these identities to cheat the detection. Figure 2 illustrates the SA behavior when performing the personification of legitimate nodes, where identities contained between the keys have been manipulated, damaging the identification of legitimate identities on the process of identifying the legitimate identity of network, and harming the confidentiality of the contents disseminated.

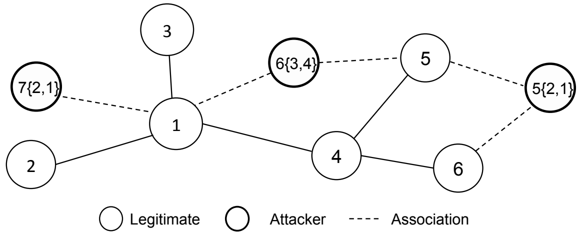


Fig. 2: Sybil Attack over IoT network

A SA seeks advantages by the identity manipulation, where it can cheat the outcome of voting systems, obtain unauthorized resources, access and publish unauthorized information. The attack affects as confidentiality as privacy since an attacker personifies a legitimate identity inflicting the confidentiality of the network. It then discloses the information acquired from a legitimate node, resulting in loss of privacy.

The manipulation of identities by an attacker occurs through identity fabrication and stolen. On the fabrication of identities technique, a attacker node generates its false identity. This technique considers the fabrication of identities through random lists, vectors, and logs. The fabrication of identities using a random list is equivalent to the list F. On the stolen of identities, the attacker can get a list of identities through the promiscuous mode. This list consists of the set of identities R. For simplicity, the F lists and R will be represented by the list $S = F \cup R$, where S is the union of sets of fabricated and stolen identities. Soon after the process of manipulation, the attacker selects the legitimate ones obtained and requests association to a network node. Identity Manipulation aims to achieve the personification of a legitimate network node for resources and private information from members of the dissemination.

A Sybil attacker can use the behavior *churn* to request association to the network. Figure 3 illustrates the action of an attacker with this behavior. In a given time t an attacking node chooses a list of identity S. Then it requests association to the network from that identity. If it can not get access, the attacker disassociates from the network and choose a new identity to start the attack. In the following moments $t + 1$, $t + 2$, the attacker does the same procedure of association and dissociation. This behavior aims the access of a network and also causes the exhaustion of resources due to requests from associations in a given short time.

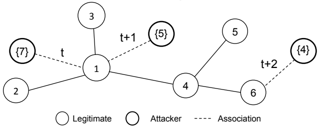


Fig. 3: SA - *Churn* behavior

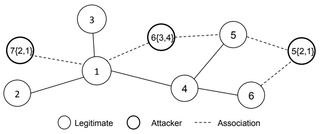


Fig. 4: SA - Multiple ID's

A Sybil attacker can also request association through of multiple identities. In this case, Figure 4 illustrates the conduct of an attacker with multiple identities. So, the attackers request association to the network exhibiting more than one identity. This behavior aims to inflict the confidentiality of the contents disseminated. The result of a vote, for example, can be affected by an attacker who acts maliciously through their identities, changing and getting information on this service. The usage of multiple identities with SA minimizes the effectiveness on the quality of service in a network.

B. Detection Based on Network Features

Approaches based on network characteristics employ both network and nodes attributes to detect a SA. The use of network features becomes more suitable for networks with constrained resources since such approach does not require additional techniques or mechanisms in order to collaborate in detection. However, in general, the networks characteristics are vulnerable to electromagnetic interference, and this reduces the identification of network nodes and requires a lot of RSS analysis of a node. The approach proposed by abbas [6] takes into account RSS, mobility, and coverage area. This approach called, LSD (Lightweight Sybil Attack), is detailed below, highlighting its weaknesses and strengths when applied to the IoT data dissemination.

Figure 5 illustrates the operation of LSD to detect SA on network. The node 1 identifies an attacker by the value of its coverage area and the RSS of a requesting node. Its coverage area is represented by dotted and white areas. The authentication procedure performed by node 1 always happens inside the dotted area, and a requesting node should send its identity while in the dotted area of node. This node validates the access to the network, storing both the RSS value and the identity of the requesting node in a list of tuples $< RSS, ID >$. The list must be updated and shared with the legitimate nodes by broadcasting its over the network. The authenticated requesting node can then communicate with other nodes. Thus, the detection of an attacker occurs when a node requests association on dotted area exhibiting more than one identity, as well as when it already associated presents to legitimate nodes a different identity.

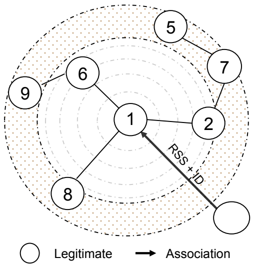


Fig. 5: Detection by the analysis of the network features

LSD presents some weaknesses that reduce its effectiveness in detection of the SA. As nodes have random mobility behavior, the LSD detection rate may present high false positives since the mobility impacts in the location of a node. The electromagnetic interferences are harmful the detection effectiveness, requiring a longer period of time to estimate the node's RSS. Thus, that estimation can further increase the amount of false positives rate. Another factor that reduces the detection rate of LSD is the absence of non-repudiation, where it can only identify SAs with fabricated identities.

However, the approaches based on network features allow the detection of SA at low resource usage. LSD, for example, dispenses extra hardwares, such as GPS and more powerful antennas, to locate an attacker. This fact simplifies the detection process because requiring less computational resources for the detection. Moreover, LSD is portable to other wireless networks contexts since it only requires data about the characteristics of nodes. Hence, LSD initially appears to be suitable for application on networks that with resource constraints as the IoT.

C. Detection Based on Cryptography

The detection approaches based on cryptography of asymmetric and symmetric keys, in general, require devices without restriction of resources. Cryptography may limit the efficiency of a detection approach since it requires high cost for generating secure keys, and needs to keep updated identity lists. Next, we will describe the operation of works [7], [13] on wireless networks that employ these techniques to detect sybil attacks.

Figure 6 shows the operation of Lightweight Sybil Resistance (LSR) [7], which employs symmetric cryptography to identify SA on network. The cryptography scheme adopted by LSR requires certification authority (CA) to grant and revoke the keys to network nodes. Each node gets a pair of keys Pk_i generated by CA, and the communication between nodes only starts when it has distributed keys for all nodes, as well as their respective lists of identities are already updated. As a new node joins to the network, nodes already associated receive of CA a new key Pk_i in order update their identity lists. The SA detection made by LSR is event-driven, in which an event consists of an action performed by a node, ranging from an association request to a task for a set of nodes. Thus, the SA identification happens when at the time two or more events are associated to a given node.

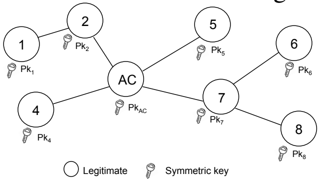


Fig. 6: Detection with SC

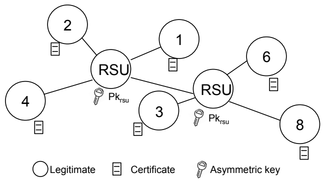


Fig. 7: Detection with AC

As LSR requires a key management through CA to identify SA, this can limits the network scalability, once LSR needs to generate a pair of keys to all nodes. In addition, it causes network overhead due to constant updates on the identity lists. Hence, LSR is not suitable for networks composed of devices with limited resources. The symmetrical cryptography employed by LSR allows the identification of SA with stolen and fabricated identities. Further, since LSR is event-driven, it only requires the verification of the identity associated with the event, obtaining a high detection rate. The mobility of the nodes does not limit the LSR detection of LSR, because the authentication process ignores the node location.

Figure 7 illustrates the operation of the mechanism Defense Against Sybil Attack (DAS) proposed by [13]. DAS employs cryptography with asymmetric keys, and makes use of road side units (RSU) to perform the authentication of nodes through temporary certificates. When a node requests access to the network, RSU authenticates this node generating temporary certificates signed by its private key certificates pk_rsu. Next, RSU shares with other RSUs the new certificate. This mechanism explores space capabilities, time aspects and the correlation between nodes to determine if a given node is an attacker, assuming that this node cannot be in two places at the same time. A detection occurs if a certificate is exhibited in two or more places on the networks at the same time.

DAS needs the constant maintenance of legitimates certificates to detect attackers correctly. So, if an attacker gets a new certificate before all RSUs revoke it, this attacker can be successful. To avoid this fault, the time of synchronization between RSUs must be smaller than the node authentication time. In IoT networks, synchronization between RSUs becomes a more difficult task due to the network density, composed of heterogeneous devices. However, DAS identifies both types of SA. By employing asymmetric keys, DAS ensures non-repudiation, required to detect SA with stolen identities. Thus, a certificate signed by the private key of an RSU guarantees its truthfulness. Further, DAS assumes RSUs distributed to provide a scalable detection, but it does not solve the issue of synchronization since its certificates are temporary.

D. Detection Based on the Relationship Between Neighbors

The detection approaches based on the relationship between neighbors can be limited by the number of messages exchanged between neighbors of a node in order to know information about its behavior. Below, we will describe the operation of work [8] on wireless networks that employs this technique to detect sybil attacks.

Figure 8 illustrates the Mobid scheme [8], that identifies SA through of relationship between neighbors. In this scheme, the nodes keep two identity lists about legitimate nodes and attackers. Each list contains the node identities and its reputation on the network at given time. A node get its reputation by its cooperation measurement with other nodes. In this way, If a node doesn't cooperate, it loses reputation. The Mobid detections requires a constant updating of the reputation of nodes from the views of their neighbors. Thus, the SA identification happens when the majority of a given node neighbors classify it with a low reputation.

In Mobid, the relationship between nodes provides a node classification through of neighbors opinion. However, this classification requires an additional communication, which can reduce the battery life. In Mobid [8], for example, for each task performed by a node is necessary the opinion of its neighbors about the behavior of that node. In addition, an attacker node can cheat the Mobid detection by acting maliciously, cooperating with other nodes to achieve high reputation. In this way, an attacker with high reputation keeps triking on the network, obtaining confidential information. On the other hand, as Mobid considers only the neighbor opinions, and such parameter is simple, Mobid can applied to several networks. However, although Mobid has a high accuracy, it requires an offline training about the malicious behavior to calibrate the detection mechanism, and this issue becomes inadequate for IoT networks.

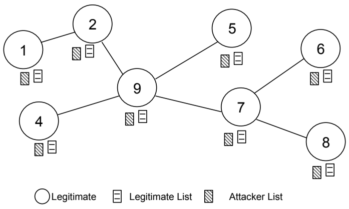


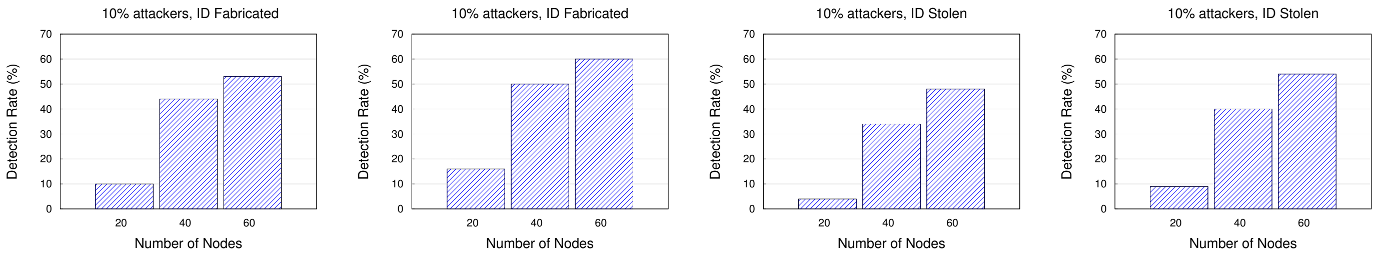
Fig. 8: Detection by neighbors relationship

IV. Evaluation

This section describes the evaluation of the mechanism LSD (Lightweight Sybil Attack Detection Framework) proposed by [6]. This mechanism was chosen because it takes into account IoT network characteristics, such as scalability and low computational complexity. LSD was implemented in the network simulator (NS3), as well as the Sybil attack (SA) with fabricated and stolen identities. The scenario set for the evaluation comprises a residential environment, in which thenodes correspond to objects as refrigerator, stove, television, and computer devices. These nodes act sequentially disseminating a data flow to a destination. A data flow consists of sending a collection of messages of 256-bytes. The choice of source and destination nodes happens randomly and the source node cannot be the destination. Thus, the source disseminates a data stream to its neighbors to forward to the destination. A new dissemination starts only when all previous data is delivered to the destination. Further, an attacker requests association in a network through fabricated and stolen identities. An attacker by requesting an association can have behavior *churn* or exhibit multiple identities ranging from two to five identities.

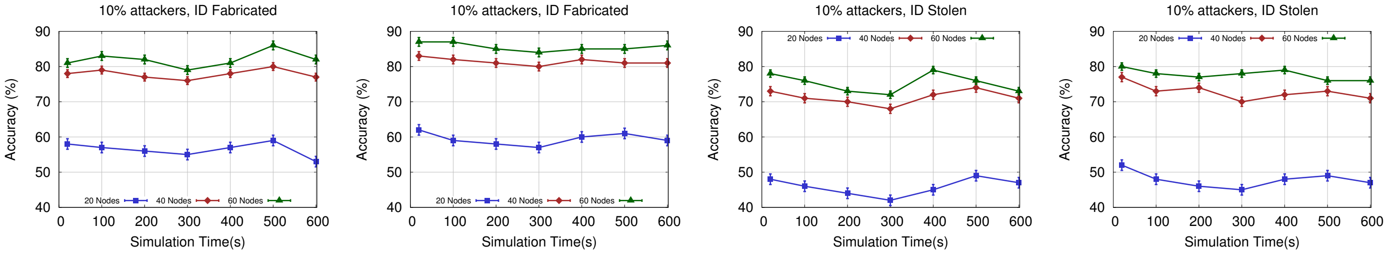
The simulation parameters used in the IoT network configuration consider the number of nodes ranging from 20, 40, and 60. These nodes can be fixed or mobile, where fixed comprise 25%. They also emit the RSS for up to 100 seconds and moving the network through the random mobility model at speeds between 0.2m/s to 2 m/s. The communication among nodes uses 802.15.4 standard. The results were obtained from thirty simulations times with 95% of confidence interval. For the SA parameters, the number of attackers was set at 10%, and SA with multiples identities requests association to network with up to five identities for attack.

The metrics to evaluate the LSD mechanism are arranged in terms of effectiveness and efficiency. The ***Detection Rate***



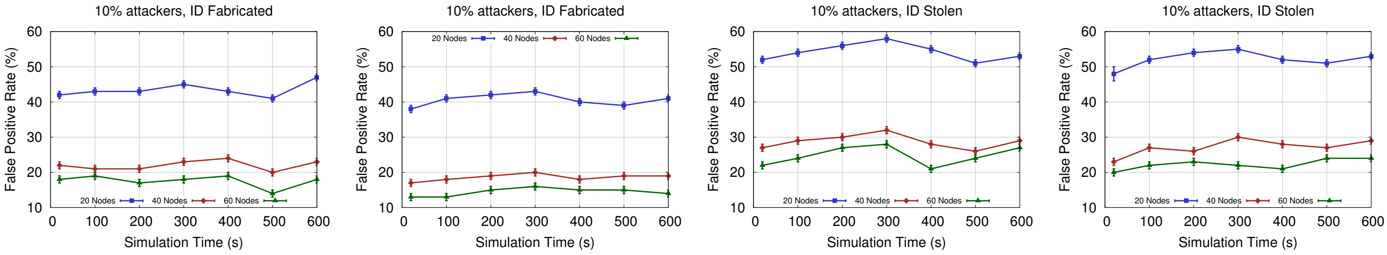
(a) *Churn* (b) Multiple Identities (c) *Churn* (d) Multiple Identities

Fig. 9: LSD's detection rate



(a) *Churn* (b) Multiple Identities (c) *Churn* (d) Multiple Identities

Fig. 10: LSD's accuracy



(a) *Churn* (b) Multiple Identities (c) *Churn* (d) Multiple Identities

Fig. 11: LSD's false positive rate

(T_{det}), ***Accuracy*** (A_c),***False Positives*** (T_{fp}), and ***Effectiveness of Attack*** (T_{efat}) are adopted for effectiveness, and the ***Flow Dissemination Time*** (C_{diss}) for efficiency.

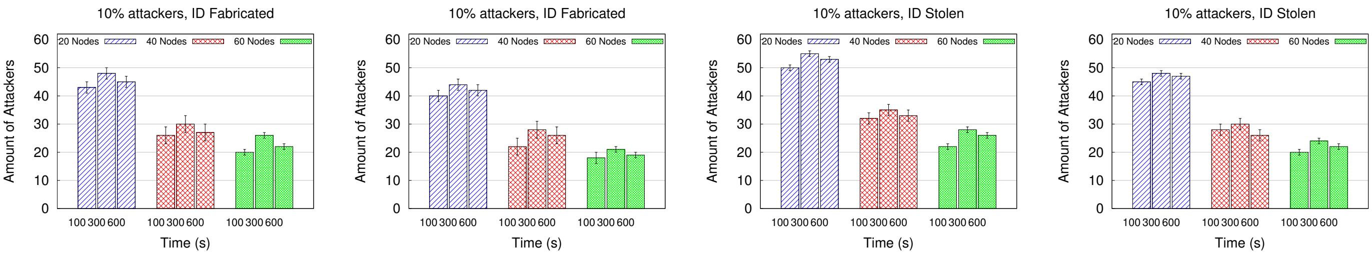
A. Results

Figures 9(a) and 9(c) show the LSD T_{det} in a network under SA with behavior *Churn* and fabricated and stolen identities, respectively. LSD under SA with stolen identities had lower detection on sparse network - 20 nodes due to it requires the cooperation of neighbors to locate a node. Further, for all networks, LSD achieved lower T_{det} under SA with stolen identities than with fabricated. Figures 9(b) and 9(d) illustrate the LSD effectiveness under SA with multiples identities, and fabricated and stolen identities, respectively. Even on sparse network, LSD was more effective to detect SA with multiples identities than under *Churn* because the multiple identities does not make continuous associations and dissociation.

Figures 10(a) and 10(b) show the LSD accuracy under SA with fabricated identities. Under a denser scenario, 60 nodes, it demonstrated better accuracy reaching 81% and 88% respectively. When the network becomes more sparse, the detection rate decreases since there is fewer nodes to aid the detection process. In addition, the precision of detection for 20 and 40 nodes is lower when compared to denser scenario. This is due to the variation in the confidence interval. Figures 10(c) and 10(d) illustrate the LSD accuracy under SA with stolen identities. The detection rate is lower than Figures 10(a) and 10(b), and this reduction occurs due to the detection technique employed by LSD disregard the verification of legitimate identities network. Hence, LSD has compromised its accuracy when the SA employs stolen identities.

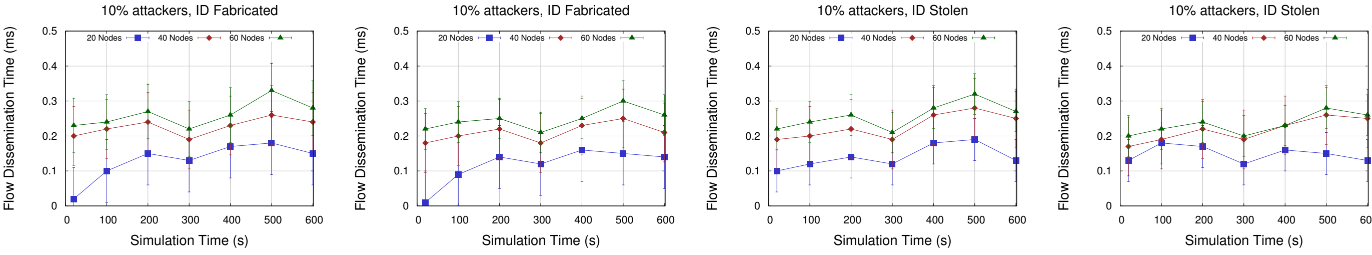
Figures 11(a) and 11(c) show the LSD T_{fp} under the SA with the behavior *Churn*. As the network becomes dense, this rate decreases, showing the inefficiency of LSD in sparse networks. Figures 11(b), and 11(d) show the LSD T_{fp} under the SA with multiple identities, where LSD is more effective, particularly when the attack uses fabricated identities. Hence, the *Churn* behavior in presence of LSD leads to higher losses for the detection of SA on the dissemination of content.

Figures 12(a) and 12(b) show the LSD effectiveness under SA with fabricated identities. The attack is more effective in a sparse network than denser one, 40 and 60 nodes, due to the lower number of nodes that perform the detection. The instant 300 (s) is the peak of attacks, justifying a reduction of the LSD accuracy, see Figure 10(a). The effectiveness of



(a) *Churn* (b) Multiple Identities (c) *Churn* (d) Multiple Identities

Fig. 12: Effectiveness of Sybil attack over dissemination



(a) *Churn* (b) Multiple Identities (c) *Churn* (d) Multiple Identities

Fig. 13: Flow dissemination time

SA with stolen identities, Figures 12(c) 12(d), is greater when compared to Figures 12(a) and 12(b). The SA had greater success in the sparse scenario, as LSD does not check the veracity of an identity and the number of neighbors is less.

Figure 13 shows the impact caused by the SA behavior in the performance of the dissemination of content. The increase of C_{diss} caused by the behavior *churn* is due to the LSD authentication, which requires a timely manner to identify an attacker as it carries out constant association and dissociation on the network. This causes overhead and increases the cost to disseminate a flow. Thus, the SA with the behavior *churn* reduces the efficiency of the dissemination of content causing a reduction in the quality of services.

V. Conclusion

This paper presented a study of effectiveness and efficiency about the techniques of Sybil attack (SA) detection to support the content dissemination of IoT. The detection techniques have been classified on network features, cryptography, and relationship between neighbors, highlighting their strengths and weaknesses. The effectiveness of *Lightweight Sybil Attack Detection Framework* (LSD) was evaluated under SA with fabricate and stolen identities, and behaviors *churn* and multiple identities. LSD showed low efficacy in sparser scenarios, especially when the attacker uses stolen identities. The dissemination of content has been harmed when the attacker has behavior *churn* behavior. Therefore, it is necessary the development of an effective SA detection technique to support the quality of dissemination of content in IoT.

References

[1] S. Li, L. Da Xu, and S. Zhao, "The internet of things: a survey," *Information Systems Frontiers*, pp. 1–17, 2014.

[2] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (iot): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[3] L. Wallgren, S. Raza, and T. Voigt, "Routing attacks and countermeasures in the rpl-based internet of things," *International Journal of Distributed Sensor Networks*, vol. 2013, 2013.

[4] S. Agrawal and D. Vieira, "A survey on internet of things-doi 10.5752/p. 2316-9451.2013 v1n2p78," *Abakós*, vol. 1, no. 2, pp. 78–95, 2013.

[5] S. Park, B. Aslam, D. Turgut, and C. C. Zou, "Defense against sybil attack in the initial deployment stage of vehicular ad hoc network based on roadside unit support," *Security and Communication Networks*, vol. 6, no. 4, pp. 523–538, 2013.

[6] S. Abbas, M. Merabti, D. Llewellyn-Jones, and K. Kifayat, "Lightweight sybil attack detection in manets," *Systems Journal, IEEE*, vol. 7, no. 2, pp. 236–248, 2013.

[7] X. Lin, "Lsr: mitigating zero-day sybil vulnerability in privacy-preserving vehicular peer-to-peer networks," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 9, pp. 237–246, 2013.

[8] D. Quercia and S. Hailes, "Sybil attacks against mobile users: friends and foes to the rescue," in *INFOCOM, 2010 Proceedings IEEE*, pp. 1–5, IEEE, 2010.

[9] P. R. Vamsi and K. Kant, "A lightweight sybil attack detection framework for wireless sensor networks," in *Contemporary Computing (IC3), 2014 Seventh International Conference on*, pp. 387–393, IEEE, 2014.

[10] E. da Silva, A. L. Santos, L. C. P. Albini, and M. N. Lima, "Quantifying misbehavior attacks against the self-organized public key management on MANETs," in *Proceedings of International Conference on Security and Cryptography (SECRYPT '08)*, (Porto, Portugal), pp. 128–135, INSTCC Press, Jul 2008.

[11] A. Blazquez, V. Tsiatsis, and K. Vandikas, "Performance evaluation of openid connect for an iot information marketplace," in *Vehicular Technology Conference (VTC Spring), 2015 IEEE 81st*, pp. 1–6, IEEE, 2015.

[12] X. Wang, J. Zhang, E. M. Schooler, and M. Ion, "Performance evaluation of attribute-based encryption: Toward data privacy in the iot," in *Communications (ICC), 2014 IEEE International Conference on*, pp. 725–730, IEEE, 2014.

[13] S. Park, B. Aslam, D. Turgut, and C. C. Zou, "Defense against sybil attack in vehicular ad hoc network based on roadside unit support," in *Military Communications Conference, 2009. MILCOM 2009. IEEE*, pp. 1–7, IEEE, 2009.