Cluster Comput (2019) 22:S1739–S1749
https://doi.org/10.1007/s10586-017-1105-z

A distributed IDS architecture model for Smart Home systems

Mariusz Gajewski[1] · **Jordi Mongay Batalla**[1,2] · **George Mastorakis**[3] · **Constandinos X. Mavromoustakis**[4]

Received: 3 May 2017 / Revised: 2 August 2017 / Accepted: 8 August 2017 / Published online: 30 August 2017

Abstract The common use of smart devices encourages potential attackers to violate privacy. Sometimes taking control of one device allows the attacker to obtain secret data (such as password for home WiFi network) or tools to carry out DoS attack, and this, despite the limited resources of such devices. One of the solutions for gaining users' confidence is to assign responsibility for detecting attacks to the service provider, particularly Internet Service Provider (ISP). It is possible, since ISP often provides also the Home Gateway (HG)—device that has multiple roles: residential router, entertainment center, and home's "command and control" center which allows to manage the Smart Home entities. The ISP may extend this set of functionalities by implementing an intrusion detection software in HG provisioned to their customers. In this article we propose an Intrusion Detection System (IDS) distributed between devices residing at user's and ISP's premises. The Home Gateway IDS and the ISP's IDS constitute together a distributed structure which allows spreading computations related to attacks against Smart Home ecosystem. On the other hand, it also leverages the operator's knowledge of security incidents across the customer premises. This distributed structure is supported by the ISP's expert system that helps to detect distributed attacks i.e., using botnets.

Keywords Smart Home · Home Gateway · Intrusion detection system · Internet of Things

✉ Jordi Mongay Batalla
jordim@tele.pw.edu.pl

Mariusz Gajewski
m.gajewski@itl.waw.pl

George Mastorakis
gmastorakis@staff.teicrete.gr

Constandinos X. Mavromoustakis
mavromoustakis.c@unic.ac.cy

1 National Institute of Telecommunications, Warsaw, Poland

2 Institute of Telecommunications, Warsaw University of Technology, Warsaw, Poland

3 Department of Informatics Engineering, Technological Educational Institute of Crete, Crete, Heraklion, Greece

4 Department of Computer Science, University of Nicosia, Nicosia, Cyprus

1 Introduction

The Smart Home concept aims at creating a cohesive ecosystem which consists of rising number of smart devices. It makes use of a dozen of technologies and standards which are also not compatible in many cases. The vast majority of the smart devices uses radio communication in unlicensed bands, so that they are easy to deploy and are ready for further expansion. On the other hand, this heterogeneity raises many security problems in the Smart Homes environment. Even if all the manufacturers follow a certain group of standards in the future, people will exploit also existing smart devices which already present a number of issues in regard to security and privacy. Many of these devices are practically not upgradable because of limited resources (from the point of view of processing capacity and storage). Thus, applying more advanced security mechanisms is difficult.

The recent years have also shown that more and more Smart Home applications support the cloud-based management model [1,2]. For that purpose, cloud providers act as collaborators for smart devices suppliers, i.e., many Internet of Things (IoT) systems make use of the cloud for data analysis, storage, and management. In this way, cloud providers are partially responsible for the security of applications and devices.

In this context, the comprehensive approach to Smart Home security requires that also smart devices in their home environment should be professionally secured and that's taking into account their limitations. For this purpose, Home Area Network (HAN) should be equipped with sufficiently powerful devices to perform security functions. In particular, this may include local detection of attacks against HAN connected devices. Control over this process can be carried out by the user by himself or another entity—particularly, the service provider. The latter also has the advantage that it is able to control traffic and data processing conducted by smart devices (to some extent) for attack detection coming from inside the HAN (i.e., by attaching a fake device). It can also take care of security updates, what obviously only applies to supported devices.

This article presents an Intrusion Detection System (IDS) architecture model for Smart Home. The presented architecture assumes that security data processing is distributed between hardware at user premises and the service provider data center. Moreover, this approach allows also service provider to share security inspection processes with professional companies which offer security expertize analytics for detecting extremely advanced threats. The major contributions of this research include proposals for,

- a strategy of distribution security functions related to Smart Home applications efficiently dealing with false positives and negatives;
- distributed anomaly detection scheme for HAN;
- an internal structure and details of distributed IDS architecture for Smart Home.

In this paper we describe the business model for Smart Home which is the base for distributed IDS as shown in Sect. 2. Next, we compare existing IDS solutions, focusing on those which are suitable for solutions used in the Internet of Things and the Smart Home environments. Finally, we present the concept of the distributed IDS for Smart Home system (Sect. 3).

2 Context

Unlike to the incumbent service providers, market entrants usually try to explore new niches offering highly specialized services. In turn, existing providers try to increase their incomes through expanding the range of delivered services [3,4]. In consequence, they often decide to enter new markets such as maintenance for Smart Home systems. For this purpose ISPs build alliances with smart device manufacturers to enrich their offer and open new distribution channel for high-tech products. Figure 1 illustrates the basic business model of the service provider who maintains the platform for Smart Home management.

This model assumes that the service provider cooperates with the smart device manufacturer who is responsible not only for supplying devices but also for commissioning the software both for smart devices and management services offered by the service provider. In turn, the service provider assures hardware distribution among customers (uses its own sales channels, or makes use of a network of distributors). Moreover, the service provider takes over the responsibility for secure service delivery. It includes exposing of secure Application Programming Interfaces (APIs), monitoring of user activity and controlling the packet traffic between the user premise and the provider's servers.

The service provider maintaining the platform for Smart Home management is usually able to detect some anomalies in behavior of these systems through traffic/event analysis. But in some cases the service provider may engage an external entity (company) for deeper security analysis, because security expertise level required in the analytics for detecting advanced threats can be beyond the capabilities of service providers. For such reasons, many companies are turning to

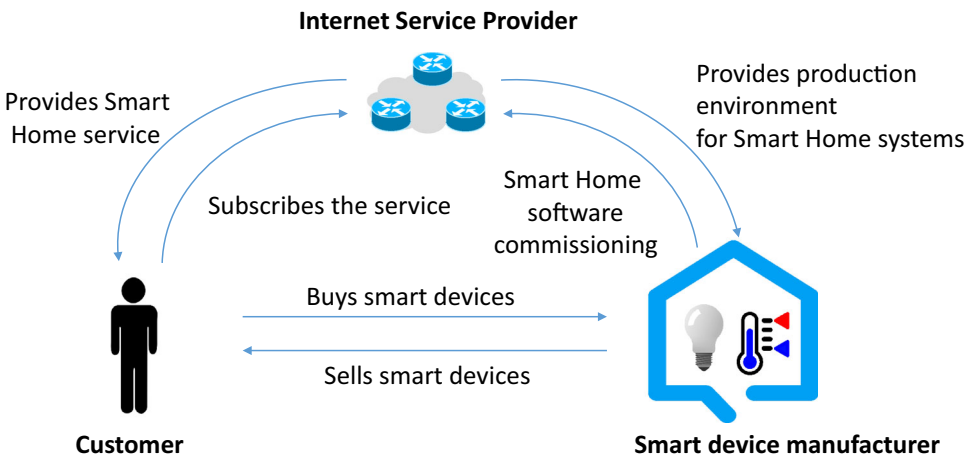


Fig. 1 Basic business model for Smart Home services maintenance by the service provider

external managed security solutions so that they can count on experts doing the monitoring and advanced security analytics. Some companies conducting security analytics are even exposing APIs to enable such sharing for their clients. Obviously, such sharing encompasses only relevant data and is done only with the selected partners, and ensuring that each partner's access is appropriately restricted.[1] This approach is also economically justified, because not every operator has the necessary resources to maintain their own high skilled Computer Emergency Response Team (CERT). The presented model follows an generic communication Back-End Data-Sharing Model presented in RFC 7452 [5]. This generalized model assumes that IoT devices upload data only to a single application service provider [6], however, users often want to be able to analyze data in combination with data from other sources. Hence, the desire for granting access to the uploaded Smart Home data to third parties arises. Such method of combining web data is known as mashup and therefore might be applied to the smart object context. To move the popularity of web tools to IoT ground, typically a RESTful API design together with the proper authentication and authorization technologies are reused.

2.1 General classification of IDS

The conventional security countermeasures like user authentication, data encryption, and security network tools (firewalls, Network Address Translations/NATs) act as first line of defense against external threats. The problem arises when an unauthorized user compromises these countermeasures and is able to use smart devices connected to HAN in unnoticed way. Since many of HAN connected devices use radio transmission, the potential attacker may easily harm or misuse the smart objects and that way influence Smart Home systems (e.g., heating, monitoring, etc.). Therefore, in addition to traditional protection methods, there should be used also security tools which provide protection against both external and internal attacks. One of the potential solution known mainly from enterprise IT systems is an IDS. It aims to detect the intrusions (or anomalies) in real time by protecting nodes from inside and outside threats.

Basically, there are two data sources for IDS—monitored network traffic and data describing events on individual machines connected to the network, including data derived from:

- log files,
- tracing systems (system tools which let trace all system calls made by other processes),
- tools for checking file integrity checksums and registry entries,
- audit system, (system tool which generates log entries to record information about the events that are happening on OS based on pre-configured rules).

Common criteria for classification of IDS solutions include the data collecting mechanisms to enable intrusion analysis. In that case we distinguish between: host-based (HIDS) and network-based (NIDS) intrusion detection systems. In the first case, the raw data collection is based on system/application logs. This category of IDS is also known as Log-based Intrusion Detection Systems (i.e., OSSEC.[2]) The second approach assumes that network traffic is treated as the source of events which may trigger intrusion detection processes (i.e., Snort,[3] Suricata[4]). For this reason, NIDS is designed to monitor all traffic between network entities and capture suspicious events.

The method used to detect attack by IDS system is an another criteria for classification. Essentially, there are three basic methods used for an inspection: (i) based on signature detection which attempts to detect abnormal behavior matching the observed behavior against pre-defined attack patterns, (ii) anomaly based, which identifies malicious activities by analyzing the events (firstly, it defines the normal behavior of the network, then, if any activity differs from normal behavior, it is marked as an intrusion), (iii) hybrid IDS, which is a combination of both anomaly-based and signature-based approaches. Moreover, this classification is extended in [7] to the case of an cross layer IDS (iv), which has the capability to monitor and detect intrusions at multiple OSI layers by analysis of exchanged data across different layers.

Further classification criteria are related to the IDS architecture (e.g., monolithic, hierarchical, distributed, agent-based), the area where these systems are applied (e.g., enterprise networks, Industrial Control System (ICS), wireless networks, etc.) and possible reaction (passive, active).

2.2 IDS for Internet of Things

The broadcast nature of radio communication within the HAN makes is susceptible to various security threats typical for Wireless Sensor Networks (WSNs) [8]. For this reason, further analysis of HAN suitable IDS is focused on solutions adopted for WSNs. In this context, efforts of researchers concentrate mainly on limitations of smart devices which make the implementation of full functionality of IDS difficult. In particular, theoretical work and simulations in this area are carried out to: (i) distribute attack detection tasks between smart devices, (ii) decrease computational complexity of detection algorithms, and (iii) limit the set of

[1] https://www.symantec.com/solutions/internet-of-things.

[2] https://ossec.github.io/.

[3] https://www.snort.org.

[4] https://suricata-ids.org.

attacks which are detectable by the IDSs based on smart devices.

In the first research area we have given particular attention to separate less computational complex tasks performed by constraint nodes (i.e., traffic/events monitoring and reporting) from more complex (i.e., analysis and attack detection) and performed by advanced devices. For this purpose IDS might operate in cooperative cluster mode [9]. It means that every node monitors its neighbors and surrounding nodes activities and operation; in case of any malicious activity detection, the cluster head is informed.

In paper [10] authors have proposed solution based on mobile agents (software modules) which are responsible for anomaly detection in wireless Smart Home sensor networks. This approach assumes that the mobile agent may be installed on each sensor node using a middleware. Middleware used for launching the mobile agent, can also be used for variety of other tasks. The authors mention here mainly the maintenance tasks such as: updating node's firmware, network management, checking, status of the node, etc. According to the authors, anomalies are detected by IDS modules located on so called Cluster Heads (CHs) which play a role of data aggregators and data forwarder to base station in WSN. For that reason, the CHs should provide sufficient processing power for all the assumed tasks. The Cluster Head performs analysis of the data received from sensors/actuators. The anomalous reading triggers anomaly agent in CH which checks if the suspicious node has been compromised. It is accomplished using the mobile agent launched on victim's node.

According to the authors, this approach eliminates the need of installing IDS (anomaly detection) software on each sensor node. This results in moving the responsibility of tracking and alerting onto nodes which have more powerful resources. On the other hand, these nodes might not be able to receive anomalous readings and consequently they will not launch the mobile agent.

A similar IDS approach based on mobile agent concept was proposed in [11]. However, unlike the concept presented in [10], where the mobile agent tasks were focused mainly on data collection and reporting to the WSN cluster heads, this solution provides specific task-oriented mobile agents. Namely, it defines following different agents responsible for performing strictly assigned tasks: Collector Agent, Misuse Detection Agent, Anomaly Detection Agent, and Alert Agent. Two of them play the crucial role as IDS components—the Misuse Detection Agent, which detects known attacks in network on the traffic data received from the Collector Agent, and the Anomaly Detection Agent which is used to detect the attacks on basis of anomaly detection algorithm. In case of an attack detection, the both detection agents trigger the Alert Agent which propagates this information.

The second research area concentrates mainly on adjusting the IDS algorithms to constrained nodes properties. These works were published in several papers. The exemplary research results were presented in [12], where authors described modifications of the anomaly based IDS exploiting genetic k-means algorithm. Authors have improved algorithm efficiency and increased attack detection rate compared to basic algorithm. Also results described in [13] present algorithms, optimized for cluster based WSNs. In that case, authors adopted machine learning approach based on selected supervised learning model—support vector machine (SVM). It was exploited for data analysis and misuse detection in a distributed environment. The learning algorithm is used to drive SVM to distinguish between normal and malicious patterns. It is designed to operate in cluster based WSNs, where all nodes monitor their neighbors.

The third research area shows that the majority of the existing intrusion detection solutions are capable of handling only a few security attacks. Particularly, the signature-based IDS solutions make use of this assumption, since they consume more resources for computations as compared to anomaly-based IDS. In this context, authors of [14] enumerate list of security threats typical for IoT and specifically WSNs, and among the following: Sinkhole Attack, Wormhole Attack: Selective Forwarding Attack, Sybil Attack, Hello Flood Attack, and the Denial of Service (DOS) Attack. Following this list, several IDS examples are described in the literature. Specifically, authors of [15] proposed an IDS detecting black hole attacks in WSNs. In this proposal, sensor node and base station are exchanging control packets containing the node id and number of packets sent to the cluster head. This information is propagated to the base station which additionally monitors all passing traffic. According to the authors, this approach decreases energy consumption of nodes. Another proposed IDS concept, described in [16], aims at detecting Sybil node attack in WSN. In their work, authors proposed two stage method for solving this problem. Namely, the first stage is that cluster head polls slave nodes for their identities and position data. Received data are stored in the table maintained by the cluster head. In the second stage, all authorized nodes reply to the cluster head with their identities and current position data including the Sybil node. Finally, the cluster head matches received and stored data to discover the Sybil node. The authors claimed that proposed system improves the energy efficiency and it detects the Sybil node with reasonable accuracy. In this context, it is worth to note that energy efficiency and securing data transmission is a new challenging area in IoT applications. Several research efforts have been published in [17,18]. In a hybrid approach joining signature- and anomaly-based methods, a good example is a conceptual IDS called SVELTE described in details in [19]. Authors decided that the monitoring part (which is computationally lightweight) is to be implemented into

resource-constrained nodes. The resource demanding functionality is placed onto the Border Router (BR) which is an edge node connecting 6LoWPAN network with the Internet (acts as a technology gateway). The basic IDS functionality of SVELTE aims to detect attacks targeted to the routing mechanisms, in particular spoofed or altered information, sinkhole, and selective forwarding attacks

As devices in IoT are resource-constrained and anomaly-based IDS requires computationally intensive operations, placement of IDS modules in a IoT network becomes a critical issue. In this context, SVELTE described in [19] follows this approach and proposes the lightweight monitoring functionality to be implemented into constrained nodes. More resource demanding IDS processes are performed by the Border Router (BR) of the 6LoWPAN network.

A generalized approach for separating the network monitoring part and the detection part is known as Cooperative Autonomous Attack Detection and was described in [20]. The proposed idea introduces a multi-hierarchy monitoring environment for capturing packets and performing flow statistics. It assumes that this functionality is spread through the network but it builds up one detection system that analyzes data monitored at different points of the network. Furthermore, an output of the detection system can become an input of other detection system by exporting aggregated monitoring data.

An similar approach was proposed in [21] however authors have emphasized that this solution has been adjusted to grid networks specificity. Smart Grid Distributed Intrusion Detection System (SGDIDS) is a hierarchical and distributed IDS dedicated for smart grids. It divides monitored network into three layers: HAN, Neighborhood Area Network (NAN), and Wide Area Network (WAN). Each of them includes dedicated nodes with Analysis modules (AM) responsible for packet flow analysis. These modules use classification techniques such as Support Vector Machines (SVM) and Artificial Immune System (AIS) to inspect network traffic to efficiently classify malicious events. According to the authors, achieved results suggest that the proposed approach employing both techniques can considerably improve detection effectiveness.

2.3 Intrusion detection in cloud systems

A cloud infrastructure operated by the service provider extensively uses virtualization techniques which enables much more flexible resource utilization and is able to serve much more users at that same time. Moreover, all components of the cloud infrastructure run through standard Internet protocols. These may encourage potential attackers to violate security of provided services. That is the reason, why extremely different challenges are faced by the service provider that secures its infrastructure. First, it has to take into account more network-oriented groups of threats, which are aimed at disrupting network operations. According to [22], cloud computing platforms might suffer from attacks such as IP spoofing, Address Resolution Protocol spoofing, Routing Information Protocol attack, DNS poisoning, Flooding, Denial of Service (DoS), Distributed Denial of Service (DDoS), etc. Service providers deploy different security solutions across their networks but in case of datacenters the general approach is similar to enterprise networks. It assumes that the first line of defense is built up from firewalls which prevent outside attacks. Providing the cloud based services requires also that service provider ensures the proper set of security countermeasures against insider attacks. For this purpose it deploys highly efficient IDSs and intrusion prevention systems (IPSs) which, in turn, are used to mitigate these attacks. Also the integration of IoT and cloud computing technologies raises new challenges for securing virtual assets and data coming from smart devices. Current trends in this area have been described in [23].

Generally, there are similar classification criteria to those used for HAN, but the scale of solutions must be proportional to the scale of data being processed and the traffic being handled. For this reason, research on cloud-oriented IDS solutions is focused mainly on efficiency and accuracy. For this purpose new methods are being developed to detect attacks, when analysis is based on huge amounts of data. In this context, authors of [24] extend the basic IDS classification by adding the: Artificial neural network based IDS (ANN), Fuzzy logic based IDS, Association rule based IDS, Support Vector Machine (SVM) based IDS, and Genetic algorithm (GA) based IDS. The above mentioned IDS types are strictly related to techniques used for high volume data analysis for attack detection purposes.

The high volume data analysis and packet traffic is a primary driver for distribution of data processing. It aims to accelerate computations on live traffic data. For that purpose, an distributed IDS model plays an important role. This concept assumes that Distributed IDS (DIDS) consists of several IDS instances and often of both types host- and network based IDSs distributed in the operator's network. All of them are able to communicate with each other and with a dedicated server responsible for aggregated data analysis and decision making. Each IDS instance collects data (and performs initial analysis or aggregation, depending on the concept), and then sends it to the central server. The IDS instance that collects data is often known as a probe or sensor. An exemplary solution following this approach is described in [25], where authors proposed that IDS instances located in different regions of the provider's network are able independently to detect attacks and "warn" proper network devices operating in other regions.

Event Monitoring Enabling Responses to Anomalous Live Disturbances (EMERALD) [26] is an example of an another

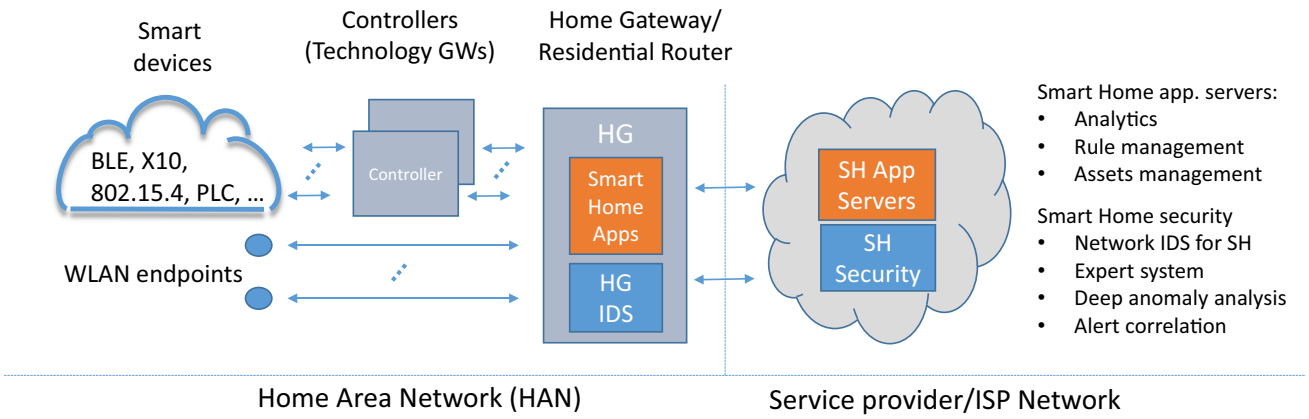


Fig. 2 Architectural framework for distributed IDS for Smart Home services

distributed network IDS that encompasses both monitoring and analysis components. This approach assumes that computation related to traffic analysis is spread throughout monitoring network nodes. Each node encompasses analysis component which uses both signature-based and Bayesian methods to detect intrusions. Completely another approach was adopted by originators of Dshield service.[5] They proposed centralized community-based firewall log correlation system which accepts logs from NIDS and firewalls around the Internet, aggregates them and then reports summaries on detected intrusions and possible attack activity. Information about detected attacks is available to network administrators in order to give them the ability to reconfigure and tune their security infrastructure.

An another approach assumes that IDS instances run not only on dedicated machines but also use hypervisor layer (a hypervisor is a platform to run VMs) for monitoring and analyzing communications between VMs, between hypervisor and VM and within the hypervisor based virtual networks. The so-called VM introspection based IDS (VMI-IDS) architecture was described for the first time in [27]. It should be noted that attack detection might be performed by applications residing on VM or at the host machine layer—as a hypervisor-based IDS.

3 Co-responsible distributed IDS for Smart Home

To meet users' security expectations and minimize the impact on the HG performance, we proposed two-layer network architecture for IDS in the Smart Home environment. This approach assumes that the preliminary attack detection is performed at user's premises however the deeper and broader analysis (including packet traffic inspection) are performed in the ISP's infrastructure.

[5] https://www.dshield.org/.

Generally, the HAN connects three types of devices (see Fig. 2): smart devices, controllers and network devices responsible for connecting the Smart Home to the Internet. Smart devices, which act as active or passive elements, directly interact with the environment. Controllers are technology dependent gateways ensuring communication between smart devices and IP based HAN. Home Gateways known also as Residential Routers, are devices mediating communication between the operator's IP network and the HAN. The HG transfers data from smart devices to Smart Home application servers (mostly maintained by the device manufacturer)—components of Smart Home applications are shown in orange in the picture. It enables the smart services manipulation from Internet connected devices. An alternative scenario assumes that all the connected smart devices are controlled via the locally hosted management server. In that scenario, possible use cases may include locating management functions in dedicated controller or Home Gateway. Also the security functions (for a Smart Home as a whole) are performed by the Home Gateway and/or Residential Router—they are shown in blue in the picture. As stated in [5], the most extensive task for the provision of Smart Home Security are carried out by the Home Gateway (or the Residential Router). This set includes the basic security functions to be preferentially applied in the Smart Home system. We do not consider network security functions and the security features related to a specific service, and also a remote service environment in this paper.

3.1 Overall architectural framework

One of the main problems with implementation of IDS on hardware platforms with limited resources is that they are able only to detect known malware using signatures. To generate these signatures, the security experts have to identify attacks, analyze them, and then describe relevant behavior. The description has to be conscious and prepared using stan-

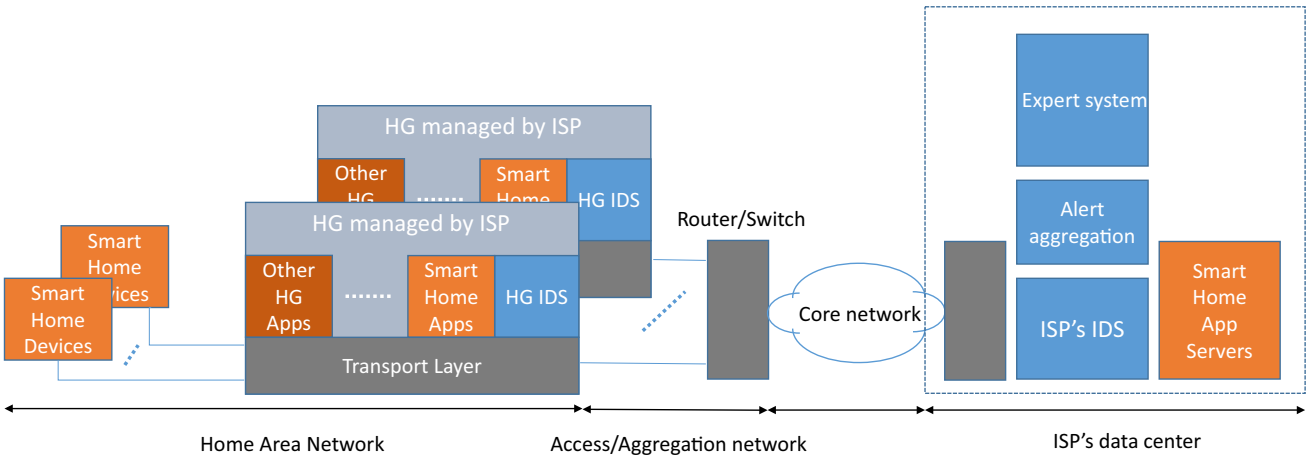


Fig. 3 The two-layer network architecture for IDS in Smart Home environment

dardized methods so that it can be the basis for detection of the same attack next time. This reactive approach has limitations, since it does not work to detect new threats and new attacks that do not have assigned signatures yet. To ensure the access to the all currently known attack signatures it is convenient to run this process in the service provider-managed environment. The most suitable solution for that seems to be the cloud platform.

A characteristic feature of all constrained devices is a limited range of potential actions and usually continuous working. For that reason, machine learning methods seem to be the best way to use data related to the operation of these devices. This form of artificial intelligence is much more reliable than a human and also is much better for scaling. In consequence, behavior monitoring of constrained devices may be the basis for anomaly-based IDS and signature-based IDS but limited to selected attacks.

The proposed solution meets expectations in following key areas of secure Smart Home: (i) fast preliminary analysis and processing of security related data obtained locally from the smart devices within the HAN, (ii) advanced data processing carried out by the service provider, and (iii) making use of the service provider's knowledge on the security incidents across maintained network. This approach does not exclude deep offline analysis which may be performed by the team of security professionals.

The basic idea of the DIDS for Smart Home system assumes at least that the service provider has to have the ability to manage the HG for the monitoring and control of smart devices within the Smart Home ecosystem. This is possible if the service provider acts e.g., as an ISP which provides Internet access together with smart services to be installed at home. An alternative solutions assumes that HG administrator allows the user to install applications which can be accessed by the service provider and used for smart devices management. The DIDS for a Smart Home combines distributed monitoring and data aggregation with data analysis carried out in ISP's cloud environment.

The presented approach assumes also that the service provider supports the Smart Home system and maintains the Smart Home applications servers which allow the user to access the Smart Home ecosystem functions over the Internet. Providing access to management functions over the Internet may introduce some new threats to the security of the Smart Home. For this reason, both the activities taken within the HAN and actions taken through the Smart Home platform should be monitored against suspicious activity.

Moreover, assuming that ISP's monitoring of the suspicious activities is limited to the supported SH applications and should minimize the impact on the HG performance, the IDS functionality should be shared between HG equipment and ISP's infrastructure. Following this conclusion, the minimal set of intrusion detection functions maintained by the provider encompasses three components: a Home Gateway IDS (HG IDS) residing on a Home Gateway (or technology gateway) at user premises, an ISP's IDS residing in the ISP's infrastructure, and the expert system located in the ISP's data center. All above mentioned functional entities constitute two-layer network architecture for intrusion detection in the Smart Home environment as depicted in Fig. 3.

In this approach, the HG IDS is responsible for local resource monitoring (i.e., CPU, memory and network bandwidth utilization, outages in communication with smart devices, etc.) and performs preliminary log analysis. It acts as a Host-based IDS (HIDS) and performs rule-based detection of Smart Home devices misuse and policy violations. The above limited set of tasks is performed by the HG IDS because they require relatively small computational effort and may be also implemented in relatively less powerful nodes (Home Gateways or technology gateways responsible for communication between constraint devices and the Internet).

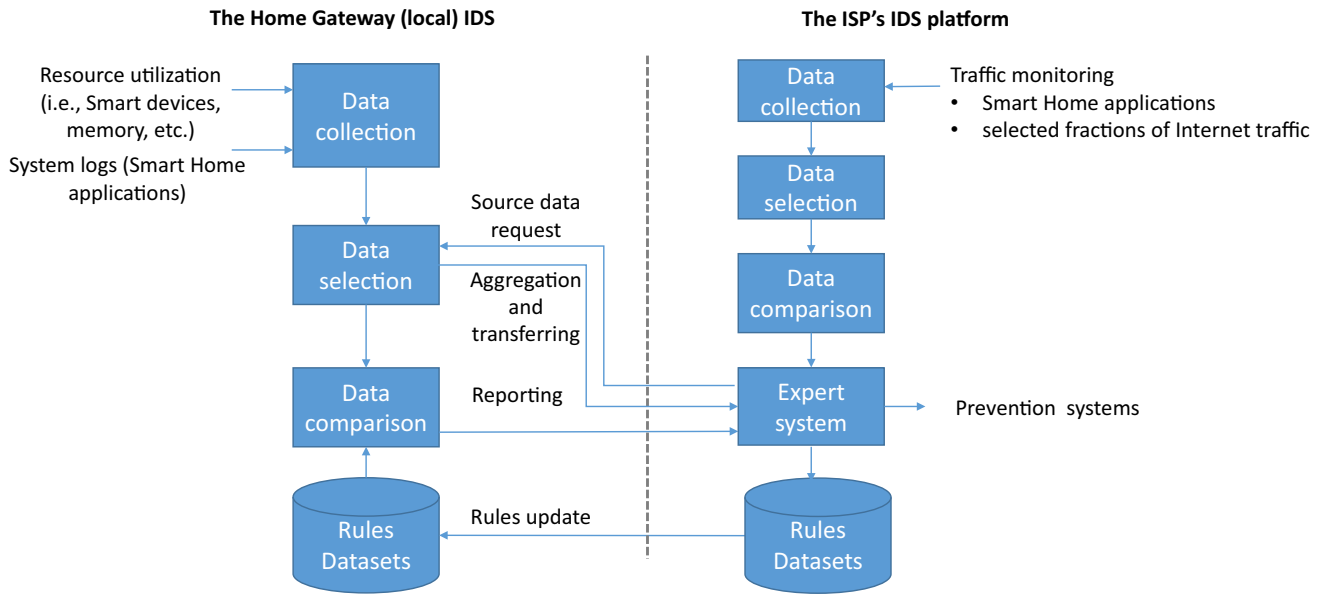


Fig. 4 The IDS communication model for Smart Home

In turn, the ISP's infrastructure is usually ready for processing the long term anomaly analysis of the user's behavior. For this reason, results of the preliminary analysis at HG IDS are exported to the expert system located at ISP's premises for further analysis. Moreover, the ISP is also responsible for inspection of the network traffic related to Smart Home applications. This means that the IDS at the ISP's data center acts as network-based system (NIDS). This two-step IDS structure requires mutual communication between local (HG IDS) and central (IDS expert system) functional entities and is described in details in next paragraph.

3.2 Functional architecture

The IDS communication model for Smart Home assumes that data for analysis is collected on the basis of the systems logs related to Smart Home applications and resource utilization related to smart devices which use HG/GW for event logging (see Fig. 4). It performs preliminary analysis of available data based on locally stored rules. This analysis exploits: log files and audit records produced by the operating system for every action taken on the HG/GW system related to Smart Home devices and applications. Examples of the audit record include file accesses, system calls, process executions, logins, etc. Moreover, preliminary analysis may encompass also the results of inspections performed by other processes, i.e., integrity checksums, registry entries, monitoring of smart devices, etc. First, the local IDS has to filter desired log files, system call traces (i.e., related to resource utilization) or output data from auxiliary processes (i.e., checking file checksums) and then these data are verified against rules provided by the rules dataset.

The applied rules dataset should fulfill requirements which face fast and effectively (by minimizing the number of false positive and false negatives) intrusion detection process. This scenario-based approach should encompass examples such as: number of logins during an short period of time, system/network reconfigurations, etc. Results of the events auditing are reported to the ISP's platform for further analysis. This entity may require more information desired for deeper analysis and for this purpose a source data might be retrieved from the HG. The source data are transferred upon the request from the ISP analytical platform.

The reported behavior is analyzed to detect anomalies in behavior by the expert system located at the ISP's premises and supported by the Smart Home solution manufacturer. This approach assumes that reported security vulnerabilities as well as standard events are observed during the training phase. These data parameterize a model of normal behavior which is additionally supplemented by traffic analysis but in this case traffic analysis is limited to the Smart Home applications and selected fractions of the Internet traffic related to these applications. In this context, access from Internet connected mobile devices seems to be the majority of use cases.

The applied rules should take into account the product life cycle and specifically bug fixes reported by the users and updates applied by the manufacturer. It should also take into account the specificity of Smart Home applications and packet traffic generated by them. These rules are introduced by SH devices manufacturer and should complement the knowledge base used in anomaly detection.

The HG IDS decodes system logs and audit records, and performs analysis against locally stored rules. The results

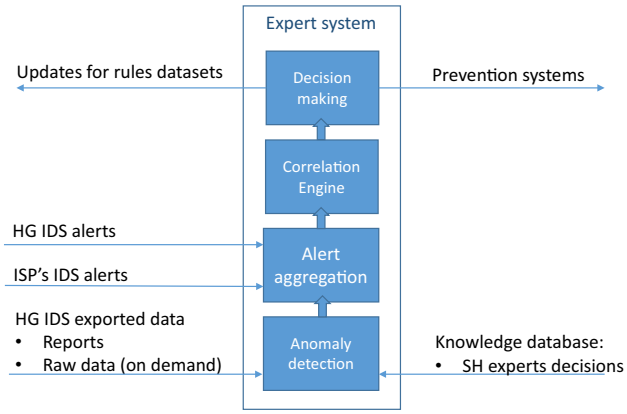


Fig. 5 The expert system concept for Smart Home IDS

of analysis are periodically transferred to the ISP's expert system. Particularly, the HG IDS sends alerts triggered on the basis of actually available rules. Depending on the needs, the aggregated logs may be also transferred to expert system for further analysis. However, this action is initiated on demand, as a result of the decision made by the expert system which may desire deeper analysis of raw data.

The expert system provides correlation engine that takes as input aggregated results of IDS instances. Its main task is to correlate detected security flaws for individual Smart Home deployments as well as Smart Home service platform. Moreover, it raises alerts for prevention purposes as well as triggers updates of rules datasets in the HGs. The overall view of the main components of the Expert system is presented in Fig. 5.

Fundamentally, the expert system is responsible for decision making aiming at increasing of intrusion detection efficiency. It deals with correlated events (mainly alerts) and its role is twofold: (i) detection of massive security breach of Smart Home area across the ISP network, (ii) improving efficiency of intrusion detection for SH IDS. The first role is fulfilled by the Correlation Engine that deals with aggregated alerts coming from different sources. Basically, the Correlation Engine performs correlation across two groups of datasets:

- alerts from heterogeneous sources are correlated, i.e., from HG IDS and ISP's IDS;
- alerts from homogenous sources are correlated i.e., alerts coming from various HG IDS instances (from different user premises).

As an output, the ISP should be able to obtain security breach alerts related to individual Smart Home ecosystem and secondly alerts related to massive suspicious activities in ISP's network coming from SH platform.

The second role is fulfilled by extending the rules datasets with new rules defined by security experts. These new security rules are result of analysis and experience of jointly security experts supported by ISP's CERT teams and SH devices manufacturer. Particularly, manufacturers role is important because of the obligation to support the product (i.e., taking into account reported bugs).

The second data source for the expert system is ISP's IDS which performs network traffic inspection at the ISP's premises. Acting as a NIDS, the ISP's IDS performs inspection of incoming and outgoing traffic to/from Smart Home

Table 1 Task description of the HG IDS and the ISP's IDS platform

Functionality	The HG IDS tasks	The ISP's IDS platform tasks
Data collection	Collection of data coming from monitoring of system resources, operating system audit trails and application logs. This task is independent of the network speed since it is based on monitoring only a single HG	Collection of data coming from network traffic monitoring at the ISP's data center. This task is accomplished with two assumptions: (i) incoming and outgoing traffic to/from SH application servers, (ii) packet streams are mirrored to decrease the impact on performance of SH network infrastructure
Data selection	Selection of collected data required for security analysis. Selected data should be related to events that are important from the security point of view	Selection of traffic related to Smart Home applications
Data comparison	Matching selected events with locally stored rules for detection of security breaches	Matching selected traffic against rules
Rules dataset	Provision of the database of events treated as security breach	Provision of the database of network attack signatures
Expert system		1. Analysis of reports coming from users' HGs for anomaly detection. Possibly a deeper analysis is made on the basis of source data (imported from HGs) 2. Correlation of alarms coming from HGs (directly or anomalies detected by the expert system) and alarms coming from traffic analysis

Application servers. The choice of its location is justified considering that SSL or TLS inspection of encrypted connection is difficult and an IDS can only perform limited inspection based on packet headers. Since most of Smart Home applications services maintained in the cloud offer an APIs access over HTTPS, the only place for decryption of traffic to and from servers hosted by ISP is the ISP's data center. Moreover, the full inspection requires access to the private key used by the SH applications services as well as computational cost associated with decryption. Therefore, locating this node at the ISP's data center is an optimum solution in terms of performance and reliability.

A comparative description of Smart Home IDS functional components is provided in Table 1. This table lists tasks performed by each component.

Distributed intrusion detection based both on IDS residing in Home Gateways and at ISP's premises is an optimal approach to the detection of many attacks. Particularly, quickly spreading attacks can be detected and will generate an increasing number of alerts. However, without the centralized infrastructure, it is not possible to have a look at the larger picture of a spreading attacks .

4 Summary and conclusions

When considering the large communication network that exists in the case of the Smart Home deployments hosted by the ISP, we propose a two-layer network composed of Home Gateways IDS entities located at users premises, and the IDS entities located within the ISP infrastructure. Moreover, the communication is assured by the use of WAN delivered by the ISP.

In this concept, the HGs collect data, carry out the preliminary analysis which is limited in scope and are based on locally stored rules. Alerts as well as cyclic reports on Smart Home users activity are exported to the expert system for further analysis. The data required for deeper analysis should be aggregated by the HG, converted into unified format and exported to the expert system. The Expert system performs analysis and correlates results coming from different HGs as well as from NIDS located at the ISP's premises.

The distributed IDS combines host based (remote) monitoring and data reduction with the system maintained in the ISP's cloud. In other words, multiple host based IDS are interworking with the central node—expert system that performs anomaly-based data analysis. This approach is adjusted to Smart Home deployment requirements that are run on ISP datacenter infrastructure.

Acknowledgements This work was undertaken under the PolLux IDSECOM project and the PolTur FUSE project supported by the National Centre for Research and Development (NCBiR) in Poland.

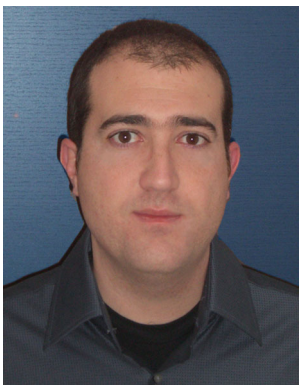
References

1. Mongay Batalla, J., Mastorakis, G., Mavromoustakis, C.X., Zurek, J.: On cohabitating networking technologies with common wireless access for home automation systems purposes. IEEE Wirel. Commun. **23**, 76–83 (2016)
2. Mongay Batalla, J., Krawiec, P.: Conception of ID layer performance at the network level for Internet of Things. Pers. Ubiquitous Comput. **18**(2), 465–480 (2015)
3. Mongay Batalla, J., Krawiec, P., Mavromoustakis, C.X., Mastorakis, G., Chilamkurti, N., Négru, D., Bruneau-Queyreix, J., Borcoci, E.: Efficient media streaming with collaborative terminals for smart city environment. IEEE Commun. Mag. **55**, 98–104 (2017)
4. Krawiec, P., Sosnowski, M., Mongay Batalla, J., Mavromoustakis, C.X., Mastorakis, G.: Dynamic adaptive streaming over CoAP. Multimed. Tools Appl. (2017)
5. Tschofenig, H., Arkko, J., Thaler, D., McPherson, D.: Architectural Considerations in Smart Object Networking. RFC 7452. https://tools.ietf.org/html/rfc7452 (2015)
6. Mongay Batalla, J., Gajewski, M., Latoszek, W., Krawiec, P., Mavromoustakis, C.X., Mastorakis, G.: ID-based service-oriented communications for unified access in IoT. Comput. Electric. Eng. **52**, 98–113 (2016)
7. Alrajeh, N.A., Khan, S., Shams, B.: Intrusion detection systems in wireless sensor networks: a review. Int. J. Distrib. Sens. Netw. **9**, 167575 (2013)
8. Khanafer, M.: Intrusion detection system for WSN-based intelligent transportation systems. In: Global Telecommunication Conference (GLOBECOM2010), IEEE, pp. 1–6 (2010)
9. Siddiqui, M.S., Choong, S.H.: Security issues in wireless mesh networks. In: Proceedings of the International Conference on Multimedia and Ubiquitous Engineering (2007)
10. Usman, M., Muthukkumarasamy, V., Wu, X.-W., Khanum, S.: Wireless smart home sensor networks: mobile agent based anomaly detection. In: Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC) (2012)
11. Mourabit, Y.E.L., Toumanari, A., Bouirden, A., Zougagh, H., Latif, R.: Intrusion detection system in wireless sensor network based on mobile agent. In: Second World Conference on Complex Systems (2014)
12. Sandhya, G., Julian, A.: Intrusion detection in wireless sensor network using genetic k-means algorithm. In: IEEE International Conference on Advanced Communication Control and Computing Technologies (2014)
13. Alsadhan, A., Khan, N.: A proposed optimized and efficient intrusion detection system for wireless sensor network. Int. J. Electr. Comput. Energ. Electron. Commun. Eng. **7**, 1621–1624 (2013)
14. Sherasiya, T., Upadhyay, H., Patel, H.B.: A survey: intrusion detection system for Internet Of Things. Int. J. Comput. Sci. Eng. (IJCSE). **5** (2016)
15. Athmani, S., Eddine Boubiche, D., Bilami, A.: Hierarchical energy efficient intrusion detection system for black hole attacks in WSNs. In: 2013 World Congress on Computer and Information Technology (WCCIT) (2013)

16. Babu Karuppiah, A., Dalfiah, J., Yuvashri, K., Rajaram, S., Khan Pathan, A.: A novel energy-efficient sybil node detection algorithm for intrusion detection system in wireless sensor networks. In: 3rd International Conference on Eco-friendly Computing and Communication Systems (2014)
17. Memos, V.A., Psannis, K.E., Ishibashi, Y., Gupta, B.B.: An efficient algorithm for media-based surveillance system (EAMSuS) in IoT smart city framework. Future Gener. Comput. Syst. doi:10.1016/j.future.2017.04.039
18. Kokkonis, G., Psannis, K.E., Roumeliotis, M., Schonfeld, D.: Real-time wireless multisensory smart surveillance with 3D-HEVC streams for internet-of-things (IoT). J. Supercomput. **73**, 1044–1062 (2016)
19. Raza, S., Wallgren, L., Voigt, T.: SVELTE: Real-time intrusion detection in the Internet of Things. Ad Hoc Netw. **11**, 2661–2674 (2013)
20. Dressler, F., Münz, G., Carle, G.: Attack detection using cooperating autonomous detection systems (CATS). In: 1st IFIP International Workshop on Autonomic Communication (WAC 2004), Poster Session (2004)
21. Zhang, Y., Wang, L., Sun, W., Green, R.C., Alam, M.: Distributed intrusion detection system in a multi-layer network architecture of smart grids. IEEE Trans. Smart Grids **2**(4), 796–808 (2011)
22. Modi, C., Patel, D., Borisaniya, B., Patel, H., Patel, A., Rajarajan, M.: A survey of intrusion detection techniques in Cloud. J. Netw. Comput. Appl. **36**, 42–57 (2013)
23. Stergiou, C., Psannis, K.E., Kim, B.-G., Gupta, B.: Secure integration of IoT and Cloud computing. Future Gener. Comput. Syst. doi:10.1016/j.future.2016.11.031 (2016)
24. Lo, C.C., Huang, C.C., Ku, J.: Cooperative intrusion detection system framework for cloud computing networks. In: First IEEE International Conference on Ubi- Media Computing (2008)
25. Dastjerdi, A.V., Bakar, K.A.: Distributed intrusion detection in clouds using mobile agents. In: III International Conference on Advanced Engineering Computing and Applications in Sciences (2009)
26. Porras, P.A., Neumann, P.G.: Emerald: event monitoring enabling response to anomalous live disturbances. In: Proceedings of the 20th National Information Systems Security Conference, pp. 353–365 (1997)
27. Garfinkel, T., Rosenblum, M.: A virtual machine introspection based architecture for intrusion detection. In: Procceeding of Network and Distributed Systems Security Symposium (2003)

Mariusz Gajewski received his M.Sc. degree in Telecommunications from the Warsaw University of Technology. He has been employed at the National Institute of Telecommunications (NIT) since 1998. In 2010, he joined the *Internet Architectures and Applications* Department in NIT. He specializes in technical aspects of network architecture, IPv6 protocol testing, Future Internet architectures as well as Internet of Things.

Jordi Mongay Batalla received his Ph.D. degree from Warsaw University of Technology, where he still works as an assistant professor. He is head of department in the National Institute of Telecommunications. His research interest focuses on QoS in IPv4/v6 and SDN networks, future architectures (information-centric networks) and applications (IoT, smart cities, IPTV).



George Mastorakis received his B. Eng. degree from the University of Manchester, his M.Sc. from University College London, and his Ph.D. from the University of the Aegean. He is Associate Professor at the Technological Educational Institute of Crete and Research Associate at the Centre for Technological Research of Crete. His research interests include cognitive radio networks, network traffic analysis, and radio resource management.

Constandinos X. Mavromoustakis received his dipl. Eng. degree from the Technical University of Crete, his M.Sc. from University College London, and his PhD from the Aristotle University of Thessaloniki. He is leading the Mobile Systems Lab of the University of Nicosia. He is vice-chair of IEEE/regional Cyprus section, and serves as the Chair of Computer Society Chapter of the Cyprus IEEE section.