IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

SPECIAL SECTION ON SECURITY AND TRUSTED COMPUTING FOR INDUSTRIAL INTERNET OF THINGS

Received May 30, 2018, accepted July 9, 2018, date of publication July 31, 2018, date of current version August 28, 2018.

*Digital Object Identifier 10.1109/ACCESS.2018.2861421*

# A Secured Data Management Scheme for Smart Societies in Industrial Internet of Things Environment

**MUHAMMAD BABAR[1,2], FAZLULLAH KHAN[3], (Member, IEEE), WASEEM IQBAL[2], ABID YAHYA[4], (Member, IEEE), FAHIM ARIF[2], (Member, IEEE), ZHIYUAN TAN[5], (Member, IEEE), AND JOSEPH M. CHUMA[4], (Member, IEEE)**

[1]Department of Computing and Technology, Iqra University, Islamabad 44000, Pakistan
[2]Computer Software Engineering Department, National University of Sciences and Technology, Islamabad 44000, Pakistan
[3]Computer Science Department, Abdul Wali Khan University, Mardan 23200, Pakistan
[4]Department of Electrical, Computer and Telecommunication, Faculty of Engineering and Technology, Botswana International University of Science and Technology, Palapye 016, Botswana
[5]School of Computing, Edinburgh Napier University, Edinburgh EH10 5DT, U.K.

Corresponding author: Muhammad Babar (babar.phd@students.mcs.edu.pk)

**ABSTRACT** Smart societies have an increasing demand for quality-oriented services and infrastructure in an industrial Internet of Things (IIoT) paradigm. Smart urbanization faces numerous challenges. Among them, secured energy demand-side management (DSM) is of particular concern. The IIoT renders the industrial systems to malware, cyberattacks, and other security risks. The IIoT with the amalgamation of big data analytics can provide efficient solutions to such challenges. This paper proposes a secured and trusted multi-layered DSM engine for a smart social society using IIoT-based big data analytics. The major objective is to provide a generic secured solution for smart societies in IIoT environment. The proposed engine uses a centralized approach to achieve optimum DSM over a home area network. To enhance the security of this engine, a payload-based authentication scheme is utilized that relies on a lightweight handshake mechanism. Our proposed method utilizes the lightweight features of the constrained application protocol to facilitate the clients in monitoring various resources residing over the server in an energy-efficient manner. In addition, data streams are processed using big data analytics with MapReduce parallel processing. The proposed authentication approach is evaluated using NetDuino Plus 2 boards that yield a lower connection overhead, memory consumption, response time, and a robust defense against various malicious attacks. On the other hand, our data processing approach is tested on reliable datasets using Apache Hadoop with Apache Spark to verify the proposed DMS engine. The test results reveal that the proposed architecture offers valuable insights into the smart social societies in the context of IIoT.

**INDEX TERMS** Demand side management, home area network, industrial Internet of Things, security, smart societies, trust.

## I. INTRODUCTION

We are living in a time on this planet when 54% population is living in urban areas; with a prediction of an increase up to 66% by 2050 [1], [2]. This urbanization will bring numerous challenges for the decision makers in providing various services and fulfillment of infrastructure needs. The Internet of Things (IoT) covers the everyday devices of physical world by enabling them to network with other devices [3]. It was predicted that the connected things will increase 31% in the year 2017, in comparison to 2016, by reaching 8.4 billion interconnected devices and will cross 20 billion by year 2020 [4]. With the advent of IoT, all the dumb devices at homes will be able to hear, listen and communicate with each other by forming a Home Area Network (HAN). Moreover, this information will be shared over the Internet to make it ubiquitous [5]. With this information sharing, the dumb devices will transform into smart devices to reform the human life style. IoT is a mesh network of home/everyday devices, and these devices may be equipped with pervasive intelligence [6].

The critical sensors and devices in essential industry infrastructure with existing IoT are interconnected to form the


2169-3536 © 2018 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Industrial IoT (IIoT). Usually, IIoT exploitation permits businesses and clients into industrial practices and accomplish sky-scraping production by reducing cost. Meanwhile, IIoT renders the industrial systems to malware, cyber attacks, and a number of vulnerabilities. The incorporation of devices within the Internet needs a variety of communication models. This necessity will likely add a number of novel and ingenious adversarial models to the IIoT [7]. In IIoT, security provisioning is a cumbersome task because each device has its own unique features. The distinctiveness of each device, person, and system linked to the Internet needs to be verified. The products and solutions available in the market are lacking secured features and they are vulnerable to a wide range of security breaches.

The huge number of smart devices in IIoT environment will produce data of high volume, high velocity, and/or high variety information (3V) assets, that require new forms of processing to enable enhanced decision making, insight discovery and process optimization called Big Data. Sharda *et al.* [8], [9] added veracity, variability, and value to these three Vs. The application of advanced analytic techniques on large data sets is known as Big Bata analytics that consists of two parts, i.e., the Big Data set and analytics. Big Data analytics is a specific approach of analytics that is used to extract the hidden features and useful insights from huge and gigantic data. The applications of Big Data analytics include but not limited to banking and securities, communications, media, healthcare, education, insurance, retail, importation etc. Advancement in IIoT is resulting in a large amount of valueable data. With the help of Big Data technologies and efficient machine learning algorithms, there is a great prospective of analytical services to the urban citizens and decision makers [10], [11]. Some of the Big Data technologies include Hadoop designed for distributed processing of large datasets, Spark used for real-time stream processing, R is designed for working with statistics, and NoSQL is an unstructured database. A smart society aims to raise the quality of service (QoS) and optimize these services for its residents [12].

With the advancement of technology and increased urbanization, power management has a vital role in the development of the nations. The increased energy consumption with inability to meet the rising demand, efficient energy utilization and load management is the main point of focus in many countries around the globe. Energy management aims to optimize the generation and distribution of energy [13]. The traditional perception about electric energy system is one way and is top-down oriented, i.e., the power is generated in power plants and transmitted to the industries and homes for usage through grids. Power plants generate electric energy and feeds it to the grid that try to equalize the demand and supply balance at all time. Balancing the demand and supply of energy is a critical aspect in operating an electric energy system. Demand side management (DSM) is used to change the consumer behavior towards the energy usage. Palensky and Dietrich [13] divided the DSM depending upon the time and impact, into following categories: 1) efficient usage of energy, 2) time of energy usage, and 3) response to the demand.

In view of the above discussion, we propose a centralized DSM engine for smart social society, using IIoT and Big Data analytics. The major contributions of this paper are as follow.

1) The proposed DSM engine has two parts. The first part is responsible for the authentication of HAN, using a novel payload-based mutual authentication mechanism. Using the lightweight CoAP protocol of IoT, requests and responses are exchanged among the clients and server for secured session establishment within HAN.
2) The second part is responsible for data stream processing and decision making for the DMS. In addition, data stream processing is performed using Big Data analytics with MapReduce parallel processing mechanism. The selection of running devices is made with the help of 0/1 Knapsack Algorithm in the decision-making stage.

The remaining sections are organized as follows. A review of related works is presented in Section II. Prospective of Demand Side Management in Smart Society: A Birds Eye view is given in Section III. The proposed Demand Side Management Engine and technology stack is presented in Section IV. Section V elaborates the results. Finally, section VI presents the conclusion of the proposed work.

## II. RELATED WORK

Internet of Things (IoT) is a merger of different technologies to connect and make use of smart devices for a better change in human lives. Riggins and Wamba [14] proposed a framework based on the idea of IoT evolution. According to them, the evolution will take place from monitored things to the network of things. Sun *et al.* [15] endorse an innovative concept of smart connected communities (abbreviated as SCC). The concept evolved from the smart cities to enhance the present living standard and meet the future demands in a technical and innovative way. A case study was also presented for smart tourism using the integration of IoT and Big Data analytics. Rathore *et al.* [16] proposed a four-tier architecture to empower the decision making of the societies. The implementation was performed using Hadoop and Spark to provide real-time data processing. The system performance was tested for real-time processing and throughput. Katal *et al.* [17] defined the Big Data as a huge amount of data that need new technologies and techniques for processing. They explained the associated properties, i.e., Variety, Volume, Velocity, Variability, Complexity and Value. They also discussed the challenges and issues related to Big Data analytics including privacy, security, data access, information sharing, analytical, skill requirement and few technical challenges. Barbato and Capone [18] classified the Demand side optimization techniques, based on three criteria 1) user interaction: it deals with the modeling of user behavior and is further categorized into individual and cooperative users

2) optimization approach: how the problems are dealt either by deterministic or stochastic approaches and 3) time scale defines the planning scope either as futuristic or real-time. The authors also formulated a DSM optimization Model. The proposed model tries to achieve various objectives such as billing, discomfort minimization, and maximization of locally generated energy. These objectives are subject to various constraints such as, electric devices (fixed, shift-able, and elastic), local energy generators, and energy storage and balance, and energy tariffs.

There are various authentication schemes in literature for securing IIoT devices to better manage communications between them. Securing these devices against adversaries ensures that energy consumption of various appliances and devices remain intact. The HTTP-based web technology utilizes the Representational State Transfer (REST) architecture [19]. The Internet Engineering Task Force (IETF) [20] has created a particular group known as Constrained RESTful Environments (CoRE). This group is accountable for designing lightweight protocols to offer web resources to IoT devices [21]. The Constrained Application Protocol (CoAP) is one such product of CoRE that relies on a subset of HTTP features to fulfill the requirements of resource-starving IoT/IIoT devices [22]. This protocol utilizes a lightweight request/response model for interaction among clients and servers, to share the resources. Each client has the choice to register itself with a particular server for resource observation [23]. CoAP is a perfect substitute for the current IoT protocols such as XMPP [24] and MQTT [25].

In view of the aforementioned discussion, CoAP is implemented in a variety of applications such as home automation system [26], transport logistics [27], freight supervision [28], and smart cities [29]. To provide security, Datagram Transport Layer Security (DTLS) is required by CoAP at the transport layer to build it more friendly toward the resource-constrained networks [29]. There are a number of proposals in literature that are based on DTLS. A robust security scheme was proposed [30] and the performance of DTLS handshaking was investigated for resource-starving sensors (devices) [31]. In addition, the DTLS implementation for smart phones (INDIGO) using CoAP, was proposed [32]. Moreover, a lightweight authentication scheme for a CoAP-based IIoT environment was proposed for resource observation [33]. The proposed scheme provides security for resource observation in a smart home environment. Although, DTLS-based mechanism sustains a great variety of cipher suites, it was initially developed for networks with large amount of resources. The resource-intensive multifaceted cipher suites of DTLS do not take into account the message length into consideration as a critical criterion for network security.

## III. PROSPECTIVE OF DEMAND SIDE MANAGEMENT IN SMART SOCIETY: A BIRD'S EYE VIEW

Demand Side Management deals with the efficient utilization of energy by selecting the high-priority and high-demand devices at right time while remaining in the prescribed load and cost limits in accordance to the specified parameters or constraints [13], [34], [35]. The DSM illustration is depicted in **Fig. 1**. The parameters which are taken into consideration in a DSM include:

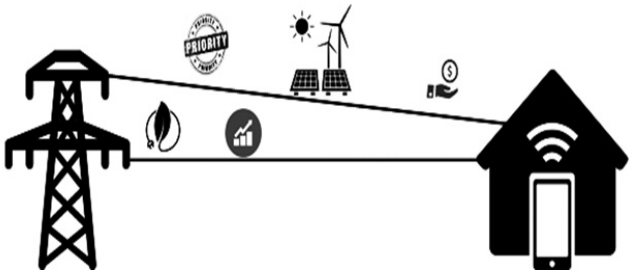


**FIGURE 1.** Demand side management: an illustration.

1) **Device Load:** This is the minimum load required to operate a specific device.
2) **Cost:** This is the most critical parameter from consumer point of view. It is the correlation of time and demand and can be different at different time of a particular day but in generally price tariffs from energy supplier are considered authority in this regard.
3) **Load Limit:** It is the most critical parameter from supplier point of view; generally it is a limit specified by the supplier and it may change according to consumer hours and tariffs.
4) **Time:** The time may be the deciding factor for many of the above parameters because the cost and the load is measured according to the time unit.
5) **Device Priority:** Priority can be defined as the measure of use or importance at a particular time of the day. Hence, priority of a device can vary during different hours of a particular day. A specific device could be in one and only one priority among the following at a given time and could change dynamically according to various situations.
    a) **Real Time:** The device is absolutely required regardless of the load and cost limits or according to the trade-offs.
    b) **High:** The device is in high demand while remaining in the load but not cost limit or vice versa, however, not both.
    c) **Medium:** The device is required but when it is feasible to remain in the load and cost limit. If no other constraint is specified.
    d) **Low:** The device is at low priority and can be shut down for high-priority device in demand.
    e) **Not Required:** The device will remain off until the priority is changed.

### A. DSM TECHNIQUES

From supplier prospective, load is the primary point of concern and from consumer side, it is the secondary concern.

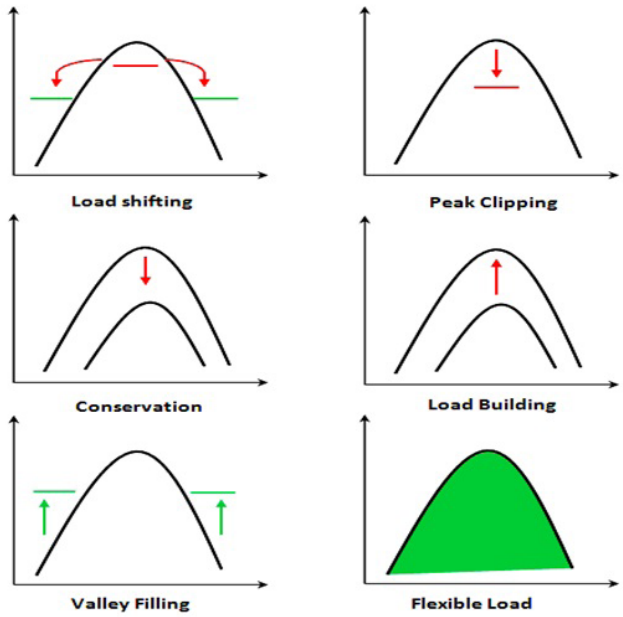


**FIGURE 2.** DSM techniques.

The main objective of DSM is to remain in the load and cost limits. This can be achieved with the techniques shown in **Fig. 2**.

1) **Load Shifting:** In this technique, the load is shifted from peak hours to off-peak hours. In other words, it is the way to segregate the devices, based on their priority level. The low-priority devices will be shifted to other suitable times, while the high-priority devices can be kept on running. The main drawback of this technique is that we have to switch off some of the devices.
2) **Peak Clipping:** This is a direct load control technique to reduce peak load and achieve an optimum level.
3) **Conservation:** This technique is used to reduce the energy consumption.
4) **Load Balancing:** This is a technique to balance the load.
5) **Valley Filling:** It is the opposite of peak clipping. It directly increases the load in off-peak hours.
6) **Flexible Load:** It is a technique that provides the facility of flexibility in the energy load.

To achieve the desire results, consumer needs to make several trade-offs between different desires. Some of them are as follow.

1) **Availability vs. Cost:** It's totally a consumer decision what exactly is important for him/her, be it the 24/7 availability or the cost optimization. More availability means a higher cost.
2) **Availability vs. Load:** High availability means increase in the load. If high availability is chosen, its the responsibility of supplier to provide as much load as needed at the consumer end. This can be achieved through alternate sources of energy.
3) **Cost vs. Load:** Cost and load are directly proportional to each other. If more load is needed, the cost will increase but not in all cases, e.g., the cost may decrease in off-peak hours. If these trade-offs are not chosen by the consumer, then DSM technique should be capable enough to choose optimum values.

### B. COMBINATORIAL OPTIMIZATION

Combinatorial optimization is about finding the optimal set based on some criteria, given a finite set of devices. Combinatorial optimization is the process of searching for maxima (or minima) of an objective function F whose domain is a discrete but large configuration space. In the proposed engine, the choice of devices is based on the parameters defined above. In addition, 0/1 Knapsack algorithm is used in the proposed engine for the selection of devices which can be described as:

$$Max \sum_{i=1}^{n} DeviceCost_i \tag{1}$$

$$\text{subject to} \sum_{i=1}^{n} DeviceLoad_i \leq Loadlimit \tag{2}$$

## IV. PROPOSED METHODOLOGY

In this section, the comprehensive and detail description of the proposed architecture is discussed. The proposed DMS engine is a multilayered centralized engine which is composed of 1) Data/Message Receiving and Pre-Processing, 2) Data Stream Processing, and Decision-Making and User Interface Layers. The position of the DSM engine in the smart social society is shown in **Fig. 3**. The HANs shown in the area are connected to this engine through LAN or MAN. Before going into the detail description of each layer of the proposed DSM engine, the overview of the architecture is given first.

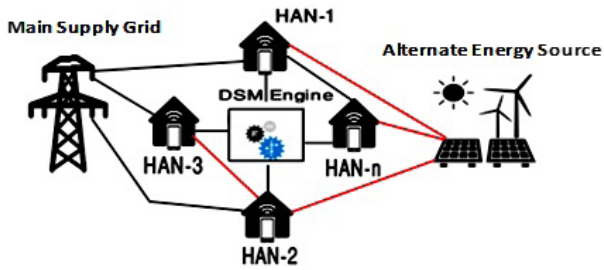


**FIGURE 3.** DSM engine for smart social society.

### A. OVERVIEW OF PROPOSED DSM ENGINE

A high level architectural overview of the proposed DSM engine is shown in **Fig. 4**. This architecture consists of multiple layers and includes a number of components. First of all, the data produced by different HANs connected to the proposed engine are collected by the message receiver. The message identification and authentication is performed at this stage.

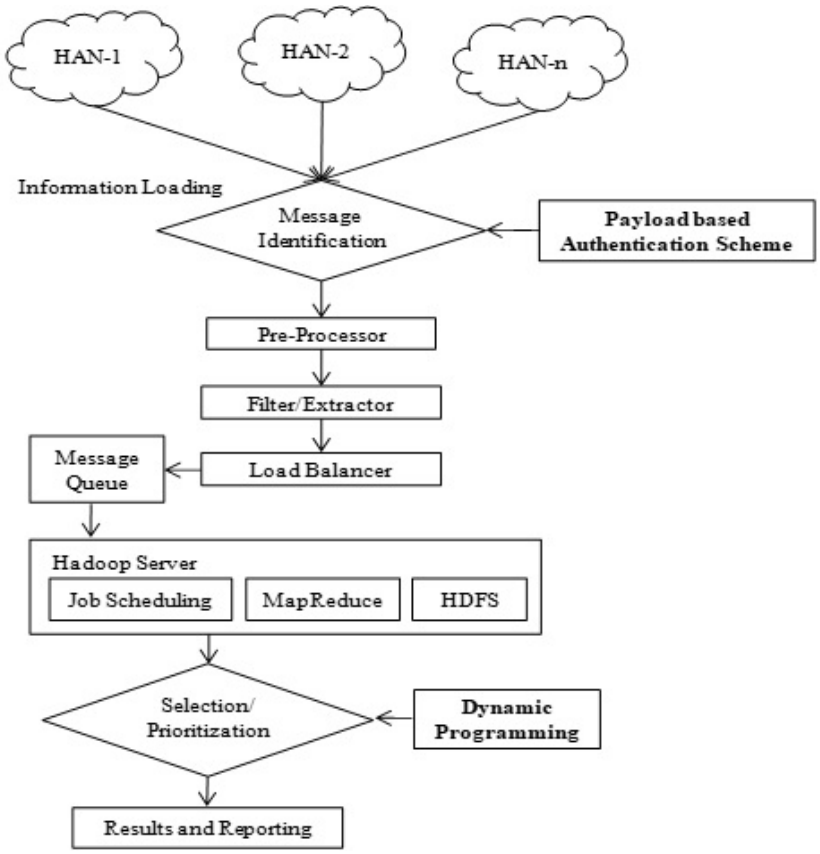


**FIGURE 4.** Overview of DSM engine.

The proposed architecture utilizes a payload-based mutual authentication scheme in order to identify the messages and provide secured connection. The payload authentication uses CoAP as the underlying protocol to meet the requirements. The message receiver is responsible for keys exchange and overall authentication. Afterwards, this data is forwarded to the pre-processing for removing the anomalies from the raw data acquired from outer IoT data sources. These anomalies may include the missing values, invalid data, unreliable data etc. The message identification process is also performed at this stage. Afterward, the filtration and extraction is performed in order to filter the data and remove the noise. Once the pre-processing is performed, the data is ready to be processed.

### B. PROPOSED DEMAND SIDE MANAGEMENT ENGINE

The DMS engine is a multilayered centralized engine for smart social societies and is graphically depicted in **Fig. 5**. The exhaustive explanation of every layer of the proposed engine is given in the forthcoming segment.

#### 1) DATA/MESSAGE SECURITY AND PRE-PROCESSING LAYER

This is the first layer of the proposed DSM engine. It is directly connected to the IoT data sources, i.e., HANs. The data is first collected and authenticated at this stage. The data from HANs is received in the form of messages which can be one of the two types: 1) HAN configuration message, and 2) device message. HAN configuration message contains the setting/ preferences information of a particular HAN. This type of message may contain HAN ID, initial priorities of all devices, max load limit, and max cost limit. Device message contains the setting information of a particular device containing the information about a particular device such as device id, load, priority, and status. Therefore, once the data is collected from data sources, the message identification process is carried out to provide secured and trusted data.

#### 2) PAYLOAD-BASED MUTUAL AUTHENTICATION SCHEME

The payload authentication scheme uses CoAP protocol as the underlying protocol to meet the load requirements. However, the CoAP-based implementations for IoT are dependent on DTLS for the protected transfer of resources among the devices. Though, the DTLS-enabled CoAP stack produces an additional protocol layer for security provisioning which increases the computational and communication cost in term of energy consumption. In our proposed authentication, there is no additional protocol layer in the proposed scheme. Furthermore, it does not compromise the security of data messages transferred among the clients and server. Authentication is offered at the time of request-response

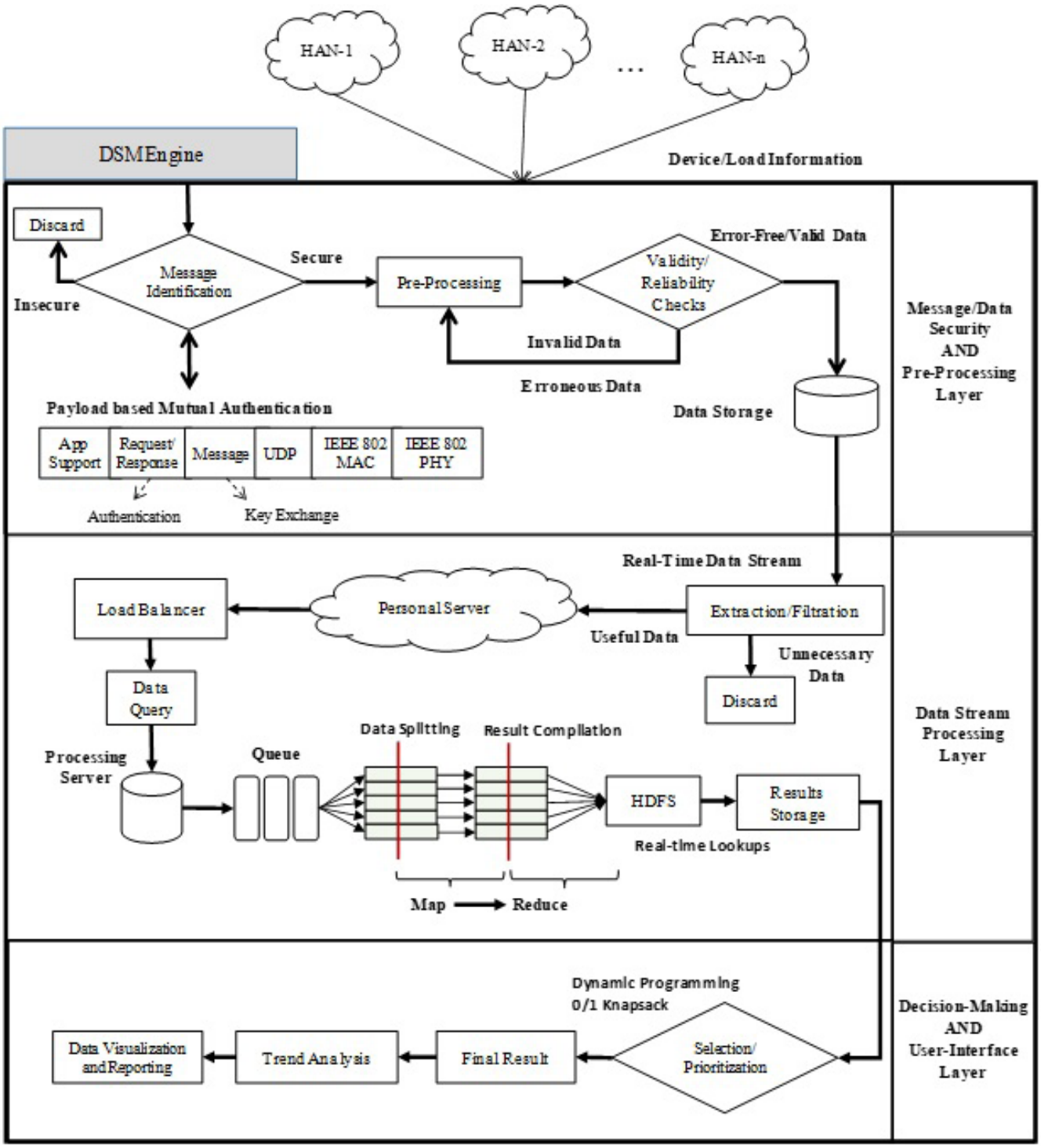


**FIGURE 5.** Proposed DSM engine.

communication among the clients and server while the session key is exchanged within the payload of transmitted data (messages). This scenario is depicted in **Fig. 5** in the security layer. For authentication, the security features are incorporated into CoAP to make it more efficient, robust, and secured against various malicious activities. Contrasting the encryption techniques based on DTLS, the proposed method offers authentication using the payload of messages exchanged among the clients and the server. Both the server and client confront each other during the process of authentication. The whole process is carried out using 4 handshake messages where, the payload of each message

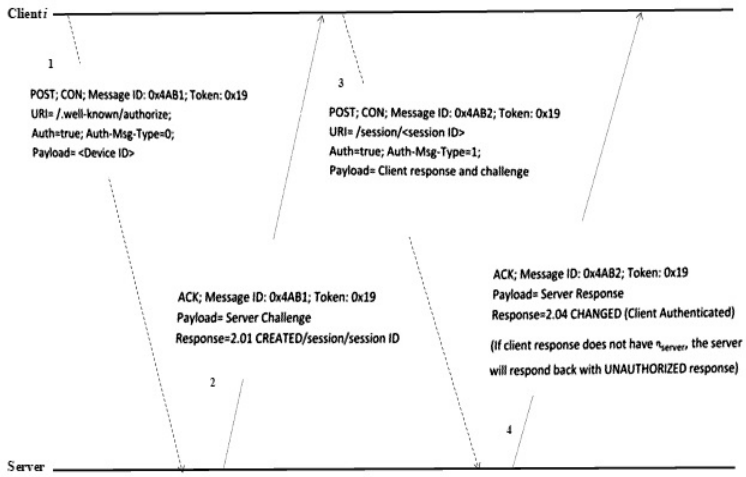


**FIGURE 6.** Authentication handshake procedure.

is kept a maximum of 256 bits as shown in **Fig. 6**. The proposed scheme is accomplished using the four steps which are 1) Session launching, 2) Server challenge, 3) Client reply and challenge, and 4) Server reply.

The session launching is headed by the provisioning stage. It is a prerequisite stage (offline) during which the clients share a secret key, with the server which is only known to the server and the owner client of the key. The server preserves a record of these keys with unique identifier (ID) linked with it. Upon successful authentication, the communication of session key between both parties takes place. It assumes that the secret keys are embedded with sensors in each device at manufacturing and deployment. If an intruder tries to temper with physical devices, an alarm is generated to inform about the breach. In our proposed work, the devices are tamper-safe [33].

Each client sends a request message to the server alike to a Hello Client message in the session initiation phase. The said request is sent to the server URI, /.well-known/authorize. The server gets the ID of the devices from the message payload during the server challenge phase. The server carries out a table look-up for a corresponding key using this ID. The server responds back with an encrypted payload if a match is found. To generate the challenge, the server produces a nonce (pseudo-random), and a impending session key which is a temporary number that is used only once by a device. An encrypted payload is generated by the server at this stage. The client requires deciphering the encrypted payload to get back the session key in the client response and challenge phase. Finally, the server deciphers the encrypted payload of the client challenge. If it is there, the server comprehends that the client has successfully authenticated itself.

### 3) PRE-PROCESSING

Upon successful authentication, the data is transmitted to the server. This data is in raw form and have a lot of discrepancies. Therefore, we need to perform the pre-processing techniques before the actual processing. In the pre-processing phase, mostly validity and reliability checks are performed to make the data clean and error-free for processing. Today's real-world Big Data are extremely susceptible to inconsistencies, missing values, different formats, and noise, due to their characteristically massive size and their prospective origin from multiple, heterogeneous sources. The low-quality data will lead to low-quality processing results. Data pre-processing methods, when applied before the actual processing, can noticeably improve the overall quality of the data processing and/or the time required for the actual processing [36]–[38]. Prior to processing, the proposed architecture performs the following: 1) data reduction, 2) data cleaning, and 3) data transformation. Data reduction can be made practical to get a reduced representation of the datasets that is smaller in quantity, yet closely preserves the integrity of the original data. In addition, processing and analysis on the reduced dataset is much more efficient yet produce almost the same analytical results. Data reduction is typically applied on massive data that do not present many valued data. Likewise, data transformation is performed. The basic purpose of the transformation is to limit the value of data to a certain limited range. We preferred the Min-Max technique to transform a specific value X to another value Y. Finally, the data cleaning is carried out. The collected data could be incomplete lacking attributes values or certain attributes of interest. It may be noisy that is a random error or variance in a measured variable, and inconsistent that containing discrepancies in codes or names. In this research, the incomplete (missing values) and noise are taken into consideration.

### 4) DATA STREAM PROCESSING LAYER

Stream processing layer is performed to process the data coming from different IoT sensors at real-time, after proper pre-processing. This processing is based on the notion of parallel processing in which various execution of processes are performed concurrently. Parallel processing formulate a program by executing it quicker due to many running engines (CPUs). The processing of the data starts with data

filtration. The filtration process is used to discard any useless data, i.e., which is not useful for the processing and does not affect its results. In the proposed architecture, the Kalman Filter (KF) is used to speed up the data processing and separate the valuable and noisy data. KF is utilized to carry out the data filtration in order to filter the noise from the data. It is a statistical method and plays a noteworthy part in sensing of the real-world data [39], [40]. Once the filtration and extraction is performed, the data is stored in a server for load balancing. Load balancing is carried out as we perform parallel processing. A load balancer is used to distribute the traffic across a number of parallel processors in order to distribute the load. It increases the reliability of processing along with the concurrent users' capacity.

To be very precise in the case of data processing, the implementation is achieved using Hadoop framework with MapReduce mechanism. At this phase, the identical formation of MapReduce and HDFS is utilized. Apart from HDFS, we can also use SQL-supposed HBASE, and HIVE, for the administration of database (offline or in-memory) to store historical information. Because, real-time streams of data are processed, a third party tool is used for real-time stream processing to merge with the Hadoop to facilitate the real-time stream execution. To provide real-time realization, Apache Spark is used with the Hadoop which is a general-purpose engine for wide-ranging stream data processing. It provides speedy stream processing for Big Data and allows re-usability across streaming applications. The Spark components are depicted in **Fig. 7**. Apache Spark provides stream processing, speedy, releasable, provides in-memory computing, offers advance analytics, and fault tolerant.

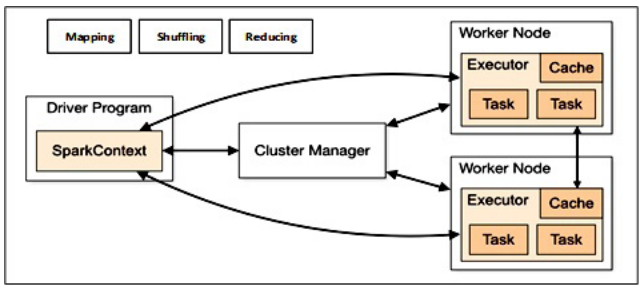


**FIGURE 7.** Apache spark working.

### 5) DECISION-MAKING AND USER-INTERFACE LAYER

After processing, results are communicated to a respective HAN for compliance and data is stored in database for future trend analysis. In this phase, first of all the selection and prioritization is taken place. In selection and prioritization, the devices are selected based on the parameters provided by the consumers. The devices are selected in two different phases. In phase I, the HANs are prioritized based on their requirements and criteria provided by them and in phase II, the devices are selected for a particular HAN. The input contains $s$ set of finite number of devices with associate cost and load and output having a set of selected devices. The I/O can be represented as follow:

$$\begin{gathered}\textbf{Input:}\\ \mathrm{D} = \{\mathrm{d}_1, \mathrm{d}_2, \mathrm{d}_n\}\\ \mathrm{C} : \mathrm{D}\ \bar{a}\ \mathrm{R}+\\ \mathrm{L} : \mathrm{D}\ \bar{a}\ \mathrm{R}+\\ \textbf{Output:}\\ \mathrm{S} = \{\mathrm{s}_1, \mathrm{s}_2, \mathrm{s}_3\}\\ \mathrm{S} \subseteq \mathrm{D} \mid \mathrm{Min} \sum_{s\in S}^{n} C_s\\ \text{and} \sum_{s\in S}^{n} L_s \leq \mathrm{MaxLoad}\end{gathered}$$

The selection and prioritization activity is performed using dynamic programming approach though 0/1 Knapsack algorithm in the proposed DMs engine.

### 6) DYNAMIC PROGRAMMING

A dynamic programming (DP) is a technique (algorithm) which is based on recursion,i.e., divide-and-conquer, with some variation. Basically, DP is the combination of recursion and some common sense. Recursion allows us to describe a function's value in the form of other values, whereas, the common sense describes the implementation of function in such a way that recursion is performed in advance and results are stored to be accessed easily that eventually makes the program faster. It is also known as memorization because it memorizes the outcome of some specific states which can be accessed in later stage to solve sub-problems. A sub-problem is solved by a particular existing sub-solution that makes this technique very fast in comparison to the existing techniques. Most of the DP problems can be classified into two different types: 1) Optimization problems, and 2) Combinatorial problems. The optimization problems anticipate you to choose a viable result to minimize or maximize the value of the required function. On the other hand, combinatorial problems anticipate you to discover the number of possible and optimal ways to accomplish something. Each DP problem has a plan to be tracked:

- To highlight that the problem can be divided into most favorable sub-problems
- Repeatedly describe the solution by stating it in the form of best possible solutions for sub-problems (smaller)
- Calculate the value of best possible solution in bottom-up manner
- Design most favorable solution from the calculated information.

As we need to find an optimal object solution from a finite set, this scenario is included in combinatorial optimization. In such problems, the traditional exhaustive searching is not feasible because the optimal solutions are discrete or tends to be discrete. Therefore, we prefer the dynamic programming approach and the knapsack problem which is in the combinatorial optimization.

### 7) DYNAMIC PROGRAMMING BASED 0/1 KNAPSACK ALGORITHM

One of the general problem is 0-1 knapsack problem that limits the number $xi$ of replicas of all type of entry to 0 or 1. For example, various items set the knapsack to find the max

total value. Every item has particular weight and value. Total weight will always be less than fixed weight called W. Hence, weights and values of items must be taken into consideration, as shown in Table 1.

**TABLE 1. Knapsack problem.**

| Item No. | Weight | Value |
|---|---|---|
| 1 | 1 | 8 |
| 2 | 3 | 6 |
| 3 | 5 | 5 |

Given a set of $n$ items from $1 - n$, all with a value $v_i$ and a weight $w_i$, with a capacity of maximum weight W, can be mathematically represented as:

$$Max \sum_{i=1}^{n} v_i w_i \tag{3}$$

$$\text{subject to,} \sum_{i=1}^{n} w_i x_i \leq W \quad and\ x_i \in \{0, 1\} \tag{4}$$

where, $x_i$ denotes the instances of $i$ consist of the knapsack.

Once the optimal solutions are obtained, then the final results are produced and store in a particular storage area for event management. These results can be used to perform the trend analysis and reporting.

## V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, the experimental results for our proposed system are provided. Our authentication module is implemented using NetDuino Plus 2 boards for the client and server interaction model. We develop our own library, **CoAP Micro**, to enable basic communication among various entities. On the other hand, the stream processing module of our proposed DSM engine is implemented using Apache Spark with Hadoop single node setup on UBUNTU 16.04 LTS. Input libraries are used for processing and producing Hadoop Readable form (sequence file) at collection and aggregation unit so that it can be processed by Hadoop. First, we provide experimental results for secured authentication followed by stream processing of DSM.

### A. SECURITY AUTHENTICATION USING PROPOSED ARCHITECTURE

Here, we provide the experimental results with regard to security layer authentication. In **Fig. 8**, the comparison of our proposed authentication scheme is provided against CoAP-based DTLS implementation (INDIGO) for smartphones. DTLS* symbolizes the handshake between a standard computer and a smartphone, in which the computer operates as a server and the smartphone acts as a client. On contrary, DTLS+ symbolizes the computer as a client and the smartphone as a server.

Similarly, the CoAP messages are transferred asynchronously over the UDP sockets in our proposed architecture. Each client keeps a record of the transferred confirmable (CON) requests to maintain their record. When a matching acknowledgement(ACK) response is received for such messages, the transmission is considered successful.

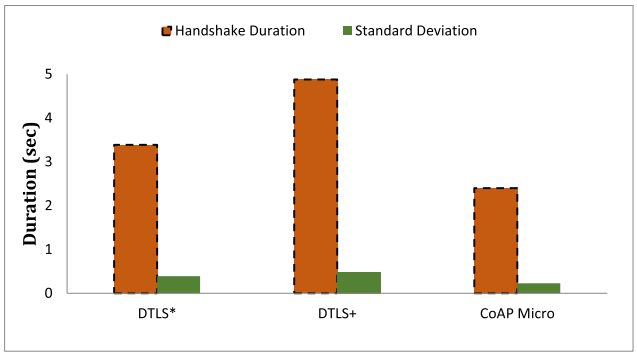


**FIGURE 8. Handshake duration.**

The average response time for a single CON message having a payload of one byte is compared against DTLS and the CoAP protocol (with no security), in **Fig. 9**.

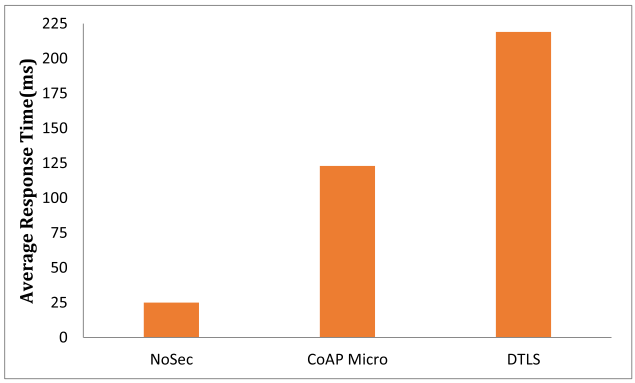


**FIGURE 9. Average response time.**

The average memory consumption of a request message at the time of compilation is acquired using the Microsoft.SPOT.Native assembly. The proposed scheme is compared with existing ones for a CON message of 500 bytes, as shown in **Fig. 10**. Unlike our proposed scheme, a considerable memory to the messages is allocated by CoapBlip [35] and its variant TinyCoAP [36] at the compile time. Both these schemes are an adaptation of the standard C language libraries that need TinyOS element for installation on a sensor node. HTTP/UDP stack, on the other hand, has a very low memory foot-print as it does not provide a reliability mechanism or a request/response matching.

### B. STREAM PROCESSING AND DECISION-MAKING USING PROPOSED ARCHITECTURE

In this section, the experimental results with regard to stream processing using big Data analytics is provided. Hadoop is used for distributed storage and parallel proceeding with MapReduce progra,ming paradigm. Hadoop performs batch processing, therefore; we utilized the Apache Spark with Hadoop in order to achieve real-time stream processing. The proposed DSM engine stream processing is experimentally validated in a smart home scenario of different sensors. The energy consumption of the sensors is tested by turning-on the devices randomly. Our smart home environment is composed

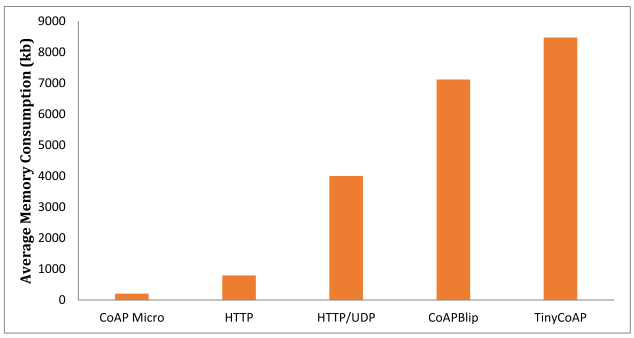


**FIGURE 10.** Memory consumption.

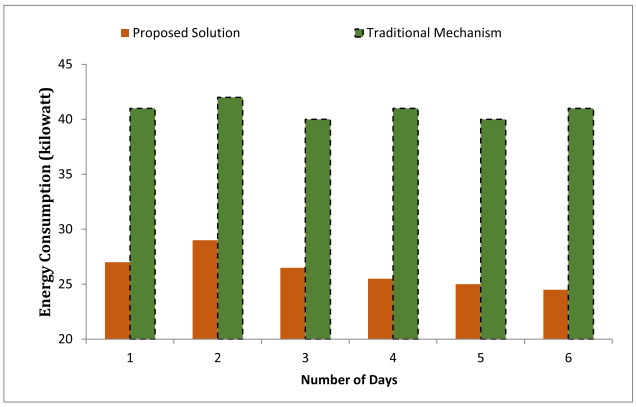


**FIGURE 11.** Energy consumption of air conditioner.

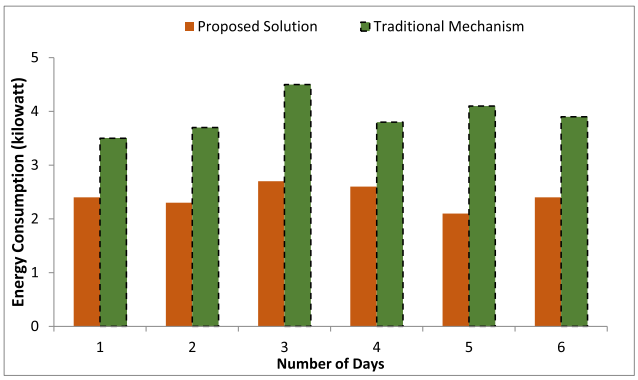


**FIGURE 12.** Energy consumption of a computer.

of a single kitchen, four rooms, and two toilets. Every room has 4 devices and the kitchen is allocated with 6 appliances ranging from high to low energy consumption. Similarly, the toilet has 2 appliances. Therefore, we have 26 sensors in the whole smart home scenario. Furthermore, some of the devices that are lastingly present in the on-state, for instance refrigerator, etc. are not taken into account in the proposed scheme. The results show that when the consumers are regularly using the devices, their energy consumption is visibly higher and increases instantly. On the other hand, when applying the proposed architecture to the smart home environment the energy utilization considerably decreases. **Fig. 11-14** depict the energy consumption of different appliances such as computer, air-conditioning system, television, and light. The consumption of the devices is significantly

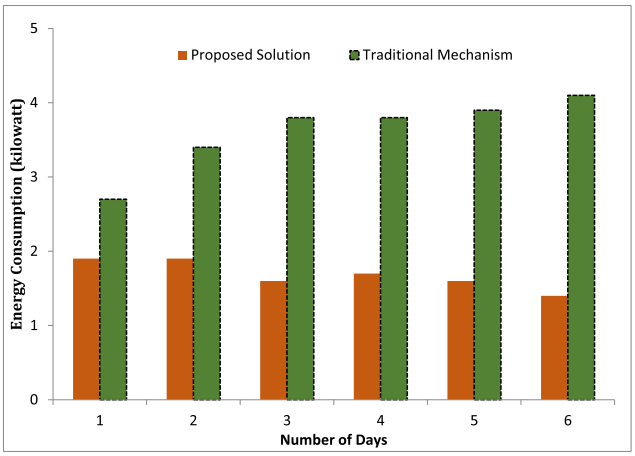


**FIGURE 13.** Energy consumption of light source.

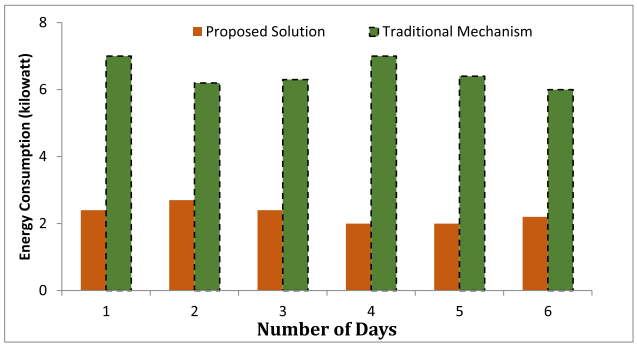


**FIGURE 14.** Energy consumption of a television.

reduced over a period of one week. Furthermore, the energy consumption is optimized as well. In the traditional mechanism, energy consumption is not optimized that is clearly depicted in these graphs. Likewise, the non-suitable energy consumption is managed and controlled in case of our proposed architecture. Therefore, the consumer can efficiently manage the needless power consumption of the machines using the proposed system.

## VI. CONCLUSION

This paper proposed a secure and trusted multi-layered DSM engine for a smart social society using IIoT and Big Data analytics. The major objective is to provide a generic secured solution for smart societies in IIoT environment. The proposed engine uses a centralized approach to achieve an optimum DSM over a Home Area Network (HAN). This centralized Engine takes input from IIoT sensors of HANs and performs authentication and Big Data analytics for the selection and prioritization of devices of a particular HAN. A payload-based authentication scheme is utilized by the DSM engine to provide an adequate level of security. The payload-based encryption method utilizes a 4-way handshake to authenticate the participating devices of the IIoT. In addition, data stream processing is performed using a MapReduce parallel processing mechanism. The security perspective of our proposed system is evaluated using NetDuino Plus 2 boards which reveals that the proposed

authentication incurs less connection overhead, computationally-efficient, and offers a robust defense against various cyber attacks. On the other hand, the stream processing is performed using Apache Hadoop with Apache Spark to verify the proposed DMS engine and the examination reveals that the proposed architecture offers valuable insights into the smart social societies in the context of IIoT.

## REFERENCES

[1] P. Bocquier, "World urbanization prospects: An alternative to the UN model of projection compatible with the mobility transition theory," *Demograph. Res.*, vol. 12, pp. 197–236, Jan. 2005.
[2] B. Cohen, "Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability," *Technol. Soc.*, vol. 28, nos. 1–2, pp. 63–80, 2006.
[3] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.
[4] Gartner Incorporation. *Gartner Press Release*. Accessed: Apr. 7, 2018. [Online]. Available: http://www.gartner.com/newsroom/id/3598917
[5] D. Zeng, S. Guo, and Z. Cheng, "The web of things: A survey," *J. Commun.*, vol. 6, no. 6, pp. 424–438, 2011.
[6] F. Xia, L. T. Yang, L. Wang, and A. Vinel, "Internet of Things," *Int. J. Commun. Syst.*, vol. 25, no. 9, p. 1101–1102, 2012.
[7] J. Granjal, E. Monteiro, and J. S. Silva, "Security for the Internet of Things: A survey of existing protocols and open research issues," *IEEE Commun. Surveys Tut.*, vol. 17, no. 3, pp. 1294–1312, 3rd Quart., 2015.
[8] Gartner Incorporation. *The Importance of Big Data: A Definition*. Accessed: May 8, 2018. [Online]. Available: https://www.gartner.com/doc/2057415/importance-big-data-definition
[9] R. Sharda, D. Delen, and E. Turban, *Business Intelligence: A Managerial Perspective on Analytics*. Upper Saddle River, NJ, USA: Prentice-Hall, 2013.
[10] P. Russom *et al.*, "Big data analytics," Transforming Data Intell., Renton, WA, USA, TDWI Best Pract. Rep., 2011, pp. 1–34, vol. 19, no. 4.
[11] M. Strohbach, H. Ziekow, V. Gazis, and N. Akiva, "Towards a big data analytics framework for IoT and smart city applications," in *Modeling and Processing for Next-Generation Big-Data Technologies*. Berlin, Germany: Springer-Verlag, 2011, pp. 257–282.
[12] M. Babar and F. Arif, "Real-time data processing scheme using big data analytics in Internet of Things based smart transportation environment," *J. Ambient Intell. Hum. Comput.*, pp. 1–11, May 2018.
[13] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE Trans. Ind. Informat.*, vol. 7, no. 3, pp. 381–388, Aug. 2011.
[14] F. J. Riggins and S. F. Wamba, "Research directions on the adoption, usage, and impact of the Internet of Things through the use of big data analytics," in *Proc. 48th Hawaii Int. Conf. Syst. Sci. (HICSS)*, Jan. 2015, pp. 1531–1540.
[15] Y. Sun, H. Song, A. J. Jara, and R. Bie, "Internet of Things and big data analytics for smart and connected communities," *IEEE Access*, vol. 4, pp. 766–773, Mar. 2016.
[16] M. M. Rathore, A. Ahmad, A. Paul, and S. Rho, "Urban planning and building smart cities based on the Internet of Things using big data analytics," *Comput. Netw.*, vol. 101, pp. 63–80, Jun. 2016.
[17] A. Katal, M. Wazid, and R. Goudar, "Big data: Issues, challenges, tools and good practices," in *Proc. 6th Int. Conf. Contemp. Comput. (IC3)*, Aug. 2013, pp. 404–409.
[18] A. Barbato and A. Capone, "Optimization models and methods for demand-side management of residential users: A survey," *Energies*, vol. 7, no. 9, pp. 5787–5824, 2014.
[19] R. T. Fielding and R. N. Taylor, "Principled design of the modern Web architecture," *ACM Trans. Internet Technol.*, vol. 2, no. 2, pp. 115–150, 2002.
[20] IETF. *Applicability of the QUIC Transport Protocol, the Internet Engineering Task Force (IETF)*. Accessed: May 27, 2018. [Online]. Available: https://www.ietf.org
[21] J. Kim, J. Lee, J. Kim, and J. Yun, "M2M service platforms: Survey, issues, and enabling technologies," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 61–76, 1st Quart., 2014.
[22] Z. Shelby, K. Hartke, and C. Bormann, *The Constrained Application Protocol (CoAP)*, document RFC 7252, 2014.
[23] K. Hartke, *Observing Resources in the Constrained Application Protocol (CoAP)*, document RFC 7641, 2015.
[24] P. Saint-Andre, *Extensible Messaging and Presence Protocol (XMPP): Core*, document RFC 6120, 2011.
[25] D. Thangavel, X. Ma, A. Valera, H.-X. Tan, and C. Keng-YanTan, "Performance evaluation of MQTT and CoAP via a common middleware," in *Proc. IEEE 9th Int. Conf. Intell. Sensors, Sensor Netw. Inf. Process. (ISSNIP)*, Apr. 2014, pp. 1–6.
[26] S.-C. Son, N.-W. Kim, B.-T. Lee, C. H. Cho, and J. W. Chong, "A time synchronization technique for coap-based home automation systems," *IEEE Trans. Consum. Electron.*, vol. 62, no. 1, pp. 10–16, Feb. 2016.
[27] M. Becker, T. Pötsch, K. Kuladinithi, and C. Goerg, "Deployment of coap in transport logistics," in *Proc. 36th IEEE Conf. Local Comput. Netw. (LCN)*, Bonn, Germany, Oct. 2011, pp. 1–3.
[28] M. Becker, K. Kuladinithi, T. Pötsch, and C. Görg, "Wireless freight supervision using open standards," in *Proc. 5th Int. Workshop Cold Chain Manage.*, Jun. 2013, pp. 10–11.
[29] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, "Long-range communications in unlicensed bands: The rising stars in the IoT and smart city scenarios," *IEEE Wireless Commun.*, vol. 23, no. 5, pp. 60–67, Oct. 2016.
[30] S. Raza, L. Seitz, D. Sitenkov, and G. Selander, "S3K: Scalable security with symmetric keys—DTLS key establishment for the Internet of Things," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 3, pp. 1270–1280, Jul. 2016.
[31] T. Kothmayr, C. Schmitt, W. Hu, M. Brünig, and G. Carle, "DTLS based security and two-way authentication for the Internet of Things," *Ad Hoc Netw.*, vol. 11, no. 8, pp. 2710–2723, 2013.
[32] D. Trabalza, S. Raza, and T. Voigt, "Indigo: Secure coap for smartphones," in *Wireless Sensor Networks for Developing Countries*. Berlin, Germany: Springer, 2013, pp. 108–119.
[33] M. A. Jan, P. Nanda, X. He, Z. Tan, and R. P. Liu, "A robust authentication scheme for observing resources in the Internet of Things environment," in *Proc. IEEE 13th Int. Conf. Trust, Secur. Privacy Comput. Commun. (TrustCom)*, Sep. 2014, pp. 205–211.
[34] G. Strbac, "Demand side management: Benefits and challenges," *Energy Policy*, vol. 36, pp. 4419–4426, Dec. 2008.
[35] T. Logenthiran, D. Srinivasan, and T. Z. Shun, "Demand side management in smart grid using heuristic optimization," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1244–1252, Sep. 2012.
[36] M. Babar and F. Arif, "Smart urban planning using big data analytics to contend with the interoperability in Internet of Things," *Future Gener. Comput. Syst.*, vol. 77, pp. 65–76, Dec. 2017.
[37] M. Babar, A. Rahman, F. Arif, and G. Jeon, "Energy-harvesting based on Internet of Things and big data analytics for smart health monitoring," *Sustain. Comput., Inform. Syst.*, to be published.
[38] M. Babar and F. Arif, "Smart urban planning using big data analytics based Internet of Things," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. Int. Symp. Wearable Comput.*, 2017, pp. 397–402.
[39] D. Simon, "Kalman filtering with state constraints: A survey of linear and nonlinear algorithms," *IET Control Theory Appl.*, vol. 4, no. 8, pp. 1303–1318, Aug. 2010.
[40] D. Li, S. Kar, J. M. F. Moura, H. V. Poor, and S. Cui, "Distributed Kalman filtering over massive data sets: Analysis through large deviations of random Riccati equations," *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1351–1372, Mar. 2015.

**MUHAMMAD BABAR** received the B.Sc. degree (Hons.) from the University of Peshawar, Pakistan, in 2008, and the M.Sc. degree in computer software engineering from the National University Sciences and Technology (NUST), Islamabad, Pakistan, in 2012. He is currently pursuing the Ph.D. degree in computer software engineering with NUST. He has authored his research work in various IEEE and ACM/Springer international conferences and journals. His research areas include but not limited to big data analytics, Internet of Things, smart city design and planning, and social Web of things.

**FAZLULLAH KHAN** received the B.S. degree (Hons.) from the University of Peshawar, Peshawar, Pakistan, in 2007, the M.S. degree in computer software engineering from the National University of Sciences and Technology, Islamabad, Pakistan, in 2011, and the M.S. degree in electrical electronics and information engineering from the Nagaoka University of Technology, Japan, in 2013. He was a recipient of the MEXT Scholarship. He is currently with the Department of Computer Science, Abdul Wali Khan University Mardan, Pakistan, as an Assistant Professor.

His research interests are cross-layer signaling and performance analysis of mobile Ad Hoc networks, vehicular Ad Hoc networks, cognitive radio networks, wireless sensor networks, wireless sensor and actor networks, and cognitive radio sensor networks. Recently, he has been involved in latest developments in the field of underwater sensor network localization and secured ticketing, Internet of vehicles security and privacy issues, software-defined radio VANET, Internet of vehicles, and big data analytics.

He has authored his research work in various IEEE and ACM/Springer international conferences and in the Institute of Electronics, Information and Communication, Japan. He edited two books *Applications of Intelligent Technologies in Healthcare* and *Recent Trends and Advances in Wireless and IoT-enabled Networks* (Springer, 2018). He is the Chair of various conferences and special sessions such as the IEEE EAI Future5v-2017, CCODE-2017, and EAI IoT-BC2. He has been invited to serve as a Technical Program Committee Member for various IEEE international conferences. He has been an active Reviewer for high-cited and highly ranked international journals, including Springer MONET, the *IET Wireless Sensor Systems*, the *IEEE Communications Magazine*, and the IEEE Transactions on ITS.

**WASEEM IQBAL** received the bachelor's degree in computer science from the Department of Computer Science, University of Peshawar, and the master's degree in information security from the National University of Sciences and Technology, Islamabad, Pakistan, where he is currently pursuing the Ph.D. degree. His areas of research include wireless network security, digital forensics, and Internet of Things. He is a certified EC-council instructor for various certifications. He is also a Certified Ethical Hacker, a Certified Hacker, and a Forensic Investigator. He was a recipient of the Merit-Based Scholarship.

**ABID YAHYA** received the bachelor's degree in electrical and electronic engineering majoring in telecommunication from the University of Engineering and Technology, Peshawar, Pakistan, and the M.Sc. and Ph.D. degrees in wireless and mobile systems from Universiti Sains Malaysia, Malaysia. He is currently with the Botswana International University of Science and Technology. He has applied this combination of practical and academic experience to a variety of consultancies for major corporations.

He supervised a number of Ph.D. and master candidates. He has authored or co-authored over 115 research publications to his credit in numerous reputable journals, conference articles, and book chapters. He received several awards and grants from various funding agencies. Recently, he has authored two books *Steganography Techniques for Digital Images* (Springer International Publishing, 2018) and *LTE-A Cellular Networks: Multi-Hop Relay for Coverage, Capacity and Performance Enhancement* (Springer International Publishing, 2017), which are being followed in national and international universities.

Dr. Yahya was an External and Internal Examiner for postgraduate students. He has been invited a number of times to be a Speaker or Visiting Lecturer at different multinational companies. He sits on various panels with the government and other industry-related panels of study.

**FAHIM ARIF** received the bachelor's degree in telecommunication from the College of Telecommunication Engineering, UET, Lahore, in 1995, and the M.S. degree in computer software engineering from the National University Science and Technology, Islamabad, in 2003. In 2007, he joined the System and Computer Engineering Department, Carleton University, Ottawa, Canada, as an International Research Scholar. He participated in numerous research and academic activities. He is currently a Principal Investigator for a project funded by NUST. Recently, his biography has been published by the South Asian Publication Who's Who in the World 2008 Edition and received the Star Laureate Award 2008 in recognition to his contributions to knowledge and research. He received the NUST Endowment Fund Scheme Scholarship from NUST in 2003 and the International Research Support Initiative Program Fund from HEC in 2007.

**ZHIYUAN TAN** (M'13) received the B.Eng. degree in computer science and technology from Northeastern University, China, the M.Eng. degree in software engineering from the Beijing University of Technology, China, and the Ph.D. degree in computer systems from the University of Technology Sydney, Ultimo, NSW, Australia. He held different research positions with three research intensive universities. In 2016, he joined ENU. He was a Post-Doctoral Researcher of cybersecurity with the University of Twente, The Netherlands, from 2014 to 2016; a Research Associate with the University of Technology Sydney, in 2014; and a Senior Research Assistant with La Trobe University, Australia, in 2013. He is currently a Lecture of cybersecurity with the School of Computing, Edinburgh Napier University, U.K. His research interests include cybersecurity, machine learning, pattern recognition, data analytics, virtualization, and cyber-physical system. He is an EAI and BCS Member.

He was the Chair of international workshops and conferences, such as SECSOC, SITN, EAI Future 5V, and EAI BD:TA 2018. He serves on the Editorial Board for the *International Journal of Computer Sciences and its Applications*. He is an Associate Editor of the IEEE Access and an Organizer of Special Issues for the *Ad Hoc and Sensor Wireless Networks Journal*, the *International Journal of Distributed Sensor Networks*, the *Computers and Electrical Engineering*, and the IEEE Access.

**JOSEPH M. CHUMA** received the B.Eng. degree in electrical and electronic engineering from the University of Nottingham, U.K., in 1992, and the M.Sc. degree in telecommunications and information systems and the Ph.D. degree in electronic systems engineering from the University of Essex, U.K., in 1995 and 2001, respectively. He is the Head of the Department of Electrical, Computer and Telecommunications Engineering, Botswana International University of Science and Technology, where he is also an Associate Professor of telecommunication and wireless communications. His main areas of research are in the design of compact single and dual-mode dielectric resonator filters for mobile and wireless communications. His research work has led to successful results in the development of dielectric loaded filters and shown that the size of the conventional combline filter can be reduced in excess of 75%. In this example, a dielectric ring has been used to load a conventional combline resonator. Dr. Chuma has served as the Dean and the Deputy Dean for the Faculty of Engineering and Technology, University of Botswana. He also serves as the Vice Chairman for the Botswana Communications Regulatory Authority Board, the Chairman of the Professional Assessment Committee of the Electrical Discipline of the Botswana Engineers Registration Board, and a member of the Registration Committee of Botswana Engineers Registration Board. He also served as a postgraduate and undergraduate external examiner in a number of universities.