Team 3 End-of-Semester Report:

Anomaly detection in BACnet protocol systems

Team Members: Sofian Ghazali

Muhammed Zahid Kamil

Rahul Balamurugan

Project Mentors: Dr. Hussein Al Nuweiri

Mr. Salah Hessien

Submission Date: 08/12/2019

"An Aggie does not lie, cheat, or steal, or tolerate those who do."

TABLE OF CONTENTS

ABSTRACT

Building Automation Networks/Systems (BAN/BAS) are networks that control and monitor utilities and equipment in smart buildings. BACnet is a widely used device-to-device communication standard that enables building automation and allows for interoperability between devices from different vendors. The issue is that the most commonly used implementations of the BACnet standard, BACnet/IP (Internet Protocol) and BACnet/WS (Web Services) are entirely unencrypted. As more and more BACnet devices are being exposed to the internet, the risk of cyber-attacks is becoming increasingly prominent in addition to the always present risk of more physical attacks against building networks. Our solution is to create a continuously adapting robust anomaly detection mechanism to protect BACnet systems from cyber-physical threats. The project aims to understand and emulate a real local area BAN implementing the BACnet/IP standard, and create an anomaly detector for the same using machine learning. The detector will be trained on the emulated network itself and tested by injecting synthetic attacks into the system. This report details the literature review and customer needs analysis done to arrive at the initial design, performance and market analysis, system flowchart, a detailed discussion of the system design and an update on the progress of the project.

CHAPTER I

INTRODUCTION

With the advent of the smart grid and distributed energy resources, most office spaces and buildings in urban areas have adopted some measure of automation. The demand for devices to control and automate the control of building utilities has led to many competing vendors with KNX, Lonworks and Modbus being a few of the more popular ones. In order to ensure interoperability, consumer convenience, and a fairer market, the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) developed the BACnet Standard in 1955. The BACnet Standard (ISO 16484-6, 2003) has evolved to be one of the most widely used communication protocols over the years due to the flexibility it affords to the management and application layers of the Building Automation Networks. Recently, there is a prevalence of users connecting their Building Networks to the Internet for remote access and monitoring, using the web services implementation of the BACnet Standard.

However, there are clear risks associated with increasing internet connectivity extended to building equipment. The current number of Building Automation Devices, another term extended to BMS systems, amounts to 48,112 [18]. This exposure is mainly since IP addresses of the building facilities using web services can be discovered via search engines such as censys.io and shodan.com that provide a list of such devices and relevant statistics. There are two types pertaining to BACnet protocol: BACnet/SC (secure connect) and BACnet/IP (Internet Protocol), of which the latter is the standard implementation used. BACnet secure ensures messages are encrypted and is a secure form of communication. However, it has interoperability issues with other devices []. On the other hand, BAC/IP allows for ease of interoperability but at a cost- the messages transmitted are in plaintext and completely unencrypted. This is made worse when these devices are exposed to the internet, making them more vulnerable to cyber-attacks.

The objective of this project is to introduce an automated, robust intrusion detection algorithm that detects the presence of anomalous data in BACnet traffic. The scope of our project does not extend towards securing the BACnet communication such as creating a firewall or secure encryption algorithm. Our task starts from the onset of a cyber-physical attack; We notify the management layer of an ensuing attack without exacerbating the situation. We aim to have our algorithm detect any anomalies in a responsive manner and absorb new datasets to become more intelligent. Machine Learning will be implemented to ensure algorithm labels the new dataset accurately with low false positive rate and adapts to new changes in data patterns. To design our project, we took inspiration from existing research into anomaly detection in building networks and also conducted a customer needs survey to understand our project needs from the perspective of security experts and the potential users of the BANs.

Literature Review

Current solutions and research have been focused on using anomaly detection and Intrusion Detection Systems (IDS) to improve system resilience and device-level security management. Some existing anomaly detection methods are the timing-based detector for cyber-physical systems (CPS) by Zimmer et. al [14], and the finite states-based detector for Modbus networks by Goldberg et. al [15]. The major challenge faced by researchers in anomaly detection is to minimize the number of false alarms and improve their efficiency. In this regard, Zheng and Reddy, researchers at TAMU College Station, formulated a THE-driven anomaly detector for BACnet that uses frequency analysis [16].

Zheng et al. designed and developed this detector in 2017 based on a novel network traffic classification model they called the ‘THE-driven classification method’. In this technique, all communication traffic is classified into: 1. Time-driven, 2. Human-driven, and 3. Event-driven categories. Time-driven traffic is

normally generated by scheduled control programs that trigger service requests according to different timers. Such traffic presents time regularity and is not affected by real-time events of the network. Human-driven traffic includes requests that are directly generated by humans or through control programs. Event-driven traffic includes service requests that are not generated by timers or humans, but as responses to certain programmed events. The classified network traffic, in this method, is then parsed for suspicious data packets and all such anomalies are flagged.

Although Zheng et. al's detector model is novel and effective, current rate of increase in data size of BACnet networks [17] warrant a need for an automated and unsupervised system. We planned to follow up on this research by using machine learning techniques instead of Fourier frequency analysis to capture normal data patterns and detect anomalous traffic. Previous research in the usage of machine learning to detect network anomalies is rather extensive. In this regard, we studied the work of Tonejc et. al [20] in characterizing BACnet network traffic data by means of unsupervised machine learning techniques. The model was trained from two different datasets. One of which is from two days' worth of network traffic from a BAS lab setup that contains about 20 different BACnet devices. The other was a dataset from artificially generated network traffic by obtaining network traffic from a different day and adding variations in the synthetically generated anomalous traffic. Clustering methods such as k-means are used to detect known attacks from the training data. Random Forests and Support Vector Machines (SVM) were used to detect new attacks.

Thus, our aim became to use the classification system proposed by Zheng et. al and the methods discussed by Tonejc et. al to create our anomaly detection system. This would fall under the category of semi-supervised [19] machine learning techniques, eliminating both the higher error probabilities of unsupervised methods [21] and the need for an enormous amount of data and time resources that characterize supervised learning methods [22].

Customer Needs Analysis (See Appendix for the survey questionnaire and results)

This section discusses the results of a survey with multiple foci- the need for security in unencrypted IP-based BAS, cyber-physical security features that customers believe they need/don't need, what information the audience believes could be obtained from a hacked BAS, how someone could possibly go about getting access to BAS controls, and finally, their opinions on the intrusion of privacy due to data collection for security purposes. The audience for the survey was multifaceted- including subject experts, Building Service operators, BACnet users (via User Forums), Building Automation companies, and the general public (potential users of BAS). From the survey, we wanted to get an idea of what attacks and devices we would need to model to emulate a real BACnet system, and what features our anomaly detection algorithm would need to have to appeal to a wider customer base. Additionally, we wanted to find out how feasible automated detection of cyber-physical threats was in the eyes of experts in the field.

The methods used in the project are surveys aimed towards two different audiences. One for the general public assuming no/partial knowledge of security on Building Automation Systems. The other (Technical survey) is for the representatives of companies, professors or any individuals who have knowledge about Building Automation Systems and BACnet. The survey forms sent to the companies were sent to individuals with prior knowledge of BACnet security and their input will be needed on how to improve BACnet security. We also reached out to people in online BACnet community forums, specifically The BACnet Institute and the BACnet Protocol Stack/Discussion on sourceforge.net.

From the responses of the general survey, we found out that the public realizes the importance of security in buildings and therefore supports the goal of our project to detect anomalies in network traffic. The majority of the public also believe that the buildings they go to school, work or live have building automation systems. This shows the versatility of BACnet protocol and we aim to implement this model

using 4 towers (raspberry pi stacks) which could represent university buildings, apartments, company buildings, hospitals etc. simply by varying the modelled sensors, services and traffic patterns. The responses from the general survey also showed that the public believe that BACnet devices need security. This shows that the public is aware of the danger in these devices being vulnerable to attackers and which in turn makes us develop BACnet devices to be secure and require authentication to access these devices. However, from the technical survey we received responses that they would rather not have too many authentication procedures as this could become too cumbersome, and to avoid collecting too much data. Thus, a middle ground solution would be to allow multiple levels of access with general users having only a single authentication upon connection and those with higher authority to access the control layer having to go through stricter authentication mechanisms. Therefore, any breach in the building can potentially be a breach of trust through one of the buildings' employees. We also wanted to know how the public would react to data being used for cyber security purposes. With data collection being a sensitive topic currently, almost 1/4th of the responses believe that it is a violation of their privacy. This result was useful is helping us decide the level of data collection we should undertake.

Regarding the veracity of the results themselves, it has to be mentioned that this survey would have been served much better by larger numbers. It can be assumed that there is uncertainty in data measurement in the case of objective questions. However, since we based our observations and analysis on the general trends rather than concrete numbers, the probability of us making erroneous inferences should be much lower. If the number of responses had been in the range of 500-1000, this survey could have been a tool to assess the state of BAS security and public awareness regarding the same.

CHAPTER II

BENCHMARKING

In the course of envisioning our final product, we discuss in this section the existing solutions pertaining to our project and highlight some criteria that can be used to evaluate the performance of our project with that of existing solutions. Many advancements have already been made in the field of anomaly detection and we aim to explore how well other solutions perform and how we can benefit from it.

Performance Criteria

Public Safety and Privacy Protection

This is an essential criterion for our project since the main idea behind designing a predictive anomaly detection algorithm is to protect users from theft and hacking in building systems. BACnet protocol lacks in implementing effective security measures because hackers can access any building automation systems through a public domain. Our goal is to alert the building system of a potential breach without human intervention. Our metrics of comparison will be based on the level of protection each of the products offers a BAN overall. A score of 1 implies that the method is ornamental at best and a score of 5 means that the product is enough to defend against any threat to the system, be it a cyber attack or a physical threat.

Responsiveness

Although anomaly detection methods might prove useful to secure BACnet systems, quick responsive approach to real-time attacks is valuable and of great challenge. The anomaly detection algorithm we are trying to implement has to train with multiple sets of data to recognize anomalous data patterns at a fast-enough rate. A delay of just a few minutes might prove fatal to cyber-physical systems, resulting in financial loss. A score of 1 in this metric indicates that the method is too slow to deal with most threats to the system, and the mechanism leaves the system wide open to external access. On the other hand, a score of 5 implies that the system is fast enough to deal with all possible threats and prevent any and all information leaks. Note that since we were unable to test each of the mentioned products first-hand, this criterion is only a rough measure of overall responsiveness based off of test results reported by the creators of the techniques.

Adaptability

Cyber threats are evolving at a very fast rate with many doing so in real time. Due to this, the capability of the security tool to adapt itself to the threats is very relevant in assessing the overall effectiveness of the method. A score of 1 in this implies that the product is fixed and cannot be updated to deal with future or evolving threats. A score of 5 means that the system can evolve with the threats perfectly, and updates very frequently. Similar to the above criterion, Adaptability cannot really be measured without exhaustive first-hand testing. However, this metric serves the objective of comparing the mentioned products loosely based on their descriptions.

Comprehensiveness

This criterion indicates the range of anomalies or threats detected by the technique as a function of the number of uncorrelated network parameters monitored by the detection mechanism. A score of 1 implies very poor range, with the tool only analysing a single network parameter. 5 indicates an all-encompassing range, with the tool analysing all possible parameters. Note that while the number of parameters measured may not accurately indicate the detection range of all devices, the majority of detection techniques can be successfully classified based on this criterion.

Measures for failure

This criterion is a measure of the product's capability to cope with system failures and to support the system in such situations. A score of 5 in this implies that the program has robust fallback and backup routines and can aid in system recovery. On the other hand, a score of 1 indicates the absence of any failsafe or recovery options in the program. A middling score in this regard indicates good failsafe options, but no recovery support.

Economics

The impact of the product on the economy is a major factor which we thought should be taken into consideration when designing our solution. The market for anomaly detection tools is not that lucrative for developers as of right now, mainly because of the low demand for such advanced means. However, security tools for BACnet are very much in demand due to the increased exposure of building networks to the internet. This way, we were able to rate the impact of each product on the market from 1 to 5, with 1 meaning it doesn't have much of one, while 5 means the product was able to dominate the market.

Cultural

Since our solution requires the periodic collection of large volumes of data (which may sometimes be personal), it was reasonable to expect that people from more private cultures could react adversely to the product. However, security tools in general do not have much of a cultural impact, and thus this criterion did not come into play during the comparison.

Global

BACnet is an international standard and is implemented in its varied forms worldwide. As a result, the impact of developing a security tool for BACnet on the global stage is quite high. Such methods if developed could reflect quite positively on the country as the research would be beneficial regardless of the place.

Environmental

As such, BACnet security does not have any direct impact on the environment. However, if the implementation of higher security measures leads to increased automation of buildings, and thus better energy efficiency across the board, this could lead to a net positive impact on the environment. Note that since this is purely speculation on our part, we did not use this as a criterion for comparison.

Product Comparison

Mentioned below are a few products like our proposed solution and how they compare:

- Password protection
- AC2000 BACnet Interface
- THE-driven anomaly detector by Zheng et al.
- Timing based detector for cyber-physical systems by Zimmer et al.
- Finite-state based detector for Modbus detectors by Goldberg et al.

Password Protection

The main method used in securing building networks is password protection, with either a full access or multi-level access with different permission levels. The issue with this method, according to engineers in the field, is that the control of the passwords is not strict during the setting up and maintenance of the system. Also, due to mismanagement, the passwords may easily get leaked. As most, if not all, BAN devices are made to be easily discoverable and accessible, the method of password protection is not all that strong against a targeted attack.

The systems engineers responsible for setting up building automation networks are often contracted to monitor the network for any anomalous status alerts, alarms or malfunctions. This is a task that is becoming extremely tedious as the size and scale of BANs are going up fast.

AC2000 interface

This interface from Tyco Security Products enables alarms to be sent in BACnet protocol to third party systems including building management systems, HVAC, fire and any other systems that support BACnet communication. The interface is bi-directional, allowing for both the sending of AC2000 alarms and the receipt of third-party Change of Value (COV) BACnet messages, which can then be displayed on the AC2000 Security Hub alarm and event management application. As a security tool, this interface is quite versatile and is able to detect many common threats to BANs. However, from the parameters of data it monitors, the tool would be unable to do much in the face of a Denial of Service attack or a Man-in-the-Middle attack. The former depends exactly on the bandwidth being occupied by COV messages or alarms to break into BACnet systems, while the latter would in all probability go unidentified.

THE-Driven Anomaly Detector

As mentioned in the literature review, this detector was a major inspiration for the project. In terms of performance, the testing information for this as reported by Zheng et al. put its accuracy above 96% across all the tested attacks. It also purportedly has a near 100% anomaly capture rate with the exception of DoS attacks, which can be identified via other network volume-based detection techniques.

Timing-based detector

This detector works on the principle of utilizing information from static timing analysis to identify unauthorized instructions in real-time cyber-physical environments. In the case of building automation systems, the timing bounds in code segments are said to be easily available, thereby facilitating this

particular method. The paper by Zimmer et al. [14] describes this mechanism which works by checking those bounds and implements the detection techniques either by itself in a self-checking manner, or through the operating system scheduler.

Finite-states based Detector

Goldenberg et al.[15] modelled a detector meant for the Modbus network protocol (and can be used in BACnet as well) based on identifying the unique deterministic finite automata "state" (DFA) of each individual channel between a HMI-PLC (Human-Machine Interface & Programmable Logic Controller) pair. The DFA is a finite-state machine that can accept or reject a given string of symbols and jumps from one state to another. It takes around 100 captured messages to identify the DFA for a given channel. This method purportedly is highly sensitive while having very low false positives. The paper mentions that Goldenberg et al. found zero false alarms over 111 hours of continuous operation.

Table 1 (below): General Comparison of each solution

1. General comparison:

	Password Protection	**AC2000 Interface**	**Timing- based detector**	**Finite-state detector**	**THE driven anomaly detector**	**Our Solution**
Technique used	Access protection using passwords; multi-level access with permissions assigned by the manager	Receive third-party Change-of-Value notifications and monitor all events via the security hub.	Statistical Attributes of data packets (Mean, Range)	Construct a deterministic finite automaton (DFA). Flag anomalies according to periodicity	Autocorrelation, THE Classification, and Fast Fourier Transform	THE-Driven traffic classification with semi-supervised ML techniques
Accuracy (%)	Not Applicable	Information Not Available	>99%	>99%	~96%	-Can't be measured-
Advantages	Easy to implement, most cost-effective; enough for non-critical systems	User-friendly patented application to configure required BACnet alarm or event outputs.	Works both by itself and also through the program scheduler	Very high accuracy and highly sensitive. Able to access the deeper network layers	Detector can be used online.	Online Traffic Classification for better understanding of traffic source Adaptive

Disadvantages	Need to be changed periodically, Easy to leak out to external parties, prone to being hacked	Not open to novel detection mechanisms.	Depends entirely on timing bounds.	Only two statistical parameters measure the anomalies.	Frequency analysis needs to be done manually; not accurate in more dynamic real life situations	Lack of data = less sophisticated algorithm
Limitations	Can be deciphered using brute force cracking technique. Does not implement user alert mechanism after multiple login attempts	Only limited to certain events or alarms that take place.	Only considers the timing bounds of a data packet. Data packets emerging from human driven activities will not be distinguished	Performance degrades for multi-period traffic patterns - slower traffic patterns increase false positive rate.	Impossible to have a 0% false positive rate. Classifies network traffic into only three different categories.	Impossible to have a 0% false positive rate Works well with more data fed into the system. Unable to introduce third party vendor detection.
Standards	Already patented therefore it does comply with BACnet standards defined by NIST	Follows their own AC2000 standards therefore can be considered credible	Complies with most of ASHRAE's standards from NIST	Complies with most of ASHRAE's standards from NIST	Complies with most of ASHRAE standards by training the model to recognize most of the standard attacks.	The testing model should be able to recognize most of ASHRAE's security assessments defined from the NIST

Table 2 (below): Comparison of each solution according to Performance Criteria

2. Performance comparison (All scores are ranked from 1-5, with 1 being the least and 5 being best):

	Password Protection	**AC2000 Interface**	**Timing- based detector**	**Finite-state detector**	**THE driven Anomaly Detector**	**Our Solution**
Responsiveness	1- Basic form of network protection and no automated capabilities to alert user of breach.	2- Only known threat patterns identified, needs to be manually updated	4- Real time response to threats is very good as this measure's differences in code time bounds	Speed varies according to periodicity of traffic patterns. Multiple traffic patterns mean algorithm slows down.	4- Low latency for detection of most traffic anomalies.	Speed of algorithm depends mainly on the amount of data the algorithm has analyzed.
Adaptability	1-Passwords can be made sophisticated but are only as good as the user's creativity	1- Not easily adaptable to new intrusion detections	3- Can adapt to multi-periodicity traffic patterns, but additional work needs to be done to achieve good performance. This includes testing on new data patterns.	4- Highly sensitive mode of detection means the algorithm can adapt to subtle changes in data packets.	4- Allows the detection of common types of attacks. Depends on labelled data.	5- Can adapt to new attacks due to unsupervised aspects and also better at identifying threats due to supervised learning methods
Comprehensiveness	1- None because this is a static method of protection.	2- Only known threat patterns identified, needs to be manually updated	3- Only applicable for time driven anomalies.	5- Highly sensitive to all changes in the interface-controller channel, analyzes deeply into the network	4-Trained to recognize different intrusion detection mechanisms through the BAS networks and synthetic generated attacks.	5- Should be able to detect any anomalous data packets due to comprehensive classification of normal traffic

Measures for failure	2- Alerts User regarding breach, ask for change password, lock system operations until authorized.	3- Alert management about successful attacks, infiltrators, and other network data red flags.	1-Difficult to check for failure as some breaches may occur undetected due to limited detection parameters	4- Any anomalies in the channel are found easily due to changes in the DFA of data packets	3 - Methods to alert user can be added to the mechanism. Built-in measures not available.	1- We have not planned any failsafe as of now

Table 3 (below): Comparison of Macro-Level parameters

3. Macro-level comparison:

	Password Protection	**AC2000 Interface**	**Timing-based detector**	**Finite-state detector**	**THE driven anomaly detector**	**Our Solution**
Economic	Does not cost anything and is available to all as part of the default settings.	A pretty competitive product on the market. Similar products are some of the most widely used notification software in current BMS.	Generally applicable across BACnet and SCADA networks. Due to its scalability and high efficiency, this method had a high impact on the building networks market.	Designed for Modbus devices, this is nevertheless compatible with general networks. Due to its more focused development, this may have a higher impact in that market	The simplicity of the algorithm and high efficiency, with its ability to be used across BACnet systema, it is anticipated that this method if released into the market could have a high impact.	Could potentially be in demand depending on the trade-off between performance and data collection volume/frequency.
Global Compatibility	Compatible for any systems.	Compatible with most BMS systems	Compatible with SCADA systems	Compatible with SCADA systems	Compatible with SCADA and BACnet systems	Compatible with SCADA and BACnet systems

CHAPTER III

FUNCTIONAL MODELLING

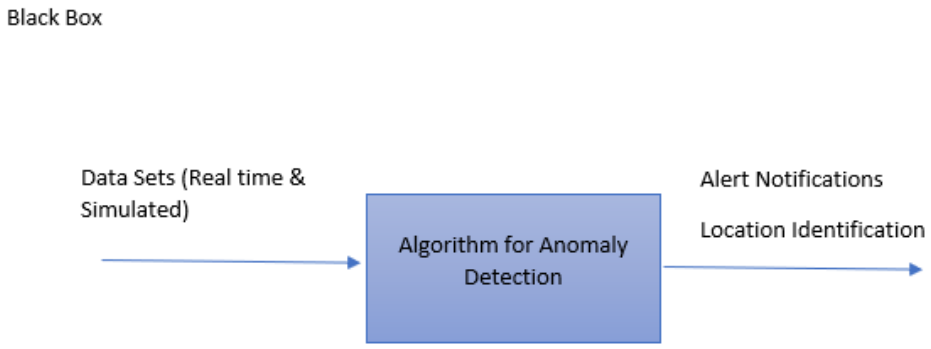


Figure 1: Black Box Design of Project

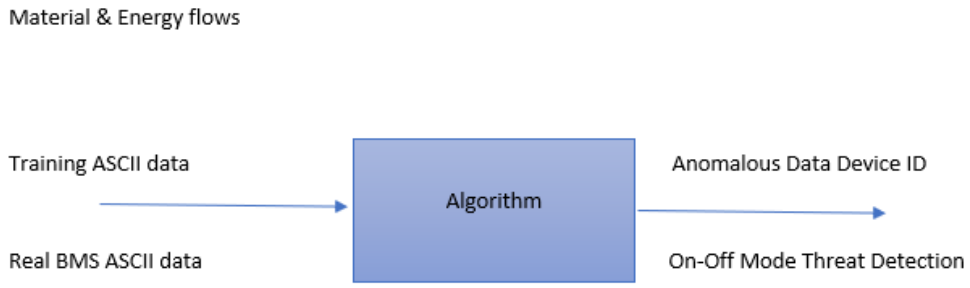
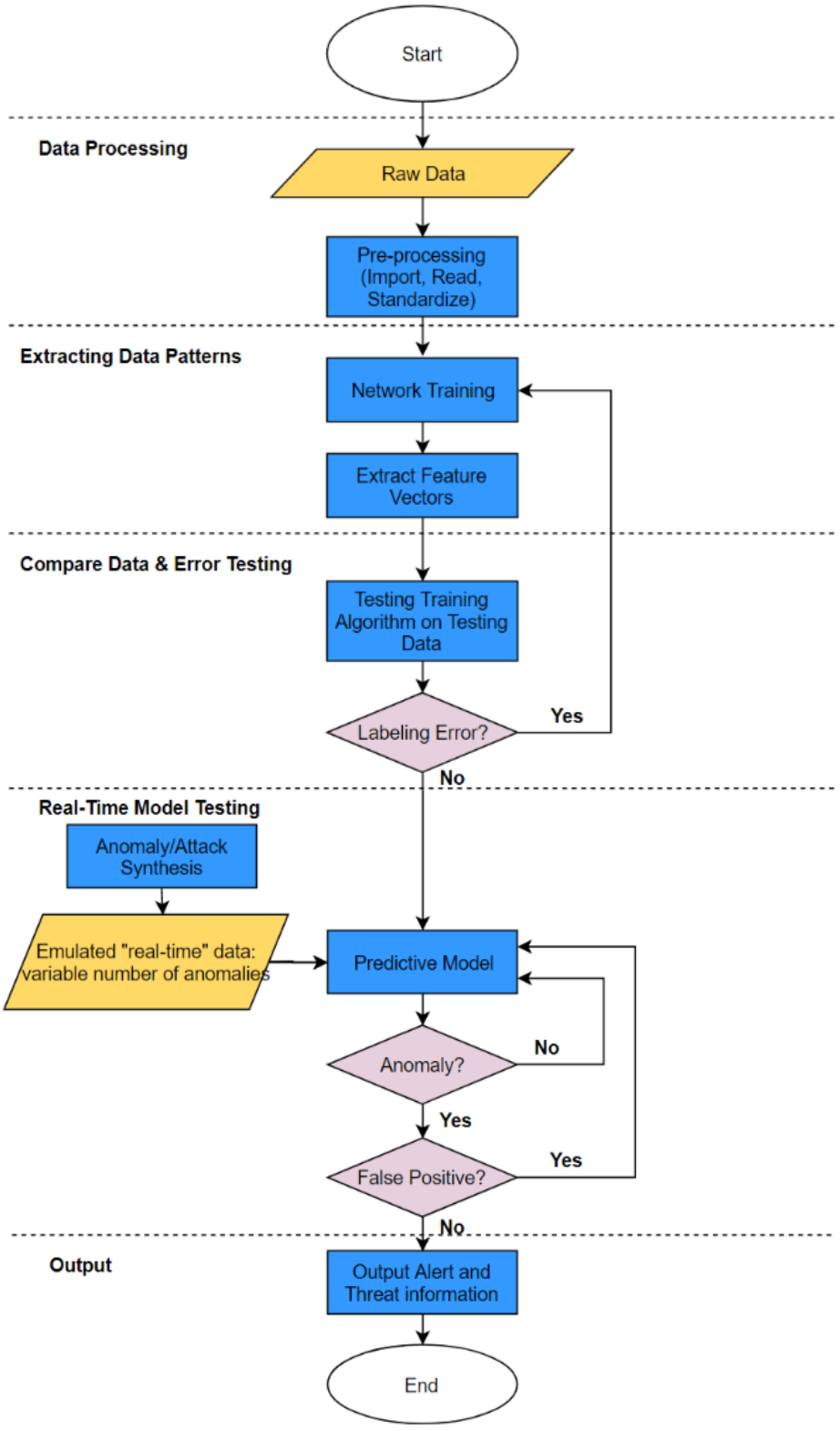


Figure 2: Detailed Information Flow in a Black Box

We developed a functional structure of our project design using a flowchart model. It was a great tool for us to convey our ideas through a visual depiction of the processes that go into designing our project. It allowed us to elucidate the intricate details of our project and aid in team productivity. For a complicated project that involves huge data collection and running algorithms, it would've been difficult for us to comprehend the number of components that needs to be implemented to make this algorithm run successfully. Using graphical data representations techniques like flow-charts and black-boxes, we can now use it as a reference point to review and understand the components throughout the duration of our senior design project. In addition, using flowcharts allowed us to consider troubleshooting issues that may occur during the project. This includes recognizing false positive results and how we can handle it by making decision boxes to determine flow of process. This flow leads to a possible solution and this assignment has allowed us to consider the possible solutions proactively, thereby allowing us to be productive in the future.

Although the type of ML technique is still under review, below is the basic structure of how any ML algorithm implementation would work.

Program Flow



Data Processing

This is the initial phase for our project. This step involves collecting the raw data from the relevant sources or higher authorities which is the first challenge we experienced. Usually for BACnet systems, there is a scarcity of data available due to building operators denying access to the BACnet traffic. This is mainly due to security purposes and so we had to think of ways to overcome this problem. Firstly, our mentor provided us BACnet traffic data that was obtained from QU. This data was labelled but was collected for a meagre timeframe of 2 hours. This timeframe is insufficient to discern any variations in the dataset. Nevertheless, enough data was collected (~100,000) for us to kick-start our network training algorithm. This dataset should be enough to act as a guiding anchor for us to generate a sensible dataset from Raspberry pi stacks. Currently this data is unlabelled and so we are yet to decide about the type of ML algorithm to be used.

An alternative method of data collection that we planned was to simulate our own BACnet dataset. This can be done by first observing normal patterns of traffic and then implementing Wireshark software. Wireshark collects simulated traffic data from Raspberry Pi microcontrollers that act as BACnet devices. In the future, we aim to approach a local company that specializes in BACnet related BMS systems and obtain real-time datasets. In case, data collection from the local company doesn't work, we will still be equipped with the simulated data from our Raspberry Pi stacks.
Once raw data is collected, we will pre-process or clean up the data for Deep Learning Networks to make sense of it. Pre-processing involves correcting raw data that is characterized by inconsistencies, lack of certain trends and outliers. Through parsing techniques, we can resolve many of the above-mentioned issues to make life easier for the Machine Learning algorithm. To start off parsing, we first import the dataset into a programming software such as Python and categorize the dataset into a manageable format for the algorithm. Sometimes, we may be missing values and, in this situation, we can choose to delete it provided that there is enough dataset. In some cases, we can calculate statistical values such as mean, mode or median to replace those missing values since this can be a good approximation of the missing dataset.

Extracting Data Patterns

The dataset that we obtained from QU was unlabelled network traffic. This means that the data doesn't possess any outcome variable such as classification of a data as 'normal' or 'anomaly' for instance. However, since this data was collected for only 2 hours, not much inferences can be made regarding anomalous results. With regards to this situation, we have decided to test out unsupervised and semi-supervised ML technique.

Unsupervised ML is when the algorithm accepts raw data as input and extracts patterns to make association between variables. Since we do not know what type of features a dataset from BACnet traffic might possess, we will use unsupervised techniques to learn about the data features. Once the data features are extracted, the algorithm is put under test conditions to label or classify the data correctly. The validity of the labelling is determined by pre-set conditions such as an accuracy threshold which categorizes a labelling as either correct or incorrect. If the accuracy level is low, then this suggests that the model is erroneous and did not find valuable patterns. In this case, we would have to manually label some dataset with outcome variable and feed it into the model, thereby facilitating semi-supervised ML technique.

Semi-supervised is a hybrid of supervised and unsupervised ML method where a small amount of dataset needs to be labelled and the rest can be fed as unlabelled data. The main ideology behind this type of training is that the algorithm has a head start in knowing the input and target variables to label the rest of the dataset based on what it learned from the labelled data.

Given the two types of ML techniques, we will determine which technique gives us a better accuracy rate or low false positive results to decide the most suitable ML tool for further training.

Compare Data and Testing

As already mentioned above, we will determine the accuracy of the model based on the level of labelling errors. Labelling error occurs due to insufficient dataset that causes the algorithm to be overfitting. Overfitting occurs when a function is made to closely fit within a limited set of points. So, we will ensure a decent amount of dataset is generated so that algorithm can learn more data features. If a labelling error arises, this data will be sent back to the model to retrain to prevent similar faulty behaviour in the future. After multiple trainings, we will observe if the model improves its labelling performance. If it doesn't, that is when we will consider switching to semi-supervised technique.

Real-Time Model Testing

The predictive algorithm refined by the preceding function blocks is tested by running "real-time" emulated data having both regular data and synthesized anomalies or attacks through the detection algorithm. As there are many kinds of anomalies, such as DoS attacks where an abnormality would show up in the volume of requests but not the requests themselves, we plan to implement an array of techniques to deal with many different anomalies. In general, however, the algorithm will compare the input "real-time" data's feature vectors with the ones identified previously as "normal" data patterns and alert the user if data segments show a significant variation from the norm. Those data segments are then read to identify the point of entry of the anomaly (sender's IP address) and the type of anomaly (if known). The information will be the output of the system.

Output

The response we aim to get will be the location of the intrusion and type of attack. The location of the intrusion will be determined by our raspberry pi's since each raspberry pi device is a unique BACnet device and therefore will have unique BACnet IP address. We aim to include an LED in each BACnet device to show the location of the intrusion when the prototype is built. The type of attack to be determined will be compared with the existing attacks that we have trained our deep learning model with so that an immediate response will be taken.

CHAPTER IV

SYSTEM DESIGN

The algorithm we develop should be able to capture any potential breaches into the Building Managements Systems and alert the building managers of a breach. Also, as discussed in the introduction, the algorithm will use Machine Learning to detect new anomalies and extract features that cannot be easily determined by the human eye. However, because of the lack of BACnet data from IT entities due to security reasons and our inability to test cyber-attacks on the actual BACnet systems when it is operational in buildings, we first needed to create an emulation of a real BACnet system. Only then could we move on to the training and testing of our model.

Data Generation

The first step of our project and one of the main deliverables is the setup to emulate a real building network using the BACnet protocol and the subsequent generation of data. This could be shared on a public domain to allow other researchers to improve upon BACnet security without the need for months of grovelling for permission to use building data or spend resources in setting up an actual building network.

The emulator will mimic 4 towers sending BACnet messages to the other towers using BACnet/IP protocol. The towers are Virtual Local Area Networks (VLANs) that will be connected to each other using an inter-VLAN router (CISCO Catalyst 3550). Each tower will have 8 floors (8 raspberry pi's as 1 stack) with each raspberry pi acting a BACnet device having a unique ID as shown below.

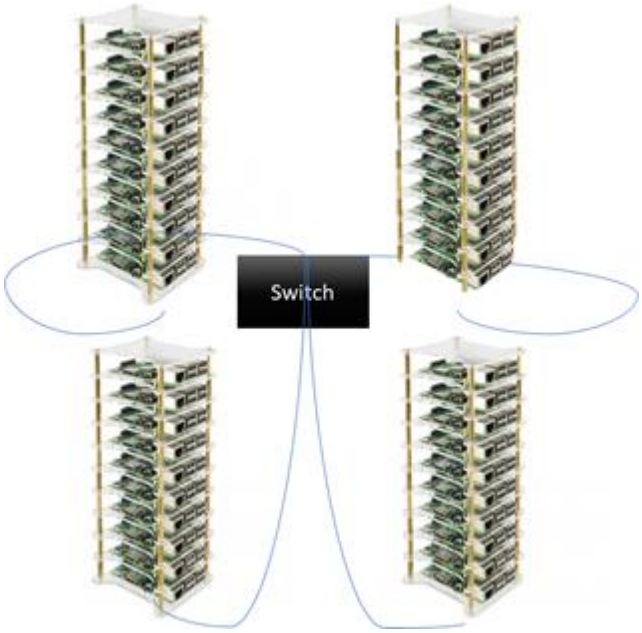


Figure 3: BACnet Emulator

NOTE: International standards require the ethernet connectivity to follow the standard IEEE 802.3 [2], which is assumed to be the case as we are using standards-compliant hardware at every stage of our project.

The emulated BACnet devices will use the BACnet standard as described by the ASHRAE 135 and all the relevant amendments [1]. For example, [10] informs when the server should send a COV (Change-of-Value) notification in order to reduce the burden on the server.

The different scenarios and vulnerabilities such as Energy-demand shock, building driven to extreme temperatures, HVAC (Heating Ventilation and Air Conditioning) failure will be mimicked for detection implementations. However, fine-tuning an IDS (Intrusion Detection System) is a manual process and therefore, machine learning algorithms can be applied for data packet classification and profiling [9]. Machine Learning will use datasets from acquired databases to identify and label the data packets that will be pre-processed to parse network packets into a CSV file.

Data Labelling & Network Training

We are still deciding whether to pursue unsupervised or semi-supervised machine learning for our project. Unsupervised machine learning will allow us to find unknown patterns in data that cannot be detected by human observation. However, with this learning technique we cannot predict the outcomes nor determine how accurate they could be [3]. The unsupervised machine learning technique that we plan to use in this case is an autoencoder neural network. This neural network has three layers: an input layer, hidden layer and a decoding layer. The autoencoder is trained to copy its input to its output [3]. An autoencoder generates the representation of data as close to its input. We plan to use the Undercomplete Autoencoder because it collects the most important features that are present in the data [4]. However, if there is a lack of enough training data can create overfitting [4]. For this project the goal is not to create an autoencoder algorithm as it is another project on its own. Therefore, we will use an existing algorithm and finely tune to suit our needs of modelling an unsupervised neural network. Existing models are available in GitHub repositories such as by Abel G. [8]

Therefore, if the unsupervised machine learning technique extracts invalid features, captures the wrong features in BACnet data packets or generates an overfitting model then we will use semi-supervised learning. Semi-Supervised learning is the hybridization of Supervised and Unsupervised learning techniques. Supervised learning makes use of labelled training dataset and trains to map the input to the output using specialized computing techniques. The problem with supervised learning is that we need to manually label data. We will train the model using the THE-driven classification model by Zheng et. al [5]. THE stands for Time, Human and Event -driven traffic. Network traffic on BACnet systems either fall in one of the categories mentioned above or in a combination of two or all of them:

Time-Driven

This traffic is generated by scheduled services, and as such is not supposed to be affected by real-time network events. Classifying this data type separately allows us to examine if there is a delay in the network packets and extract timestamps (the time at which BAC device delivers the data). Delays generally occurs due to some data tampering from the BACnet devices.

Human-Driven

This type of network traffic is generated by humans and is often non-periodic. It can happen at any stage and it only constitutes 5% of the total BACnet traffic [6]. Human-driven network traffic typically involves change of value on objects; in other words, a malicious user tries to alter the data in physical devices such as temperature or light intensity. A service request called WP (Write Property) allows data to be changed and this must be immediately identified to prevent malfunction of devices.

Event-driven

This category consists of network traffic generated due to change-of-state events, such as alarms, device feedback upon change in value (ON/OFF/set) and system status alerts.

After training has been completed, a new set of unlabelled data is fed, and the algorithm is made to discern patterns and produce a labelling accordingly. If it produces error labelling, then this data is fed back into the algorithm to learn more about the dataset. This process continues until we achieve a suitable degree of accuracy, close to 95% (rough estimate).

Anomaly Detection

Once the features are extracted and the model is trained. The trained neural network algorithm will be used to detect anomalies and flag data breaches. There is an important issue of false-positive results and this can impact the reliability of our algorithm. For this process, we plan to measure the false-positive rate by using a confusion matrix shown in figure x below. This technique is composed of a table that describes the performance of the classification model and gives a visual representation of the performance of the algorithm. This technique will be valuable in making us aware of the degree of false-positiveness and fine-tune our program as required.

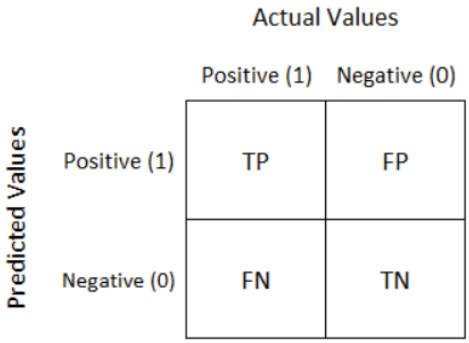


Figure 4: Confusion Matrix [7]

The accuracy will be calculated using the formula in figure xx below

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Figure 5: Accuracy Calculation Formula from Confusion Matrix [7]

Testing

The BACnet system designed for data generation will be subjected to synthetic attacks to test the resilience and response of the anomaly detector. The attacks will encompass the main threats to BACnet systems, such as DDoS attacks, Data tampering, and Data-exfiltration.

The types of attacks that a BACnet system could face can be found by the National Institute of Standards and Technology [11]. For example, snooping attacks can be detected by intruders using the 'Read Property' service to obtain knowledge of the device and the network. Other types of attacks include Application of Service attacks such as using the 'Write Property' service to change any property of any BACnet objects and disrupt the Building Management System. We will carry out similar attacks on the system, thereby testing the robustness of our algorithm if it possesses the capability to detect deviations in the network traffic. There are many other types of attacks that can be generated that extends to time-based attack, frequency-based attack due to human input. We will also closely study the dataset and learn to inject various attacks to make sure our algorithm detects the most basic threat information.

Discussion

The final anomaly detection program should be capable of analysing network data packets and optimizing itself where it is implemented. The program should then be able to identify any anomalies in the analysed data with as low of a false positive rate as is possible. This product is limited by our low level of expertise in software development and the use of generic machine learning tools. Also, the false positive rate is unavoidable to a certain extent as the network data is inherently random. However, it must be mentioned that the use of machine learning has been aimed at minimizing the known errors and figure out as many other sources of error as we can through trial and error over a very high number of iterations. If possible, we are also looking at adding a self-updating feature to the program, the difficulty of which makes it something that will only be possible if we are well ahead of schedule. Compared to other products on the market though, our program should be more flexible and adaptive upon implementation.

The biggest constraint we anticipate for this project is the lack of BACnet data. Mainly, this would impede our training and testing process and could lead to overfitting if the issue persisted. To overcome this problem, we built an emulator as mentioned above in the Data Generation section. Another constraint we will face is the lack of expertise and time to implement a machine learning algorithm from scratch and so, we will resort to using existing solution available in GitHub repositories. The other major constraint we are facing is the method of Machine Learning to implement on the BACnet dataset. There are two subsets in data mining field that includes Machine Learning and Deep Learning. Machine Learning methodology mostly requires labelled dataset and needs to be spoon-fed with answers to predict outcomes on unknown data. This method has high accuracy and needs less oversight when looking at performance of the model. However, the scope of Machine Learning is limited to the information that is fed into algorithm and uses that as setpoint to predict outcomes. On the other hand, Deep Learning usually involves an algorithm containing a plethora of neuron-like structures, where untagged or unintelligible data can be fed. The hierarchical set of neurons will each learn a different feature of the dataset, thereby extracting useful patterns for building a detailed analysis of the data. Deep Learning is not limited to any dataset since it is flexible to any type of data. However, care needs to be taken when predicting labels since Deep Learning can also learn wrong patterns and so human intervention at the beginning is recommended.

This project uses application of knowledge from ECEN 210 (Computation and Algorithms) and ECEN 303 (Probability) in the development of the machine learning part of the project. ECEN 210 is used extensively because of the Neural Network implementation that requires systematic programming experience in C-language to fine-tune the model. The probability course is important in order to understand the intense math since almost all Deep Learning or Machine Learning algorithms require prerequisites in understanding mathematical notations and concepts such as Bayes theorem, expectation, variance, correlation, covariance etc. Training models are tuned with a probabilistic framework [13]. Another course that is essential is linear algebra because we will be dealing with BACnet data packets as matrices when importing csv files into the python interpreter. Linear Algebra is very vital in machine learning algorithms because they massively rely on matrix manipulation with transposing, multiplication, inverse matrices, singular and non-singular matrices should all be considered [12].

CHAPTER V

CONCLUSION

In this End-of-Semester Report, we have described all the analyses that resulted in our current system design. The literature review and subsequent market analysis provided many inspirations regarding the final product idea, and the customer needs survey helped us understand the requirements of the market from a consumer/user's perspective. Towards the end, the functional modeling assignment helped us tie up everything into an implementable plan, the one we are currently following.

Project Timeline

So far, we have completed two major milestones in our project. One is the visualization of the final product and the system design, and the other is the setting-up of one Raspberry Pi stack to test and fine-tune data generation. The remaining work that we aim to accomplish is shown in the table below:

	December				January					February				March		April
Generation of BACnet Traffic Dataset																
Finalize Machine Learning Techniques																
Pre-Processing of Dataset																
Extracting Data Features + Labeling unknown dataset																
Error Correction + BACnet device implementation for more Raspberry Pi stacks																
Error Correction + App Development/IoT notification																
Education on Machine Learning																
Final Presentation																
Final Report																

Next Semester Plans

The next semester, the team will focus on developing and implementing the software (machine learning) component of the project. Below is the general timeline of how we will progress through the next semester.

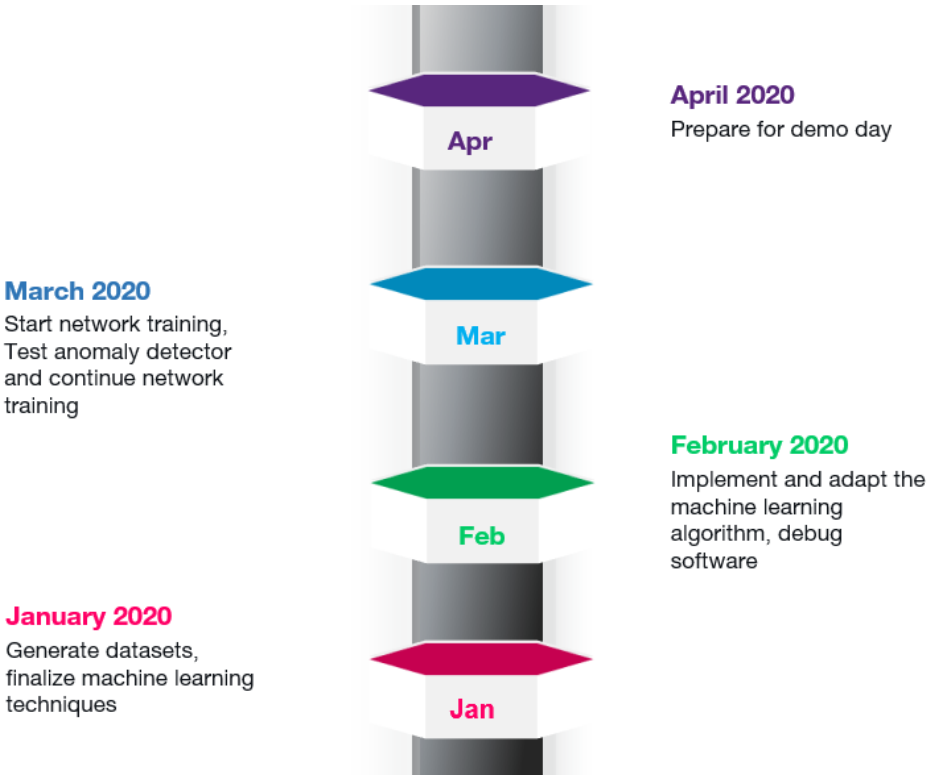


Figure 6: General Timeline of Next Semester

Comments

This project, according to an expert we interviewed for an ethnographic study, has a lot of potential now that the research into BACnet security is a hot topic. The prospect of being able to do something new and contribute something worthwhile to the field of research was, as it is now, very enticing to us three undergraduates. As such, we put in some effort into learning all the prerequisite skills to implement the machine learning and network manipulating aspects of the project. However, visible progress has been slow. With the semester coming to an end, we realized that the ability to learn and research on new topics is vital to the development of our project. We have learnt new skills from editing videos, researching new topics, distributing surveys and writing reports. The skills learned in this course can be described as invaluable that will be essential in our future development. The challenge of understanding machine learning algorithms, neural networks, extracting features and labels for the BACnet system will also be useful knowledge as we move towards an era of Artificial Intelligence.

REFERENCES

[1] Ashrae.org. (2019). *Standard 135-2016, BACnet™ -- A Data Communication Protocol for Building Automation and Control Networks*. [online] Available at: https://www.ashrae.org/technical-resources/standards-and-guidelines/standards-addenda/standard-135-2016-bacnet-a-data-communication-protocol-for-building-automation-and-control-networks [Accessed 6 Dec. 2019].

[2] Standards.ieee.org. (2019). *IEEE 802.3-2018 - IEEE Standard for Ethernet*. [online] Available at: https://standards.ieee.org/standard/802_3-2018.html [Accessed 6 Dec. 2019].

[3] Chandrayan, P. (2019). *Deep Learning: Autoencoders Fundamentals and types*. [online] Medium. Available at: https://codeburst.io/deep-learning-types-and-autoencoders-a40ee6754663 [Accessed 6 Dec. 2019].

[4] OpenGenus IQ: Learn Computer Science. (2019). *Different types of Autoencoders*. [online] Available at: https://iq.opengenus.org/types-of-autoencoder/ [Accessed 6 Dec. 2019].

[5] Z. Zheng and A. Reddy, "Safeguarding Building Automation Networks: THE-Driven Anomaly Detector Based on Traffic Analysis," 2017. [Online]. Available: http://cesg.tamu.edu/wp-content/uploads/2018/01/ICCCN-Zhiyuan-Zheng-Paper-2017-July-1.pdf. [Accessed: 08- Sep- 2019].

[6] Cimetrics. (2019). *uBACstac - BACnet Protocol stack for small devices*. [online] Available at: https://www.cimetrics.com/products/products-bacnet-ubacstac [Accessed 8 Sep. 2019].

[7] Ferreira, H. (2019). *Confusion matrix and other metrics in machine learning*. [online] Medium. Available at: https://medium.com/hugo-ferreiras-blog/confusion-matrix-and-other-metrics-in-machine-learning-894688cb1c0a [Accessed 6 Dec. 2019].

[8] G, A. (2019). *abelusha/AutoEncoders-for-Anomaly-Detection*. [online] GitHub. Available at: https://github.com/abelusha/AutoEncoders-for-Anomaly-Detection [Accessed 6 Dec. 2019].

[9]"Hands-on Machine Learning on Google Cloud Platform", *Subscription.packtpub.com*, 2019. [Online]. Available:https://subscription.packtpub.com/book/big_data_and_business_intelligence/9781788393485/6/ch06lvl1sec39/supervised-and-unsupervised-machine-learning . [Accessed: 08- Sep- 2019].

[10] Dms.hvacpartners.com. (2019). *Bacnet Basics User's Guide*. [online] Available at: https://dms.hvacpartners.com/docs/1000/Public/04/11-808-417-01.pdf [Accessed 3 Sep. 2019].

[11] Bushby, S. (2003). *BACnet(trademark): A Standard Communication Infrastructure for Intelligent Buildings.*. [online] NIST. Available at: https://www.nist.gov/publications/bacnettrademark-standard-communication-infrastructure-intelligent-buildings [Accessed 6 Dec. 2019].

[12] Donges, N. (2019). *Basic Linear Algebra for Deep Learning*. [online] Medium. Available at: https://towardsdatascience.com/linear-algebra-for-deep-learning-f21d7e7d7f23 [Accessed 6 Dec. 2019].

[13] Brownlee, J. (2019). *5 Reasons to Learn Probability for Machine Learning*. [online] Machine Learning Mastery. Available at: https://machinelearningmastery.com/why-learn-probability-for-machine-learning/ [Accessed 6 Dec. 2019].

[14] C. Zimmer et. al, "Time-based intrusion detection in cyber-physical systems," in *ICCPS'10 Proceedings of the 1st ACM/IEEE International Conference on Cyber-Physical Systems, Stockholm, Sweden, April 13-15, 2010*, ACM, New York, USA, 2010, pp. 109-118.

[15] N. Goldberg et. al, "Accurate modeling of Modbus/TCP for intrusion detection in SCADA systems," *International Journal of Critical Infrastructure Protection*, vol. 6, issue 2, June 2013, pp. 63-75. [online]. Available at: https://www.sciencedirect.com/science/article/pii/S1874548213000243. [Accessed 8 Sep 2019].

[16] Z. Zheng and A. Reddy, "Safeguarding Building Automation Networks: THE-Driven Anomaly Detector Based on Traffic Analysis," 2017. [Online]. Available: http://cesg.tamu.edu/wp-content/uploads/2018/01/ICCCN-Zhiyuan-Zheng-Paper-2017-July-1.pdf. [Accessed: 08- Sep- 2019].

[17] S. Lyons, "Current Status of Cyber Security in the BAS Industry," *Cimetrics February 2019 Newsletter*, Mar 1, 2019. [online]. Available at: https://www.cimetrics.com. [Accessed 8 Sep 2019].

[18] **Metadata, https://censys.io/ipv4/metadata?q=bacnet&**

[19] D. Bhattacharyya and J. Kalita, "Introduction," Network Anomaly Detection: A Machine Learning Perspective, New York: CRC Press, 2014, pp. 1-13.

[20] J. Tonejc et al, "Machine Learning Methods for Anomaly Detection in BACnet Networks," *Journal of Universal Computer Science*, vol. 22, no. 9 (2016), 1203-1224. [online] Available at: https://pdfs.semanticscholar.org/d823/6a08011ad5f33e5d5c8f20d87c85a08bf784.pdf [Accessed 8 Sep 2019]

[21] P. Liang and D. Klein, "Analyzing the Errors of Unsupervised Learning," CS Division, EECS Department, Univ. of Cali., Berkeley, CA 94720, USA, June 2008. [online] Available at: https://pdfs.semanticscholar.org/038a/fe82cc61215e9087e572d8aab9663c1bdb0f.pdf. [Accessed 8 Sep 2019]

[22] S. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," *Informatica* 31, 2007, pp. 249-268. [Online]. Available at: http://www.informatica.si/index.php/informatica/article/viewFile/148/140. [Accessed 8 Sep 2019]

APPENDIX

BACnet Survey Form

Building Automation Systems (BAS) are softwares that can connect and automate relevant devices/sensors like a/c vents, electronic locks, lighting, supply lines and other utilities in a building. Of these systems, the BACnet (Building Automation Control network) is a widely used standard protocol for BAS systems. BACnet is preferred because it employs basic encryption methods as compared to most others which are completely in clear text, but it is still very vulnerable to cyber and physical attacks.

Our project is to identify/detect hackers who have gained access to the Building Automation Systems by implementing data collection and machine learning on the BACnet protocol.

Please help us understand the general opinion on BAS security by taking a few minutes of your time to complete this survey. Thank you.

* Required

1. What is your profession? *

_____

2. Does any building you regularly use have a Building Automation System?

Mark only one oval.

- ◯ Yes, one of them does
- ◯ Yes, more than one of them does
- ◯ No
- ◯ Maybe

3. Why do you think BACnet systems are important?

Check all that apply.

- ☐ To allow for interoperable (compatible communication without restrictions) communications between devices
- ☐ To detect hackers in a Building Management System
- ☐ To improve the automation capabilities of Building Management Systems

4. Do you need security in BACnet devices?

Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Other: _____

5. Do you think authentication is required every time you connect to a Building Automation System (BAS) device?

Mark only one oval.

- ◯ Yes
- ◯ No

1/2

6. Are automated detection of threats on BACnet systems a feasible option?
Mark only one oval.

- ◯ Yes
- ◯ No
- ◯ Other: _____

7. What information would you target if you get access to a BACnet device?

_____

8. What is your opinion on data collection for cyber-security measures in BAS?
Mark only one oval.

- ◯ It is a violation of my privacy
- ◯ It is necessary, but I don't like it
- ◯ It is necessary and I don't really mind it
- ◯ It is unnecessary
- ◯ Other: _____

BACnet survey

Building Automation Systems (BAS) are softwares that can connect and automate relevant devices like a/c vents, electronic locks, lighting, supply lines and other utilities in a building. Of these systems, the BACnet (Building Automation Control network) is a widely used standard which provides a wide platform for network integration and supports a wide selection of devices. It is preferred because it uses some basic encryption methods as compared to most other BAS which do not, but the BACnet is still very vulnerable to cyber and physical attacks. This is especially the case as the number of BAS networks is increasing these days.

Our project is aimed at implementing data collection and machine learning to teach a software how to detect attacks or threats on a BACnet system.

Please help us understand the general opinion on BAS security by taking a few minutes of your time to complete this survey. Thank you.

1. What is your profession?

_____

2. Do you think authentication is required every time you connect to a Building Automation System (BAS) device?

Mark only one oval.

- ◯ Yes
- ◯ No

3. What information do you think can be obtained from hacking into Building Automation Systems? Check all that apply:

Check all that apply.

- ☐ Credit card/Debit card information
- ☐ Personal information about employees
- ☐ Individual/Team's working hours or schedule
- ☐ Information of critical services like gas line monitoring and air conditioning
- ☐ IP addresses of BAS devices
- ☐ Other: _____

4. What sensors do you think a BACnet connected building includes? Check all that apply:
Check all that apply.

- [ ] Motion Sensors
- [ ] Temperature sensors (control over the temperature of a room instead of one central Air Conditioner)
- [ ] Pressure sensors
- [ ] Fine dust sensors/Air quality sensors
- [ ] Water tank level and quality sensors
- [ ] Energy consumption monitors
- [ ] Other: _____

5. Which of these attacks do you think a BACnet is most likely to face?
Mark only one oval.

- ◯ Physical attacks on BAS devices
- ◯ Man in the middle (data interception and impersonation) attacks
- ◯ Denial of service (brute force spamming of signal traffic) attacks
- ◯ Password hacking
- ◯ Other: _____

6. What security features do you have in your BAS(if you use one)? Check all that apply:
Check all that apply.

- [ ] Automated regular data collection
- [ ] password protection to control access
- [ ] Data Encryption and distribution of trusted keys
- [ ] Hired security guards to monitor physical activity around BAS devices
- [ ] Access authentication using challenge-response mechanism
- [ ] Not applicable. I do not use/have a BAS.

7. Of the above, which features would you like to have(if you don't already)? Check all that apply:
Check all that apply.

- [ ] Automated regular data collection
- [ ] password protection to control access
- [ ] Data Encryption and distribution of trusted keys
- [ ] Hired security guards to monitor physical activity around BAS devices
- [ ] Access authentication using challenge-response mechanism
- [ ] Other: _____

8. **What would you like Building Automation Systems to not include?**

9. **Are automated detection of threats on BACnet systems a feasible option?**
Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Strongly agree
- Agree

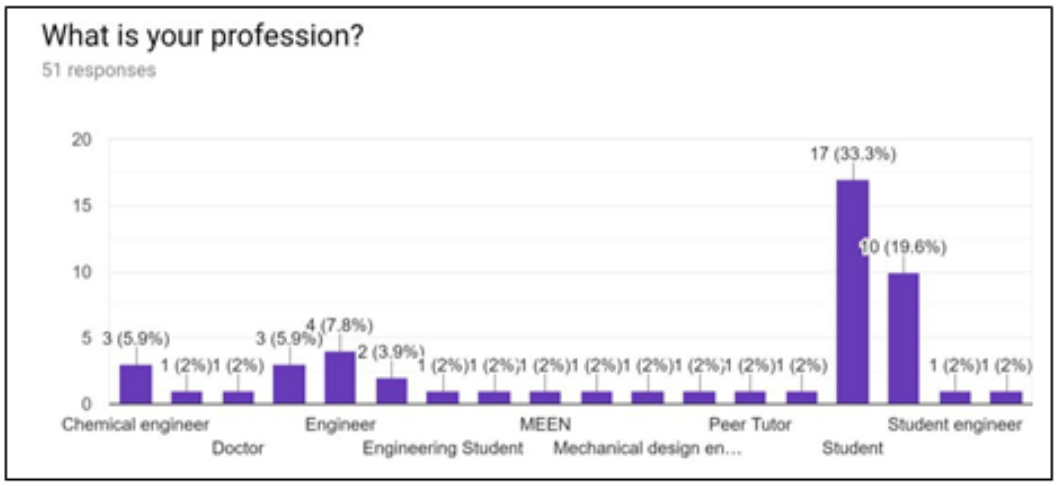
10. **What is your opinion on data collection to prevent cyber-physical attacks on BAS?**
Mark only one oval.

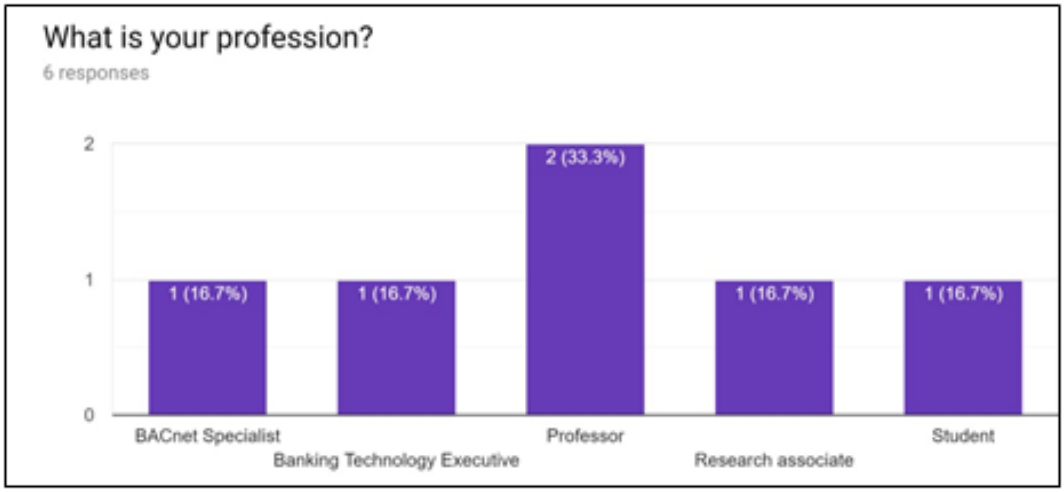
- It is a violation of my privacy
- It is necessary for security
- It is necessary but I don't like it
- It is necessary, but I would like to have minimal amounts of data collection in people's working spaces

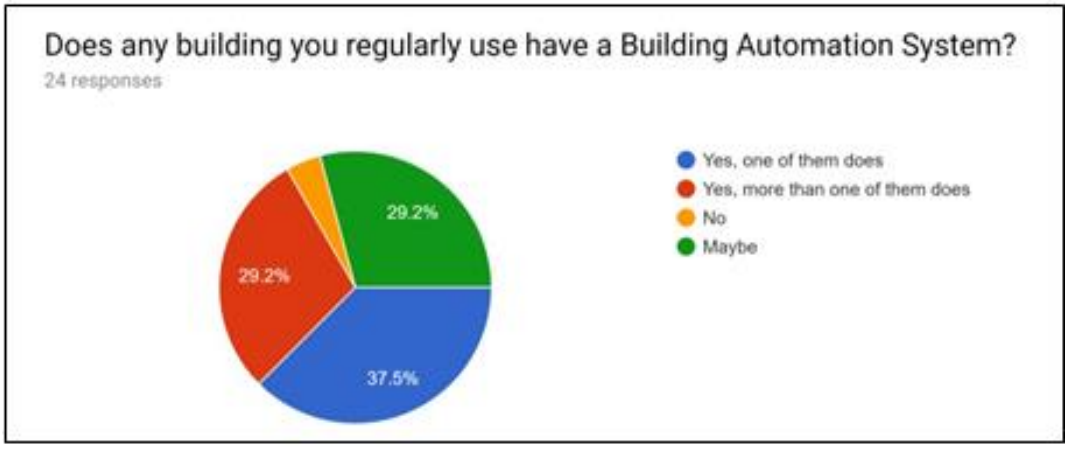
11. **On a scale of 1 to 10, how severe is your need for security measures in your building automation system?**
Mark only one oval.

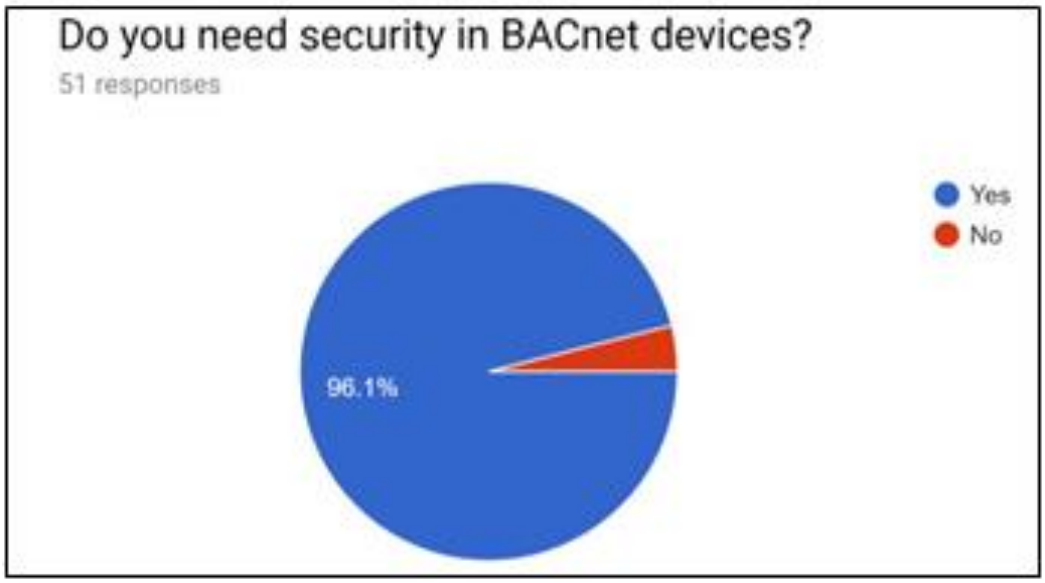
	1	2	3	4	5	6	7	8	9	10	
Not really necessary											We need it right now

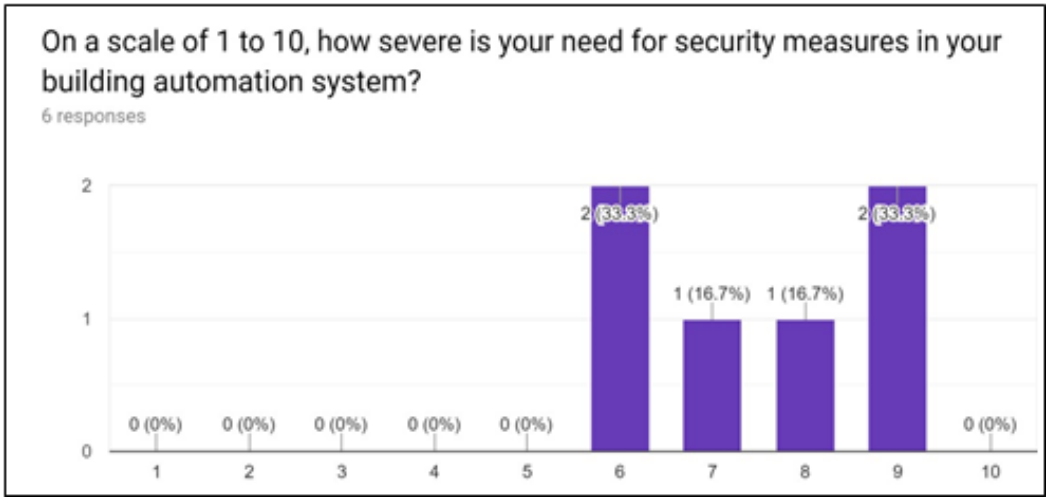
12. **Thanks for giving this survey your time. We appreciate your opinions and would like to hear your thoughts on this survey and/or on BACnet (BAS) security in general. Please do tell us here:**

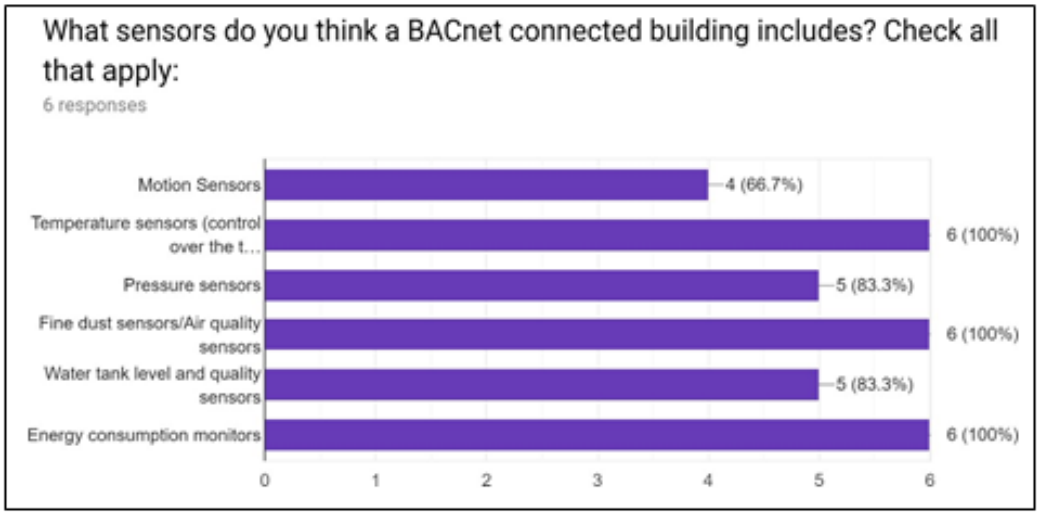

What is your profession?
51 responses
20
15
10
5
0
3 (5.9%)
1 (2%)
1 (2%)
3 (5.9%)
4 (7.8%)
2 (3.9%)
1 (2%)
1 (2%)
1 (2%)
1 (2%)
1 (2%)
1 (2%)
1 (2%)
1 (2%)
17 (33.3%)
10 (19.6%)
1 (2%)
1 (2%)
Chemical engineer
Doctor
Engineer
Engineering Student
MEEN
Mechanical design en…
Peer Tutor
Student
Student engineer

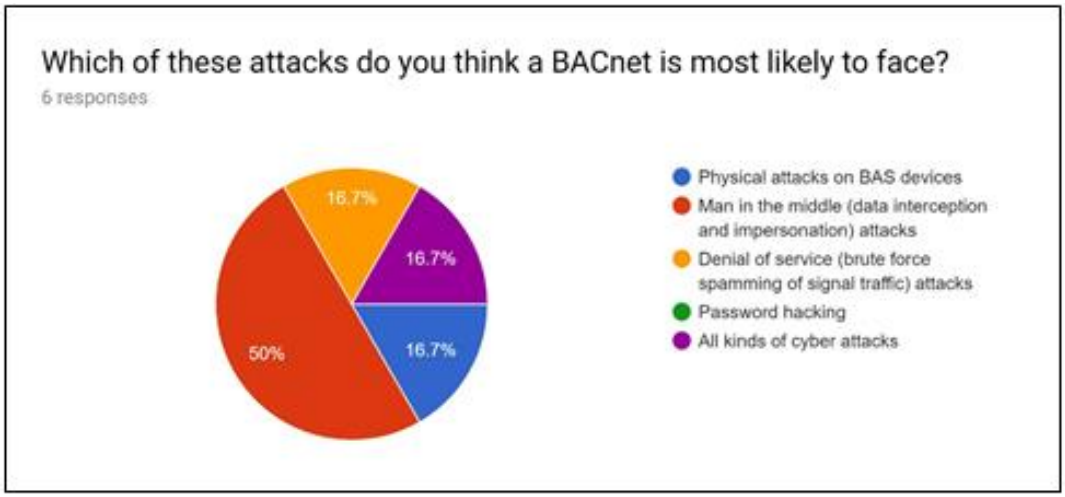

What is your profession?
6 responses
2
1
0
1 (16.7%)
1 (16.7%)
2 (33.3%)
1 (16.7%)
1 (16.7%)
BACnet Specialist
Banking Technology Executive
Professor
Research associate
Student

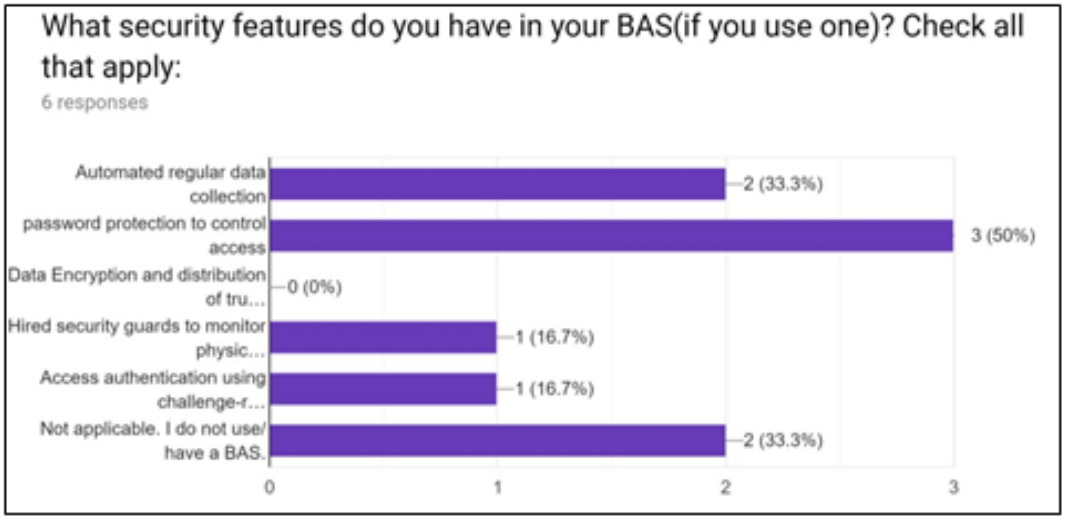

Does any building you regularly use have a Building Automation System?
24 responses
29.2%
29.2%
37.5%
Yes, one of them does
Yes, more than one of them does
No
Maybe

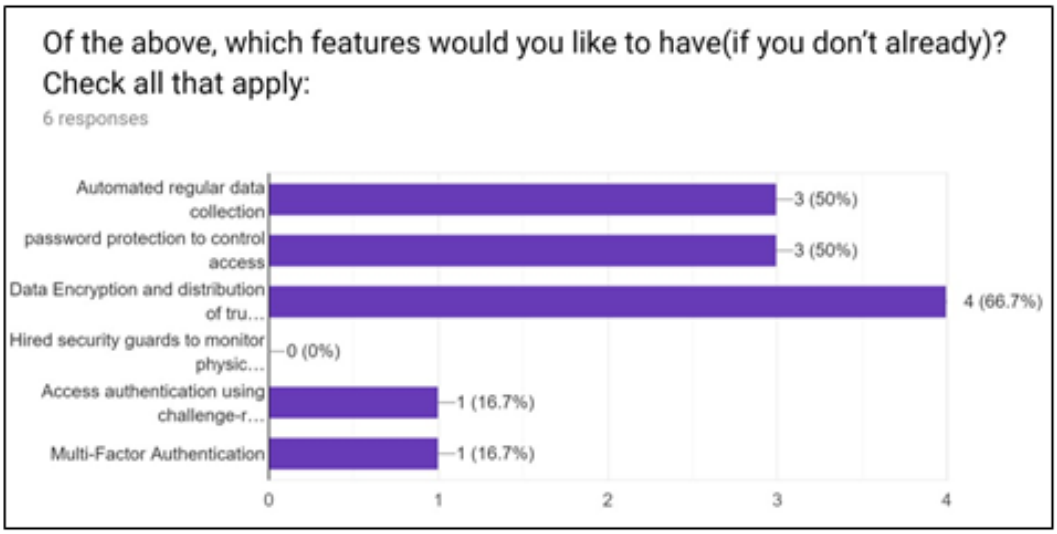

Do you need security in BACnet devices?
51 responses
Yes
No
96.1%


On a scale of 1 to 10, how severe is your need for security measures in your building automation system?
6 responses
2
1
0
0 (0%)
0 (0%)
0 (0%)
0 (0%)
0 (0%)
2 (33.3%)
1 (16.7%)
1 (16.7%)
2 (33.3%)
0 (0%)
1
2
3
4
5
6
7
8
9
10


What sensors do you think a BACnet connected building includes? Check all that apply:
6 responses
Motion Sensors
4 (66.7%)
Temperature sensors (control over the t...
6 (100%)
Pressure sensors
5 (83.3%)
Fine dust sensors/Air quality sensors
6 (100%)
Water tank level and quality sensors
5 (83.3%)
Energy consumption monitors
6 (100%)
0
1
2
3
4
5
6

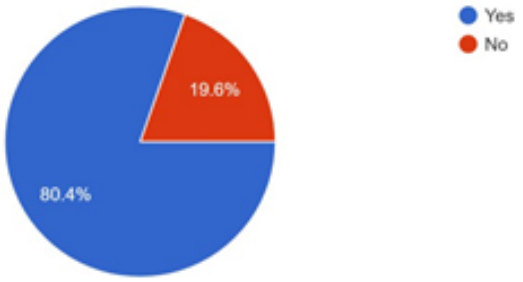

Which of these attacks do you think a BACnet is most likely to face?
6 responses
16.7%
16.7%
50%
16.7%
Physical attacks on BAS devices
Man in the middle (data interception and impersonation) attacks
Denial of service (brute force spamming of signal traffic) attacks
Password hacking
All kinds of cyber attacks


What security features do you have in your BAS(if you use one)? Check all that apply:
6 responses
Automated regular data collection
2 (33.3%)
password protection to control access
3 (50%)
Data Encryption and distribution of tru...
0 (0%)
Hired security guards to monitor physic...
1 (16.7%)
Access authentication using challenge-r...
1 (16.7%)
Not applicable. I do not use/ have a BAS.
2 (33.3%)
0
1
2
3


Of the above, which features would you like to have(if you don't already)? Check all that apply:
6 responses
Automated regular data collection
3 (50%)
password protection to control access
3 (50%)
Data Encryption and distribution of tru...
4 (66.7%)
Hired security guards to monitor physic...
0 (0%)
Access authentication using challenge-r...
1 (16.7%)
Multi-Factor Authentication
1 (16.7%)
0
1
2
3
4

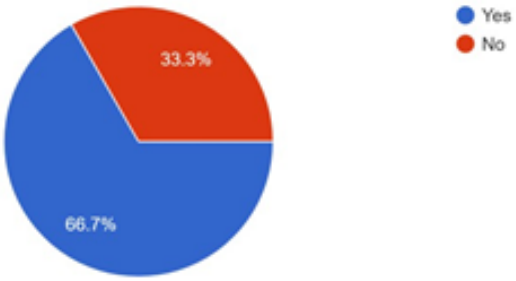
Do you think authentication is required every time you connect to a Building Automation System (BAS) device?

51 responses



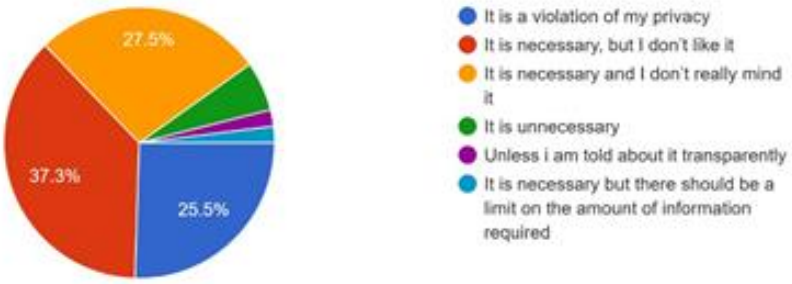
Do you think authentication is required every time you connect to a Building Automation System (BAS) device?

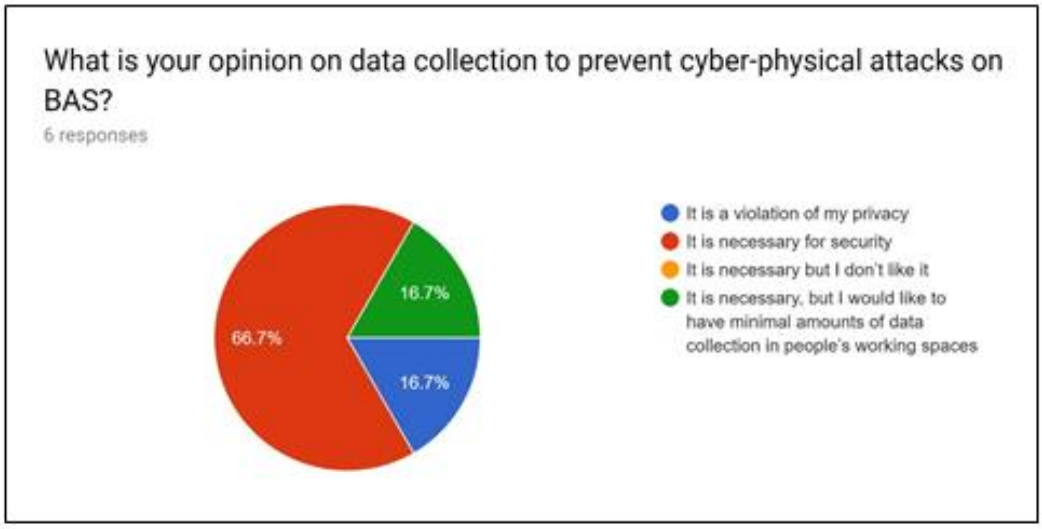
6 responses

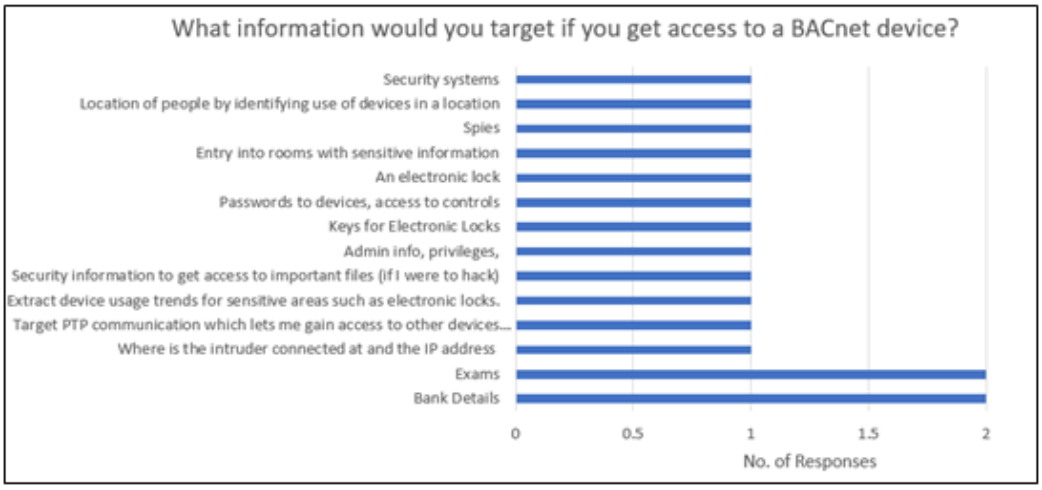


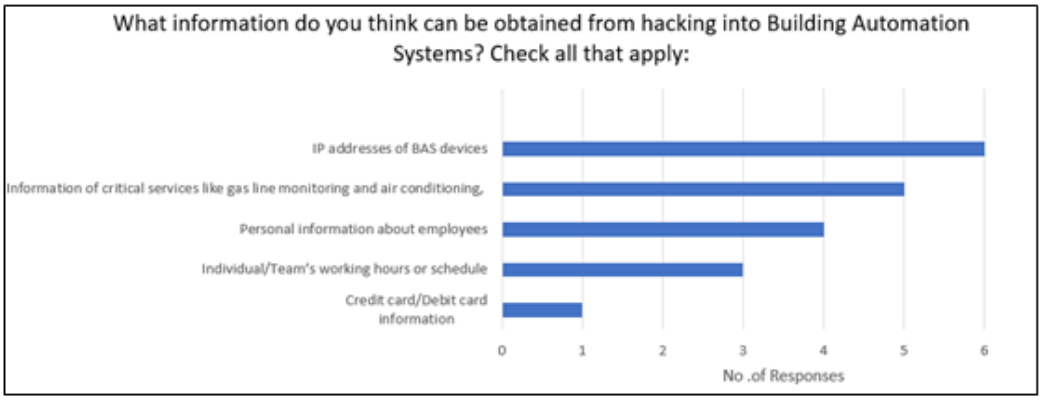
What is your opinion on data collection for cyber-security measures in BAS?

51 responses




What is your opinion on data collection to prevent cyber-physical attacks on BAS?
6 responses
16.7%
66.7%
16.7%
It is a violation of my privacy
It is necessary for security
It is necessary but I don't like it
It is necessary, but I would like to have minimal amounts of data collection in people's working spaces


What information would you target if you get access to a BACnet device?
Security systems
Location of people by identifying use of devices in a location
Spies
Entry into rooms with sensitive information
An electronic lock
Passwords to devices, access to controls
Keys for Electronic Locks
Admin info, privileges,
Security information to get access to important files (if I were to hack)
Extract device usage trends for sensitive areas such as electronic locks.
Target PTP communication which lets me gain access to other devices...
Where is the intruder connected at and the IP address
Exams
Bank Details
0
0.5
1
1.5
2
No. of Responses


What information do you think can be obtained from hacking into Building Automation Systems? Check all that apply:
IP addresses of BAS devices
Information of critical services like gas line monitoring and air conditioning,
Personal information about employees
Individual/Team's working hours or schedule
Credit card/Debit card information
0
1
2
3
4
5
6
No .of Responses