12-8-2021

The Distributed Trust Model Applied to the Energy Grid of Things

Narmada Sonali Fernando
Portland State University

The Distributed Trust Model Applied to the Energy Grid of Things

by

Narmada Sonali Fernando

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Electrical and Computer Engineering

Thesis Committee:
John M. Acken, Chair
Robert B. Bass
Roy Kravitz

Portland State University
2021

Abstract

Electric power system operators can manage distribution system utilization and usage by coordinating end customer usage of distributed energy resources. The end customers in this regard are Service Provisioning Customers, who provide their energy resources to a Grid Service Provider, which in turn dispatches large aggregations of distributed energy resources to provide reliable service to the power system. The security of this system relies upon information protection mechanisms, as described in IEEE 2030.5. However, in addition to preventive security measures, a monitoring function is required to ensure trustworthiness.

Trust models are a method to detect and respond to both expected and unexpected behavior. Different trust models are required for various types and characteristics of each situation. This thesis describes the topics that must be considered when developing a trust model as it applies to distributed energy resources. This thesis also provide the creation and application of a Distributed Trust Model applied to distributed energy resources. A key feature of Distributed Trust Model is to evaluate and alert an authority of any abnormalities. The decision to send an alert at the right time is critical to avoid possible disasters caused by intruders of the communication system. Major contributions of this thesis are to introduce a method for the Distributed Trust Model(DTM) to set the correct thresholds to send alerts at the appropriate times and evaluate the specified threshold values. Additionally, a method is defined to assess the decision-making equations that send those alerts. Overall, the

hypothesis tool can provide the statistical probability of failing to send an alert or false alerts based on the selected threshold. The hypothesis analysis provided by this tool helps a decision maker to understand statistical impacts of a specific threshold.

Dedication

To My Dear Mom, Dad, Brother, Coral Jean Cotterell, My Family in Sri-Lanka
and Shelly the dog.

Acknowledgements

This work is not possible without the guidance, infinite help, and patience of my advisor, Dr. John M. Acken. I appreciate you always believe in me. Thank you for always providing invaluable mentorship in countless hours and always being available to meet with me and provide sound advice.

Many thanks go to Dr. Robert B. Bass for providing me with endless opportunities, help, support, and resources to learn to grow and always providing a positive attitude. Thank you for letting me be part of the Power Engineering Group and collaborate in this Energy Grid of Things Project. I thank you both from the bottom of my heart. This project has had a lasting impact on me.

I want to express my gratitude to Professor Roy Kravitz. I appreciate your support, invaluable advice, and always willingness to help to clarify a complex topic. Thank you for the inspirational classes that helped me gain a better insight into the Internet of Things devices.

Thank you, Abdullah Barghouti, for developing the Trust Model Simulator, which made this research possible. I appreciate your contribution and collaboration that helped create many technical documents related to the Energy Grid of Things Distributed Trust Model.

I appreciate the help and support from fellow researchers in the Power Engineering Group at Portland State University. Thank you, Tyler Slay, for your guidance and explanation of the working of the energy grid and Smart Energy Profile, and numerous other technical topics.

Thank you very much, Mohammed Alsaid, for your valuable input during many, many technical meetings from the beginning on figuring out the Energy Grid of Things and associate specifications. Thank you, Blue Spitzer, for your technical insight and knowledge of how the Distributed Control Module is designed for implementation and clarifying when needed.

I also would like to thank the members of Dr. John M. Ackens's research group, where I had the privilege to engage in many, many valuable discussions on various research topics, including the Distributed Trust Model. Thank you, Aurelien Mozipo and Chenyang Li, for your valuable feedback and questions regarding this research.

Thank you, my dear friend, Coral Jean Cotterel, for your endless support and the guidance that helped me through this journey. Thank you for always being available to clarify any doubts I have and being a sounding board.

Finally, I would like to thank my parents for their support, accommodation, and financial support to get me through graduate school. In addition, I appreciate you and my brother and my family in Sri Lanka for always being supportive and providing me with any resources I needed to help me in this journey. Last I would like to thank Shelly, the dog, for keeping me company day and night while working on this project.

Contents

List of Tables

List of Figures

Acronyms

ACL Access Control List

AES-CCM Advanced Encryption Standard Cipher Block Chaining - Message Authentication Code

CA Certificate Authority

CDTA Central Distributed Trust Aggregator

CTA Consumer Technology Association

DCM Distributed Control Module

DER Distributed Energy Resources

DS Distrust Score

DTM Distributed Trust Model

DTMC Distributed Trust Model Client

EER Equal Error Rate

EGoT Energy Grid of Things

ESI Energy Service Interface

FN False Negative

FNR False Negative Rate

FP False Positive

FPR False Positive Rate

GO Grid Operator

GSP Grid Service Provider

HTTP

HTTPS

IEEE The Institute of Electrical and Electronics Engineers

LAN Local Area Network

MVoT Metric Vector of Trust

PEG Power Engineering Group

PSU Portland State University

RFC Relative Factor of Certainty

SDTT Standard Deviation of Transit Time

SFDI Short-Form Device Identifier

SPC Service Provisioning Customer

TCP Transmission Control Protocol

TLS Transport Layer Security

TN True Negative

TP True Positive

TS Trust Score

TSLC Time Since Last Communication

WAN Wide Area Network

XML Extensible Markup Language

1 Introduction

Trust is essential to secure the Energy Grid of Things (EGoT). The utilization and usage of an electrical distribution system may be managed by intelligently adjusting end-user demand. The term used for such a system is called an “EGoT". Existing information security solutions are based upon confidentiality, integrity, and accessibility. Confidentiality prevents message content exposure to unauthorized parties. Integrity is when the message is not altered in transit or by an unauthorized party. Finally, accessibility is when the message is accessible only to the authorized party. The current security solution is vulnerable to attacks which might go undetected. A Distributed Trust Model (DTM) can detect abnormalities in network communications [1]. One of the problems faced by the Distributed Trust Model(DTM) is setting the correct threshold to send alerts at the appropriate times. Additionally, a method is needed to evaluate the equations used for sending those alerts. The contribution of this thesis is to provide a solution that analyzes how a set of threshold values impact the error ratio of trust alerts. Additionally, check the decision-making equations that send alerts to the authority and see those trust equations are appropriately set.

A mathematical approach to ethics presented by American mathematician George David Birkhoff in 1941 applies mathematical formulae to measure ethical behavior and tradeoffs of friendship and privileges between three entities A, B and C [2]. This is the

earliest research tied to the concept of trust found during this research.

Viriyasitavat et al. [3] mentioned one of the earliest research done in computational trust is S.P Marsh's doctoral dissertation [4]. Marsh presented a mechanism to set and compare threshold values with calculated trust value and determine the level of criticality to send alerts. Some aspects of this mechanism are used in this thesis and they are described in later chapters.

Earlier research of distributed trust models, for example by Abdul-Rahman and Halles [5], mentions that communication networks are secured via 'privacy,' 'authenticity,' and 'access control.' The EGoT maintains confidentiality, integrity, and accessibility, but, without the distributed trust model, it cannot detect abnormal activities early on that helps stop an attack. The ability to gain and maintain 'trustworthiness' is not sufficiently implemented as part of existing energy grid security solutions.

Proposed solution: This thesis presents a data monitoring system designed to augment existing security technology that detect abnormal activities in the energy grids connected to networked intelligent consumer devices. Such devices include solar panels, water heaters, and battery inverter systems. I provide a distributed trust model suitable for the EGoT that detects abnormalities that can corrupt data, thereby enhancing the stability and trustworthiness of the energy grid.

Sakhnini et al. conducted a bibliometric survey explained in the article "Security aspects of Internet of Things aided smart grids. "They observe "how the published journal articles on smart grid cybersecurity grew exponentially from 1998 to 2018," as shown in

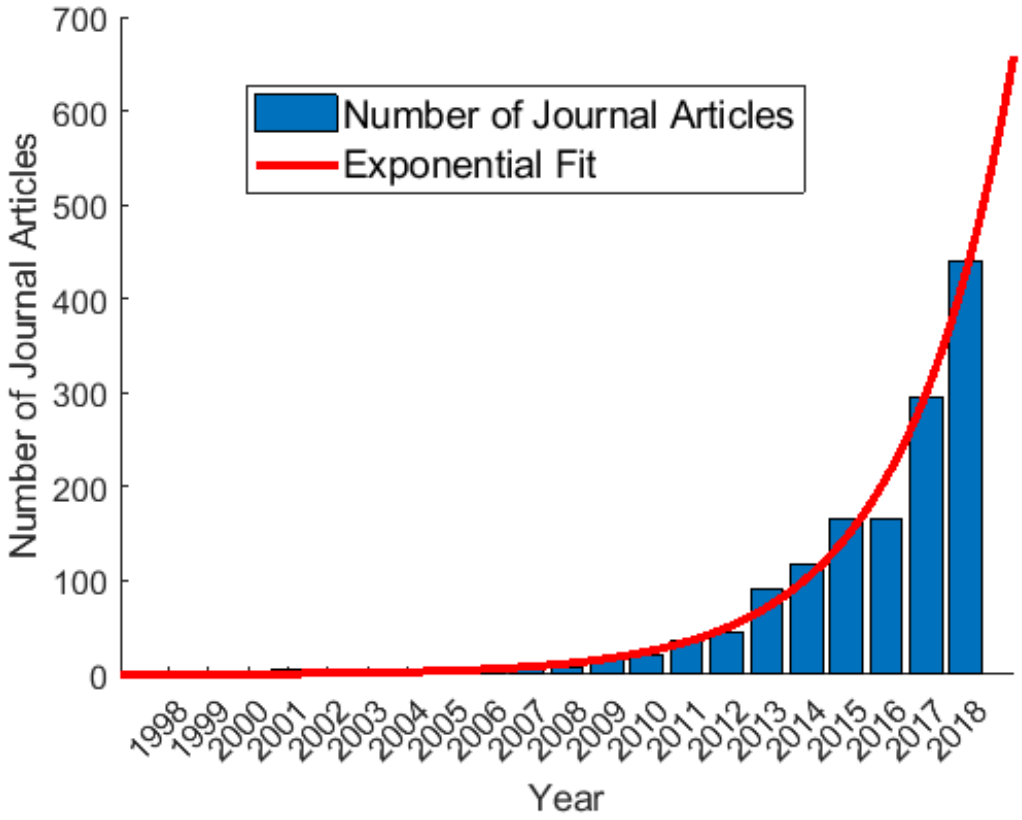


Figure 1.1: Number of journal articles on security systems of the smart grid published every year [6].

Figure 1.1 [6]. Their bibliometric survey findings showed very few journal articles on mitigating cybersecurity attacks on the EGoT, but this may be due to the concept being relatively new, and therefore yielding a too-small sample size. They also claim most journal articles focus on detecting anomalies. The bibliometric survey also presents the need for EGoT cyber-threat detection systems to have higher accuracy in abnormal activity detection and reduce the irregular activity detection time. This thesis intends to measure accuracy in detecting an anomaly in the EGoT via the Distributed Trust Model.

My research proposes a distributed trust model that augments the existing security by monitoring any abnormalities in the EGoT. My research also defines a method that allows a

DTM user to set evaluation thresholds suitable for their system. The distributed trust model matters to the Grid Service Providers (GSP) and customers of energy service providers. Early detection of abnormal activities caused by malicious attackers helps enhance grid stability and trustworthiness. The DTM results in higher customer trust, trust and retention, and participation in distributed intelligent energy devices. In addition, the trust model increases the trust by Grid Operator (GO) and Grid Service Provider (GSP).

When evaluating the cost factor of the distributed trust model, there are two contributors, the added cost to implement and the slight delay in system communication. However, the added cost of the DTM is worth it because the DTM augments the existing security of the energy grid of things (EGoT). The ability to gain and maintain 'trustworthiness' is not part of current energy grid security solutions. The distributed trust model verifies when information is converged normally and raises an alarm when abnormal activities are detected.

This thesis proposes a series of assessments or calculations of trust and distrust to show how the system performs. Accuracy was measured by comparing calculated trust to actual trustworthiness as true positive, true negative, false positive, and false negative. Central Distributed Trust Aggregator (CDTA) alerts the GSP when there is an information anomaly. True-positive is when the CDTA correctly sends out an alert message. On the contrary, if the CDTA sends out an alert with no attacks, the correct term for such an event is a false-positive. False-negative is when the CDTA fails to alert the GSP in case of an attack. Finally, True-negative is if there is no attack and CDTA decides not to send an alert. The

error rates for false positive and false negative demonstrates the measure of accuracy. Equal Error Rate (EER) is a measurement of accuracy where the false-positive rates equal the false-negative rate.

An EGoT uses aggregations of customer-owned Distributed Energy Resources (DER) to provide essential energy services through large-scale, coordinated dispatch of DER. Because of its role in ensuring power system reliability and efficiency, EGoT is considered a critical infrastructure, belonging to both the Energy Sector and the Communications sector. Critical infrastructure is defined as "a system and assets, whether physical or virtual, so vital to the U.S. that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters [7]." An EGoT is dependent on internet communication; hence, security is a significant concern. When devices are connected to the internet and are part of a network, security vulnerabilities increase. This thesis discusses different features of the DTM that help select the most suitable DTM for a specific EGoT Architecture.

This thesis describes the integration of a distributed trust model to Energy Grid of Things architecture. Chapter two describes the distributed trust model, energy grid of things, security vulnerabilities of the communication network, different types of trust models and their characteristics, and IEEE 2030.5 protocol security features. Chapter three describes the implemented distributed trust model system and energy grid of things architecture of the ongoing project. Chapter three also presents the Energy Service Interface (ESI) for the

EGoT, developed by the Power Engineering Group (PEG) at Portland State University (PSU). Chapter four describes hypothesis testing, how it applies to the distributed trust model system, and the hypothesis testing methodology used in this research. Chapter five analyzes the research results, chapter six discusses the research findings, and chapter seven provides the derived conclusion of the conducted research.

2 Background

2.1 Overview

As discussed in Chapter 1, the distributed trust model monitors and detect abnormalities in an EGoT communication channel. This chapter describes the importance of securing the EGoT and provides examples of historical cyberattacks imposed on digital communication networks. This chapter explains what an EGoT is and its architecture. This chapter also describes the DTM in detail and how it fits into the EGoT. Also included in this chapter are the examples of types of trust models, trust model components, and possible attacks and potential defense mechanisms integrated into the digital communication network. The IEEE 2030.5, the communication protocol used by the EGoT design, is applied to this thesis and described in this chapter, including the security features of IEEE 2030.5.

2.2 Energy grid of things applied to the power grid

An EGoT manages the reliability and efficiency of an electrical distribution system by intelligently adjusting end-user demand of customer-owned Distributed Energy Resources (DER). An EGoT provides reliable services through a large-scale, coordinated dispatch of DER. An EGoT ensures power system reliability and efficiency, and it is part of both the energy sector and the communication sector. Both sectors are Critical Infrastructure. An

EGoT depends on internet communication. Hence, there is a dire need for security to protect information. When devices are connected to the internet and are part of a network, security vulnerabilities increase.

2.2.1 Cyberattacks examples

This section describes some examples of cyberattacks. The first example of a cyberattack is a computer worm called Stuxnet, which attacked Iranian nuclear facilities. The worm infected PLCs (Programmable Logic Controller), that controlled centrifuges. The worm spoofed data communication and changed the centrifuge rotation speed, causing them resonate destructively. The attack resulted in significant damage to thousands of centrifuges. The infectious worm propagated over shared networks, by exploiting a printer-spool vulnerability. Inside the victim's network, the manipulated PLCs provide spoofed sensor input data instead of real-time data in order to conceal the attack.

The second example occurred in December 2017. An imposed cyberattack occurred on the Triconex safety system of an unidentified power station, believed to be in Saudi Arabia. Hackers used a malware called "Triton" to gain control of the industrial control system. The intruders obtained possession of the Triconex safety system and changed settings to turn off any alarms. However, this attack failed due to the culprits not knowing how to control the industrial safety system completely; they accidentally triggered a shut-off and caused an alert to go out, making the system operators aware of the assault [1]. The unexpected signals

[1] Jim Finkle. Hackers halt plant operations in watershed cyber attack, Dec 2017.

alerted the Triconex customers [2] of the compromised system. Schneider Electric SE, the manufacturer of the Triconex industrial safety system, did not confirm the attack.

The third example occurred in 2000. A cyberattack resulted in a sewage spill into Queensland Australia's Maroochy river and coastal waters. The attacker highjacked the Maroochy sewage control system and redirected the sewage into the waterways of Australia. The dumping of this sewage contaminated the clean water available for public consumption.

The fourth example occurred in the years 2015 and 2016. An attack on the Ukraine power grid impacted 30 substations, lasting somewhere between one to six hours. The intruders hacked power plant networks and Supervisory Control And Data Acquisition (SCADA) systems and shut off power distribution to consumers. The Ukrainian power grid attack was conducted via spear phishing, distributing, and installing malware in employee computers. This led to the compromised control system that sent out open commands to open substation circuit breakers, which ultimately caused the power system to fail [8].

Described above are several examples of cyberattack incidents in digital communication channels in different parts of the world. With this in mind, let's look at understanding the properties of the EGoT, described in the next section.

2.3 Energy Grid of Things

Portland State University's (PSU) Power Engineering Group (PEG) is developing a prototype EGoT system with two principal objectives: 1. Develop an Energy Service

[2]The guardian 2017. Triton: hackers take out safety systems in 'watershed' attack on energy plant, Dec 2017.

Interface (ESI), and 2. Develop a Distributed Trust model.

As defined by Widergren et al., "an ESI is a bidirectional, service-oriented, logical interface that supports the secure communication of information between entities inside and entities outside of a customer boundary to facilitate various energy interactions between electrical loads, storage, and generation within customer facilities and external entities" [9].

An ESI provides a set of rules that govern information exchange between a Grid Service Provider (GSP) and Service Provisioning Customer (SPC) to facilitate the exchange of energy services. The GSP contracts with a Grid Operator (GO) to provide essential reliability services. The GSP then coordinates with SPCs for individual contributions to the requested services. Figure 2.1 shows an example of an electrical distribution system topology. Figure 2.2 shows the EGoT concept applied to that topology. PSU's EGoT uses the standard cybersecurity protocols defined in IEEE 2030.5. IEEE 2030.5 is a standardized communication protocol that supports the exchange of energy services. Protocol features include discovering DER and providing energy services such as demand response and flow reservation.

2.3.1 Energy Grid of Things Infrastructure

In the PSU PEG project, the EGoT infrastructure consists of several important actors: the Grid Operator (GO), the Grid Service Provider (GSP), and the Service Provisioning Customer (SPC). Figure 2.3 shows both the EGoT network connections and how the actors are generally connected. The GO is the energy provider to the grid and ensures there is enough energy provided to balance out the energy consumption of customer. In this project,

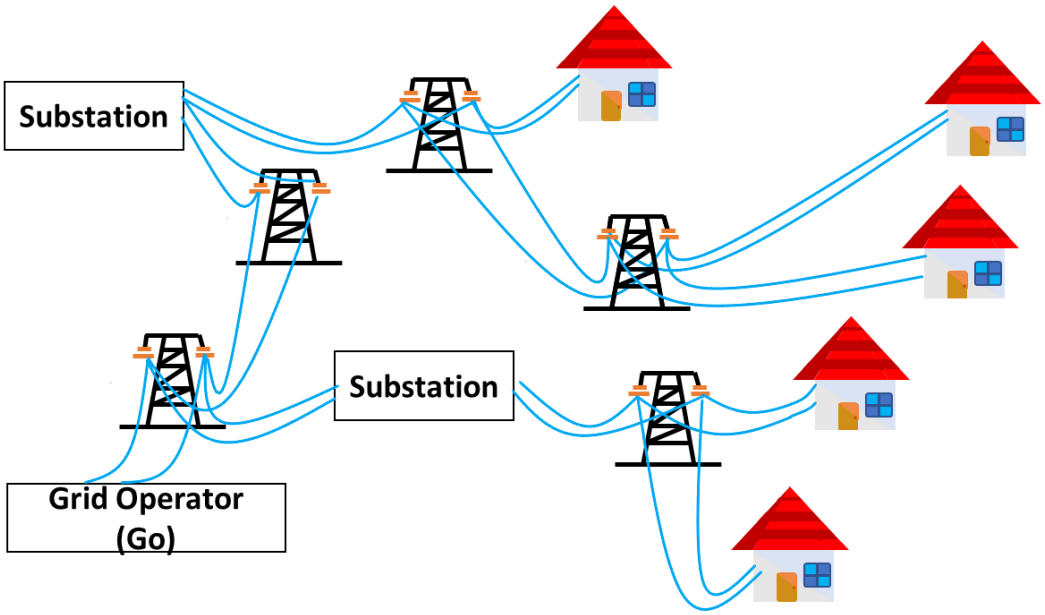


Figure 2.1: An example Power Distribution system showing Grid Operator and customers.

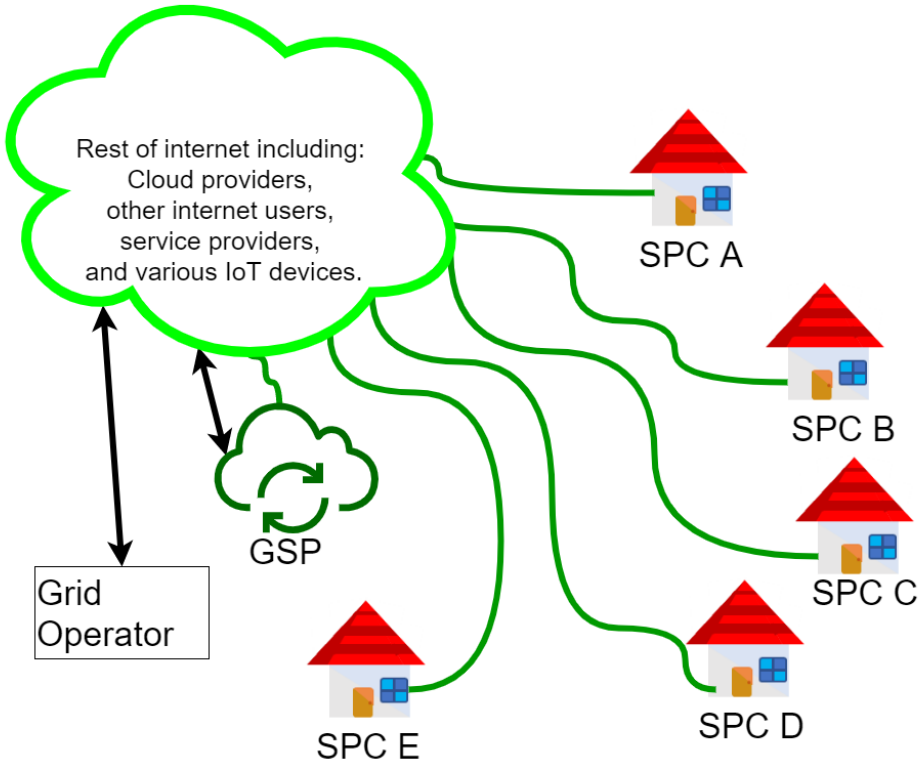


Figure 2.2: An example of the Energy Grid of Things with the GO, GSP, and the SPC communication system.

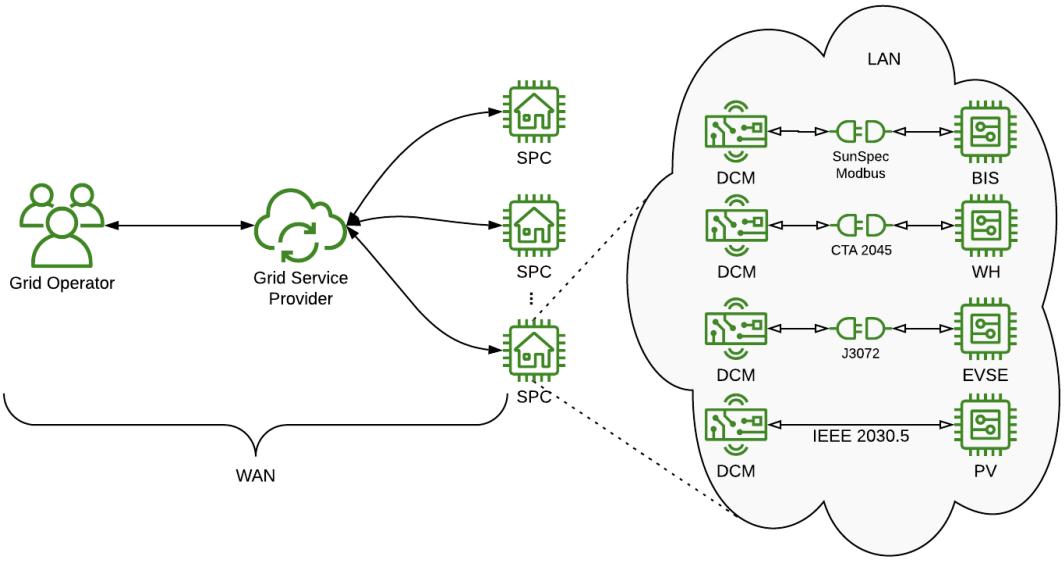


Figure 2.3: EGoT infrastructure and how its actors are connected.

there is one GO. The GSP provides grid services to the GO by allocating DER resources from the SPCs. The GSP manages aggregated DER assets to meet the operational objectives of the GO.

The GSP and the SPCs are part of a Wide Area Network (WAN). SPCs and their DER are part of Local Area Network (LAN). Located with each DER is the Distributed Control Module (DCM). Generally, the DER and the GSP do not share the same communication protocol. Hence, the DCM resides between the DER and the GSP to manage information exchange between them. The DCM is responsible for interpreting the intent of incoming messages and communicating that intent using a protocol the recipient understands.

2.3.1.1 GSP - DCM Information Exchange

The GSP and DCM exchange information request to dispatch energy services. The GSP and the DCM use protocols that define and follow strict security rules to ensure this communication channel is secure. The PSU PEG project uses the IEEE 2030.5 protocol which defines access control lists, device credentials, resource access authentication and authorization, and cipher suites. Section 2.6 detailed information about IEEE 2030.5 protocol features used in this project.

2.3.1.2 DCM - DER Information Exchange

The DCM is responsible for converging grid service information received from the GSP to the DER. In the EGoT project, the GSP uses the IEEE 2030.5 protocol, and the DER uses the CTA-2045 protocol. The DCM messages the DER regarding grid service lists, grid service scheduling information, and heartbeat signals. In contrast, the DER messages the DCM regarding the current service status and polling for the heartbeat signal of the grid.

2.3.1.3 GO - GSP Information Exchange

For this project, a simulated GO, exchanges grid service messages with the GSP. In the EGoT project, there are no specific protocols implemented between the GO and the GSP communication channel. The following section describes an actual grid service communication between the GSP server, and DCM/DER client.

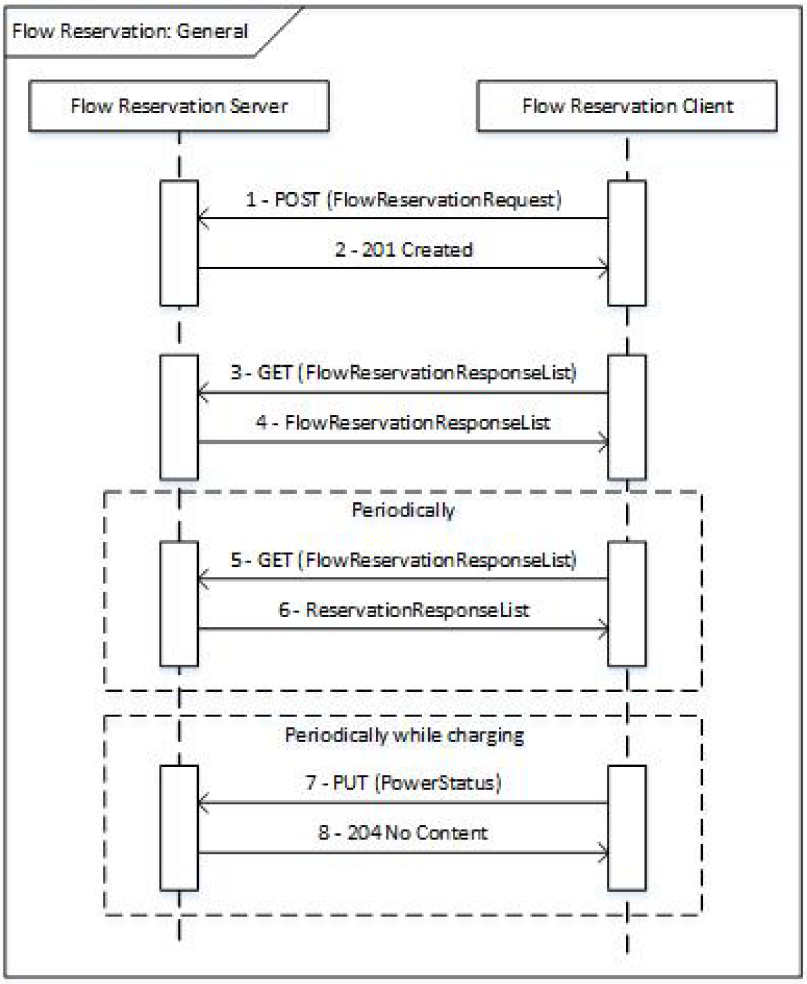


Figure 2.4: This figure illustrates the general communication exchange between the server and the client for a flow reservation service [10].

2.3.2 General Flow Reservation communication example

The IEEE 2030.5 protocol provides an example of a communication exchange between a server and a client for a general flow reservation service, as shown in Figure 2.4.

Description of the steps labeled in Figure 2.4 are:

First the client requests a $FlowReservation\ Request$ from the server in response the

server send a HTTP 201 creation of requested resource. Next, the client requests a $FlowReservation\ ResponseList$ from the server, followed by the server presenting a $FlowReservation\ ResponseList$ for client end devices.

Periodically the client requests a $FlowReservation\ ResponseList$ from the server and the server presents a $FlowReservation\ ResponseList$ for client end device.

While charging the client posts the power status of the DER and the server responds with an acknowledgment once received.

2.4 Description of Types of Trust Models

The primary responsibility of the distributed trust model is to monitor and report any anomalies to appropriate authorities. Without any interference to the communication channel, the DTM augments existing cybersecurity measures to improve system trustworthiness. The DTM evaluates the flow of messages between a client and a server, updates the trust status, and creates alerts if an anomaly is detected.

Clients and servers are called actors of the network. The evaluation of the DTM depends on a series of criteria for a specific actor, such as the history, expectations, particular events, etc. The DTM can distinguish a critical event versus a minor event and decide when and whom to inform of such abnormalities. Another feature of the DTM is to keep track of the reputation of the device it is monitoring. A DTM can help augment a network's security by monitoring and evaluating incoming/outgoing messages and notifying the appropriate parties when suspicious activities are detected.

There are a variety of parameters and styles for creating DTMs. This thesis describes the characteristics of published Distributed Trust Models and methods of seeing suspicious activities. Features of different types of trust models include its organization, characterization, and components. Upcoming subsections provides a brief description of some examples of trust model characteristics.

2.4.1 Peer-to-Peer or Mesh Trust Model

A Peer-to-Peer (P2P) or mesh trust model is a network of trust models wherein each node communicates with other peer nodes. Advantages of P2P networks include improved network strength, flexibility, and variety in available data. Disadvantages of a P2P network include a lack of accountability due to anonymity. This exposes nodes to ill-treatments from malicious nodes [11]. Moussavi-Khalkhali et al. presented an illustration of what the mesh or peer-to-peer network architecture looks like, shown in Figure 2.5 [12].

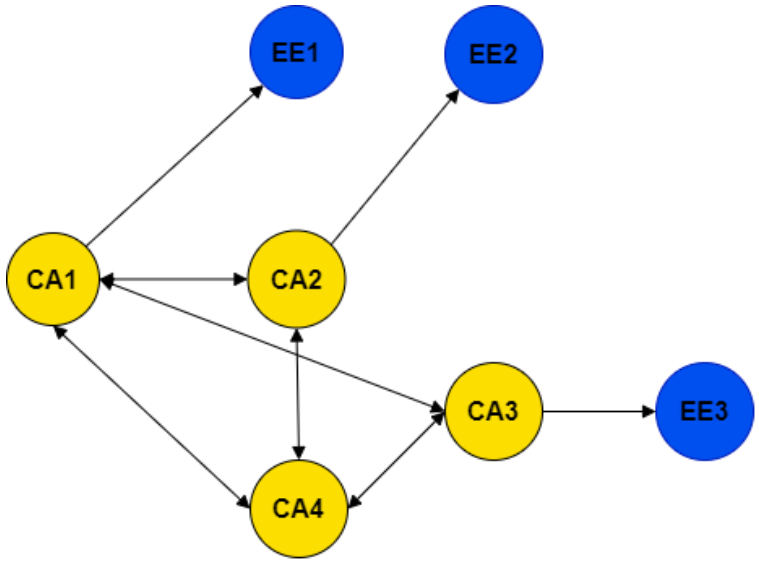


Figure 2.5: An example of a mesh network architecture [12] where Certificate Authority (CA) and EE nodes are present.

2.4.2 Hierarchical Trust Model

A Hierarchical Trust Model enables devices to cluster in tiers, allowing an upper tier to monitor known trusted clusters and use that learning to judge new groups [13]. Communication between tiers to learn or alert about threats is a possible advantage of this model.

2.4.3 Trust Model with Path Management

Khalid et al. introduced the idea of having a trust path manager that keeps a record of safe paths that do not go by malicious nodes and other routes that go by malicious nodes. When the trust monitor identifies a way that leads to a malicious node, it eliminates that path and removes it from the recommended path list [14].

2.4.4 Centralized Trust Model

A Centralized Trust Model has one central node that is the main point of trust management. All the other nodes rely on that node for getting recommendations of trust. Nunoo-Mensah et al. mention that individual nodes report all trust monitoring findings to the central node for analysis. The centralized trust node evaluates trust [15] and share the results back to the individual nodes. The authors also mentioned that the centralized trust node can become a bottle neck and slowdown decision making when evaluating trust in a large digital communication network.

2.4.5 Decentralized Trust Model

Each node in a Decentralized trust model is responsible for evaluating and determining other nodes' trust since there is no central node. Suryanarayana et al. points out that each node in a decentralized network can form its own individual defense mechanism in response to an attack [16]. In a decentralized network, nodes are allowed to self-certify, a mechanism used by nodes to authenticate the sender's identity [17].

2.4.6 Inner Circle vs. Outer Circle

In an inner circle and outer circle network structure there are closely affiliated actors and loosely affiliated actors. Closely interconnected actors are part of the network inner circle and loosely associated actors in the outer ring of the network. Inner circle vs. outer circle or community-based trust evaluation can help identify a compromised node. For example, suppose one or more nodes provide a mistrust rating about a node in the inner circle. Nodes within a circle have a significant weight [18].

2.4.7 Isolated Distributed Trust Model

Isolated DTM does not communicate with trust models outside of the communication network. Instead, it monitors directly connected devices for any anomaly. In Figure 2.6, the DTM is in an isolated location and is part of the Local Area Network (LAN). Its primary responsibility is to monitor directly-connected devices for anomalies and report them to the appropriate parties. A benefit of having isolated trust models is to ensure there are no outside network interference to the isolated distributed trust model.

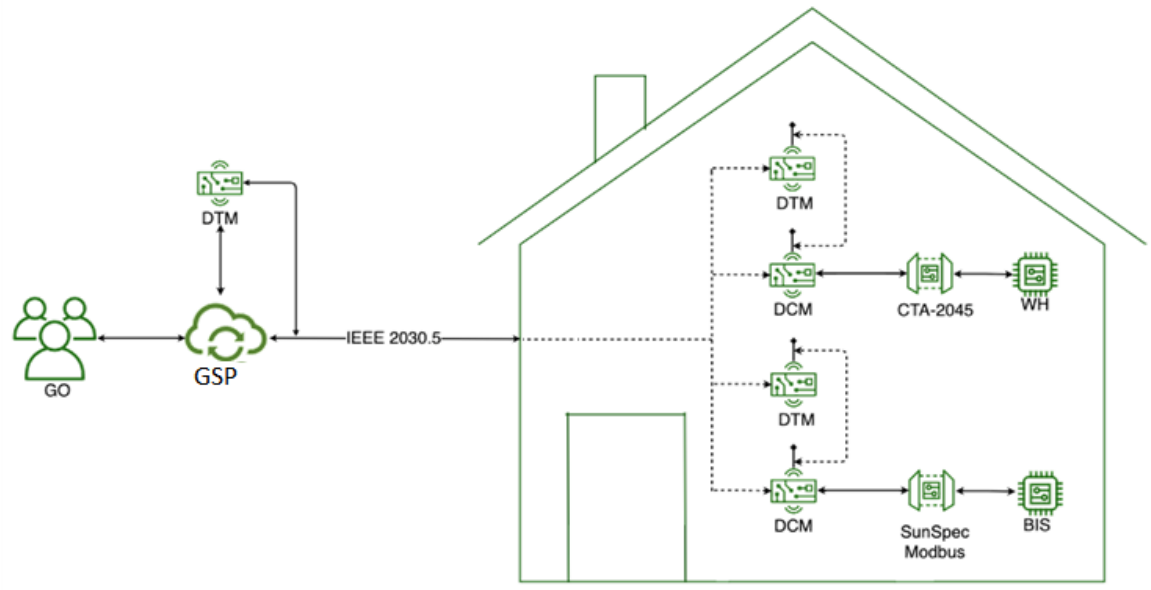


Figure 2.6: An example of an isolated Distributed Trust Model.

2.4.8 Trust Model for Federated System

Chun and Bavier focus on a decentralized trust model for a federated system. The network structure of the nodes is hierarchical with a layered trust architecture. Chun and Bavier's trust model architecture consist of three layers: Layer one, Authentication and Authorization. Layer two is accountability. Layer three is anomaly detection.

In the first layer, Chun and Bavier's design allows trust delegation from one node to another node it trust along with a mechanism to trace back the node that was responsible for delegation of trust to another node. This enables a way to trace back and identify the responsible node that delegated trust to the malicious node and easily track the chain of trust [19].

Chun and Bavier name the second layer of trust Accountability. In this layer, the

activities between nodes are monitored and logged to understand overall behavior and resource usage. Additionally, this layer monitors the trust relationship between nodes [19].

For the third layer of trust, Chun and Bavier introduce the detection of anomalies and taking appropriate actions in response to a detected anomaly. The anomaly detection occurs locally as well as across the network. The fourth layer of trust uses layer one's delegation of trust and chain of trust data and layer two's accountability layer's resource usage data to detect the anomaly and determine the warning level and associated actions [19].

2.5 Characteristics of Trust

Khalid et al. capture several characteristics of trust. As illustrated by Figure 2.7, Khalid et al. present intransitive trust decision, Figure 2.7 (a), subjective trust decision Figure 2.7 (b), and properties of trust Figure 2.7 (c).

- **Asymmetric trust** - where the trust is not symmetrical between A and B, such that A trusts B but, B may not trust A.
- **Subjective trust** - trust is subjective such that A may have a higher trust in C rather than B to do task X. This does not imply that B is less capable of performing task X compared to C.
- **Partial transitivity or intransitive trust** - the establishment of trust varies from one actor to another. Where A trusts B, and B trusts C but, A may not trust C.

- **Context Sensitive trust** - where A may trust B to do task X, but A may not trust B to do task Y. Instead, A may trust C to do task Y.

- **Reflexive trust** - represents the level of confidence A has of each trust decision it is making.

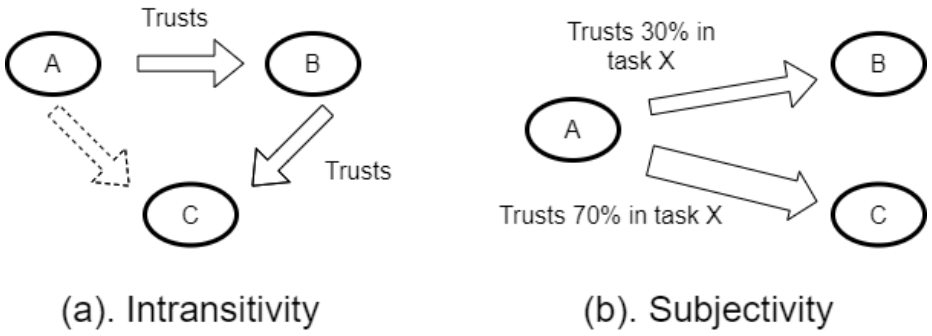


Figure 2.7: Figure illustrates characteristics of trust establishment. [14]

Li et al. provide a good example in reference to power grids where the source and the destination nodes are in different domains. This is presented in Figure 2.8, where P is the starting point of a process, and $N_0 ... N_k$ are the intermediate nodes that are present in the path from P to R.

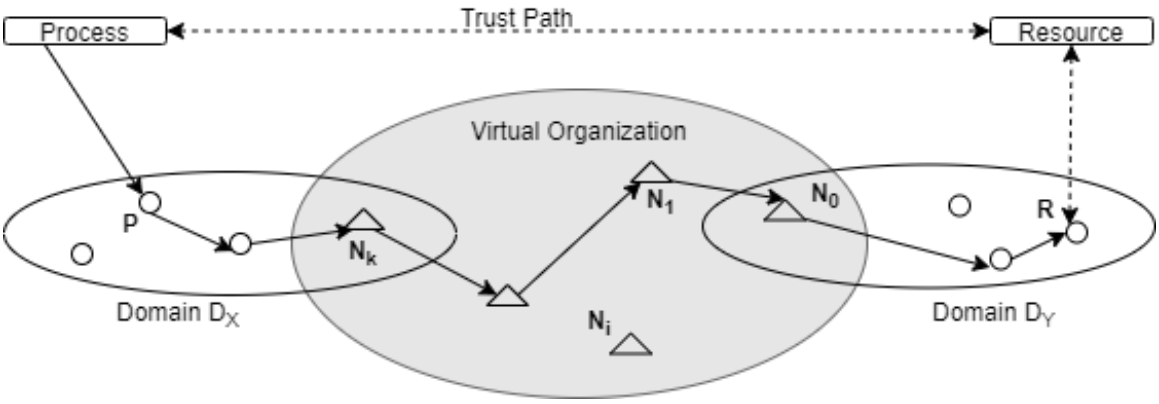


Figure 2.8: Figure illustrates the process of traversing domains and virtual organization to get access to specific resources [20].

2.6 Trust Model components

Selecting and designing a trust model and using the trust models requires analyzing the specific application. Considerations must include storage limitations, organization of the trust model system, communication, and model equations. "Storage" refers to how trust is stored and what mechanism is used to access the trust metric when it is needed. Trust storage mechanisms can either be distributed or centralized. In a distributed trust model, multiple nodes are locally responsible for trust storage. In a centralized trust model, a central node is responsible for trust storage [21].

Current and historic trust values help derive an actor's trust score. The trust metric dimensions vary based on the application. Trust metrics may have as few as one variable or as many as N entries representing total trust. In summary, anyone designing an architecture for a DTM must decide which of the previously described organization, characteristics, and components are part of their systems.

2.6.1 Vector-based Trust

Li and Zhao [18] strategically point out an apparent key characteristic of a vector-based trust management system. The vector-based trust model can have multiple variables that can independently identify different abnormalities as a sign of attack. For example, a trust vector can use a variable to evaluate the changes in the frequency of communication, where communication increase is a sign of a Denial of Service (DoS) attack.

Vector-based trust enables each node to evaluate the trustworthiness of a peer without

the dependency of a centralized network. Vector-based trust is similar to how human social networks behave. For example, humans may trust another based on past interactions. Likewise, each node keeps its historical interactions within a trust vector. Entries of the trust vector use a mathematical equation to derive its value and evaluate the network security status.

When calculating trust, the DTM must understand the influence of either certainty or uncertainty. Solhaug and Stølen specify that uncertain classification can be due to doubtfulness in future events, lack of evidence in the possibility of a future malicious attack, and ignorance of existing abnormal behavior [22].

A Sybil attack is when malicious nodes are able to create many false identities nodes to manipulate the system to their agenda. A Sybil attack can cause a series of malicious nodes a high trust score. A trust vector helps identify a possibility of a Sybil attack by keeping a local trust rating of other nodes and comparing it against the most recent trust calculations of those nodes. In case of a discrepancy between the calculated trust score and the stored trust scores, the trust vector may suspect the network is under Sybil attack.

2.6.2 Trust Architecture

Xiong and Liu designed a peer-to-peer trust architecture. Theirs is a decentralized trust system. The network peers use trust data to calculate the trust scores and recommendations and store them at each peer across the network. Each peer has a piece of global trust data stored and a data locator that indicates the location and placement of trust data [23]. Xiong

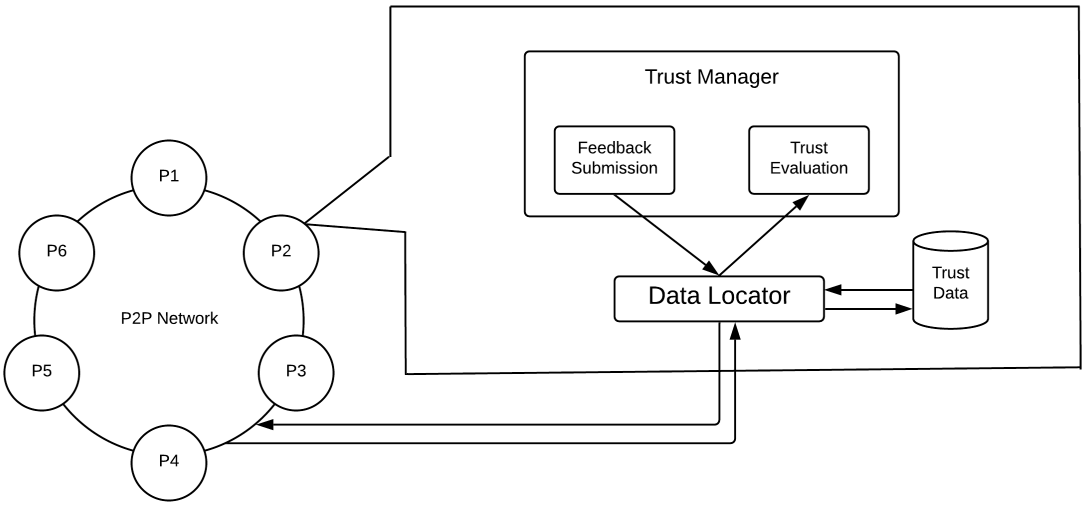


Figure 2.9: An example of a peer-to-peer trust architecture [23].

and Liu also have a trust manager at each peer to evaluate the trust data of other peers and provide feedback to communicate the reputation of the corresponding peer.

2.7 Attacks

When designing a DTM architecture, one must consider the breadth of possible attacks and unknown future attacks. The DTM is well-suited for detecting attacks within a specific context. This section discusses potential attacks on an EGoT system and possible attacks on the distributed trust model itself.

2.7.1 Attacks on Communication Channels of the Energy Grid of Things

Since power grids are critical infrastructure, attacks can result in severe ramifications. For example, malfunctioning of the power grids, communication networks cause interference to time synchronization, reliability, latency, the criticality of data delivery, support for

multicast, and interoperability [24]. Ultimately what's listed above causes a severe impact on the customer trust in participating in the EGoT caused by unreliability.

2.7.1.1 Duplicate Address Selection/ Detection

Duplicated address selection is a form of attack that could affect an EGoT system. A malicious node uses another node's ID instead of its assigned ID. This action makes it difficult to detect the location/origin of the malicious node that conducted the attack. Additionally, malicious nodes gain credibility by sending out addresses of a credible node and pretending to be a credible node.

2.7.1.2 Man-in-the-Middle Attack

An EGoT system is vulnerable to man-in-the-middle (MITM) attacks, wherein a snooping party can watch and intercept information exchange between two parties. The intruder uses the learned information to pretend to be an authentic device. This then provides the intruder a means to execute attacks. Mohapatra et al. noted that MITM attacks cause network packet alteration with add-on delay, drop of packets [25].

2.7.1.3 Denial of Service

A denial-of-service attack occurs when a node gets flooded with a series of messages from a malicious node, making it inaccessible to intended nodes. The communication exchange between the energy consumers, the energy providers, and the distributors is critical for

ensuring power system reliability and efficiency. As such, it is essential to ensure EGoT communication channels go uninterrupted.

2.7.1.4 Eavesdropping

Eavesdropping is when an attacker listens in on the communication between two nodes and forms an attack on that channel. Eavesdropping, or active reconnaissance, is a preliminary step for enabling many attacks, such as man-in-the-middle, denial-of-service, duplicate address selection.

2.7.2 Attacks on Distributed Trust Model

Attacks to the Distributed Trust Model are also possible. It is easy to think that having an anomaly detection monitoring device like the DTM is sufficient for augmenting communication cybersecurity. What if the DTM itself is under attack? Then the system anomalies go unnoticed until a disruptive attack occurs.

Smith, et al. presented the possibility where a node is exposed to sensitive information during trust negotiation [26]. The DTM can be vulnerable to such events where the malicious node obtains sensitive information stored within.

2.7.2.1 Bad Mouthing

In a badmouthing attack, the DTM reports false information about the device it is monitoring. Examples of incorrect information include announcing malicious nodes as good or non-malicious devices as malicious. When a compromised node is under a

badmouthing attack, it provides inaccurate recommendations to fit its attack plan. It gives positive recommendations to preferred nodes and negative or low recommendations to other nodes [27].

2.7.2.2 On-off attack

In an on-off attack on the Trust Model, a node occasionally behave abnormally or maliciously and, most of the time, act normally. The malicious node acts normal the majority of the time to hide from discovery. The malicious nodes can gain complete trust, continuously collect more information about the system, and drive occasional attacks with minor consequences.

2.7.2.3 Imposter attack

An imposter attack occurs when a malicious node changes its identity to hide the attack source. The malicious node can use an imposter identity to conduct the vicious attacks and remove any imposter information used to carry out attacks.

2.7.2.4 Conflicting Behavior Attack

In a conflicting behavior attack, the malicious node act differently depending on the interacting node. A malicious node behaves appropriately with some nodes and inappropriately with other nodes, maintains a good reputation, and stays on the network. An example of this type of attack is when a node provides good recommendations about the

nodes in its immediate circle and provides false recommendations of nodes outside the community.

2.7.2.5 Sybil Attack

In a Sybil attack, a malicious node creates multiple phantom nodes with imposter device IDs to confuse and conduct attacks on the network [1]. A malicious node executes an attack by creating phantom nodes that mislead the network and exhaust its resources. The Sybil attack is named after a case study subject who was diagnosed with multiple personalities disorder.

2.8 Defenses

Abdul-Rahman and Halles [5] address the absence of trustworthiness and the need to manage trust effectively. According to the authors, the current implementation of trust in security is via 'privacy,' 'authenticity,' and 'access-control.' The ability to gain and maintain 'trustworthiness' is not part of the current information security solution. The authors present a concept where in decentralized trust management systems, each agent or node makes independent decisions. For example, each node maintains its policies and make trust decisions independently. Observing certainty is a category of trust that looks at similar entities or nodes and maintain different levels of trust. Recurrent interactions helps the certainty of the recommendations.

The current solution is to authenticate and certify incoming and outgoing data packages. The limitations of the current solution include the room for viruses, malware, or corrupted data to infect EGoT without ever being detected or detected too late [5]. This

solution is implemented as access control to limit the exposure of sensitive information during a trust negotiation presented by [26].

In response to cyberattacks on a power grid integrated with computing and communication systems, Sun et al. proposed a defense mechanism that monitors packets for anomalies and comparing them with prerecorded data of past attacks. This information is used to identify the correct alert to send as well as mitigation steps to take [28].

2.9 IEEE 2030.5: Standard for Smart Energy Profile Application Protocol

provides the security aspect of the network communication channel's application layer. IEEE 2030.5 defines requirements for access control, registration, and device credentials for the devices. Before accessing any resources, there needs to be authentication, and authorization, and cipher suite usage at the application channel of the communication network.

The IEEE 2030.5 protocol enables any security measure of communication channel if the client sends a request to the destination server HTTPS port address instead of the HTTP port. A secure TLS 1.2 handshake occurs in case of an HTTPS request, and records are sent according to the Transmission Control Protocol (TCP). This process conducts mutual authentication with the device or self-signed certificates, message encryptions, message authentication with Advanced Encryption Standard Cipher Block Chaining - Message Authentication Code (AES-CCM) that outputs a ciphertext and a message authentication code. Additionally, this method provides a device Access Control List (ACL) that defines a

way to grant access to resources for devices based on Authentication level and addresses information. Finally, the IEEE 2030.5 protocol describes how to check the device registration information before giving access to participate in the EGoT. IEEE 2030.5 is an application profile that fulfills the needs of Energy Grid of Things participants to manage energy and information flow between the DERs and GSPs in a standardized form in a secure manner. Security features of the standard include device registration, access control, device access authentication, cipher suite that helps secure network connection. Security feature of the IEEE 2030.5 are discussed in Sections 2.9.1 - 2.9.4.

2.9.1 Registration Attribute

IEEE 2003.5 protocol is customized to the smart grid so that there is a device registration attribute that offers local registration. The method defined in the 2030.5 protocol allows device registration. It takes in the out-of-band registration information of Distributed Energy Resourcess (DERs) to participate in the grid services. Registration information includes the device information, individual PIN, a Short-Form Device Identifier (SFDI).

2.9.2 Access Control List Attribute

The access control attribute grants access to a particular client's resources. IEEE 2030.5 protocol describes customized access control attributes to help handle complex access policies.

2.9.3 Resource Access Authorization

The client obtains access to resources when the client registers with the host via an out-of-band. Authorization is followed immediately after successfully completing the authentication process if allowed by the existing policy rules. For devices with self-signed certificates the device SFDI (Short Form Device Identification) shall be matched.

2.9.4 Device Access Authentication & Authorization

The IEEE 2030.5 protocol enables security measures of communication channels if the client sends a request to the destination server HTTPS port address instead of the HTTP port. In case of an HTTPS request, a secure TLS 1.2 handshake occurs and sends the records according to TCP. This process includes the server and the client validating each other's device certificates to authenticate the transaction successfully. The authentication process can be conducted with self-signed device certificates and certificates by a 3rd party such as Certification Authority (CA). The device access authentication follows device authorization. The device SFDI (Short Form Device Identification) is used to compare against the SFDI provided by the device at the time of its registration.

2.9.5 Chapter 2 Summary

An electric power system can be made more reliable and efficient by implementing an EGoT system, wherein dispatched customer-owned aggregated DER provides essential grid services. Because the electric power system is an integral part of a country's critical infrastructure, security is necessary. Cybersecurity can help secure information security

measures. However, no security system can predict all possible future attacks. Trust-based monitoring augments traditional cybersecurity mechanisms. One such technique is to implement trust models. This background chapter presented considerations that help when selecting and designing a trust model system, emphasizing the application of a trust model to an EGoT system. The description of various trust models helps understand the flexibility of the DTM to fit into many different digital communication networks to evaluate trust. The report of some potential security attacks against both the system and the trust model provides a better perspective on vulnerabilities of digital communication and the DTM.

3 The Distributed Trust Model System

3.1 Overview

Chapter 1 and 2 described the DTM and EGoT, their architecture, vulnerabilities, and current security features. This thesis reflects the work done by the PSU Power Engineering Group (PEG) on implementing a Distributed Trust model to their DER aggregator system. Hence, this chapter describes the PSU PEG's design and implementation of the DTM, as shown in Figure 3.1, and its relationship to the EGoT described earlier in chapter 2. This chapter consists of two main sections, the Distributed Trust Model Client (DTMC) which is the DTM at the Service Provisioning Customer (SPC) and the Central Distributed Trust Aggregator (CDTA) which is the DTM at the Grid Service Provider (GSP).

The DTMC section includes its components, including the input message from an actor, the input classifier block, the trust equation evaluation block, and the Metric Vector of Trust (MVoT). Also provided is a detailed description of the MVoT equations and their purposes in the DTM system. Finally, the CDTA section of this chapter describes the components of the CDTA, specifically the central MVoT aggregator and the decision/actions/recommender block.

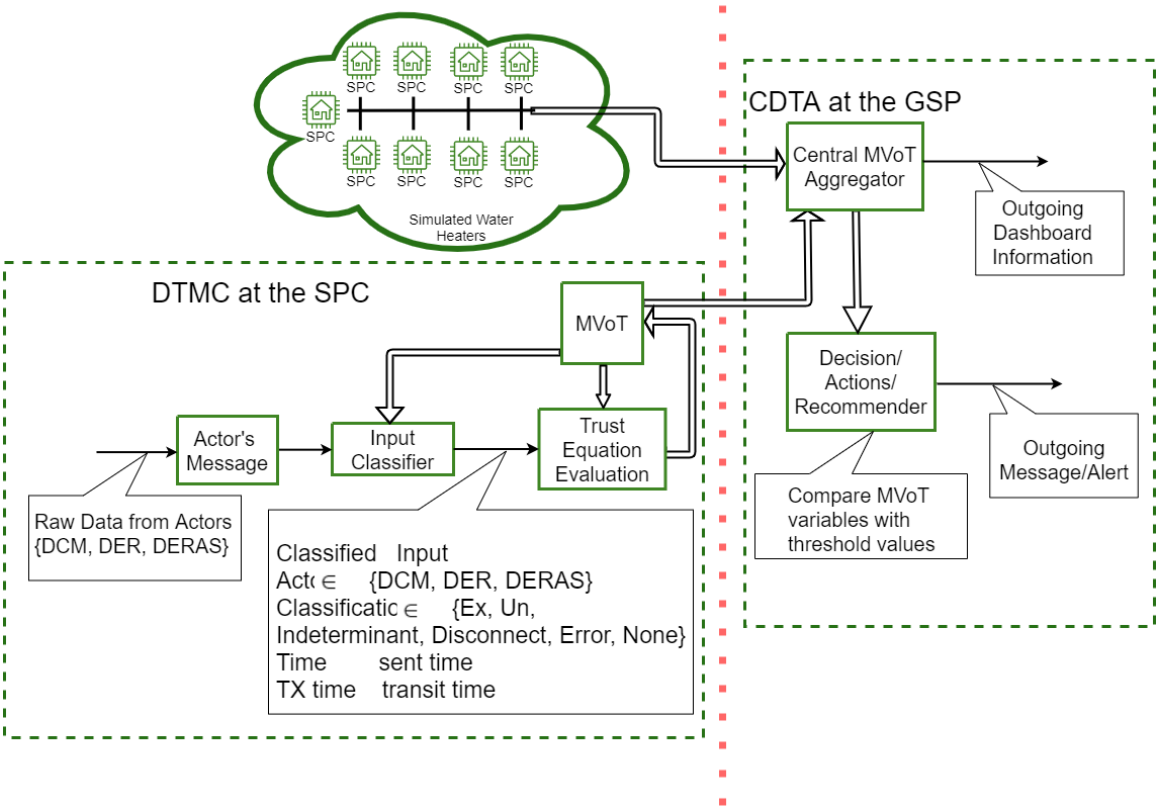


Figure 3.1: This figure shows the overall connection of the DTM system. Left of the red dotted lines are components of the DTM-Client at the SPC. This figure illustrates the setup for the initial prototype version with ten SPC and a DTMC at the GSP included.

3.2 The Architecture of the Distributed Trust Model System

Figure 3.1 illustrates the distributed trust model system architecture. The left side of the figure shows the process of the distributed trust model client classifying the actor's input message for the trust equation evaluation block to calculate the MVoT variables and update the MVoT. On the right side of the figure, the CDTA takes in the MVoT values from the DTMC and evaluates them using message thresholds to decide to send out alert messages if necessary. The central MVoT aggregator also generates dashboard information for the GSP.

The DTM system has two main components, the Distributed Trust Model Client

(DTMC) and the Central Distributed Trust Aggregator (CDTA), situated at the Service Provisioning Customer (SPC) and the Grid Service Provider (GSP), respectively. The DTM at the SPC is part of the Local Area Network (LAN), and the CDTA is part of the Wide Area Network (WAN). Upcoming sections provides a closer look at the individual components of the DTM system shown in Figure 3.1.

3.3 Distributed Trust Model Client

The Distributed Trust Model Client (DTMC), located at the Service Provisioning Customer (SPC), monitors the DCM messages to and from the connected devices such as the GSP and the DER. The primary responsibility of the DTMC is to monitor any incoming and outgoing messages and to evaluate and update the MVoT variables. The two actors at the SPC are the DCM, and the DER.

3.3.1 Input Message

Figure 3.2 illustrates the distributed trust model system integrated into the energy grid of things. The DTMC is monitoring the DCM-GSP and DCM-DER. The DTMC reports to the CDTA its MVoT calculations. When the DCM sends or receives a message from an actor, it provides a summary of that message to the DTMC. In practice, the DCM provides the message to the DTMC in the form of Extensible Markup Language (XML) file.

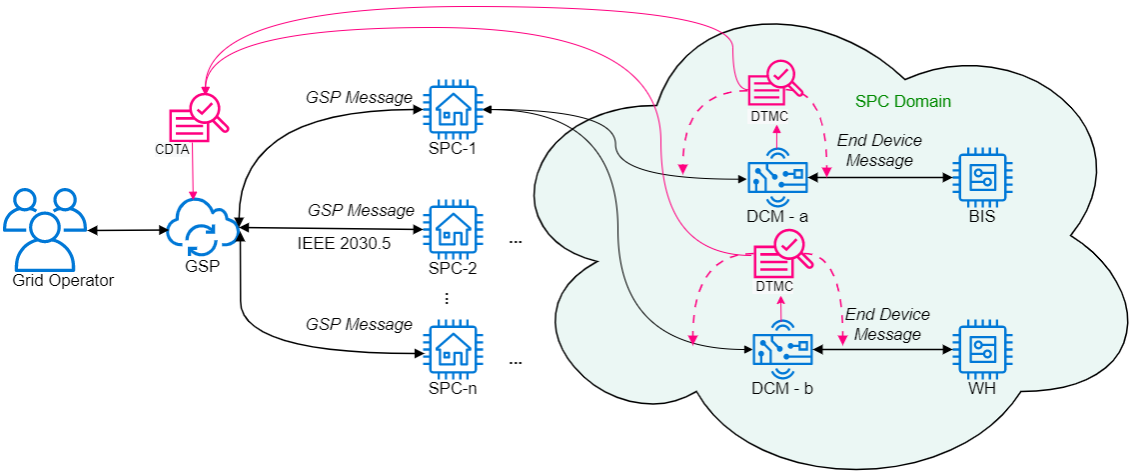


Figure 3.2: This figure shows the communication pathways within the EGoT system. The dotted lines point to communication channels the DTM is monitoring. The solid lines represent communication channels. Left: The CDTA is present at the GSP. Right: A DTMC is present at each DCM.

Classification	**Description**
Expected	All the message contents are present and not out-of-order.
Unexpected	If the message is out-of-order, repetitive, or contains extreme values.
Indeterminant	If the DTM is unable to precisely classify the message.
Disconnect	If a device, a DER, does not respond in a timely fashion.
Error	If any of the message contents are incorrect, missing, or the presence of any content discrepancy.
None	When remaining messages that does not fit into classification.

Table 3.1: This table describes the list of classifications generated from the $Input\ Classifier$ block

3.3.2 Input Classifier Block

The input classifier block classifies the raw input message to be $expected$, $unexpected$, $indeterminant$, $disconnect$, $error$ or $none$. Table 3.1 provides a possible list of classifications the $input\ classifier$ generates and the corresponding classification description. A message is classified as $Expected$ when all the appropriate message content is present, and that message is an appropriate response to a previously received message.

Another example of an expected message is when the DCM receives a message from the GSP with DER operations and scheduling. The DCM translates that message into a protocol the DER understands and sends it to the DER. There are exceptions for some anomalies in data, such as the DCM polling the GSP for resources of a specific function set or a client device requesting an available resource list of some sort. The DTMC most often classifies an input message of a DCM to be an $expected$ message.

If the messages to/from an actor are out-of-order, then the DTMC classifies the message as an $unexpected$ message. An example where a message is classified as $unexpected$ is when the GSP, an actor who shall not initiate contact to the DCM, sends a message to the DCM without receiving an initial request. Another example of an $unexpected$ message classification is when the GSP sends a flow reservation response list to a flow reservation request message from the DCM. The proper response from the GSP is to reply to the DCM informing the flow reservation response has been created with relevant event information. Finally, a message can be classified as $unexpected$ if the GSP responds with incorrect service information to a specific polling request of a DCM.

Indeterminant message is a classification assigned to a message that cannot be precisely defined as expected or unexpected or one of the other categories. The DTMC more likely classifies an input message of a DER, to be $indeterminant$. In practice there are many different input messages, and they cannot be precisely classified as expected, unexpected, etc. For example, if the GSP sent the same message twice, then the remaining messages are sent only once as expected, the input classifier acknowledges this message as

$indeterminant$. That may be due to a network glitch or other accidental cause. A message is classified as a $disconnect$ if the DER receives no hard beat response from the DCM. As a result in the DER disconnects from the DCM.

The DTMC classifies the input message as an $error$ if the message timing is a future time. Also the input message is classified as an $error$ if the content data type is incorrect, or if the actor information is missing or invalid. The input message classifies as $none$ when no classification is needed.

In addition to the input message classification, the input classifier block also output the ID of the actor who sent the message, the message send time, and the message transit time. All the output contents are formatted conveniently for the trust equation evaluation block to calculate trust variables and update the MVoT.

3.3.3 Trust Equations Evaluation Block

The trust equation evaluation block, shown in Figure 3.1, receives the classified input from the $input\ classifier$ block. Then the trust equation evaluation block pulls from the MVoT storage block, shown in Figure 3.1, the corresponding MVoT variables of the actor mentioned in the classified input. The $Trust\ Equation\ Evaluation$ block takes in the MVoT data and the input classifier content, and then recalculates the trust variables of the MVoT. The following section provides a detailed description of the MVoT variables.

3.3.4 Metric Vector of Trust Equations

Our MVoT consist of seventeen variables. Compared to other literature work, often only one or two variables are used. The MVoT consists of seventeen variables. Table 3.2 summarizes the variables of the MVoT used in this project. Each variable is adjusted using the contents of an incoming message. MVoT variables are then analyzed to detect if one or more abnormalities are present, which could be a signs that there may be an attack. These variables can be either codependent or independent based upon their relationships. The purpose of having a Trust Score (TS) and a Distrust Score (DS) is to ensure the distributed trust model decisions are looking at both the trustworthiness and the distrust worthiness of each actor. These variables help the DTM check if the message content data are out of bounds or not and whether to send an alert.

3.3.5 Trust Score

The trust score Equation 3.1 represents the overall trust score for each actor. Overall, the trust score equation considers several factors, such as the actor's historical behavior of delivering expected messages and unexpected messages and how confident the DTM is about this actor. These factors help the DTM to derive an estimate of overall trustworthiness for each actor. α is a weight that determines the influence of unexpected messages. The purpose of α is to increase or decrease the impact the unexpected message has on the overall trust score. The influence of unexpected messages has on the trust score is determined by the value of α. For a higher α, the trust score is less lenient about the impact

Item	Variable	Keyword	Definition
1	*TS*	Trust Score	Overall trust score for each actor
2	*DS*	Distrust Score	Distrust score for each actor
3	*C*	Certainty	How certain is the DTM for each evaluation
4	*CExMsg*	Count Of Expected Messages	Total count of messages that are expected for each actor
5	*CUnMsg*	Count Of Unexpected Messages	Total count of messages that are unexpected for each actor
6	*TotMsg*	Total Number Of Messages	Count of total messages
7	*Time_Stmp*	Time of the Last Message Received	Time of the most recent message received from the actor
8	*Regstr_Time*	Registration Date (Unix Time)	This can be the first time a message is received from an actor.
9	*ComFreq*	Frequency of Communication	How often an actor communicates with the DCM.
10	*TX_Tme*	Measured Transit Time	Time different for message to travel from the source to destination.
11	*Avg_TX_Tme*	Average transaction Time	Expected transaction time is average transaction time.
12	*TSLC*	Time Since Last Communication	Time delta of the last message is received.
13	*SDTT*	Standard Deviation Of Txn Time	Extent of deviation for Transit time as a whole
14	*RFC*	Relative Factor of Certainty	Certainty indicator of lean toward or against trust score or distrust score.
15	*T_Out*	Count Of Timeouts	Total count of timeouts for each actor (Protocol specific. E.g., 2030.5).
16	*C_Alrt*	Count Of Alerts	Total count of alerts sent out to each actors.
17	*C_Other*	Count Of Other Actions (TBD)	For future: designing of additional actions sent out to actors.

Table 3.2: Table providing a list of Metric Vector of Trust (MVoT) and the corresponding definitions used in this research

of adverse events.

$$TS(i+1) = [CExMSG(i+1) - (\alpha \times CUnMSG(i+1))] \times C(i+1) \tag{3.1}$$

3.3.6 Distrust Score

The distrust score is different from trustworthiness. The significance of maintaining a separate distrust score is to have a clear awareness of how untrustworthy an actor is. The distrust score, Equation 3.2, is simply a function of the number of unexpected message times the current certainty value for the actor.

$$DS(i+1) = CUnMSG(i+1) \times C(i+1) \tag{3.2}$$

3.3.7 Certainty

Each actor has a certainty score representing how confident the DTM is for each evaluation. For each actor, the certainty is a function of the actor's Relative Factor of Certainty (RFC), the current certainty value, and the communication frequency (ComFreq) value multiplied together, divided by the actor's Time Since Last Communication (TSLC). Gamma (γ) is a weight value that determines how fast the added messages influence certainty. The term $1 - e^{(\gamma - TotMsg)}$ is set up for the increase in total messages has an increased influence on the certainty. Having a certainty score helps the DTM factor in the general confidence it has in calculating MVoT variables.

$$C(i) = \left(RFC \times \left(\left(1 - e^{(-\gamma \times TotMsg)}\right) \times \frac{ComFreq}{max\_ComFreq} \right) \right) \times \frac{min\_TSLC}{TSLC} \quad (3.3)$$

3.3.8 Relative Factor of Certainty

The Relative Factor of Certainty (RFC), Equation 3.4, indicates how certain one is that the indicator leans toward or against a trust score or a distrust score. In the RFC equation, the ratio of the expected messages and the sum of unexpected messages and the count of expected messages is subtracted by 0.5 to understand how far away the results are from the 50% certainty. The derived ratio indicates that there are enough messages received to be confident about the evaluation. Suppose this ratio is much less than the 50% range there for leaning toward uncertainty. The β sets the maximum value of what RFC can be. For example, if $\beta = 1.6$, then the maximum RFC value is 0.8.

$$RFC = \left| \frac{CExMsg}{CExMsg + CUnMsg} - 0.5 \right| \times \beta \quad (3.4)$$

3.3.9 Expected, Unexpected and Total Message Count

The DTM MVoT keeps a separate count of messages classified as expected ($CExMsg$), unexpected ($CUnMsg$), and total ($TotMsg$) messages, Eq. 3.5 - 3.7. $TotMsg$ keep a count of all six different message classifications done by the DTMC. The trust score consists primarily of expected and unexpected messages. These values are part of the calculations of other MVoT variables. For example, suppose the number of expected messages is significantly larger than the total messages sent to the actor. In that case, it is a clear sign of how unreliable an actor's communication is.

$$CExMsg(i+1) = CExMsg(i) + 1 \tag{3.5}$$

$$CUnMsg(i+1) = CUnMsg(i) + 1 \tag{3.6}$$

$$TotMsg(i+1) = TotMsg(i) + 1 \tag{3.7}$$

3.3.10 Current Time and Time Stamp

The MVoT stores the time stamp ($Time\_Stmp$), Eq. 3.8, of when a message is received. This value keeps track of the last time an actor communicated. The device registration time, Eq. 3.9, quantifies how long the device has been participating in grid services.

$$Time\_Stmp(i+1) = CurrentTime \tag{3.8}$$

$$Regstr\_Time = Time\_Stmp(i) \tag{3.9}$$

3.3.11 Communication Frequency

The communication frequency, Eq. 3.10, provides the communication frequency of actor messages per unit of time. This variable helps identify if there has been a drastic change in the way an actor is communicating.

$$ComFreq = \frac{TotMsg}{CurrentTime - Regstr\_Time} \tag{3.10}$$

3.3.12 Average Transit Time

The average transaction time (Avg_TX_Tme), Eq. 3.11, represents the mean value of message transaction time. The DTM uses average transaction time to calculate expected transaction time, where the x_i represents the actor's i^{th} message transaction time, Eq. 3.11 keeps track of incremental average transaction time, Eq. 3.12, the incremental transaction time of the n+1 event, Avg_TX_Tme, evaluates if there is too much of a delay in messages from an actor by comparing the message transaction time (TX_Tme) with the incremental average message transaction time(Avg_TX_Tme).

$$\mu_n = \frac{1}{n}\sum_{i=1}^{n}(x_i) \tag{3.11}$$

$$\mu_{n+1} = \frac{n \times \mu_n + x_{n+1}}{n+1} \tag{3.12}$$

3.3.13 Time Since Last Communication

Time Since Last Communication ($TSLC$), keep track of the time since an actor last communicated. $TSLC$ helps understand if the actor is taking longer, shorter, or just the right amount of time to send messages. In addition, $TSLC$, Eq. 3.13, helps detect any anomalies in an actor's communication interval.

$$TSLC = Time\ delta\ of\ the\ last\ message\ received \tag{3.13}$$

3.3.14 Standard Deviation of Transit Time

Equation 3.14 and Equation 3.15 are the Standard Deviation of Transit Time (SDTT) and standard deviation of transaction time applied to the current time, respectively. Equation 3.15 incrementally updates the standard deviation with each new message from an actor. At the same time, Equation 3.15 shows the measured volatility of the actor's transaction time. Abnormalities in the message transaction time are a critical indicator of a compromised actor.

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_n)^2} \tag{3.14}$$

$$\sigma_{n+1} = \sqrt{\frac{n \times (\sigma_n)^2 + (x_{i+1} - \mu_n)(X_{i+1} - \mu_{n+1})}{n+1}} \tag{3.15}$$

3.3.15 Time Outs

The MVoT also keeps a count of timeouts, Eq. 3.16. An example of a time out is when a DER timed out if there is no response or a heartbeat signal from the DCM within a specific time window. The indication of timeout count is a sign of compromised nodes, primarily if other actors in the network are functioning normally.

$$T\_Out(i+1) = T\_Out(i) + 1 \tag{3.16}$$

3.3.16 Count of Alerts

The count of alerts, Eq. 3.17 helps the DTM keep track of types of alerts and the count of specific alerts sent out for each actor. Designers of the MVoT can expand the count of variable to keep track of each type of alert message sent out by the DTM.

$$C\_Alrt(i+1) = C\_Alrt(i) + 1 \tag{3.17}$$

3.3.17 MVoT Variables Summary

The Certainty (C) calculation captures how certain the Distributed Trust Model Client is of each evaluation. The frequency of communication, $ComFreq$, is how often the actor communicates with other actors. The message transit time variable TX_Tme measures the amount of time it takes for a message to travel from the source to the destination. The average transit, as the name, maintains the average message transit time per actor. The $Time\ Since\ the\ Last\ Communication$ ($TSLC$) calculates its length since the DTM received a message. Count of timeouts keeps a count of the amount of time the actor timed out.

The count of expected messages, $CExMsg$, keeps track of messages classified as expected. The count of unexpected messages, $CUnExMsg$, tracks messages classified as unexpected by the input classifier. The variable total number of messages, $TotMsg$ keeps track of all messages each actor sent individually. The timestamp, $Time\_Stmp$, variable keeps track of the most recent time each actor communicated. The variable device

registration time, $Regstr\_Time$, stores when the initial message from an actor is received. Count of alert, C_Alrt , variable keeps track of each alert sent from the Distributed Trust Model. The standard deviation of transit time, $SDTT$ keeps track of the distribution of message transit time for each actor.

The variables count of expected messages ($CExMsg$), count of unexpected messages ($CUnExMsg$), total number of messages ($TotMsg$), timestamp ($Time\_Stmp$), device registration time ($Regstr\_Time$), count of alerts (C_Alrt), and standard deviation of message transit time ($SDTT$) are variables that are codependent to the variables mentioned above. The DTMC at each DER calculates the MVoT for that DER and sends the MVoT to the CDTA.

3.4 Central Distributed Trust Aggregator

The Central Distributed Trust Aggregator (CDTA) is located at the Grid Service Provider (GSP). The CDTA is responsible for aggregating and analyzing all MVoT. It collects MVoT data from the DTMCs and sends out recommendations, and alarms to authorities based upon its evaluation to the GSP. It does not make decision or take actions. Additionally, the CDTA provides a dashboard of scores to the GSP. Components of the CDTA include the $decision/actions/recommender$ block and the $central\ MVoT\ aggregator$, and $dashboard$.

3.4.1 Central Distributed Trust Aggregator MVoT

The CDTA sends an alert to the GSP of any abnormal activities it observes. The CDTA monitors abnormalities in a more significant population of actors to a specific data category before alerting the GSP.

Consider how the CDTA decides to send an alert message to the GSP. As mentioned earlier in this chapter, when the DTMC gets a message from an actor, the $input\ classifier$ classifies the message as expected, unexpected, indeterminant, disconnects, error, or none. After the DTMC executes the $trust\ equation\ evaluation$, the CDTA $central\ MVoT\ aggregator$ collects all the DTMC MVoT for which it is responsible.

3.4.2 Decision/Action/Recommender Block

The DTMC $trust\ equation\ evaluation$ block pulls the actor MVoT data from the storage and evaluates the actor MVoT variables. DTMC $trust\ equation\ evaluation$ block then updates the stored MVoT variables with the new evaluated results. The DTMC MVoT block passes these evaluated results to the CDTA to analyze among the greater population.

The CDTA $decision/action/recommender$ block has a threshold for each alert message to check against any abnormalities and alert the GSP. If the set threshold values are exceeded, that means there are abnormalities. In that case, the CDTA $decision/action/recommender$ block sends an alert message to the GSP with a corresponding alert message.

3.4.3 Chapter 3 Summary

The distributed trust model augments security in a communication network. The DTM system architecture designed in this research is a hybrid of centralized and distributed trust models. Hybrid DTM enables individual actors to observe abnormalities and report to a central node to look for large-scale abnormalities of many actors. Additionally, the expandable MVoT with n variables enables the DTM system to observe numerous abnormalities.

4 Hypothesis Testing

4.1 Overview

The purpose of applying hypothesis testing to the DTM System is to set count and value thresholds and evaluate the equations. Hypothesis analysis of the CDTA performance helps the user understand the impact of selected count and value thresholds have on the CDTA's decision to send an alert to the GSP. The CDTA decides on whether to send an alert message to the GSP. Two thresholds help choose to send an alert to the GSP, value threshold and count threshold. The value threshold is a value selected for each MVoT variable. The value threshold checks if the reported MVoT value of an actor is greater than the set value threshold. It is an anomaly if the actor MVoT data exceeds the value threshold. For the hypothesis analysis, the count of anomalies were compared against a count threshold for each actor's corresponding MVoT variable.

The count of anomalies for each MVoT variable is flagged if it exceeds the count threshold. Then, the CDTA sends an alert message to the GSP informing of a system anomaly. The list of messages that the CDTA may send to the GSP are shown in Table 4.1. When the CDTA evaluate the MVoT data and decides to send an alert message to the GSP it is called Positive. After evaluating MVoT data and when the CDTA decides not to send an alert message it is called Negative. The essential contributions of this process are to provide

a methodology and a tool to help set effective decision-making thresholds of when an alert shall be sent to the GSP and implement a method that evaluates trust equations and MVoT equations. Table 4.2 summarizes the terminology used in this chapter and its definition.

Messages from the Central Distributed Trust Aggregator to the GSP
"Excessive time since last communication from GSP for n SPCs."
"Excessive time since last communication from DER for n SPCs."
"Trust is low for GSP from n SPCs."
"Trust is low for DCM for n SPCs."
"Trust is low for DER for n SPCs."
"Trust is low for DTM, self assessment for n SPCs."
"Communication rate is low from GSP for n SPCs."
"Communication rate is low from DER for n SPCs."
"Communication rate is low for DCM for n SPCs."
"Communication rate is excessive from GSP for n SPCs., possible DoS."
"Communication rate is excessive from DER for n SPCs., possible DoS."
"Communication rate is excessive for DCM for n SPCs., possible DoS."
"Slow transit time for messages from GSP for n SPCs."
"count discrepancy for each MVoT value except: ExMsG, Timestamp, Registration, T_out, C_Alert, and C_other)."

Table 4.1: The table representing messages from the Central Distributed Trust Aggregator to the GSP

Term	Definition
Confusion Matrix	A methodology to evaluate the performance of an classification model.
Confusion Metric	Multiple metrics derived from the confusion matrix to evaluate the performance of the system such as CDTA.
Binary Classification	A binary classifier makes a decision between two hypothesis. Such as null hypothesis and alternative hypothesis.
Positive	When the CDTA sends an alert message to the GSP.
Negative	When the CDTA does not sends an alert message to the GSP.
True Positive	The CDTA sends an alert message to the GSP at an event where an alert rightfully should be sent out.
True Negative	The CDTA did not send an alert message to the GSP at an event where an alert rightfully should not be sent out.
False Positive	The CDTA sends an alert message to the GSP at an event where an alert rightfully should not be sent out since there are no attack(s).
False Negative	The CDTA did not send an alert message to the GSP at an event where it should have due to attack(s).
Null Hypothesis	When the CDTA sends an alert message to the GSP.
Alternative Hypothesis	When the CDTA does not send an alert message to the GSP.
Count Threshold	The threshold set for the count of actors exceeding a specific MVoT variable value.
Value Threshold	The threshold selected to check if the reported MVoT value of an actor is greater than the specified value threshold.

Table 4.2: This table describes the terminology used in hypothesis testing.

4.2 Confusion Matrix

Often used in Machine Learning, a confusion matrix is used to analyze system performance. The confusion matrix is a visualization of a classification model. In this thesis, the confusion matrix analyzes the CDTA performance when sending alert messages.

4.2.1 Confusion Matrix Terminology

Figure 4.1 illustrates the confusion matrix for a binary classifier applied to the CDTA process of evaluation to send an alert to the GSP. A binary classifier decides between two hypotheses. In this thesis, the two decisions are a null hypothesis versus an alternative hypothesis, such as it is considered positive when the CDTA sends an alert message to the GSP versus negative when the CDTA did not send an alert message to the GSP. In this thesis, the confusion matrix visualizes the ability of the CDTA to correctly evaluate and alert the GSP when there is an attack present. Several keywords need to be understood to understand how the confusion matrix applies to evaluating alert messages sent to the GSP.

It is considered True Positive (TP) when the CDTA sends an alert message to the GSP at an event where an alert should be sent out. It is considered True Negative (TN) when the CDTA did not send an alert message to the GSP at an event where an alert should not be sent out. False Positive (FP) is when the CDTA sends an alert message to the GSP at an event where an alert should not be sent out since there are no attacks present. False Negative (FN) is when the CDTA did not send an alert message to the GSP at an event where it should have because of attacks present. False-positive is mistaken rejection of a trust worthy entity is known as Type I error. False-negative is mistaken acceptance of a untrustworthy entity is known as Type II error.

and false-negative is mistaken rejection are also known as Type I error and Type II error.

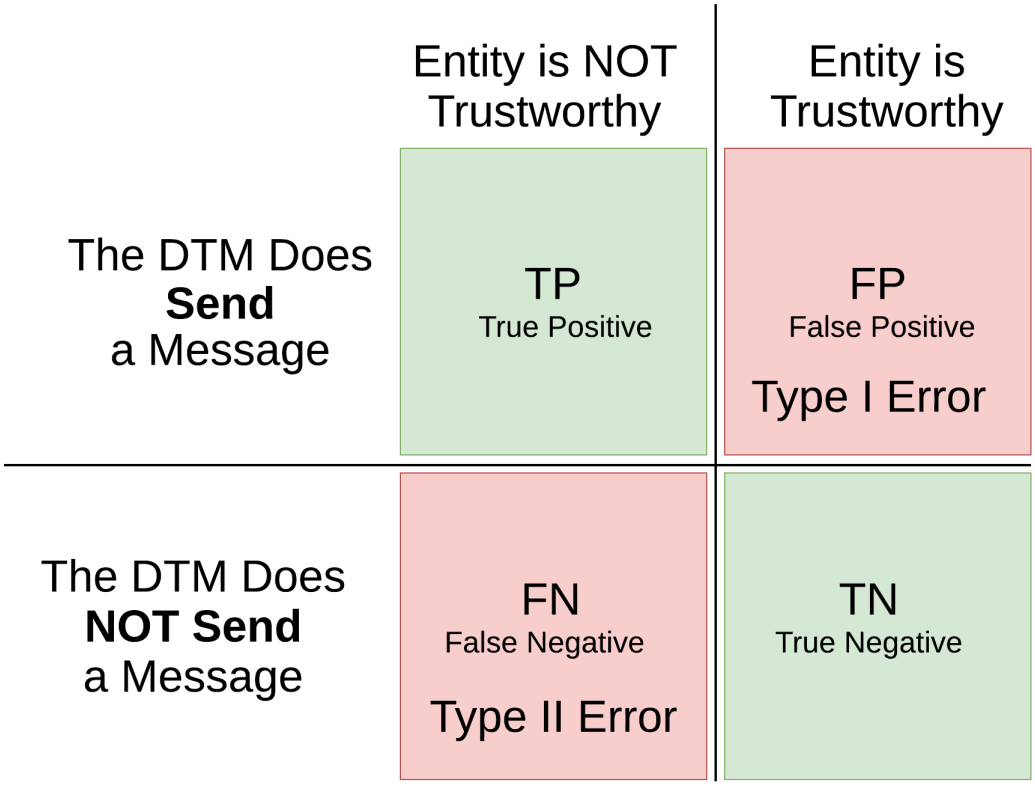


Figure 4.1: Figure showing Trust evaluation for actual trustworthiness. The rows represents the trust evaluations and the columns represents actual trust. The Green area represents correct binary classification and the red are represents error in classification.

4.2.2 Confusion Matrix Applied to the CDTA

An example where the CDTA decides whether to send a message out to the GSP helps understand the evaluation process of the CDTA and its decision-making performance. This subsection covers how TP, TN, FP, and FN scenarios for how the CDTA determines the excessive time since the last communication.

As mentioned earlier, the term "positive" applies to identify whether the CDTA sent an alert message to the GSP. The CDTA sends an alert message to the GSP based on real-time

data in a scenario where the count of actors exceeds the value threshold of time since last communication, greater than the count threshold. Thus, the CDTA has a list of alert messages, a corresponding count, and a value threshold that checks against each actor's MVoT variables. CDTA sends the message if the count of anomalies in each alert category is greater than the specific alert count threshold.

Anomalies occur when a particular measured value exceeds the value threshold. Each decision/action alert message count and value threshold of the CDTA have a selected value assigned along with a corresponding alert message. A method is needed to select the count threshold because the accuracy of the decision to send an alert to the GSP depends of the count threshold. Therefore, an essential contribution of this thesis provides a method that helps weigh the outcome of setting the value or the count threshold. The DTM System hypothesis tool helps evaluate thresholds. This tool generates plots representing confusion metrics, while changing the count or the value threshold to determining when the CDTA alerts the GSP. Generating selective confusion metric plots helps approximate and visualize when it is too early to alert the GSP and when it is too late. This scenario is like the analogy of "crying wolf." If there are not a considerable amount of alerts, then the CDTA should not alert the GSP. The CDTA shall not wait too long to alert the GSP. If the CDTA waits too long to alert the GSP, damage because of an attack may already occur.

This thesis provides a tool that applies binary classification to all the MVoT variable data to decide the rate of correctly identifying any data abnormalities and alerting the GSP. An example is how binary classification applies to the MVoT variable TSLC data. It is true

positive when the CDTA correctly measures the count of actors surpassing the value threshold for TSLC to be greater than the count threshold set for TSLC, and there is an attack. In an event of a TP the CDTA alerts the GSP with an allocated message. The designated message for this example is “excessive time since last communication” to inform the GSP about the increase in time for an actor to communicate.

It is a true negative when the CDTA correctly measures the count of actors surpassing the value threshold for TSLC to be less than the count threshold set for TSLC, and there is no attack. In such a case, the CDTA does not send an alert message.

It is a false positive when the CDTA incorrectly measures the count of actors surpassing the value threshold for TSLC to be greater than the count threshold set for TSLC. There is no attack present and alerts the GSP of excessive TSLC.

It is false negative when the CDTA measures the count of actors surpassing the value threshold for TSLC to be less than the count threshold set for TSLC. There is an attack present although the CDTA does not alert the GSP of excessive TSLC. Table 4.3 summarizes the above description of the binary classification applied to the MVoT variable in a case there is an attack or not.

Binary Classification	The count of actors exceeding the value threshold for a specific time is greater than the actor count threshold	Is there an attack for that specific time?	Did the CDTA Alert the GSP?
TP	Yes	Yes	Yes
FP	Yes	No	Yes
TN	No	No	No
FN	No	Yes	No

Table 4.3: This table describes the confusion matrix’s binary classification categories.

4.2.3 Application of Value Threshold to the Central MVoT Aggregator Block

There are three possible scenarios to evaluate the CDTA's performance using hypothesis testing. This thesis uses scenario 1 and 2.

Scenario 1. Hold the value threshold constant and vary the count threshold, as shown in detailed description of Scenario 1 in Section 4.2.3.1.

Scenario 2. Change the value threshold and have the count threshold constant, as shown in Section 4.2.3.2.

Scenario 3. Vary both the count threshold and the count threshold.

The MVoT variable TSLC data for 30 actors in a set period of 40 hours is used to understand scenarios one and two. The two main variables are the count of actor threshold, called TSLC_N, with excessive TSLC delay. The second variable is TSLC value threshold, TSLC_Th, which look at the percentage of reported TSLC values surpassing the TSLC Value threshold, TSLC_Th, causing a delay. These two variables contribute to two plots wherein each plot, TSLC_N or TSLC_Th, stays constant, and the other varies.

4.2.3.1 Scenario 1

The variable TSLC_N is held constant in this scenario, and the TSLC_Th varies, as shown in Figure 4.2. Thus, each hour a record is maintained of a count of actors exceeding the value threshold, TSLC_Th, and an alert message sent to the GSP.

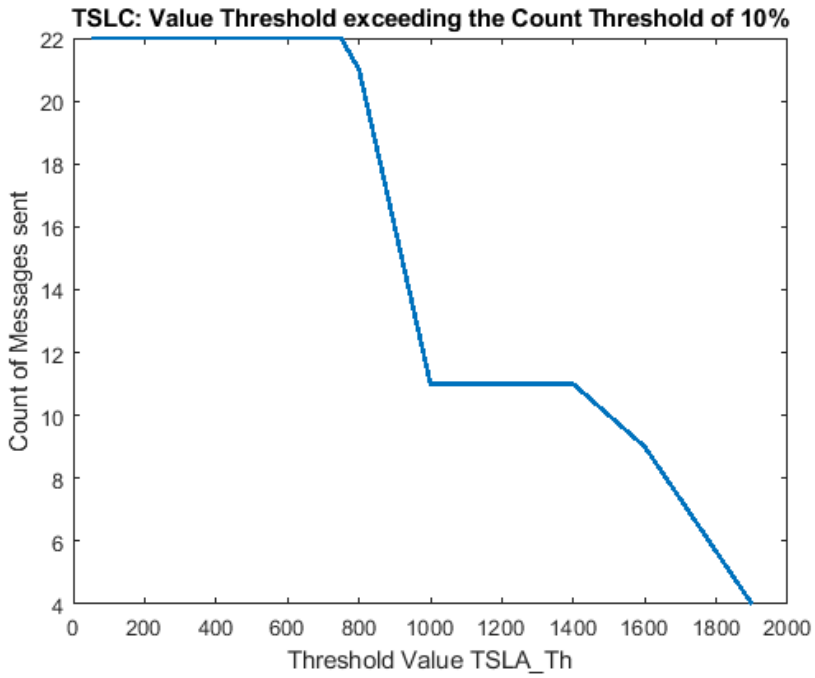


Figure 4.2: Plot showing TSLC data from the central MVoT aggregator where the TSLC_Th value varies and TSLC_N, the count of actors surpassing the count threshold, remains constant.

4.2.3.2 Scenario 2

In the second scenario, the variable TSLC_N varies, and the TSLC_Th is held constant at 1900 sec, as shown in Figure 4.3. For a given time window, such as 30 hours in this scenario, the count of actors exceeding each TSLC_N threshold varies.

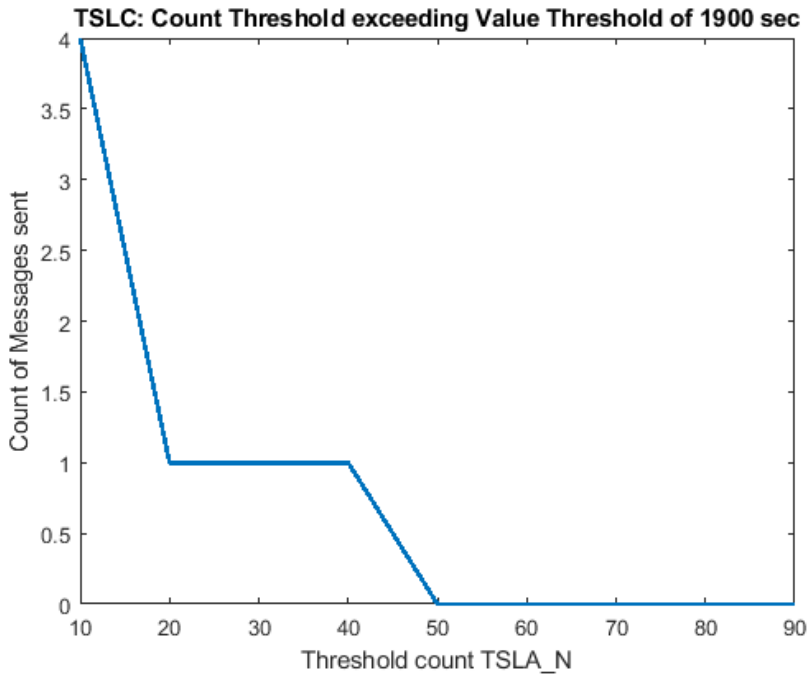


Figure 4.3: Plot showing TSLC data from the central MVoT aggregator where the TSLC_Th value is constant and TSLC_N, the count of actors surpassing the count threshold, varying.

4.2.3.3 Scenario 3

In the third scenario, both the variable TSLC_N and the TSLC_Th vary, as shown in Table 4.4. Thus, resulting plots show the outcome when both thresholds, TSLC_N and the TSLC_TH, vary.

	FPR of TSLC_Th (Sec)					FNR of TSLC_Th (Sec)				
Count Threshold	95	190	285	380	475	95	190	285	380	475
1	0	0	0	0	0	1	1	1	1	1
2	0	0	0	0	0	1	1	1	1	1
3	0	0	0	0	0	1	1	1	1	1
4	0	0	0	0	0	1	1	1	1	1
5	0	0	0	0	0	1	1	1	1	1
6	0	0	0	0	0	1	1	1	1	1
7	0	0	0	0	0.33	1	1	1	1	0.97
8	0	0	0	0	0.33	1	1	1	1	0.97
9	0	0	0	0	0.67	1	1	1	0.97	0.97
10	0	0	0	0	0.67	1	1	1	0.97	0.95
11	0	0	0	0.33	0.67	1	1	1	0.97	0.92
12	0	0	0	0.33	0.67	1	1	1	0.97	0.86
13	0	0	0	0.67	0.67	1	1	1	0.95	0.84
14	0	0	0.33	0.67	0.67	1	1	0.97	0.89	0.81
15	0	0	0.33	0.67	0.67	1	1	0.97	0.89	0.78
16	0	0	0.33	0.67	0.67	1	1	0.91	0.84	0.76
17	0	0	0.67	0.67	0.67	1	1	0.89	0.76	0.76
18	0	0	0.67	1	1	1	0.95	0.89	0.76	0.73
19	0	0	0.67	1	1	1	0.92	0.86	0.76	0.73
20	0	0.33	0.67	1	1	1	0.86	0.78	0.32	0.27
21	0	0.33	0.67	1	1	1	0.86	0.76	0.27	0.22
22	0	0.33	0.67	1	1	1	0.81	0.66	0.19	0.16
23	0	0.33	0.67	1	1	1	0.78	0.64	0.19	0.16
24	0	0.67	0.67	1	1	0.97	0.7	0.60	0.19	0.10
25	0.33	0.67	0.67	1	1	0.86	0.62	0.51	0.19	0.05
26	0.67	0.67	0.67	1	1	0.86	0.62	0.51	0.05	0.03
27	1	1	1	1	1	0.86	0.51	0.46	0	0
28	1	1	1	1	1	0.73	0.43	0.41	0	0
29	1	1	1	1	1	0	0	0	0	0
30	1	1	1	1	1	0	0	0	0	0

Table 4.4: The table shows the actors exceeding the count threshold when both the actor count threshold TSLC_N and the value threshold TSLC_Th vary.

Example of binary classification applied to TSLC data: Table 4.5 provides an example of how binary classification is derived from a tally of actors exceeding the set TSLC_Th value 900 sec in four hours.

- **Hour 1:** the count of actors surpassing the threshold value is **greater** than the set count threshold, and there is **an attack**, then the binary classification is TP.
- **Hour 2:** the count of actors surpassing the threshold value is **greater** than the set count threshold, and there is **no attack**, then the binary classification is FP.
- **Hour 3:** the count of actors surpassing the threshold value is **less** than the set count threshold, and there is **an attack**, then the binary classification is FN.
- **Hour 4:** the count of actors surpassing the threshold value is **less** than the set count threshold, and there is **no attack**, then the binary classification is TN.

Hours	**Actors exceeding the set value threshold, TSLC_Th = 900 sec, each hour**	**Input:Attack = 1, no attack = 0**	**Binary classification when Count threshold = 1**
hour 1	16	1	TP
hour 2	18	0	FP
hour 3	1	0	TN
hour 4	0	1	FN

Table 4.5: The table shows the application of binary classification to the data of MVoT variable TSLC.

4.3 Confusion Metric Equations

The list of confusion metrics and their equation are provided in the next Sections 4.3.1 to 4.3.8.

4.3.1 Accuracy

Accuracy is the ratios of correct decisions made by the CDTA to alert the GSP versus the total count of decisions made by the CDTA, Equation 4.1.

$$accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{4.1}$$

4.3.2 Sensitivity

Sensitivity is the ratio of the CDTA correctly decided to send alerts compared to the actual incidents of CDTA should send out an alert, Equation 4.2.

$$sensitivity = \frac{TP}{(TP + FN)} \tag{4.2}$$

4.3.3 Precision

Precision indicates how often the CDTA correctly evaluated and sent an alert message to the GSP, Equation 4.3.

$$precision = \frac{TP}{(TP + FP)} \tag{4.3}$$

4.3.4 Specificity

Specificity indicates how often the CDTA correctly evaluated to not send a message to the GSP, Equation 4.4.

$$specificity = \frac{TN}{(TN + FP)} \tag{4.4}$$

4.3.5 F1 Score

F1 Score is the harmonic mean of the precision, and sensitivity, Equation 4.5. Ideally the F1 score shall be at one which shows the performance of the CDTA is at 100%.

$$F1\ score = \frac{TP}{TP + 0.5(FP + FN)} \tag{4.5}$$

4.3.6 False Discovery Rate

False Discovery rate is the probability the CDTA incorrectly evaluates and sends an alert message to the GSP, Equation 4.6.

$$FDR = \frac{FP}{(FP + TP)} \tag{4.6}$$

4.3.7 False Negative Rate

False Negative Rate (FNR) is the ratio the CDTA mistakenly decided not to send alerts compared to the actual incidents of CDTA should send out an alert, Equation 4.7.

$$FNR = \frac{FN}{(FN + TP)} \tag{4.7}$$

4.3.8 False Positive Rate

False Positive Rate (FPR) is the rate of incorrect alert messages sent by the CDTA compared to the count of evaluation where the CDTA should not send out alert messages, Equation 4.8.

$$FPR = \frac{FP}{(FP + TN)} \tag{4.8}$$

Of all the confusion metric equations mentioned in section 4.3, equations 4.1-4.8, this thesis focuses on three specific confusion metrics: 1) the EER derived from plotting the FNR versus the FPR, 2) the sensitivity, and 3) the F-1 score.

4.3.9 Equal Error Rate

The EER is the balancing point where the FN rate equals the FP rate. For example, in Figure 4.4 the EER occurs when FNR equals FPR approximately at 90% rate where the count threshold is 11.

The targeted EER is to be less than 50% for the first cycle of testing for initial testing. When the error rate is closer to 0%, this indicates that the CDTA evaluates an anomaly activity with very low error. Therefore, for the final testing cycle for CDTA performance evaluation, the measured error rate shall be closer to 5% or less. For this thesis, all the data used are simulated data.

FNR represents how often the CDTA failed to send an alert message to the GSP in an instances of an attack. **FPR** represents how often the CDTA alerting the GSP when there

was no attack. In an ideal case, the EER value is close to zero since that represents a lower error rate of failing to alert the GSP about an anomaly and warning the GSP when there is no anomaly. The next chapter shows the impact of having a lower EER for a set of MVoT variable data.

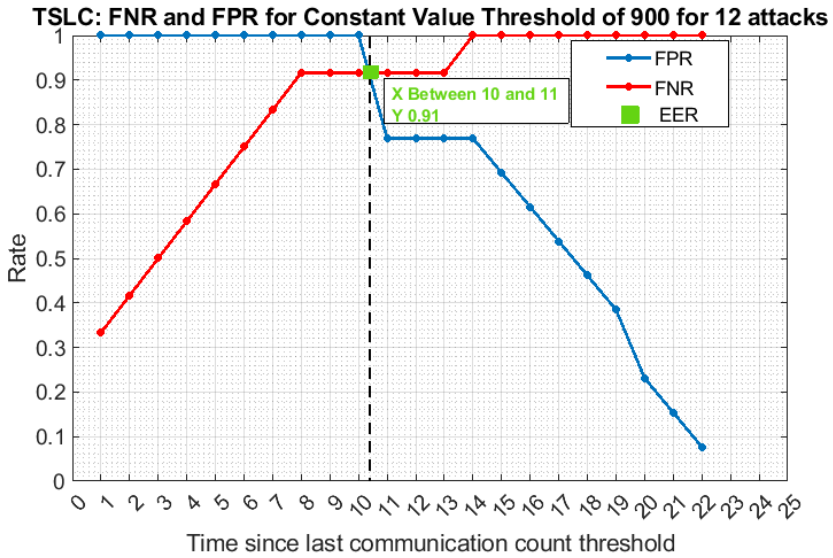


Figure 4.4: The plot of EER where the red line represents the FNR, and the Blue line represents FPR. In this case the EER is the point where the two lines are intersecting. The confusion metric variable FNR versus FPR showed the CDTA performance when the TSLC_TH held constant at 900 for 12 attacks.

4.3.10 Sensitivity

The sensitivity represents the alerts correctly measured to be true versus all the event alerts the CDTA sent to the GSP. Therefore, a lower FNR is preferred. Lower FNR indicates a lower likelihood of failing to send an alert when there is an attack.

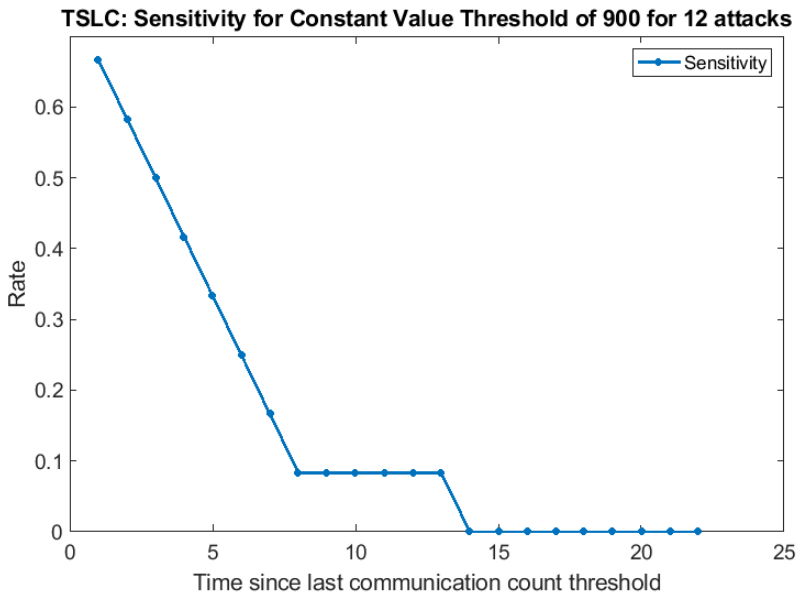


Figure 4.5: The plot of sensitivity curve showing CDTA performance when the TSLC_TH is held constant at 900 for 12 attacks.

4.3.11 F-1 Score

The F-1 score is the harmonic mean of precision and sensitivity. The F-1 score represents the likelihood that the CDTA correctly sends an alert message to the GSP versus the sum of all the messages sent and 50% of total failure to alerts and false alerts.

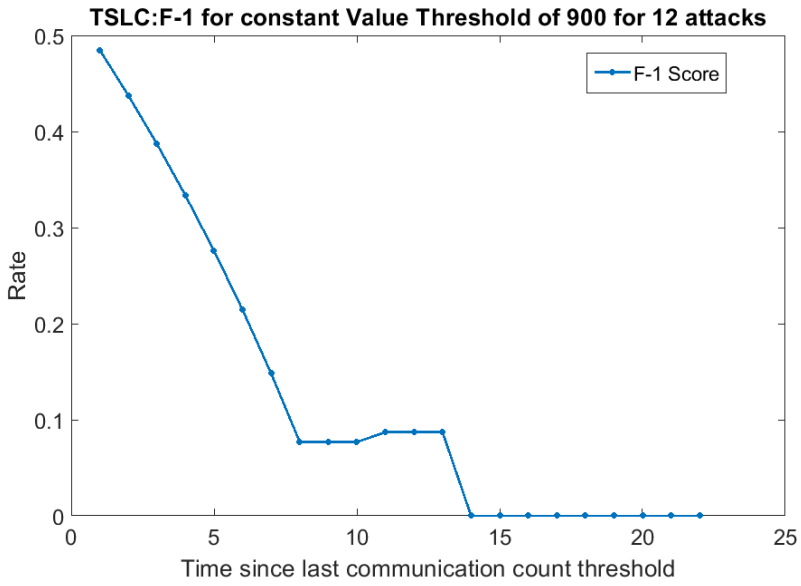


Figure 4.6: The plot of F-1 score curve shows the CDTA performance when the TSLC_TH is 900 for 12 attacks.

4.4 Dashboard or Presentation of Data

In Distributed Energy Resource Management System (DERMS), there is a dashboard that presents data to the GSP. The dashboard shows DERMS operations. Hypothesis testing is applied before the procedure to set thresholds and after operation to reevaluate the thresholds. Dashboard plots help with the reevaluation phase. For example, for each MVoT variable, there is a time-varying cumulative distribution for different actors. Each time-varying cumulative distribution plot shows the probability that a corresponding MVoT variable have a value less than or equal to the value on the y axis.

The second plot that the dashboard provides is the cumulative distribution at a snapshot of a time, such as the current hour. Dashboard data helps verify the hypothesis testing for a condition: whether the CDTA send a message to the GSP or not. For example, suppose the

CDTA evaluation results in sending a message to the GSP because of a TSLC violation. In that case, there is an attack. Both the time-varying cumulative distribution and the snapshot of cumulative distribution show that 5% or less of the actors are violating the TSLC value threshold. It is a clear sign that the selected count threshold and value threshold need to be adjusted. The example shows how the hypothesis testing is applied before the operation to set thresholds and after operation to reevaluate the thresholds.

4.5 Test Profiles

The hypothesis testing uses simulated data generated from the trust model data generator. There are several profiles the trust model data generator is capable of creating. Trust model data generator profiles and their descriptions are mentioned below.

- All_expected — a profile that generates only expected message profiles data where all the message profiles contents are occupied with valid data and not out-of-order.
- All_Unexpected — a profile that generates unexpected message profiles data such as messages that are out-of-order, repetitive, or contains extreme values.
- Almost_good — a profile that generates expected message profiles until a threshold shows where the unexpected message profiles generates.
- Almost_bad — a profile that generates unexpected messages until a user-specified threshold is reached to generate expected messages.

- Random — a profile that generates random classified message profiles for random actors.
- Mixed — a profile that generates a user-specified combination of two or more message profile classification categories.
- User_specified — a profile that generates manually entered user-specified message classification category for an actor.

4.6 Hypothesis Test Elements

According to Sudhamathy and Venkateswaran, "hypothesis testing is the theory, methods, and practice of testing a hypothesis by comparing it with the null hypothesis. The null hypothesis is only rejected if its probability falls below a predetermined significance level, in which case the hypothesis being tested is said to have that level of significance [29]."

- **Application** — The hypothesis testing applies to the CDTA evaluation and correct identification measured abnormalities surpassing the set abnormalities thresholds in the central MVoT aggregator and send alert message to the GSP.
- **Theory** — In theory, the CDTA shall send an alert to the GSP if the number of actors with abnormal values at a given time exceeds the count threshold and there is an attack present. Thus, this would be the only scenario in which CDTA sends an alert message to the GSP.

- **Practice** — There are possibilities for the CDTA to make errors and fail to send an alert or incorrectly send an alert to the GSP.

- **The Null Hypothesis** — It is expected that the CDTA correctly sends an alert message to the GSP when there is an attack.

- **The Alternative Hypothesis** — Alternate expectation of the CDTA performance is when the CDTA does not send an alert message to the GSP when there are no attacks.

- **The Independent Variable** — Variables that vary in this testing are

  1) The count attacks for a given time, such as an hour,

  2) Count threshold and or the value threshold.

- **The Dependent Variable** — There are the variables that help measure the CDTA evaluation performance and send messages to the GSP. Dependent variables are TP, FP, TN, FN, FNR, FPR, Sensitivity, F1 score.

- **The Control Group** — represents scenarios where the CDTA is evaluating all expected data without any abnormal actor counts surpassing the count threshold. Results from all expected messages situations serve as a baseline to compare against different experiments conducted. This thesis presents two scenarios that belong to the control group.

  Scenario 1: Contains ten all-expected message classification profiles with different times steps.

Scenario 2: Contains ten profiles with-all expected message classification profiles with identical time steps.

Additional Scenarios: Contains a combination of ten profile sets where each set consist of a variation of the following combinations:

- Message time step ranges between one minute and one hour.
- The time gap of each message has an identical time-step between one minute to one hour.
- Three profiles with three unexpected messages from the same actor and the remaining seven with expected messages.
- Three profiles with three unexpected messages from the random actor and the remaining seven with expected messages.
- Three profiles with six unexpected messages from the same actor and the remaining seven with expected messages.
- Three profiles with six unexpected messages from the random actor and the remaining seven with expected messages.
- Three profiles with six indeterminant messages from the random actor and the remaining seven with expected messages.
- Three profiles with six indeterminant messages from the same actor and the remaining seven with expected messages.

  - The occurrence of six or three unexpected messages randomly for each profile with the remaining messages are classified as expected.
  - The occurrence of six or three unexpected messages is at the middle section of each profile, with the remaining messages classified as expected.
  - The occurrence of six or three unexpected messages to be at the ending part of each profile with unexpected messages.

- **The Experimental Group** — The experimental group is absent from this thesis since there are no real data available from the DTM System prototype. This thesis uses data generated from the DTM Simulator.

The plot in Figure 4.7 shows a scenario where the CDTA is not sending a message to the GSP when it should not have sent a message at a zero rate. A false negative zero rate is ideal and illustrated in Figure 4.7 appropriately since the DTM system is undergoing zero attacks. A false-negative rate is at 100% until the count threshold, N, is six. The false-negative rate curve remains at zero rate since it is not possible failure to alert of an attack when there are no attacks present. The false positive rate drops to zero only when the N count threshold increases to 18, resulting in zero false alarms sent to the GSP.

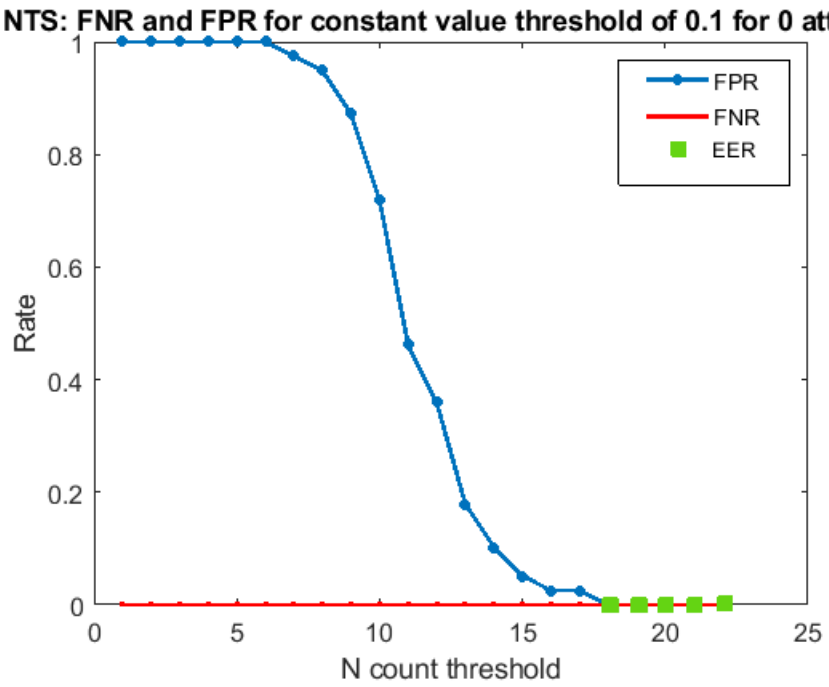


Figure 4.7: The plot shows the normalized trust score FNR and FPR plots where the trust score value is constant, and the number of actors surpassing the count threshold varies depending on where the system is not under attack.

4.7 Summary

In summary, hypothesis testing helps evaluate the performance of the CDTA thresholds. If the thresholds are too lax, then the probability of the CDTA sending an alert at the correct time is high. On the other hand, if it is too strict, then the likelihood of the CDTA sending alert messages is very low. A better resolution can be derived to set the CDTA thresholds using the confusion matrix binary classification and hypothesis testing.

5 Results & Analysis

5.1 Overview

The four possible results of binary classifications are True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). The FP (FPR) and FN (FNR) help analyze the CDTA performance when sending GSP alerts. A simple statement of FNR applied to the CDTA would be the CDTA's failure to alert the GSP in case of an attack. The idea of FPR applied to the CDTA would be the CDTA's falsely alerting the GSP without any attack present. The point where the FNR equals the FPR is known as Equal Error Rate (EER). The EER shall be lower in this analysis for a better performing system, and its occurrence shall also be at a lower threshold level. If the EER is high, then the balancing point of the rate of errors the CDTA makes when deciding to alert the GSP increased. Hence the preference for lower EER. The importance of a lower threshold value ensures the earlier detection of abnormalities of the system. This analysis looks at how the F1 score applied to evaluate the CDTA's performance on sending alerts to the GSP.

5.2 Understanding Hypothesis Analysis Plots

Figure 5.1 is an example FNR and FPR behavior. In this plot, there is more than one Equal Error Rate (EER) point. When there are many EER points, the number of choices to select a

threshold increases. Choosing a threshold value somewhere between a series of EER points causes less room for error since the left and right of the threshold value also represent EER and provide less room for error due to a shift in the set threshold caused by mistake.

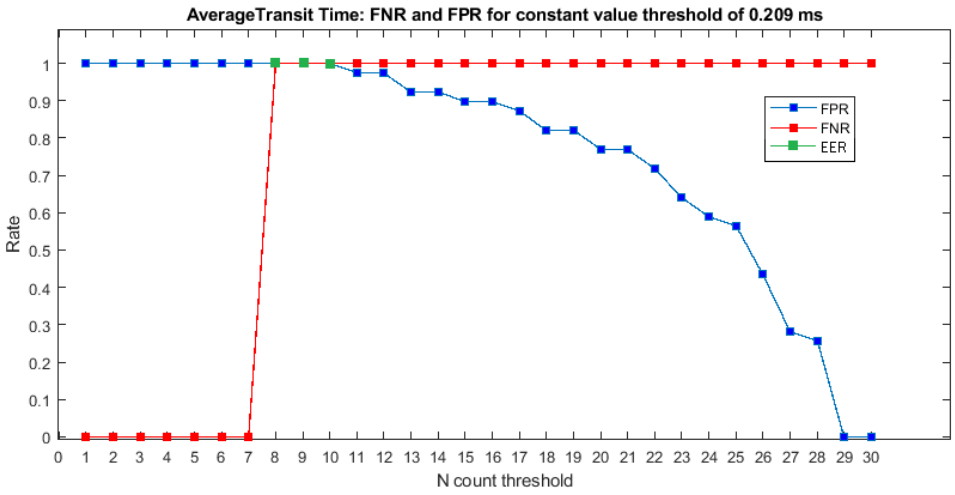


Figure 5.1: This figure illustrates an FNR and FPR plot with multiple EER points.

In Figure 5.1, left of the EER points the FNR is zero and FPR is one. Thus, a value threshold to the left of the EER points is preferred where the FNR is zero and FPR is one. Toward the right of the EER points, the false-negative rate is one, and the FPR gradually decreases to zero. Thus, a system that prefers fewer false alarms benefits from picking a threshold value to the right of the EER. To have an EER of one is not recommended, although many EER's are present in Figure 5.1. Having an EER of one shows that FPR and FNR are at 100%. Ideally, it is preferred to have an EER closer to 0%.

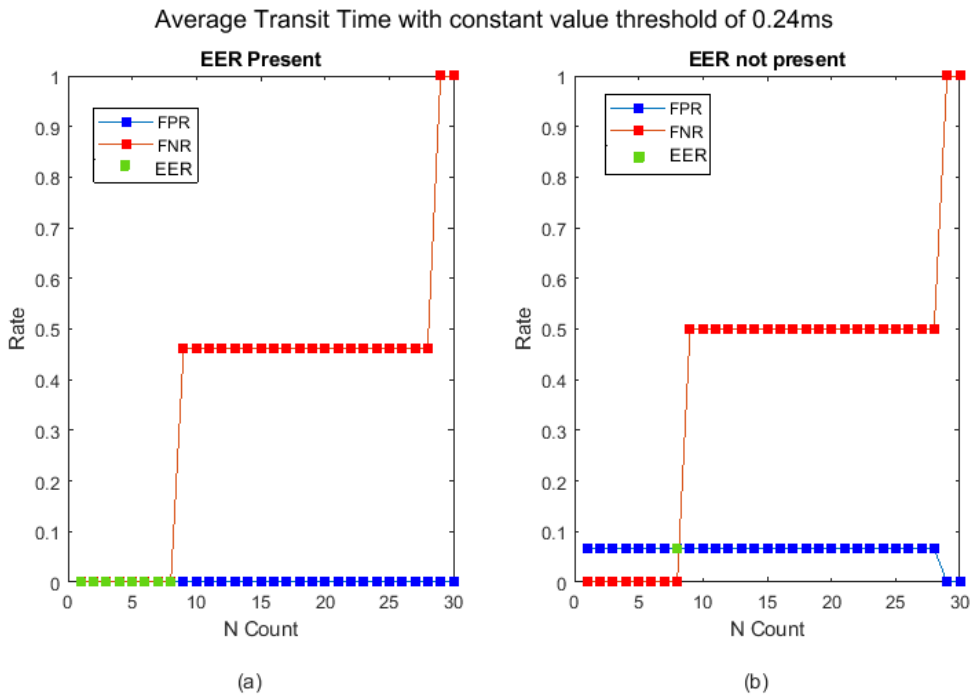


Figure 5.2: This figure is an example of an FNR and FPR plot where (a) there are multiple EER points for multiple count threshold, and (b) only one EER point present.

Figure 5.2 (a) represents an ideal scenario where EER of zero is present for N = [1...8]. Having an EER of zero is ideal due to the zero error rate, and having a lower EER for many threshold values is also preferred to reduce the impact from the set threshold slightly shifting to the right or the left.

Figure 5.2 (b) represents a scenario where the EER occurs once. In this scenario, the FNR and FPR rates are much closer to each other and at a rate close to 0%. A system that have multiple threshold points at the EER is at a disadvantage and might not have any other choice but to pick that one threshold value where FNR and FPR are equal. Having only one EER can be susceptible to errors if there is even a slight shift in threshold point chosen at the EER point. When there is a series of EER points it is recommended to pick a threshold in

the middle to avoid any errors since the EER points exist to the left and the right side of the selected threshold point. Figure 5.2 can be helpful for a system that prefers having FPR and FNR rates closer to zero instead of a definite EER point. Another critical point is that the F1 score has a drastic change after the EER but does not approach zero after the occurrence of EER; instead, it mimics the FPR curve returns to zero when the FPR curve transitions zero.

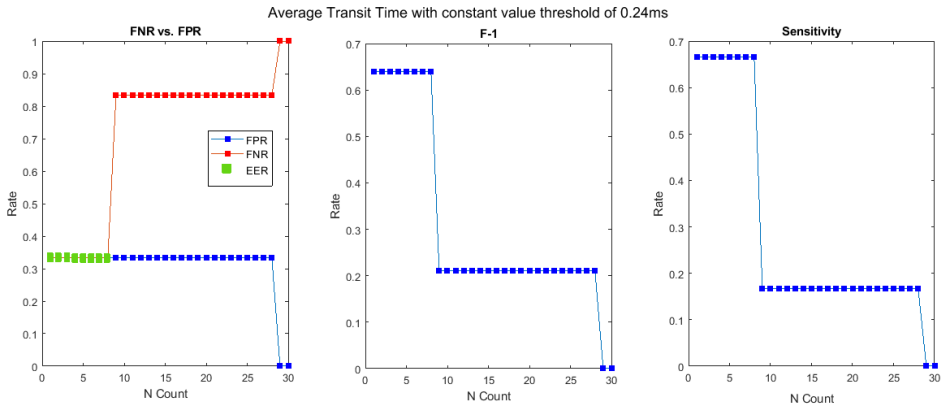


Figure 5.3: This figure shows a comparison of the EER, F1 score and the sensitivity.

5.3 Applying Hypothesis Testing to the MVoT variables

This section presents the procedure for conducting hypothesis testing to evaluate CDTA performance when sending an alert to the GSP. This section present the findings from observing metrics such as FNR and FPR plots, F1 score plots, and sensitivity plots for a particular set of simulated MVoT variable data under three types of attacks. The three different attack patterns are no attacks, one attack conducted for three separate hours, and one attack conducted for six individual hours. The specific MVoT variables observed in this analysis are Trust Score, Distrust Score, Certainty, Relative Factor of Certainty, Transit Time, Average Transit Time, Communication Frequency, and Time Since Last

Communication. Appendix B lists the 26 profile IDs used for the hypothesis testing analysis. Each profile ID consist of ten SPCs labeled SPC A through SPC J. Each SPC represents the messages sent from DTMC at that SPC to the CDTA. Each DTMC monitors and classifies message exchange between three actors: the GSP, the DCM, and the DER. The DTMC sends the message classification information and corresponding MVoT calculations to the CDTA. Message classifications observed for this analysis include expected messages, unexpected messages, and indeterminant messages. Table 5.1 - 5.3 provides a summary of each profile IDs. Tables identify the count of unexpected messages and indeterminant messages. The table also identifies if those messages are from the same or different actors. The table also includes such messages occurring in the beginning, middle, end, or random—the collection of messages are in ascending order of when the DCM received them at a specified time.

Profile ID#	Description
1	SPC A - J all expected.Each message has a 5 minute time step.
2	SPC A - J all are expected.Each message is between 1 minute to 1 hour.
3	SPC A - F are all expected. SPC's G - J has 6 unexpected messages in the beginning, middle, end, and randomly for different actors with message time step between: 1 min - 1 hour.
4	SPC A - F is all expected. SPC's G - J has 6 unexpected messages in the beginning for different actors with message time step between: 1 min - 1 hour.
5	SPC A - F is all expected. SPC's G - J has 6 unexpected messages. in the middle for different actors with message time step between: 1 min - 1 hour.
6	SPC A - F is all expected. SPC's G - J has 6 unexpected messages in the end for different actors with message time step between: 1 min - 1 hour.
7	SPC A - F are all expected. SPC's G - J has 6 unexpected messages in the beginning, middle, end, and randomly for same actors with message time step between: 1 min - 1 hour.
8	SPC A - F is all expected. SPC's G - J has 6 unexpected messages in the beginning for same actors with message time step between: 1 min - 1 hour.
9	SPC A - F is all expected. SPC's G - J has 6 unexpected messages. in the middle for same actors with message time step between: 1 min - 1 hour.
10	SPC A - F is all expected. SPC's G - J has 6 unexpected messages in the end for same actors with message time step between: 1 min - 1 hour.
11	SPC A - F are all expected. SPC's G - J has 3 unexpected messages in the beginning, middle, end, and randomly for different actors with message time step between: 1 min - 1 hour.
12	SPC A - F is all expected. SPC's G - J has 3 unexpected messages in the beginning for different actors with message time step between: 1 min - 1 hour.
13	SPC A - F is all expected. SPC's G - J has 3 unexpected messages. in the middle for different actors with message time step between: 1 min - 1 hour.

Table 5.1: The profile IDs and corresponding descriptions. Appendix B provide a set of tables with detailed descriptions of each profile

Profile ID#	Description
14	SPC A - F is all expected. SPC's G - J has 3 unexpected messages in the end for different actors with message time step between: 1 min - 1 hour.
15	SPC A - F are all expected. SPC's G - J has 3 unexpected messages in the beginning, middle, end, and randomly for same actors with message time step between: 1 min - 1 hour.
16	SPC A - F is all expected. SPC's G - J has 3 unexpected messages in the beginning for same actors with message time step between: 1 min - 1 hour.
17	SPC A - F is all expected. SPC's G - J has 3 unexpected messages. in the middle for same actors with message time step between: 1 min - 1 hour.
18	SPC A - F is all expected. SPC's G - J has 3 unexpected messages in the end for the same actors with message time step between: 1 min - 1 hour.

Table 5.2: The profile IDs and corresponding descriptions. Appendix B provide a set of table with detailed description of each profile.

Profile ID#	Description
19	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for the same actor: beginning, middle, and end. SPC I - J has an unexpected message at the end and mix for the same actor with message time step between: 1 min - 1 hour.
20	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for the same actor: at the beginning. SPC I - J has an unexpected message at the beginning for the same actor with message time step between: 1 min - 1 hour.
21	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for the same actor: at the middle. SPC I - J has an unexpected message at the middle for the same actor with message time step between: 1 min - 1 hour.
22	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for the same actor: at the end. SPC I - J has an unexpected message at the end for the same actor with message time step between: 1 min - 1 hour.
23	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for different actors: beginning, middle, and end. SPC I - J has an unexpected message at the end and mix for different actors with message time step between: 1 min - 1 hour.
24	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for different actors: at the beginning. SPC I - J has an unexpected message at the beginning for different actors with message time step between: 1 min - 1 hour.
25	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for different actors: at the middle. SPC I - J has an unexpected message in the middle for different actors with message time step between: 1 min - 1 hour.
26	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for different actors: at the end. SPC I - J has an unexpected message at the end for different actors with message time step between: 1 min - 1 hour.

Table 5.3: The profile IDs and corresponding descriptions. Appendix B provide a set of tables with detailed descriptions of each profile.

5.3.1 Key Observations of FNR and FPR Curves

Figure 5.4 shows the comparison results of FNR and the FPR curves when zero attacks are applied to the profile ID#1 Trust Score as the value thresholds increase from 12.02 to 228.05. The FNR remains a zero rate, and the FPR curve has a constant non-zero rate until the count threshold reaches 29. The observed behavior of the FNR curves of Figure 5.4 is identical to the FNR curves of all other profile ID MVoT variables. The shape of the FPR curve changes based upon the MVoT variable data. Equation 4.8 shows the FPR derived from FP divided by FP + TN. The FP occurs when the count of values exceeding the value threshold is greater, and there is an attack present; if an attack is not present, it is a TP. Therefore, it makes sense to observe unique FPR curves for each MVoTs of each provided ID with the criteria mentioned above. Changes to the shape of the curve are due to the variance in evaluated MVoT data, as shown in Figure 5.5.

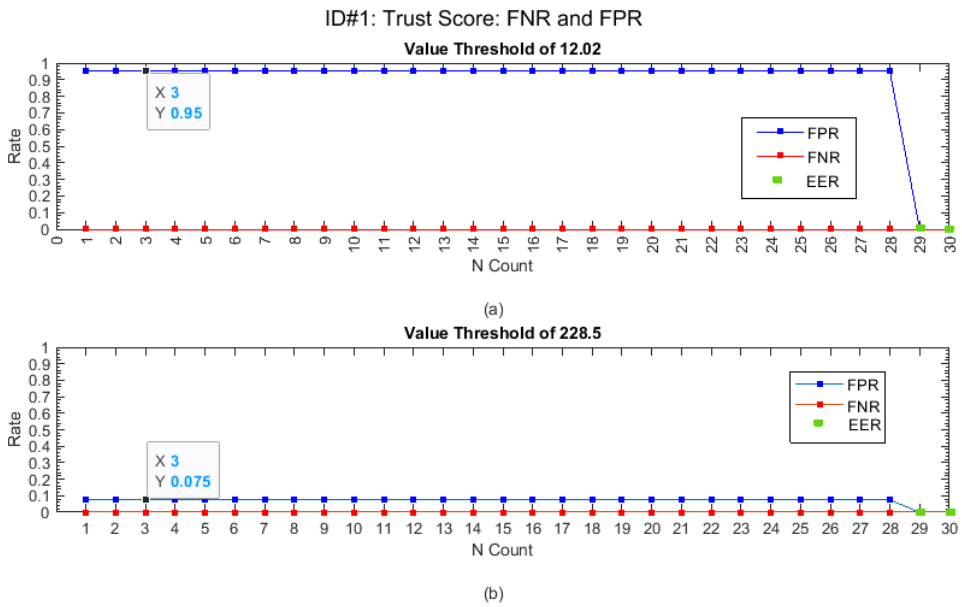


Figure 5.4: Figure illustrates the change in FNR and FPR when the value threshold changes for Profile ID#1 Trust Score when there are no attacks present. The FNR and FPR curves (a) when the value threshold is 12.02 and (b) where the value threshold is 228.5.

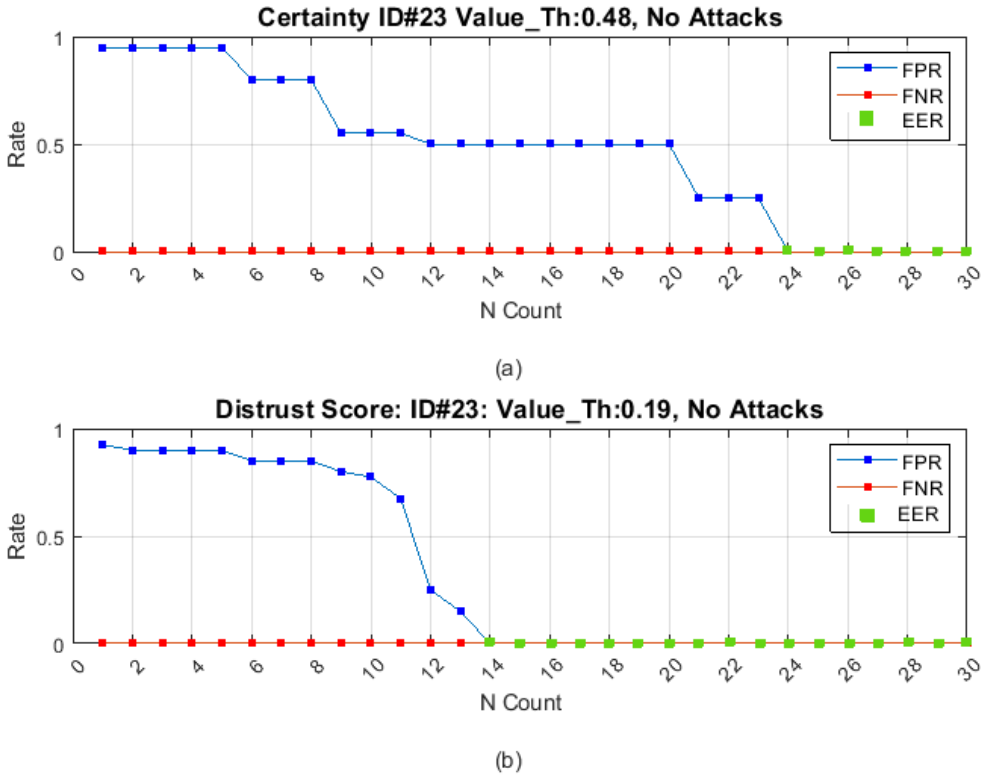


Figure 5.5: ID#23: This figure illustrates the change in the FPR curve for Different MVoT variables and value thresholds when there are no attacks present.

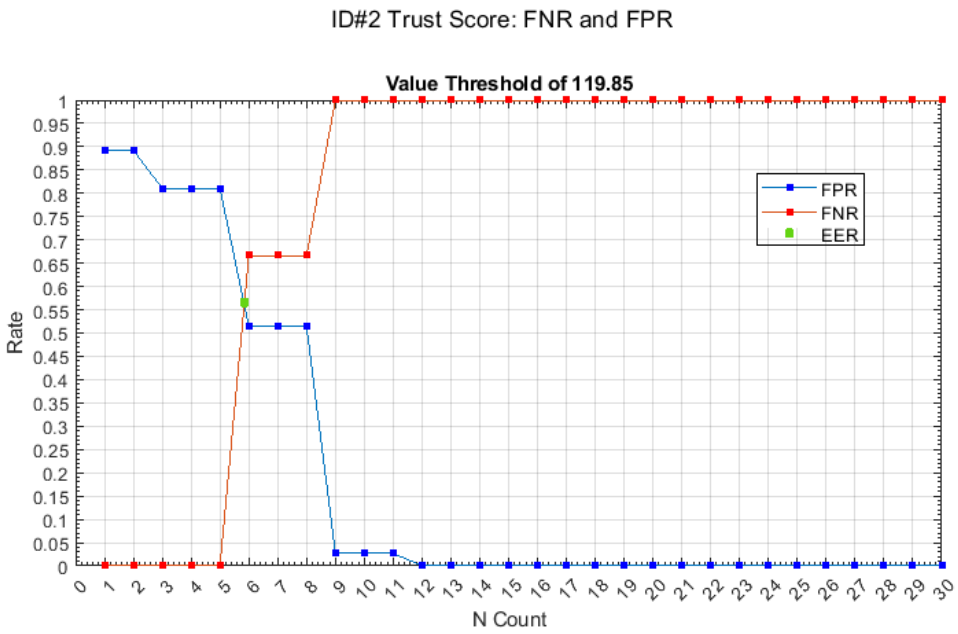


Figure 5.6: Trust Score of Profile ID#2: This figure illustrates the occurrence of one EER point where the FNR and FPR curves are overlapping.

Figure 5.6 illustrates an interesting phenomenon where multiple FPR and FNR points are parallel to each other. Even when the value threshold increases or decreases, there was never a specific value threshold that allowed those parallel points to intersect. The FNR and FPR points for count thresholds of 6,7 and 8 were approximately 0.15 points away from each other.

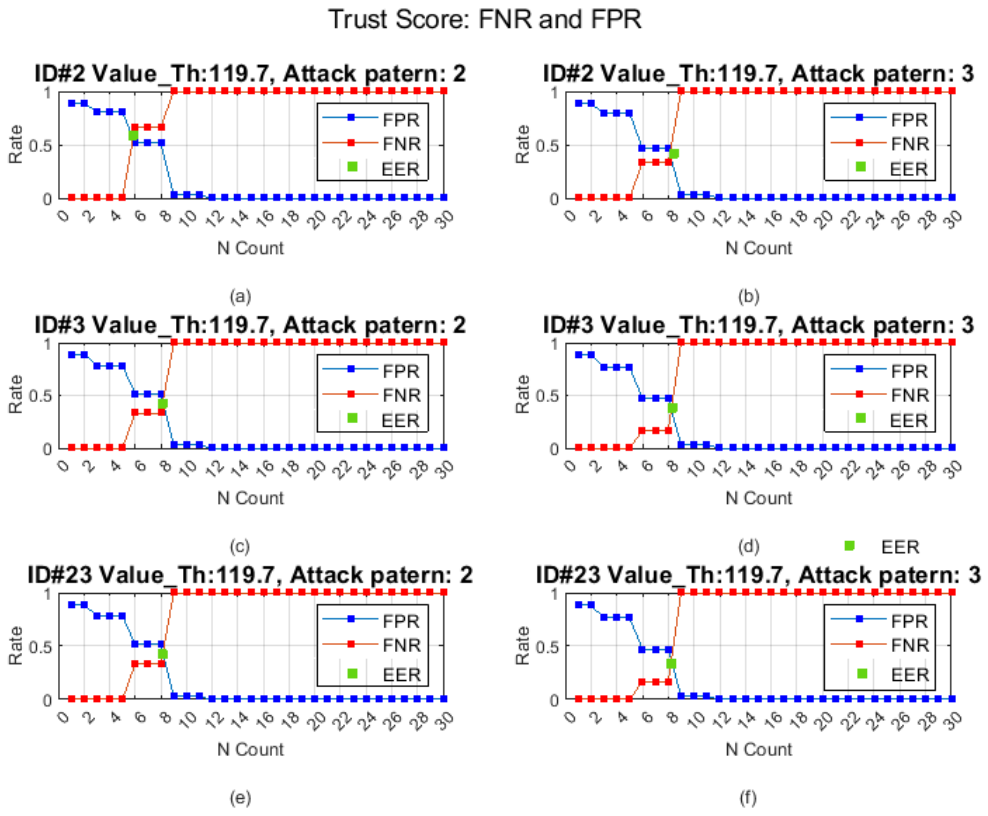


Figure 5.7: This figure illustrates the comparison of FNR and FPR curves for a selective set of MVoT variable Trust Scores for attack patterns 2 and 3 (a) attack pattern 2 applied to Profile ID#2, (b) attack pattern 3 applied to Profile ID#2, (c) attack pattern 2 applied to Profile ID#3, (d) attack pattern 3 applied to Profile ID#3 (e) attack pattern 2 applied to Profile ID#23 (f) attack pattern 3 applied to Profile ID#23.

Figure 5.7 shows the effect of attack patterns 2 and 3 on Profile ID# 2, 3, and 23's MVoT variable data for a threshold value of 119.7. Variation in attack patterns result in a slight change to the EER points and the FNR and FPR curves between each MVoT variable data.

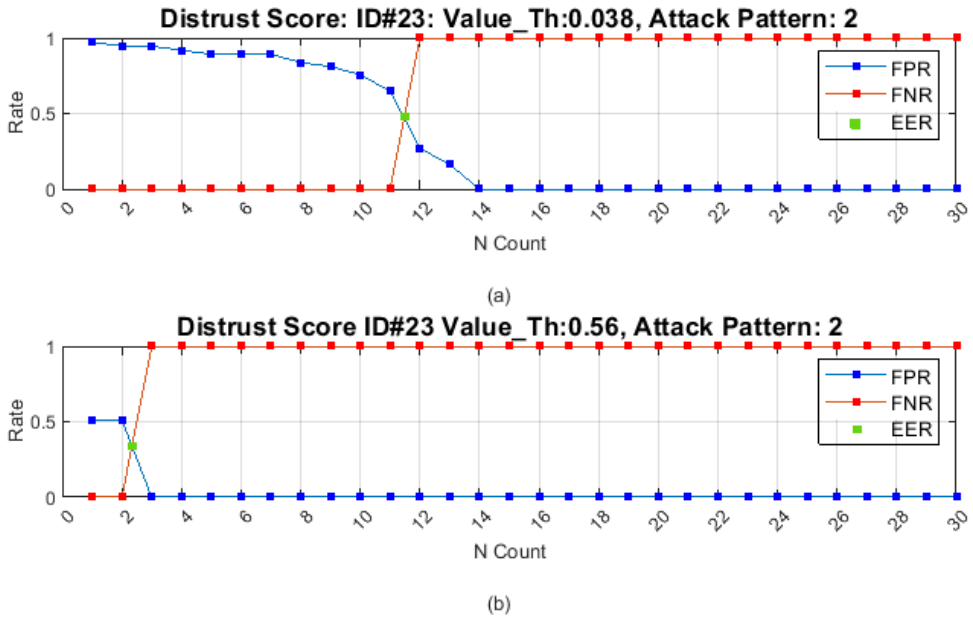


Figure 5.8: The figure shows the lowering of EER as the value threshold increase for profile ID23'S Distrust score.

Figure 5.8 shows that the EER decreases as the value threshold increases. At the same time, there is a decrease in threshold count where the EER occurs. However, the occurrence of EER changes based on the MVoT variable data. Figure 5.4 is an example where the lower was present at a higher count threshold.

5.3.2 Key Observations of F1 Score Plots

Figure 5.9 shows the resulting F1 score and sensitivity rates when no attacks are applied to the profile IDs 1 through 26. The F1 score and the sensitivity rates stayed continuously zero. Furthermore, attack pattern 1 applied to the simulated data of other MVoT variables showed that the resulting F1 score and sensitivity plots were also zero.

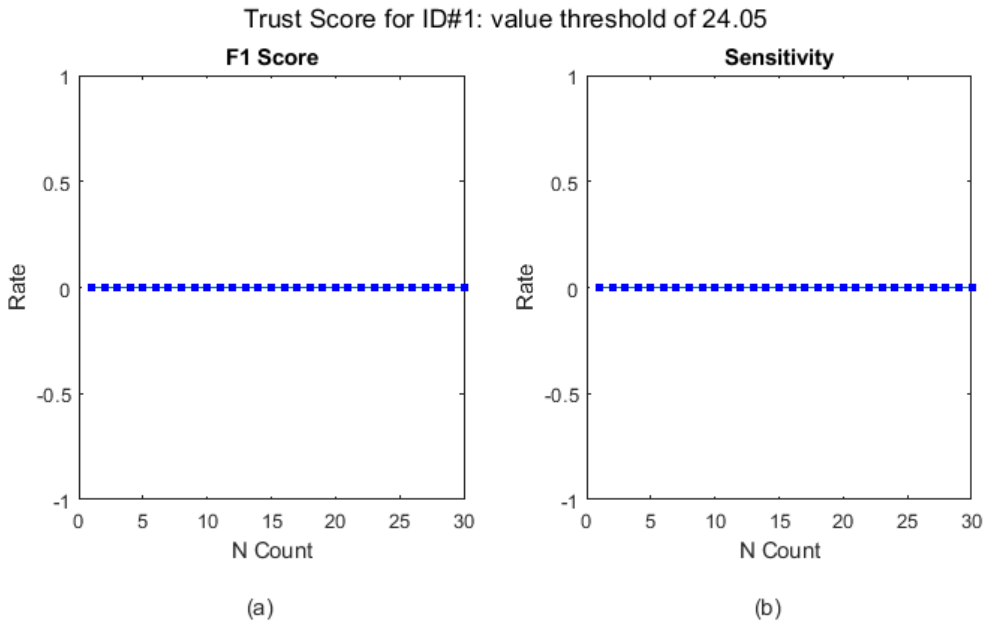


Figure 5.9: Certainty Profile ID#1: For value threshold of 24.05 without any attacks present: (a) F1 score and (b) Sensitivity.

There was no direct correlation between the lowest EER and the high F1 score. For example, a lower EER did not result in the highest F1 score. Instead, the direct correlation between the EER and the F1 score was that the F1 score drops to zero between the same count thresholds where the EER occurs, only if the FPR drops to zero, Figure 5.10. On the other hand, if the FPR did not fall to zero, then the F1 score settles where the FPR curve is, Figure 5.3. This phenomenon was present for all the MVoT variable data. This confirms the data that is used for the FNR, FPR, F1 score analysis has a direct impact on the shape of the resulting plots.

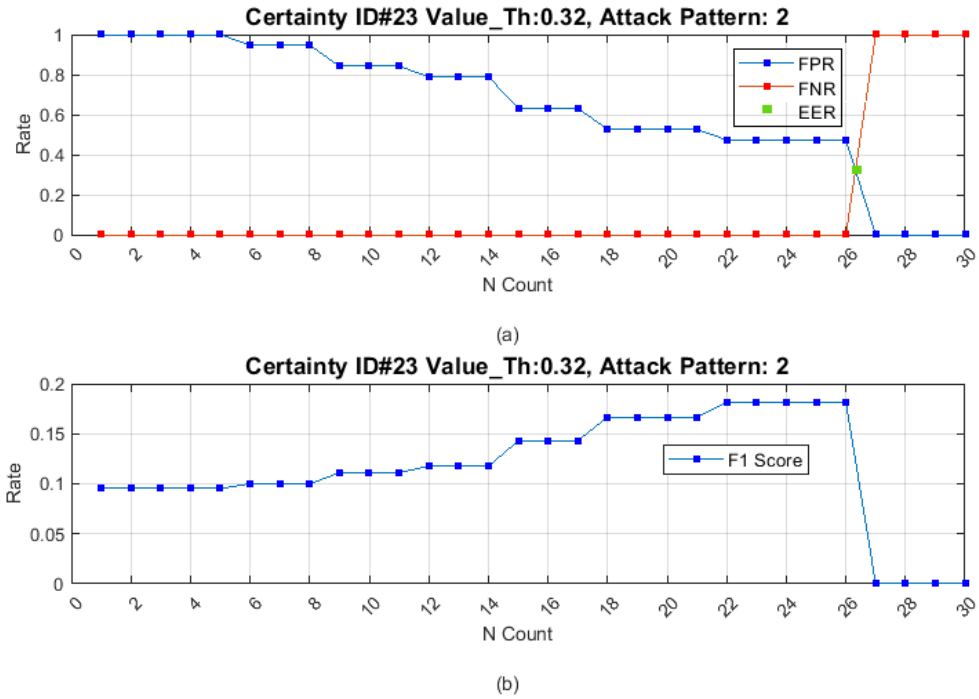


Figure 5.10: Certainty Profile ID#23: For value threshold of 0.32 with attack pattern 2: (a) FPR and FNR curve intersecting at the EER (b) F1 score transitioning to zero between the count thresholds 26 and 27 where the EER is occurring.

Figure 5.11 shows a snippet of how the F1 score curve changes as the value changes. The changes in the F1 score occurred at different locations of the chart for each plot of Figure 5.11 (a) - (b). This figure shows that the maximum F1 score does not occur at the maximum value threshold. The reason for the inconsistent in F1 score plots is due to the MVoT data that is analyzed. The analyzed data are simulated and cannot derive identical F1 score curves. Instead, they are simulated data to mimic real-world data. When real-world data are available, the behavior of the F1 score curve is unique as the MVoT variable data changes.

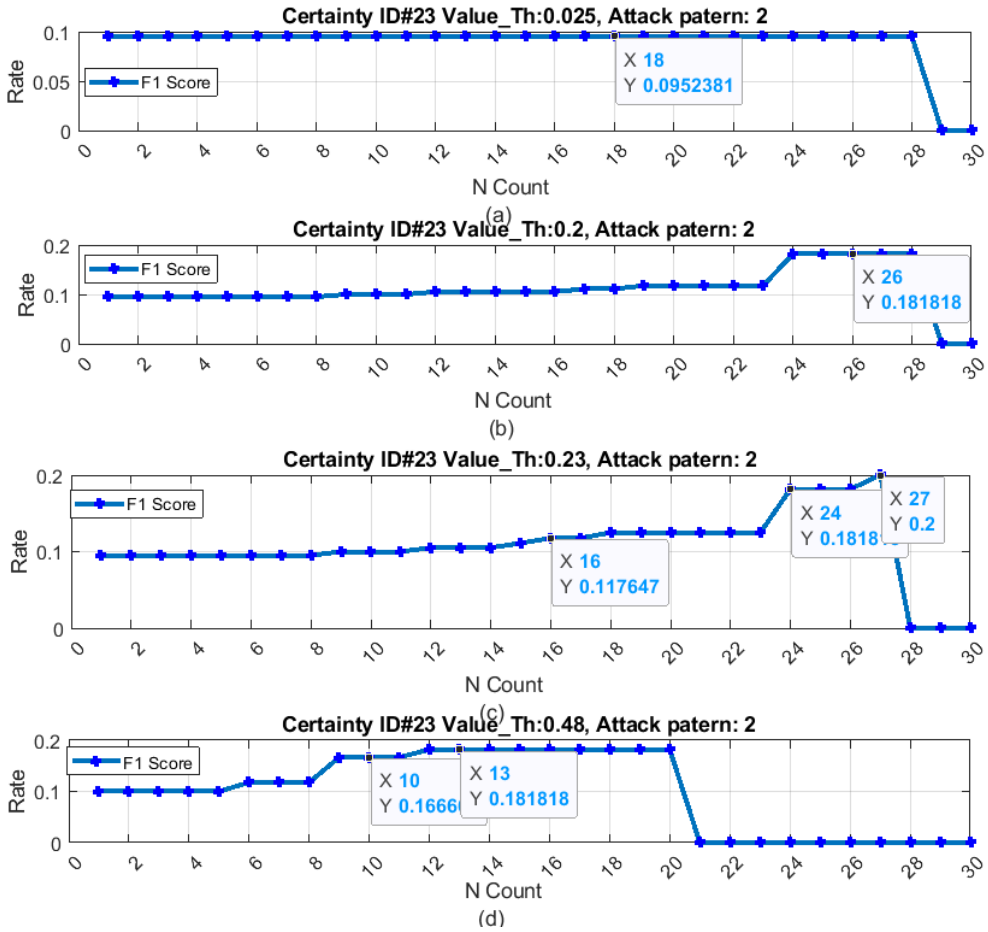


Figure 5.11: Certainty data of Profile ID#23 with attack pattern 2 applied. With the given criteria mentioned before, this figure illustrates the changes in the F1 score curve as the threshold curve increase. Certainty F1 score curve when: (a) value threshold is 0.025, (b) value threshold is 0.2, (c) value threshold is 0.23, (d) value threshold is 0.48.

Figure 5.12 (a) to (d) show a snippet of what the Maximum F1 score looks like for selective MVoT Variables of profile ID 23. The curves are not alike, and the maximum F1 score rate occurs for different value and count thresholds. It is preferred to have multiple points of maximum F1 scores to repeat consecutively or close to each other without a drastic change in the curve; this shows the CDTA, the analyzed system, is working correctly. Figure 5.12 (c) shows there are two maximum F1 rates present at count threshold 1 and 2.

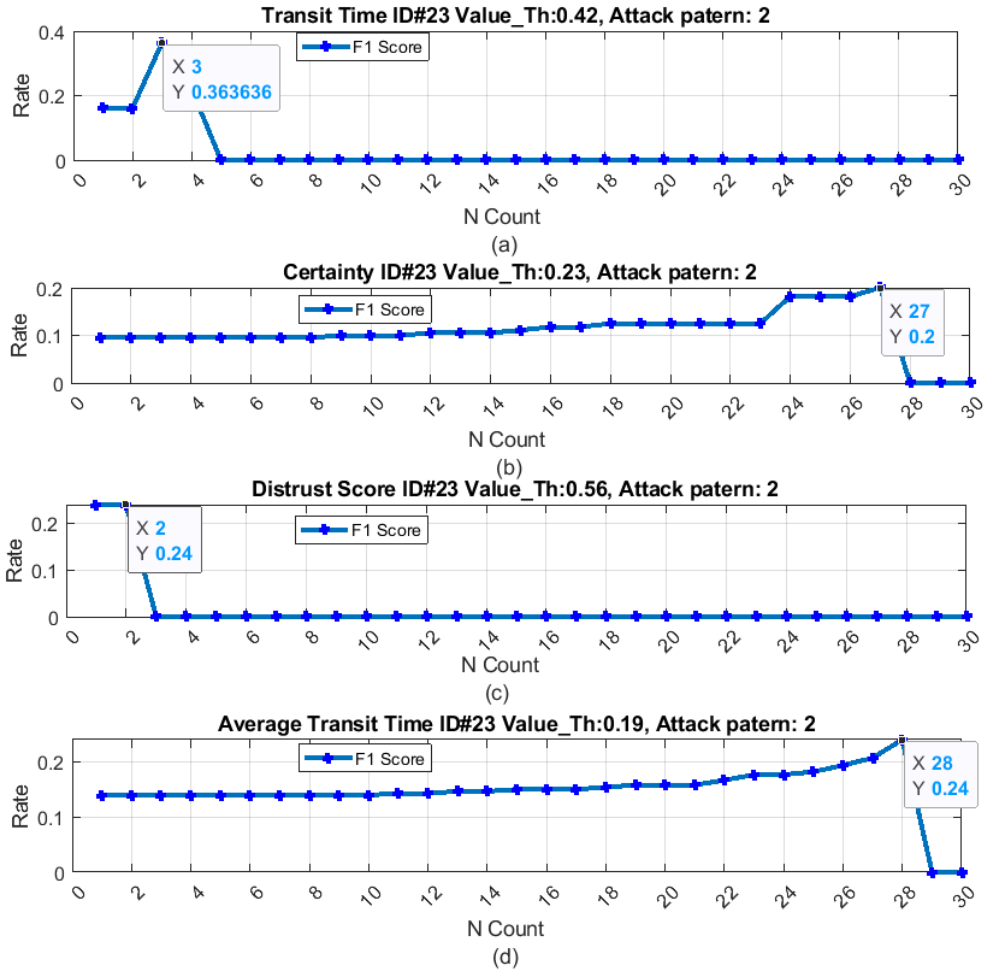


Figure 5.12: This figure illustrates the occurrence of the maximum F1 rate for different MVoT Variable data of Profile ID 23 when attack pattern 2 is applied. F1 score curve of (a) Transit Time, (b) Certainty, (c) Distrust Score, and (d) Average Transit Time.

5.4 Summary

In Summary, the derived results of EER, F1 score, and Sensitivity provide insight for understanding the behavior of a selective number of MVoT variables when using simulated data. Findings of this research justified the use of EER and F1 scores as an appropriate measurement for evaluating the performance of the CDTA.

6 Discussion

6.1 Discussion

In this thesis, hypothesis testing was applied to understand the threshold values used to send alerts and to define a methodology for setting the values. The findings of the previous chapter show that the behavior of FNR and FPR curves varies a lot from one MVoT variable to another. Varying unexpected message locations, such as in the beginning, middle, end, or randomly in a series of messages, have little impact on the hypothesis analysis. This observation is common to all the test cases performed in this analysis.

When the attack pattern is zero, the FNR curve, the F1 score curves, and the sensitivity curves all had a constant zero rate. This observation is also common to all the test cases performed in this analysis. Another observation is that the data used for analysis affects the shape of the FPR curve and the FNR curve. The findings also confirmed that the increase and decrease in the value threshold and count threshold changes the FNR and FPR curves. The observed behavior is not common across different MVoT variables. The occurrence of EER is independent of the value or the count threshold.

Another observation was that EER does not change much when there was a slight change to the attack pattern, although EER occurs at different count thresholds. However, there were significant changes to the FNR and FPR curves as the value threshold or the

count threshold changes. Figure 5.8 and Figure 5.4 show the changes to the FNR and FPR curves as the value changes, unless the FNR curve remains at zero rates when no attacks are present.

6.2 Significance of EER

EER should be low, less than 5%. A lower EER implies that the, FNR and FPR, are lower and closer to zero, and they are equal to each other to balance out the negative impact caused by errors. An example provided in Figure 5.4 shows the lower EER appears for a higher count threshold of MVoT variable Trust Score, compared to Figure 5.8, where the lowest EER occurred at a lower count threshold for the MVoT variable data of the Distrust Score.

Table 6.1 shows the resulting minimum EER for the MVoT Variable data of Profile ID 23 when attack pattern 2 is applied. Table 6.1 shows the resulting minimum EER for the MVoT Variable data of Profile ID 23 when attack pattern 2 is applied. In both tables, the first column lists the Profile ID 23 MVoT variable name. The second column shows the observed minimum EER of the MVoT variable data. The third column shows the count threshold and value thresholds EERs occurred. Finally, the fourth column shows the attack pattern applied to derive the data. Each entry of Table 6.1 and Table6.2 shows the MVoT variable name the corresponding minimum EER and the count threshold, or range of count thresholds and the value threshold or the range of value threshold and the applied attack pattern.

One observation of Table 6.1 and Table 6.2 is that the minimum EERs are not common

among MVoT variables. Therefore, if two variables are common, it is understood to be coincidental and not expected. For example, the minimum EER of Distrust Score and the Transit time are 0.33 in Table 6.1, but the minimum EER for those variables is different in Table 6.2.

Another observation derived from the minimum EER data of Table 6.1 and Table 6.2 is that there is a slight difference between the resulting minimum EER when there are three attacks vs. six attacks. The percent difference ranges from 0 to 10% for the resulting minimum EER when there are three attacks (attack pattern 3) vs. six attacks (attack pattern 3). The average percent difference was 3.6%. This analysis is essential to understand the resulting behavior of the EER when different attack patterns are applied to each MVoT. In each of the MVoT variable data observed in this thesis, there was a variance in the EER point.

Profile ID 23 MVoT Variable Name	**Minimum EER**	**(Count Threshold, Value threshold)**	**Attack Pattern**
Transit Time	0.33	(4 and 5, 0.405) (2 and 3, 0.43)	2
Average Transit Time	0.32	(Between 28 and 29, 0.19)	2
Distrust Score	0.33	(Between 2 and 3, 0.56)	2
Trust Score	0.39	(Between 2 and 3, 538.7)	2
Certainty	0.3	(Between 28 and 29, range of 0.15 to 0.2), (Between 26 and 27, range of 0.25 to 0.35), (Between 23 and 24, range of 0.37 to 0.45).	2
RFC	0.32	(Between 28 and 29, 0.85)	2
TSLC	0.5	(Between 1 to 23, range of 180 to 3420)	2
Communication Frequency	0.5	(Between 8 and 9, 0.0025 ms), (Between 5 and 6, range of 0.0033 ms to 0.0076 ms), (Between 2 and 3, range of 0.008 ms to 0.016 ms).	2

Table 6.1: The table is showing the minimum EER for MVoT variable data of Profile ID 23 for attack pattern 2

Profile ID 23 MVoT Variable Name	**Minimum EER**	**(Count Threshold, Value threshold)**	**Attack Pattern**
Transit Time	0.31	(Between 28 and 29, 0.19)	3
Average Transit Time	0.3	(Between 28 and 29, 0.19)	3
Distrust Score	0.33	(Between 2 and 3, 0.56)	3
Trust Score	0.35	(Between 5 and 6, 239.4)	3
Certainty	0.25	(Between 23 and 24, 0.4)	3
RFC	0.32	(Between 28 and 29, 0.85)	3
TSLC	0.5	(Between 1 to 23, range of 180 to 3420)	3
Communication Frequency	0.5	(Between 8 and 9, 0.0025 ms), (Between 5 and 6, range of 0.0033 ms to 0.0076 ms), (Between 2 and 3, range of 0.008 ms to 0.016 ms).	3

Table 6.2: This table is showing the minimum EER for MVoT variable data of Profile ID 23 for attack pattern 3

6.3 Significance of F1 Score

The F1 score shows statistical analysis of the harmonic mean of the CDTA correctly alerting the GSP to falsely alarming and failing to warn the GSP. As mentioned in the previous chapter, the relationship between the EER and the F1 score is that the F1 score has a sudden change at the count threshold where the EER occurs. This sudden change reflects the FPR curve. If the FPR curve drops to zero, then the F1 score rate also drops to zero. If the FPR curve drops and settles at a different pace, then the F1 score also mimics the curve shape, although the rates are not identical. When there are no attacks, the F1 score plots always have a rate of zero.

Profile ID 23 MVoT Variable Name	**Max F1 Score**	**(Count Threshold / Value threshold)**	**Attack Pattern**
Transit Time	0.36	(3,0.43)	2
Average Transit Time	0.24	(13,0.24), (25,0.209), (28,0.19)	2
Distrust Score	0.24	(1 to 2, 0.56), (From 9 to 11, 0.41)	2
Trust Score	0.24	(Between 1 and 2, 538.7)	2
Certainty	0.2	(From 24 to 26, 0.27)	2
RFC	0.23	(27 and 28, 0.85)	2
TSLC	0.14	(Ranges from 1 to 23 , 180 sec to 3420 sec)	2
Communication Frequency	0.14	(Ranges from 15 to 1 , from 0.0008 ms to 0.016 ms)	2
Transit Time	0.303	(3,0.405)	3
Average Transit Time	0.43	(13,0.24), (25,0.209), (28,0.19)	3
Distrust Score	0.43	(1 to 2, 0.56), (From 9 to 11, 0.41)	3
Trust Score	.0.43	(Between 1 and 2, 538.7)	3
Certainty	0.62	(From 24 to 26, 0.27)	3
RFC	0.42	(25, 0.95)	3
TSLC	0.26	(ranges from 1 to 23 , 180 sec to 3420 sec)	3
Communication Frequency	0.26	(ranges from 15 to 1 , from 0.0008 ms to 0.016 ms)	3

Table 6.3: This table provides the maximum F1 score observed for profile ID 23 for attack patterns 2 and 3.

Table 6.3 provides the maximum F1 score observed for the particular set of MVoT variable data of profile ID 23 when attack patter 2 or 3 applied. It showed that the maximum F1 score for a specific MVoT variable data was not close between the two attack patterns. For example, the minimum percent difference for MVoT variable Transit Time for the two attacks was 15.83%. However, the remaining MVoT variables range from 79.17% to 85.71% for the remaining MVoT variable data when attack patterns 2 and 3 applied. Such a

drastic result in the percent difference of the F1 curve is alarming.

6.4 Summary

In Summary, the derived results of EER, F1 score, and Sensitivity were insightful to understand the behavior of those plots for a selective number of MVoT variables when using simulated data. The data justified the use of EER and F1 scores as an appropriate measurement for evaluating the performance of the CDTA. The comparison of various attack patterns applied to profile ID 23 MVoT variables and the compared results show that the EER is susceptible to the different attack patterns that were used in this thesis. This does not imply the EER provides similar results for many attack patterns.

7 Conclusion

This thesis work describes a tool that shows the DTM system dependence on specific MVoT thresholds. One key feature of the Distributed Trust Model is the evaluation of messages and alerting authorities of any abnormalities. The decision to send the alert at the right time is critical to avoid possible disruptions caused by intruders.

The hypothesis tool serves as an analysis tool for the decision-maker to make a sound decision on threshold selection for decision to send alerts. The hypothesis tool explains the statistical probability of failing to send an alert or false alert based on the selected threshold. The EER feature of the hypothesis tool provides the threshold point or a range of threshold values where FPRs and FNRs are equal to each other.

The hypothesis test tool can use simulated data and real data from system actors. The hypothesis tool evaluates the data using a range of threshold values and threshold counts, which vary and perform confusion metric evaluations on the input data. The tool also evaluates a specific threshold. Either the count threshold or the value threshold is held constant while the other varies. The tool helps a user analyze the count threshold and the value thresholds impact on error using confusion metric plots.

An improvement to this hypothesis tool can be done by converting this to a Python program or any other platform. The suggested change would help advance the features of

the hypothesis testing tool to enable 3D graphics of FNR and FPR curves to have both the value thresholds and the count threshold varying simultaneously.

Security is a continuing battle in the communication system; therefore, it is essential to have an adaptable evolving security solution. The solution provided in this thesis, the DTM hypothesis tool, analyzes how a set of threshold values impact the error ratio of trust alerts. Additionally, the DTM hypothesis tool checks the decision-making equations that send alerts to the authority and checks to see MVoT equations are appropriately set. The solution provided in this thesis provides a measuring tool for an adaptable evolving security solution.

Bibliography

[1] Kamran Ahmad Awan, Ikram Ud Din, Ahmad Almogren, Mohsen Guizani, Ayman Altameem, and Sultan Ullah Jadoon. Robusttrust–a pro-privacy robust distributed trust management mechanism for internet of things. *IEEE Access*, 7:62095–62106, 2019.

[2] George David Birkhoff. A mathematical approach to ethics. *Rice Institute Pamphlet-Rice University Studies*, 28(1), 1941.

[3] Wattana Viriyasitavat and Andrew Martin. A survey of trust in workflows and relevant contexts. *IEEE Communications Surveys Tutorials*, 14(3):911–940, 2012.

[4] Stephen Paul Marsh. Formalising trust as a computational concept. *PhD thesis, University of Stirling*, 1994.

[5] Alfarez Abdul-Rahman and Stephen Hailes. A distributed trust model. In *Proceedings of the 1997 workshop on New security paradigms*, pages 48–60, 1998.

[6] Jacob Sakhnini, Hadis Karimipour, Ali Dehghantanha, Reza M Parizi, and Gautam Srivastava. Security aspects of internet of things aided smart grids: A bibliometric survey. *Internet of things*, page 100111, 2019.

[7] G. Giacomello, F.N. Moro, and M. Valigi. *Technology and International Relations: The New Frontier in Global Power*. Edward Elgar Publishing Limited, 2021.

[8] Daisuke Mashima, Binbin Chen, Prageeth Gunathilaka, and Edwin Lesmana Tjiong. Towards a grid-wide, high-fidelity electrical substation honeynet. In *2017 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 89–95, 2017.

[9] Steve Widergren, Ron Melton, Aditya Khandekar, Bruce Nordman, and Mark Knight. The plug-and-play electricity era: Interoperability to integrate anything, anywhere, anytime. *IEEE Power and Energy Magazine*, 17(5):47–58, 2019.

[10] IEEE Communications Society. IEEE Standard for Smart Energy Profile Application Protocol. *IEEE Std 2030.5-2018 (Revision of IEEE Std 2030.5-2013)*, pages 1–361, 2018.

[11] Sepandar D Kamvar, Mario T Schlosser, and Hector Garcia-Molina. The eigentrust algorithm for reputation management in P2P networks. In *Proceedings of the 12th international conference on World Wide Web*, pages 640–651, 2003.

[12] Arezou Moussavi-Khalkhali, Ram Krishnan, and Mo Jamshidi. Periodic virtual hierarchy: A trust model for smart grid devices. *International Journal of Security and Its Applications*, 10(11):249–266, 2016.

[13] W. T. Luke Teacy, Nicholas R. Jennings, Alex Rogers, and Michael Luck. A hierarchical bayesian trust model based on reputation and group behaviour. In *6th European Workshop on Multi-Agent Systems (18/12/08 - 19/12/08)*, December 2008. Event Dates: 18th-19th December, 2008.

[14] Osman Khalid, Samee U Khan, Sajjad A Madani, Khizar Hayat, Majid I Khan, Nasro Min-Allah, Joanna Kolodziej, Lizhe Wang, Sherali Zeadally, and Dan Chen. Comparative study of trust and reputation systems for wireless sensor networks. *Security and Communication Networks*, 6(6):669–688, 2013.

[15] Henry Nunoo-Mensah, Kwame Osei Boateng, and James Dzisi Gadze. The adoption of socio- and bio-inspired algorithms for trust models in wireless sensor networks: A survey. *International Journal of Communication Systems*, 31(7), 2018.

[16] Girish Suryanarayana, J.R. Erenkrantz, and R.N. Taylor. An architectural approach for decentralized trust management. *IEEE Internet Computing*, 9(6):16–23, 2005.

[17] Andreea Visan, Florin Pop, and Valentin Cristea. Decentralized trust management in peer-to-peer systems. In *2011 10th International Symposium on Parallel and Distributed Computing*, pages 232–239. IEEE, 2011.

[18] Huanyu Zhao and Xiaolin Li. Vectortrust: trust vector aggregation scheme for trust management in peer-to-peer networks. *The Journal of Supercomputing*, 64(3):805–829, 2013.

[19] Brent N Chun and Andy Bavier. Decentralized trust management and accountability in federated systems. In *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*, pages 9–pp. IEEE, 2004.

[20] Tie-Yan Li, Huafei Zhu, and Kwok-Yan Lam. A novel two-level trust model for grid. In *International Conference on Information and Communications Security*, pages 214–225. Springer, 2003.

[21] Jinfang Jiang, Guangjie Han, Lei Shu, Sammy Chan, and Kun Wang. A trust model based on cloud theory in underwater acoustic sensor networks. *IEEE Transactions on Industrial Informatics*, 13(1):342–350, 2015.

[22] Bjørnar Solhaug and Ketil Stølen. Uncertainty, subjectivity, trust and risk: How it all fits together. In *International Workshop on Security and Trust Management*, pages 1–5. Springer, 2011.

[23] Li Xiong and Ling Liu. Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities. *IEEE transactions on Knowledge and Data Engineering*, 16(7):843–857, 2004.

[24] Ataul Bari, Jin Jiang, Walid Saad, and Arunita Jaekel. Challenges in the smart grid applications: an overview. *International Journal of Distributed Sensor Networks*, 10(2):974682, 2014.

[25] Hitesh Mohapatra, Subhashree Rath, Subarna Panda, and Ranjan Kumar. Handling of man-in-the-middle attack in wsn through intrusion detection system. *International journal*, 8(5):1503–1510, 2020.

[26] Jian-yun Lei, Bing-cai Zhang, and Xiao-hai Fang. Trust vector-based sensitive information protecting scheme in automatic trust negotiation. In *Proceedings of 2011*

International Conference on Computer Science and Network Technology, volume 2, pages 735–738. IEEE, 2011.

[27] Yan Lindsay Sun, Zue Han, Wei Yu, and KJ Ray Liu. A trust evaluation framework in distributed networks: Vulnerability analysis and defense against attacks. In *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, pages 1–13. IEEE, 2006.

[28] Chih-Che Sun, Adam Hahn, and Chen-Ching Liu. Cyber security of a power grid: State-of-the-art. *International Journal of Electrical Power & Energy Systems*, 99:45–56, 2018.

[29] G Sudhamathy and C Jothi Venkateswaran. *R Programming: An Approach to Data Analytics*. MJP Publisher, 2019.

[30] N. S. Fernando. Data from: The distributed trust model applied to the energy grid of things. *Electrical and Computer Engineering Datasets*, 2021.

Appendix A: User Guide: DTM Hypothesis Testing Tool

A.1 Overview

This appendix serves as a user guide to the DTM hypothesis testing tool used in this research. By the end of this guide, the user knows how to generate hypothesis plots discussed in this thesis. The DTM hypothesis tool and all the CSV files used the application Microsoft® Excel® for Microsoft 365 (64-bit). Majority of files used in this research can be found at [30]

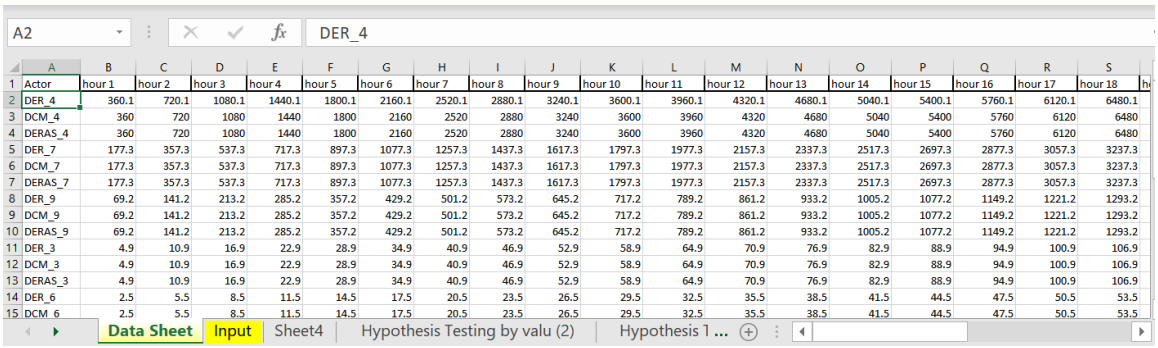
A.1.1 Types of Files

A.1.1.1 Distributed Trust Model's Hypothesis Testing Tool

The DTM hypothesis testing tool is a CSV file that is used in this research. This tool takes in generated files from the Trust Model simulator's $per\ hour\ script$ module. Each MVoT variable have a DTM hypothesis testing tool to evaluate.

A.1.1.2 Input CSV's

The input to the DTM hypothesis tool is a CSV file. The hypothesis testing tool takes in data from generated CSV files of the Trust Model simulator's $perhourscript$ module.



A2 | DER_4

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Actor	hour 1	hour 2	hour 3	hour 4	hour 5	hour 6	hour 7	hour 8	hour 9	hour 10	hour 11	hour 12	hour 13	hour 14	hour 15	hour 16	hour 17	hour 18
2	DER_4	360.1	720.1	1080.1	1440.1	1800.1	2160.1	2520.1	2880.1	3240.1	3600.1	3960.1	4320.1	4680.1	5040.1	5400.1	5760.1	6120.1	6480.1
3	DCM_4	360	720	1080	1440	1800	2160	2520	2880	3240	3600	3960	4320	4680	5040	5400	5760	6120	6480
4	DERAS_4	360	720	1080	1440	1800	2160	2520	2880	3240	3600	3960	4320	4680	5040	5400	5760	6120	6480
5	DER_7	177.3	357.3	537.3	717.3	897.3	1077.3	1257.3	1437.3	1617.3	1797.3	1977.3	2157.3	2337.3	2517.3	2697.3	2877.3	3057.3	3237.3
6	DCM_7	177.3	357.3	537.3	717.3	897.3	1077.3	1257.3	1437.3	1617.3	1797.3	1977.3	2157.3	2337.3	2517.3	2697.3	2877.3	3057.3	3237.3
7	DERAS_7	177.3	357.3	537.3	717.3	897.3	1077.3	1257.3	1437.3	1617.3	1797.3	1977.3	2157.3	2337.3	2517.3	2697.3	2877.3	3057.3	3237.3
8	DER_9	69.2	141.2	213.2	285.2	357.2	429.2	501.2	573.2	645.2	717.2	789.2	861.2	933.2	1005.2	1077.2	1149.2	1221.2	1293.2
9	DCM_9	69.2	141.2	213.2	285.2	357.2	429.2	501.2	573.2	645.2	717.2	789.2	861.2	933.2	1005.2	1077.2	1149.2	1221.2	1293.2
10	DERAS_9	69.2	141.2	213.2	285.2	357.2	429.2	501.2	573.2	645.2	717.2	789.2	861.2	933.2	1005.2	1077.2	1149.2	1221.2	1293.2
11	DER_3	4.9	10.9	16.9	22.9	28.9	34.9	40.9	46.9	52.9	58.9	64.9	70.9	76.9	82.9	88.9	94.9	100.9	106.9
12	DCM_3	4.9	10.9	16.9	22.9	28.9	34.9	40.9	46.9	52.9	58.9	64.9	70.9	76.9	82.9	88.9	94.9	100.9	106.9
13	DERAS_3	4.9	10.9	16.9	22.9	28.9	34.9	40.9	46.9	52.9	58.9	64.9	70.9	76.9	82.9	88.9	94.9	100.9	106.9
14	DER_6	2.5	5.5	8.5	11.5	14.5	17.5	20.5	23.5	26.5	29.5	32.5	35.5	38.5	41.5	44.5	47.5	50.5	53.5
15	DCM_6	2.5	5.5	8.5	11.5	14.5	17.5	20.5	23.5	26.5	29.5	32.5	35.5	38.5	41.5	44.5	47.5	50.5	53.5

Data Sheet | Input | Sheet4 | Hypothesis Testing by valu (2) | Hypothesis T ...

Figure A.1: This figure shows an illustration of MVoT data added to the $Data\ Sheet$ of the DTM Hypothesis Testing tool.

A.1.2 Adding MVoT Variable data

Step 1: Add the MVoT variable data for a X x Y table where X and Y represents total number of actors and time respectively. Note: Add the data at starting at the initial cell $A2$.

A.1.3 MVoT Data Input sheet

Step 2: Go to the $MVoT\ Data\ Input$ sheet. (shown in Figure A.2)

Step 3: Figure A.2 includes labels from 1 - 9 for user input.

User input descriptions:

1. **Enter** the MVoT Variable Name.
2. **Enter** the Threshold value.
3. **Enter** the Count of Actors threshold.

Labels that are numbered from 4 - 6 are about inputs are specific to sending a message using the value threshold.

4. **Enter** the Start Value of the $valuethreshold$ of the MVoT variable.
5. **Enter** the ending value for the $valuethreshold$ of the MVoT variable.

6. **Enter** the number of data points between the start and end of $valuethreshold$

Labels that are numbered from 7 - 9 are about inputs that are specific to sending a message using the count threshold.

7. **Enter** the minimum message count per actor per hour. Typically this value is set to 1.

8. **Enter** the maximum message count per actor per hour.

9. **Enter** the step value for count of messages.

10. **Enter** 1 for an attack and 0 for no attack for the corresponding hour.

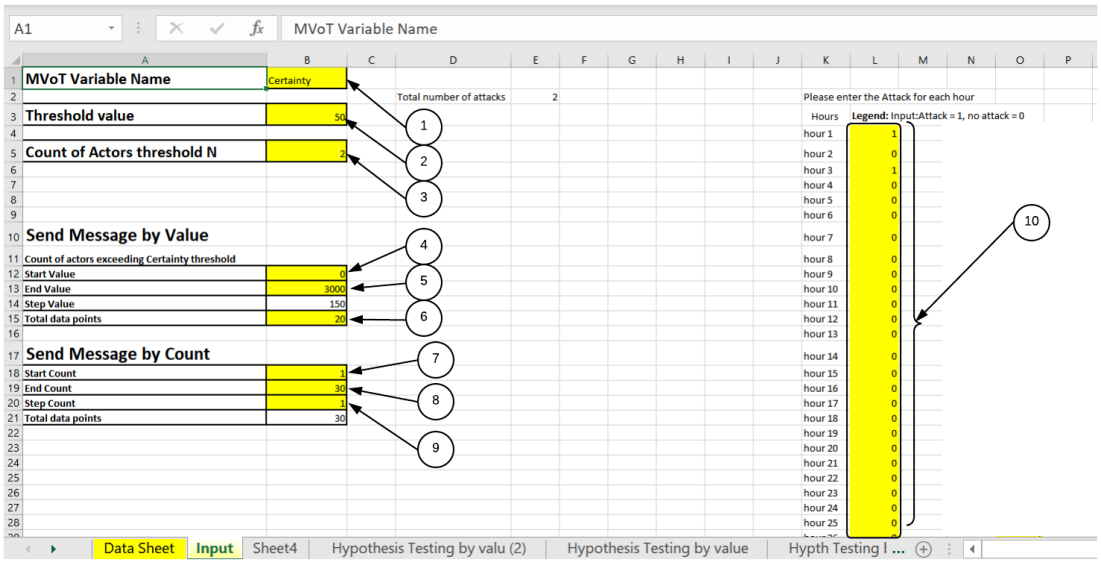


Figure A.2: This figure shows a series of inputs, highlighted in yellow. The DTM Hypothesis testing tool takes in to calculate and plot hypothesis measurements.

A.1.4 Supplemental Sheets

The Distributed Trust Model hypothesis testing tool consists of several sheets that serves as supplemental sheet that helps derive plots. These are used for calculations and are not

specifically meant for reading/inspections. Appendix A.1.4 provides snapshots of those supplemental sheets. Figures A.3 and A.4 illustrates a small sample of what hypothesis testing by value data and hypothesis testing by count data and their corresponding confusion metric plots respectively.

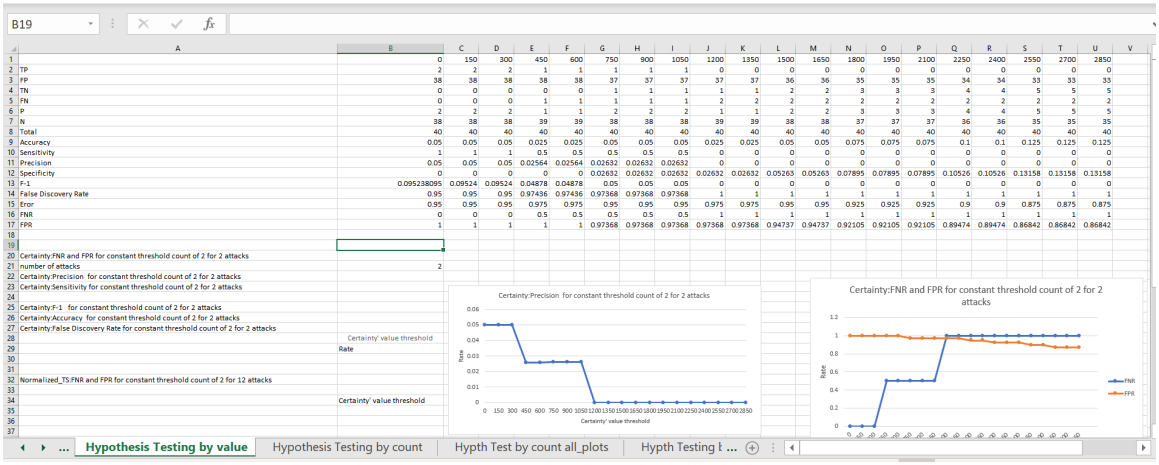


Figure A.3: This figure shows what is included in the Excel sheet Hypothesis Testing by Value calculated binary classification of the MVoT variable, in this example certainty, outlined with a black margin and a snapshot of resulting plots of classification metrics.

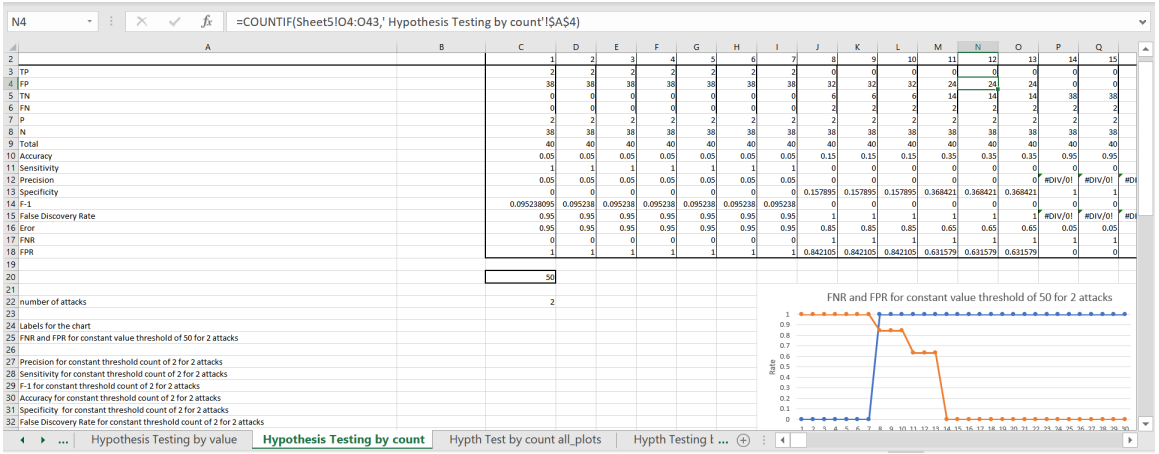


Figure A.4: This figure shows what is included in the Excel sheet Hypothesis Testing by Count calculated binary classification of the MVoT variable, in this example certainty, outlined with a black margin and a snapshot of resulting plots of classification metrics.

Figure A.5 shows a snapshot of series of hypothesis testing sheets consisting of data where threshold value and the count value varies.

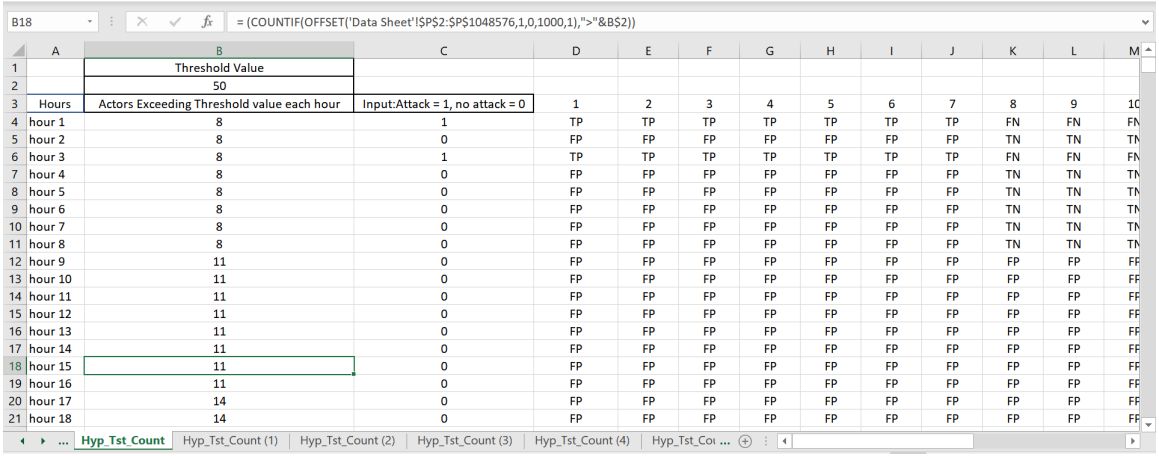


Figure A.5: This figure shows what is included in the Excel sheet Hypothesis Testing by Count calculated binary classification of the MVoT variable, in this example certainty, outlined with a black margin and a snapshot of resulting plots of classification metrics.

A.1.5 Output sheet

The $HypthTestbycountallplots$ sheet consist of confusion metric plots for all the data when both the count variable and the value variable changes. These plots serves as a visualization tool to show how the confusion metric plots adjust when the count and value thresholds changes.

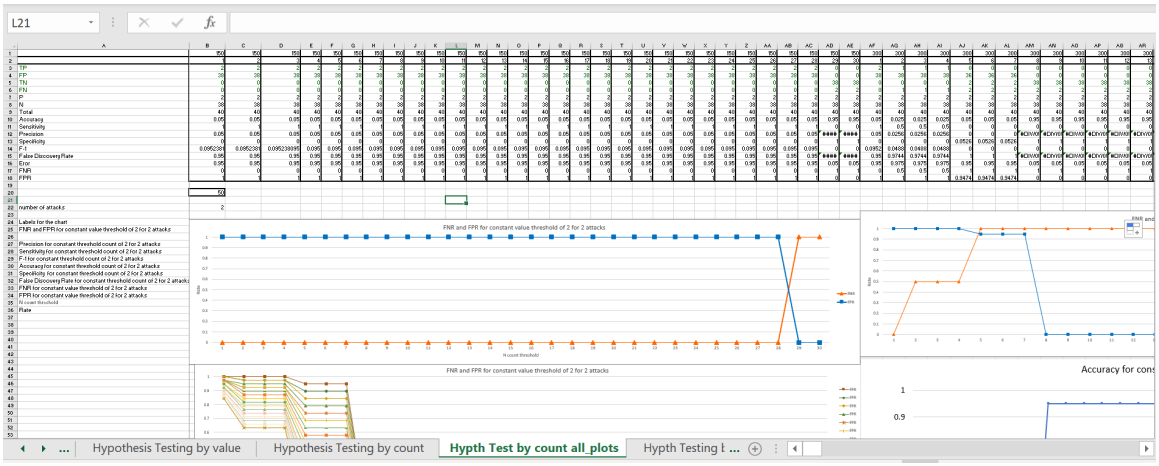


Figure A.6: This figure shows what is included in the Excel sheet Hypothesis Testing by count of the MVoT variable, outlined with a black margin and a snapshot of resulting plots of classification metrics.

Appendix B: Test Conditions

B.1 Overview

This appendix provides 26 tables consist of scenarios of messages sent from the DTMC at the SPC to the CDTA. Each SPC consists of one DTMC and observes three actors: a GSP, a DCM, and a DER. Each table shows a combination of ten SPC's with the three different message classification counts listed: expected, unexpected, and indeterminant. Additionally, the table indicates the time step of sending the messages from the DTMC to the CDTA, the approximate location of the indeterminant and unexpected messages, and whether those messages are from the same or different actors.

Profile ID#	Description
1	A -J all expected. Each message has a 5 minute time step.
2	A -J all are expected. Each message is between 1 minute to 1 hour.
3	A - F are all expected, SPC's G, H, I, J have 6 unexpected messages occurring in the beginning, middle, end, and randomly respectively from different actors, and the rest of the messages are expected. Each message is between 1 minute to 1 hour.
4	SPC A to SPC F all expected, SPC's G, H, I, J have 6 unexpected messages occurring at the beginning from different actors and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
5	SPC A to SPC F all expected, SPC's G, H, I, J have 6 unexpected messages occurring in the middle from different actors and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
6	SPC A to SPC F all expected, SPC's G, H, I, J have 6 unexpected messages occurring at the end from different actors, and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
7	SPC A to SPC F all expected, SPC's G, H, I, J have 6 unexpected messages occurring in the beginning, middle, end, and randomly from the same actors and the rest of the messages are expected. Each message is between 1 minute to 1 hour.

Table B.1: A table providing the profile IDs and corresponding descriptions.

Profile ID#	Description
8	SPC A to SPC F all expected, SPC's G, H, I, J have 6 unexpected messages occurring at the beginning from the same actors, and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
9	SPC A to SPC F all expected, SPC's G, H, I, J have 6 unexpected messages occurring in the middle from the same actors and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
10	SPC A to SPC F all expected, SPC's G, H, I, J have 6 unexpected messages occurring in the end from the same actors, and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
11	SPC A to SPC F all expected, SPC's G, H, I, J have 3 unexpected messages occurring in the beginning, middle, end, and randomly respectively from different actors, and the rest of the messages are expected. Each message is between 1 minute to 1 hour.
12	SPC A to SPC F all expected, SPC's G, H, I, J have 3 unexpected messages occurring at the beginning from different actors and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
13	SPC A to SPC F all expected, SPC's G, H, I, J have 3 unexpected messages occurring in the middle from different actors and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.

Table B.2: A table providing the profile IDs and corresponding descriptions.

Profile ID#	Description
14	SPC A to SPC F all expected, SPC's G, H, I, J have 3 unexpected messages occurring in the end from different actors, and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
15	SPC A to SPC F all expected, SPC's G, H, I, J have 6 unexpected messages occurring in the beginning, middle, end, and randomly respectively from the same actors, and the rest of the messages are expected. Each message is between 1 minute to 1 hour.
16	SPC A to SPC F all expected, SPC's G, H, I, J have 3 unexpected messages occurring at the beginning from the same actors and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
17	SPC A to SPC F all expected, SPC's G, H, I, J have 3 unexpected messages occurring in the middle from the same actors and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
18	SPC A to SPC F all expected, SPC's G, H, I, J have 3 unexpected messages occurring in the end from the same actors, and the rest of the messages are all expected. Each message is between 1 minute to 1 hour.
19	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for the same actor: beginning, middle, and end. SPC I - J has an unexpected message at the end and mix for the same actor with message time step: 1 min - 1 hour.
20	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for the same actor: at the beginning. SPC I - J has an unexpected message at the beginning for the same actor with message time step: 1 min - 1 hour.

Table B.3: A table providing the profile IDs and corresponding descriptions.

Profile ID#	Description
21	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for the same actor: at the middle. SPC I - J has an unexpected message at the middle for the same actor with message time step: 1 min - 1 hour.
22	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for the same actor: at the end. SPC I - J has an unexpected message at the end for the same actor with message time step: 1 min - 1 hour.
23	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for different actors: beginning, middle, and end. SPC I - J has an unexpected message at the end and mix for different actors with message time step: 1 min - 1 hour.
24	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for different actors: at the beginning. SPC I - J has an unexpected message at the beginning for different actors with message time step: 1 min - 1 hour.
25	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for different actors: at the middle. SPC I - J has an unexpected message in the middle for different actors with message time step: 1 min - 1 hour.
26	SPC A - E is all expected. SPC's F -H has 3 indeterminant messages for different actors: at the end. SPC I - J has an unexpected message at the end for different actors with message time step: 1 min - 1 hour.

Table B.4: A table providing the profile IDs and corresponding descriptions.

Profile	Time Step	Expected	Unexpected	Indeterminant	Location	same/ different actor
SPC A	5 min	All Expected	0	0	-	
SPC B	5 min	All Expected	0	0	-	
SPC C	5 min	All Expected	0	0	-	
SPC D	5 min	All Expected	0	0	-	
SPC E	5 min	All Expected	0	0	-	
SPC F	5 min	All Expected	0	0	-	
SPC G	5 min	All Expected	0	0	-	
SPC H	5 min	All Expected	0	0	-	
SPC I	5 min	All Expected	0	0	-	
SPC J	5 min	All Expected	0	0	-	

Table B.5: ID# 1: This table contains ten Expected profiles with an identical time step of five minutes for each message sent.

Profile	Time Step	Expected	Unexpected	Indeterminant	Location	same/ different actor
SPC A	1 min	All Expected	0	0	-	
SPC B	2 min	All Expected	0	0	-	
SPC C	5 min	All Expected	0	0	-	
SPC D	10 min	All Expected	0	0	-	
SPC E	15 min	All Expected	0	0	-	
SPC F	20 min	All Expected	0	0	-	
SPC G	25 min	All Expected	0	0	-	
SPC H	30 min	All Expected	0	0	-	
SPC I	40 min	All Expected	0	0	-	
SPC J	1 hr	All Expected	0	0	-	

Table B.6: ID# 2: This table contains ten expected and zero unexpected profiles with a time step between one minute and one hour for each message sent.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		6	Beginning	Different Actor
SPC H	30 min		6	Middle	Different Actor
SPC I	40 min		6	End	Different Actor
SPC J	1 hr		6	Random	Different Actor

Table B.7: ID# 3: This table contains a combination of six expected and four unexpected profiles with different actors, each with six unexpected messages occurring for SPC G to SPC J in a combination of beginning, middle, end, and random, respectively.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		6 Ux Messages	Beginning	Different Actor
SPC H	30 min		6 Ux Messages	Beginning	Different Actor
SPC I	40 min		6 Ux Messages	Beginning	Different Actor
SPC J	1 hr		6 Ux Messages	Beginning	Different Actor

Table B.8: ID# 4:This table contains a combination of six expected and four unexpected profiles with different actors, each with six unexpected messages occurring in the beginning.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		6 Ux Messages	Middle	Different Actor
SPC H	30 min		6 Ux Messages	Middle	Different Actor
SPC I	40 min		6 Ux Messages	Middle	Different Actor
SPC J	1 hr		6 Ux Messages	Middle	Different Actor

Table B.9: ID# 5:This table contains six expected and four unexpected profiles with different actors, each with six unexpected messages occurring in the middle.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		6 Ux Messages	End	Different Actor
SPC H	30 min		6 Ux Messages	End	Different Actor
SPC I	40 min		6 Ux Messages	End	Different Actor
SPC J	1 hr		6 Ux Messages	End	Different Actor

Table B.10: ID# 6:This table contains a combination of six expected and four unexpected profiles with different actors, each with six unexpected messages occurring in the end.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		6 Ux Messages	Beginning	Same Actor
SPC H	30 min		6 Ux Messages	Middle	Same Actor
SPC I	40 min		6 Ux Messages	End	Same Actor
SPC J	1 hr		6 Ux Messages	Random	Same Actor

Table B.11: ID# 7:This table contains a combination of six expected and four unexpected profiles with the same actors, each with six unexpected messages occurring for SPC G to SPC J in a combination of beginning, middle, end, and random, respectively.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		6 Ux Messages	Beginning	Same Actor
SPC H	30 min		6 Ux Messages	Beginning	Same Actor
SPC I	40 min		6 Ux Messages	Beginning	Same Actor
SPC J	1 hr		6 Ux Messages	Beginning	Same Actor

Table B.12: ID# 8:This table contains a combination of six expected and four unexpected profiles with the same actors, each with six unexpected messages occurring in the beginning.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		6 Ux Messages	Middle	Same Actor
SPC H	30 min		6 Ux Messages	Middle	Same Actor
SPC I	40 min		6 Ux Messages	Middle	Same Actor
SPC J	1 hr		6 Ux Messages	Middle	Same Actor

Table B.13: ID# 9:This table contains six expected and four unexpected profiles with the same actors, each with six unexpected messages occurring in the middle.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		6 Ux Messages	End	Same Actor
SPC H	30 min		6 Ux Messages	End	Same Actor
SPC I	40 min		6 Ux Messages	End	Same Actor
SPC J	1 hr		6 Ux Messages	End	Same Actor

ID# 10:This table contains a combination of six expected and four unexpected profiles with the same actors, each with six unexpected messages occurring in the end.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		3 Ux Messages	Beginning	Same Actor
SPC H	30 min		3 Ux Messages	Middle	Same Actor
SPC I	40 min		3 Ux Messages	End	Same Actor
SPC J	1 hr		3 Ux Messages	Mix	Same Actor

Table B.14: ID# 11:This table contains a combination of six expected and four unexpected profiles with the same actors, each with three unexpected messages occurring for SPC G to SPC J in a combination of beginning, middle, end, and random, respectively.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		3 Ux Messages	Beginning	Different Actor
SPC H	30 min		3 Ux Messages	Beginning	Different Actor
SPC I	40 min		3 Ux Messages	Beginning	Different Actor
SPC J	1 hr		3 Ux Messages	Beginning	Different Actor

Table B.15: ID# 12:This table contains six expected and four unexpected profiles with different actors, each with three unexpected messages occurring in the beginning.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		3 Ux Messages	Middle	Different Actor
SPC H	30 min		3 Ux Messages	Middle	Different Actor
SPC I	40 min		3 Ux Messages	Middle	Different Actor
SPC J	1 hr		3 Ux Messages	Middle	Different Actor

Table B.16: ID# 13:This table contains six expected and four unexpected profiles with different actors, each with three unexpected messages occurring in the middle.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		3 Ux Messages	End	Different Actor
SPC H	30 min		3 Ux Messages	End	Different Actor
SPC I	40 min		3 Ux Messages	End	Different Actor
SPC J	1 hr		3 Ux Messages	End	Different Actor

Table B.17: ID# 14:This table contains six expected and four unexpected profiles with different actors, each with three unexpected messages occurring in the end.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		3 Ux Messages	Beginning	Same Actor
SPC H	30 min		3 Ux Messages	Middle	Same Actor
SPC I	40 min		3 Ux Messages	End	Same Actor
SPC J	1 hr		3 Ux Messages	Mix	Same Actor

Table B.18: ID# 15:This table contains a combination of six expected and four unexpected profiles with the same actors, each with three unexpected messages occurring for SPC G to SPC J in a combination of beginning, middle, end, and random, respectively.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		3 Ux Messages	Beginning	Same Actor
SPC H	30 min		3 Ux Messages	Beginning	Same Actor
SPC I	40 min		3 Ux Messages	Beginning	Same Actor
SPC J	1 hr		3 Ux Messages	Beginning	Same Actor

Table B.19: ID# 16:This table contains a combination of six expected and four unexpected profiles with the same actors, each with three unexpected messages occurring at the beginning

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		3 Ux Messages	Middle	Same Actor
SPC H	30 min		3 Ux Messages	Middle	Same Actor
SPC I	40 min		3 Ux Messages	Middle	Same Actor
SPC J	1 hr		3 Ux Messages	Middle	Same Actor

Table B.20: ID# 17:This table contains a combination of six expected and four unexpected profiles with the same actors, each with three unexpected messages occurring in the middle.

Profiles	Step	Expected Messages	Unexpected Messages	Location	same/different actor
SPC A	1 min	All Expected	-	-	All Actor
SPC B	2 min	All Expected	-	-	All Actor
SPC C	5 min	All Expected	-	-	All Actor
SPC D	10 min	All Expected	-	-	All Actor
SPC E	15 min	All Expected	-	-	All Actor
SPC F	20 min	All Expected	-	-	All Actor
SPC G	25 min		3 Ux Messages	End	Same Actor
SPC H	30 min		3 Ux Messages	End	Same Actor
SPC I	40 min		3 Ux Messages	End	Same Actor
SPC J	1 hr		3 Ux Messages	End	Same Actor

Table B.21: ID# 18:This table contains a combination of six expected and four unexpected profiles with the same actors, each with three unexpected messages occurring in the end.

Profiles	Step	Expected Messages	Unexpected Messages	Indeterminant Messages	Location	same/ different actor
SPC A	1 min	All Expected	-	-	-	All Actor
SPC B	2 min	All Expected	-	-	-	All Actor
SPC C	5 min	All Expected	-	-	-	All Actor
SPC D	10 min	All Expected	-	-	-	All Actor
SPC E	15 min	All Expected	-	-	-	All Actor
SPC F	20 min	-	-	3	Mix	Same Actor
SPC G	25 min	-	-	3	Beginning	Same Actor
SPC H	30 min	-	-	3	Middle	Same Actor
SPC I	40 min	-	3	-	End	Same Actor
SPC J	1 hr	-	3	-	Mix	Same Actor

Table B.22: ID# 19:This table contains five expected, three indeterminant messages, and two unexpected profiles with same actors, each with three unexpected messages or indeterminant messages occurring for SPC F to SPC J in a combination of beginning, middle, end, and random, respectively.

Profiles	Step	Expected Messages	Unexpected Messages	Indeterminant Messages	Location	same/ different actor
SPC A	1 min	All Expected	-	-	-	All Actor
SPC B	2 min	All Expected	-	-	-	All Actor
SPC C	5 min	All Expected	-	-	-	All Actor
SPC D	10 min	All Expected	-	-	-	All Actor
SPC E	15 min	All Expected	-	-	-	All Actor
SPC F	20 min	-	-	3	Beginning	Same Actor
SPC G	25 min	-	-	3	Beginning	Same Actor
SPC H	30 min	-	-	3	Beginning	Same Actor
SPC I	40 min	-	3	-	Beginning	Same Actor
SPC J	1 hr	-	3	-	Beginning	Same Actor

Table B.23: ID# 20:This table contains five expected, three indeterminant messages, and two unexpected profiles with same actors, each with three unexpected or indeterminant messages occurring in the beginning.

Profiles	Step	Expected Messages	Unexpected Messages	Indeterminant Messages	Location	same/ different actor
SPC A	1 min	All Expected	-	-	-	All Actor
SPC B	2 min	All Expected	-	-	-	All Actor
SPC C	5 min	All Expected	-	-	-	All Actor
SPC D	10 min	All Expected	-	-	-	All Actor
SPC E	15 min	All Expected	-	-	-	All Actor
SPC F	20 min	-	-	3	Middle	Same Actor
SPC G	25 min	-	-	3	Middle	Same Actor
SPC H	30 min	-	-	3	Middle	Same Actor
SPC I	40 min	-	3	-	Middle	Same Actor
SPC J	1 hr	-	3	-	Middle	Same Actor

Table B.24: ID# 21:This table contains five expected, three indeterminant messages, and two unexpected profiles with same actors, each with three unexpected or indeterminant messages occurring in the middle.

Profiles	Step	Expected Messages	Unexpected Messages	Indeterminant Messages	Location	same/ different actor
SPC A	1 min	All Expected	-	-	-	All Actor
SPC B	2 min	All Expected	-	-	-	All Actor
SPC C	5 min	All Expected	-	-	-	All Actor
SPC D	10 min	All Expected	-	-	-	All Actor
SPC E	15 min	All Expected	-	-	-	All Actor
SPC F	20 min	-	-	3	End	Same Actor
SPC G	25 min	-	-	3	End	Same Actor
SPC H	30 min	-	-	3	End	Same Actor
SPC I	40 min	-	3	-	End	Same Actor
SPC J	1 hr	-	3	-	End	Same Actor

Table B.25: ID# 22:This table contains five expected, three indeterminant messages, and two unexpected profiles with same actors, each with three unexpected or indeterminant messages occurring in the end.

Profiles	Step	Expected Messages	Unexpected Messages	Indeterminant Messages	Location	same/ different actor
SPC A	1 min	All Expected	-	-	-	All Actor
SPC B	2 min	All Expected	-	-	-	All Actor
SPC C	5 min	All Expected	-	-	-	All Actor
SPC D	10 min	All Expected	-	-	-	All Actor
SPC E	15 min	All Expected	-	-	-	All Actor
SPC F	20 min	-	-	3	Mix	Different Actor
SPC G	25 min	-	-	3	Beginning	Different Actor
SPC H	30 min	-	-	3	Middle	Different Actor
SPC I	40 min	-	3	-	End	Different Actor
SPC J	1 hr	-	3	-	Mix	Different Actor

Table B.26: ID# 23:This table contains five expected, three indeterminant messages, and two unexpected profiles with different actors, each with three unexpected messages or indeterminant messages occurring for SPC F to SPC J in a combination of beginning, middle, end, and random, respectively.

Profiles	Step	Expected Messages	Unexpected Messages	Indeterminant Messages	Location	same/ different actor
SPC A	1 min	All Expected	-	-	-	All Actor
SPC B	2 min	All Expected	-	-	-	All Actor
SPC C	5 min	All Expected	-	-	-	All Actor
SPC D	10 min	All Expected	-	-	-	All Actor
SPC E	15 min	All Expected	-	-	-	All Actor
SPC F	20 min	-	-	3	Beginning	Different Actor
SPC G	25 min	-	-	3	Beginning	Different Actor
SPC H	30 min	-	-	3	Beginning	Different Actor
SPC I	40 min	-	3	-	Beginning	Different Actor
SPC J	1 hr	-	3	-	Beginning	Different Actor

Table B.27: ID# 24:This table contains five expected, three indeterminant, and two unexpected message profiles with different actors—each with three unexpected or indeterminant messages occurring in the beginning.

Profiles	Step	Expected Messages	Unexpected Messages	Indeterminant Messages	Location	same/ different actor
SPC A	1 min	All Expected	-	-	-	All Actor
SPC B	2 min	All Expected	-	-	-	All Actor
SPC C	5 min	All Expected	-	-	-	All Actor
SPC D	10 min	All Expected	-	-	-	All Actor
SPC E	15 min	All Expected	-	-	-	All Actor
SPC F	20 min	-	-	3	Middle	Different Actor
SPC G	25 min	-	-	3	Middle	Different Actor
SPC H	30 min	-	-	3	Middle	Different Actor
SPC I	40 min	-	3	-	Middle	Different Actor
SPC J	1 hr	-	3	-	Middle	Different Actor

Table B.28: ID# 25:This table contains five expected, three indeterminant, and two unexpected message profiles with different actors—each with three unexpected or indeterminant messages occurring in the middle.

Profiles	Step	Expected Messages	Unexpected Messages	Indeterminant Messages	Location	same/ different actor
SPC A	1 min	All Expected	-	-	-	All Actor
SPC B	2 min	All Expected	-	-	-	All Actor
SPC C	5 min	All Expected	-	-	-	All Actor
SPC D	10 min	All Expected	-	-	-	All Actor
SPC E	15 min	All Expected	-	-	-	All Actor
SPC F	20 min	-	-	3	End	Different Actor
SPC G	25 min	-	-	3	End	Different Actor
SPC H	30 min	-	-	3	End	Different Actor
SPC I	40 min	-	3	-	End	Different Actor
SPC J	1 hr	-	3	-	End	Different Actor

Table B.29: ID# 26:This table contains five expected, three indeterminant, and two unexpected message profiles with different actors—each with three unexpected or indeterminant messages occurring in the end.