Anomaly Detection on BACnet Systems

Sofian Ghazali
Rahul Balamurugan
Zahid Kamil

Dr. Hussein Al-Nuweiri
Dr. Ferdous Wahid
Mr. Salah Hessien

BMS – Building Management Systems

BACnet – Building Automation Control Network

Automated Building System Controls [1]

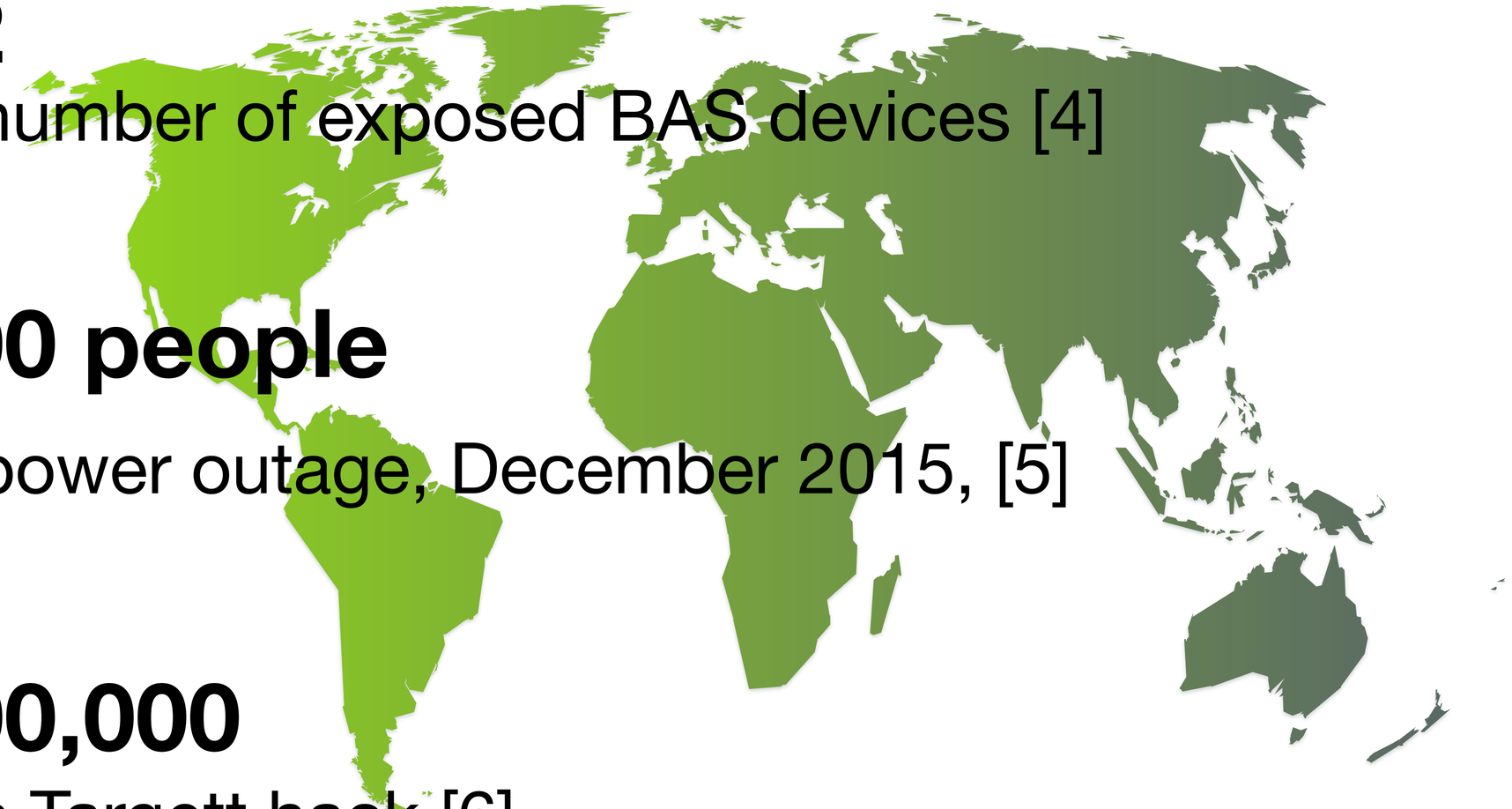
Bacnet Logo [2]

Standards

ANSI-ASHRAE 135-2016 [3]

48,112

Current number of exposed BAS devices [4]



230,000 people

Ukraine power outage, December 2015, [5]

$18,500,000

Losses in Targett hack [6]

Problem Statement

“There is a need for robust intrusion detection systems using Machine Learning to alert the user of malicious zero-day attacks”

Objectives

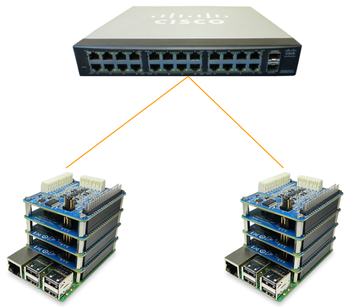
- [x] **Understand BMS BACnet data traffic**
- [x] **Classify BACnet traffic and train the anomaly detector**
- [x] **Evaluate detector performance against novel attacks**

Constraints

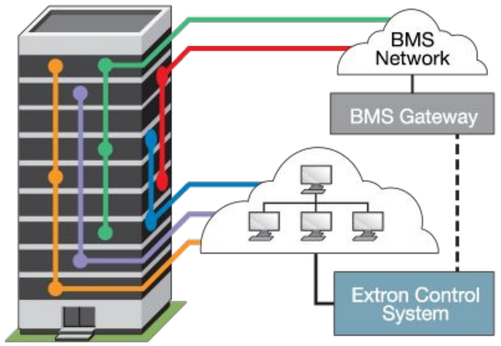
Data for training[7]

Lack of anomalous data or real system for testing [8]

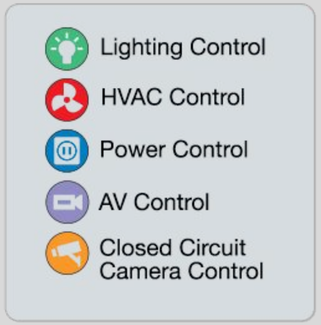
Solutions



Cyber-physical BACnet emulator



Implement model on real BACnet BMS



Addressing the Objectives

1. **Gathered BACnet traffic data set from Qatar University**

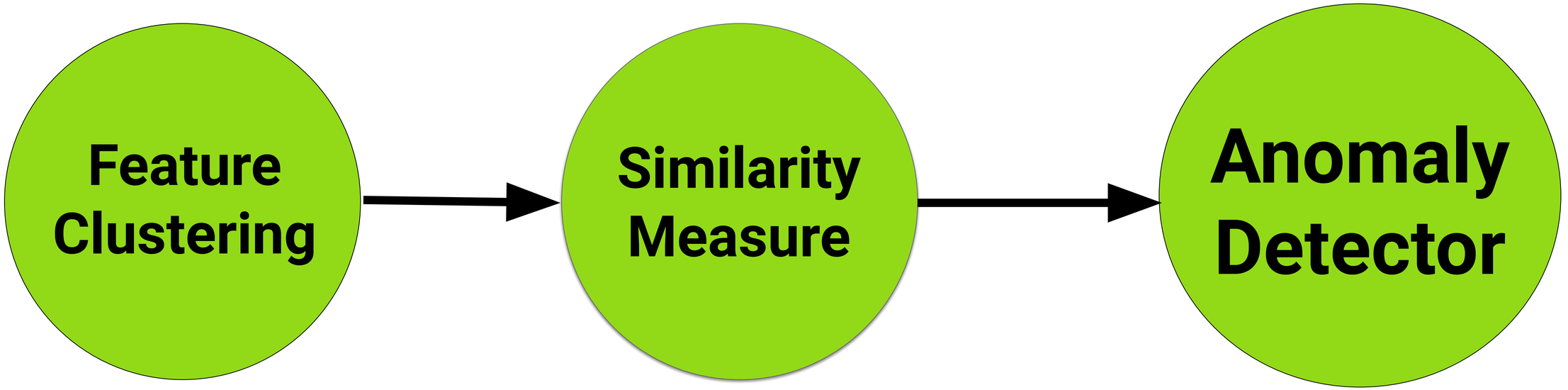
2. **Built representation of the data, trained detector**

3. **Tested on synthesized anomalous data**

Performance Criteria & Analysis

	AC2000 Interface [11]	THE-Driven Anomaly Detector [12]	Our Solution
Responsiveness	Responsive to certain attacks	Not dynamic	Highly Responsive due to unsupervised learning capabilities
Adaptability	Limited to predetermined threat models	Limited by frequency analysis Techniques	Needs human input to adapt to new patterns, but easily does so
Comprehensiveness	Threats like Man-in-the Middle or DOS can intercept the system	Can classify traffic into Time, Human based and event based	Accounts for a wide range of threats as it only learns normal patterns
Economics	Need separate device-moderately expensive	Can be run on the main host- Cheap implementation	Can be run on the main host- Cheap implementation
Global	Can be applied in any BACnet protocol systems	Can be applied in any BACnet protocol systems	BACnet network across the globe
Accuracy	Information Not Available	~96%	~98%

Anomaly Detection with ML

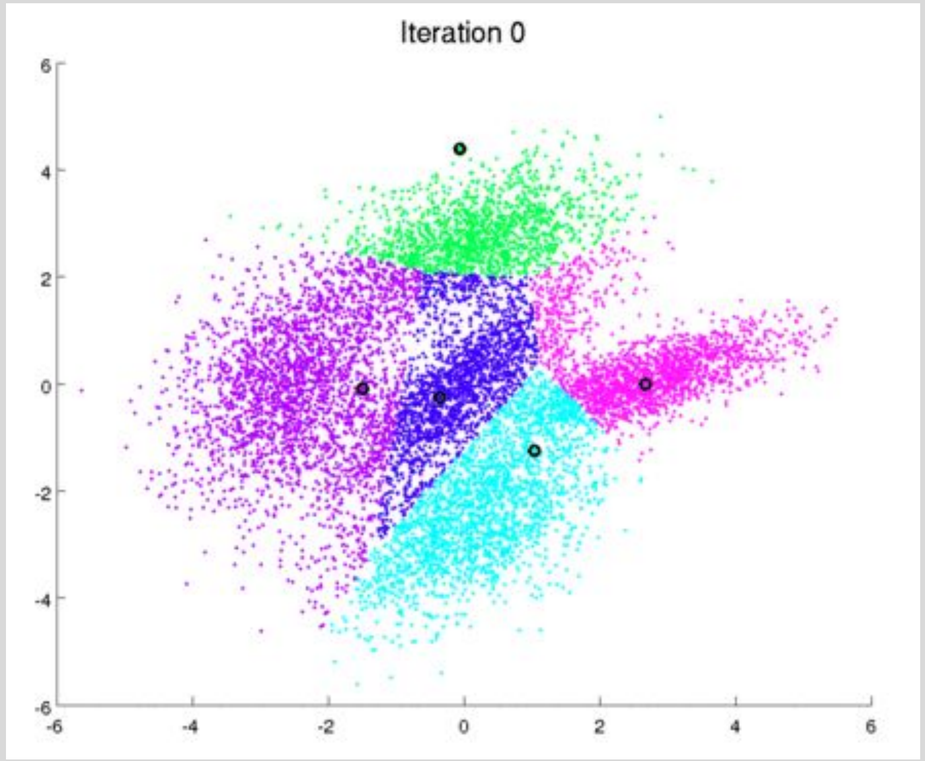


Take features X1, X2 & X3 and cluster into groups

Take inputs from user and compare with model to find dissimilarity

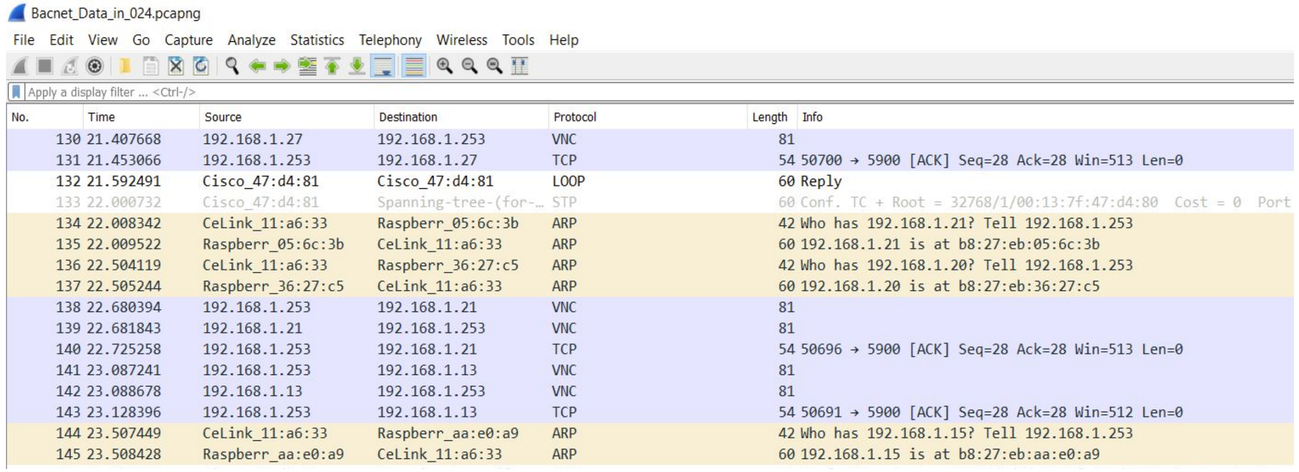
Dissimilarity Measure goes above threshold, its an anomaly

K-Modes and K-Means Algorithm



1. **Choose Number of cluster K**
2. **Pick an observation to be centroid**
3. **Compare data points to centroid & Cluster based on mode and mean of features**
4. **Assign data to the closest centroid**
5. **Update centroid & Repeat until centroid converges**

Preprocessing



No.	Time	Source	Destination	Protocol	Length	Info
130	21.407668	192.168.1.27	192.168.1.253	VNC	81	
131	21.453066	192.168.1.253	192.168.1.27	TCP	54	50700 → 5900 [ACK] Seq=28 Ack=28 Win=513 Len=0
132	21.592491	Cisco_47:d4:81	Cisco_47:d4:81	LOOP	60	Reply
133	22.000732	Cisco_47:d4:81	Spanning-tree-(for-...	STP	60	Conf. TC + Root = 32768/1/00:13:7f:47:d4:80 Cost = 0 Port
134	22.008342	CeLink_11:a6:33	Raspberr_05:6c:3b	ARP	42	Who has 192.168.1.21? Tell 192.168.1.253
135	22.009522	Raspberr_05:6c:3b	CeLink_11:a6:33	ARP	60	192.168.1.21 is at b8:27:eb:05:6c:3b
136	22.504119	CeLink_11:a6:33	Raspberr_36:27:c5	ARP	42	Who has 192.168.1.20? Tell 192.168.1.253
137	22.505244	Raspberr_36:27:c5	CeLink_11:a6:33	ARP	60	192.168.1.20 is at b8:27:eb:36:27:c5
138	22.680394	192.168.1.253	192.168.1.21	VNC	81	
139	22.681843	192.168.1.21	192.168.1.253	VNC	81	
140	22.725258	192.168.1.253	192.168.1.21	TCP	54	50696 → 5900 [ACK] Seq=28 Ack=28 Win=513 Len=0
141	23.087241	192.168.1.253	192.168.1.13	VNC	81	
142	23.088678	192.168.1.13	192.168.1.253	VNC	81	
143	23.128396	192.168.1.253	192.168.1.13	TCP	54	50691 → 5900 [ACK] Seq=28 Ack=28 Win=512 Len=0
144	23.507449	CeLink_11:a6:33	Raspberr_aa:e0:a9	ARP	42	Who has 192.168.1.15? Tell 192.168.1.253
145	23.508428	Raspberr_aa:e0:a9	CeLink_11:a6:33	ARP	60	192.168.1.15 is at b8:27:eb:aa:e0:a9

Wireshark captured data (pcap file)

Time	Source	Destination	Protocol	Length	Info	Messages
0.000000	10.10.10.43	10.30.10.12	BACnet-APDU	187.0	Complex-ACK readPropertyMultiple[87]	Read-Property

BACnet Data in csv format

	Ack-Message	At	I-am	Read-Property	Tell	Unconfirmed-Transfer	Who-Is	Who-Router	Write-Property
0	20	23	14	333	30	42	20	47	6

Frequency information (one time interval)

	bit1	bit2	bit3	bit4	bit5	...	Messages_Tell	Messages_Unconfirmed-Transfer	Messages_Who-Is
0	0	0	0	0	1	...	0	0	0

Preprocessed dataframe (normal dataset)

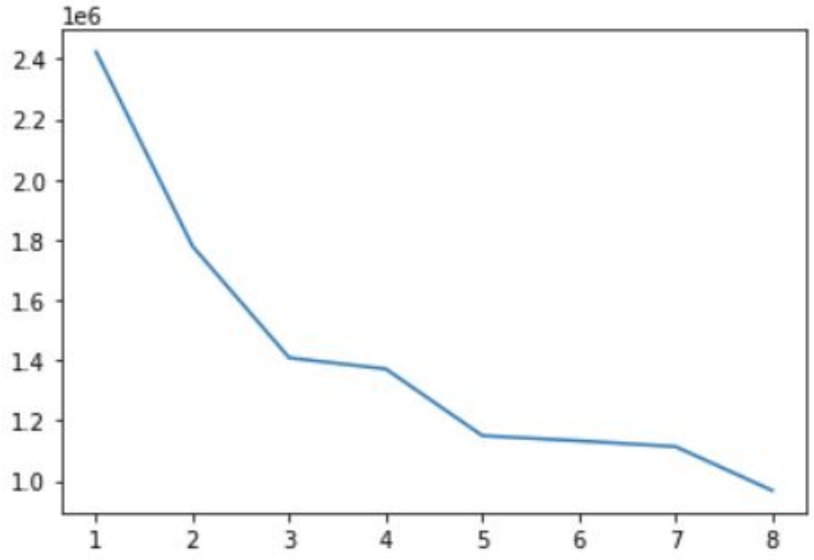
	Messages_Ack-Message	Messages_At	Messages_I-Am	Messages_Read-Property	Messages_Tell	Messages_Unconfirmed-Transfer	Messages_Who-Is	Messages_Who-Router	Message
0	0.096294	0.537352	-1.415910	0.173225	0.114256	0.227122	0.288777	-0.697657	-

Scaled frequency information (one time interval)

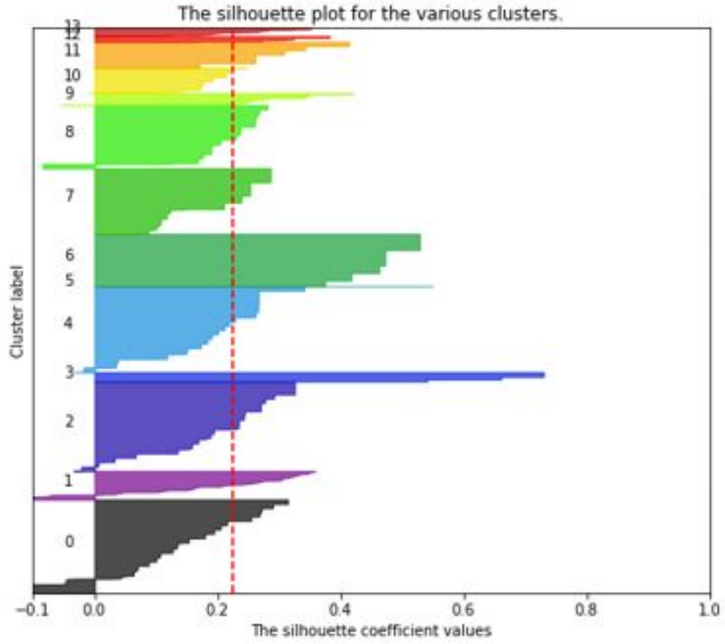
Cluster Selection

We used two methods to determine the number of our clusters that can best represent our data:

1. Elbow method
2. Silhouette method



1. Elbow Method



2. Silhouette Method

Anomaly Detection

Algorithm:
Data = input arrays, cluster arrays
Output = Dissimilarity score

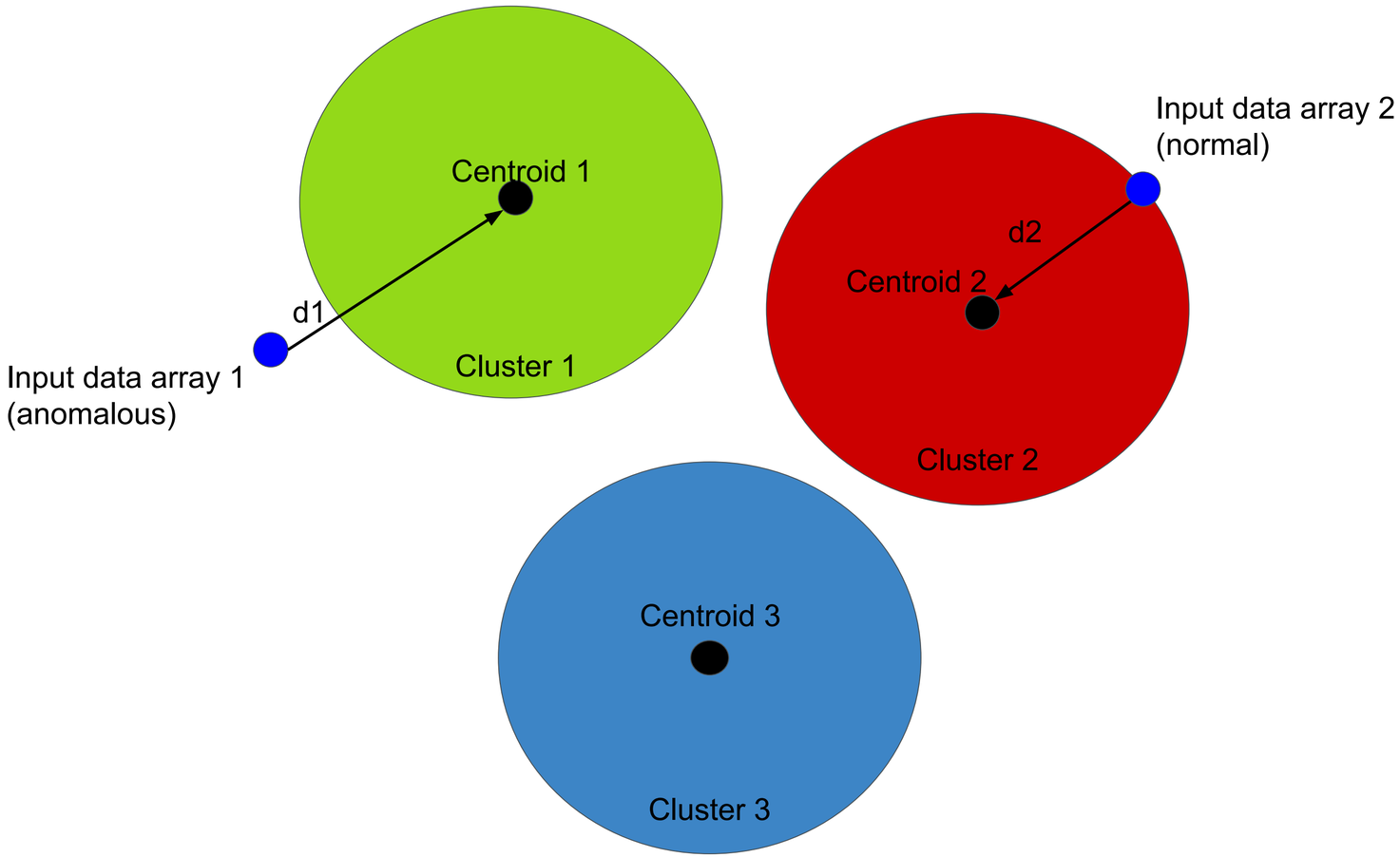
Start:
di = min{score(input,r)}, r∈Clusters
if di<Threshold, classified as normal.
Otherwise, determined to be an anomaly.
return(di)
End

Dissimilarity Score for K-modes:
Jaccard-Needham Dissimilarity

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}.$$

Distance measure for K-means:
Squared Euclidean Distance

$$\|x - y\| = \sum_{i=1}^{d} (x_i - y_i)^2$$



Threshold setting

Message & IP Anomaly Scoring Functions - Jaccard Needham Dissimilarity

	Jaccard	Dice	Hamming	Rogers-Tanimoto	Sokal-Michener	Sokal-Sneath	Yule
3	0.0454545	0.0232558	0.0131579	0.025974	0.025974	0.0869565	0

DOS Attack scoring function - Squared Euclidean Distance

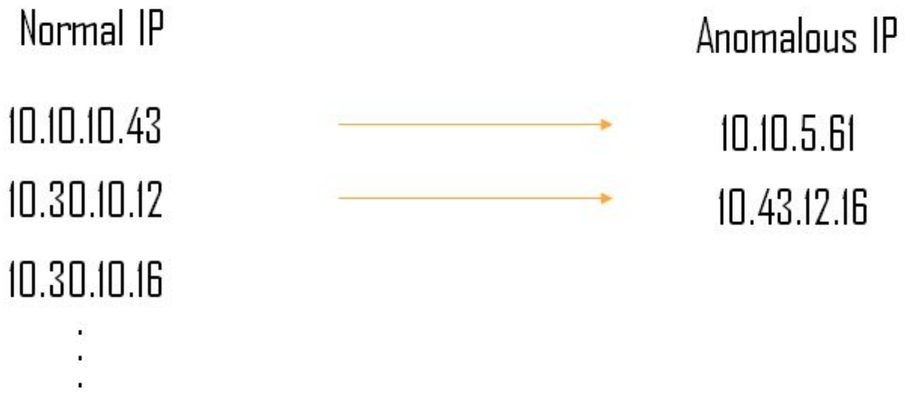
```
Threshold = 9.112510349254734

Checking Anomaly Detector:
   Anomaly Detector passed normal data check!
```

Maximum score for each respective scoring function -> Threshold

Any score above (Jaccard, 0.45) or (Euclidean, 9.3) is determined to be an anomaly.

Model Testing

Generating anomalous dataset:



Input exceeds above threshold of 0.45 (input > 0.45) -> Anomaly
If any row exceeds the threshold (Jaccard 0.45) it was considered as an anomaly.

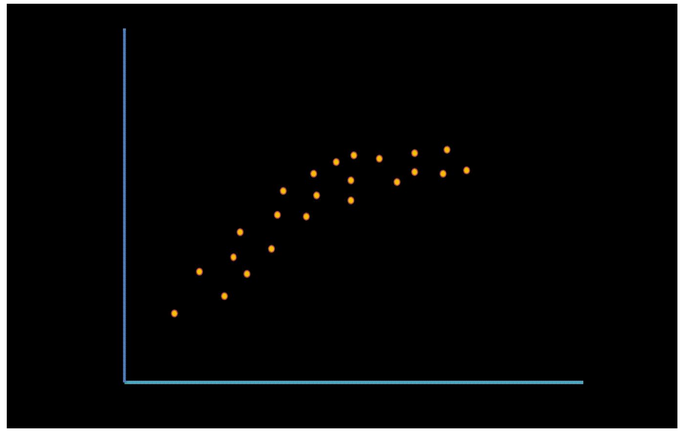
98% Accuracy

Graphical User Interface (GUI)

GUI Application for testing our model for messages and ips

GUI Application for testing our model with Frequency

Risks



Overfitting [9]

Security [10]

Ethics

Potential Breach - Transparency [12]

Data Collection - Privacy [11]

To Recap

1. Implement Unsupervised ML algorithm
2. Build an Anomaly Detector using Similarity Score Algorithm
3. Build prototype GUI to track anomalies

Thank You!

References

1. MLN Company. (2019). *Building Management Systems - MLN Company*. [online] Available at: http://www.mlncompany.com/what-we-do/building-management-systems/ [Accessed 4 Dec. 2019].
2. Dms.hvacpartners.com. (2019). Bacnet Basics User's Guide. [online] Available at: https://dms.hvacpartners.com/docs/1000/Public/04/11-808-417-01.pdf [Accessed 3 Sep. 2019].
3. Metadata, https://censys.io/ipv4/metadata?q=bacnet&
4. Pavel et al, "Ukraine's power outage was a cyber attack: Ukrenergo", Reuters, Jan 18,2017. Accessed 16 Sep 2019
5. Google Search. [Online]. Available: https://www.google.com/search?q=laptop&sxsrf=ACYBGNTFECOoj5LsWVDNZAtvfmGRDvmjTg:1575445883360&source=lnms&tbm=isch&sa=X&ved=2ahUKEwiG_vWhwZvmAhXFqlkKHUNgD4QQ_AUoAXoECA8QAw&biw=1536&bih=722#imgrc=kte5opKgrQ4V2M: [Accessed: 04-Dec-2019].
6. Google.com. (2019). [online] Available at: https://www.google.com/search?biw=1536&bih=722&tbm=isch&sxsrf=ACYBGNS8tUPvlhPpovtFekzpGNjoJJSAQw%3A1575445890149&sa=1&ei=gmXnXdvkCKHH_Qb-uYGgCA&q=python+logo&oq=python+logo&gs_l=img.3..0i67l6j0l4.172058.173659..173854...1.0..0.263.1954.2-8.....0....1..gws-wiz-img.....10..35i39j35i362i39j0i131.gFK6bEnemfw&ved=0ahUKEwjbrZSlwZvmAhWhY98KHf5cAlQQ4dUDCAc&uact=5#imgrc=0-lAWpuQBjsPGM: [Accessed 4 Dec. 2019].
7. "Download," *Wireshark · Go Deep.* [Online]. Available: https://www.wireshark.org/. [Accessed: 04-Dec-2019].
8. https://www.google.com/search?biw=1536&bih=722&tbm=isch&sxsrf=ACYBGNQ-tVlGHlieP1T0np9TKirjY3qQ9w%3A1575446119500&sa=1&ei=Z2bnXYuQHu2c_Qbfj7XoAw&q=vnc+viewer+logo&oq=vnc+viewer+logo&gs_l=img.3..0.10232.10815..11082...0.0..0.241.1666.2-7.....0....1..gws-wiz-img.....35i39j0i24.BsssxU3wkTM&ved=0ahUKEwjL38KSwpvmAhVtTt8KHd9HDT0Q4dUDCAc&uact=5#imgrc=sk6RlltId5PenM:
9. Ashrae.org. (2019). *Standard 135-2016, BACnet™ -- A Data Communication Protocol for Building Automation and Control Networks*. [online] Available at: https://www.ashrae.org/technical-resources/standards-and-guidelines/standards-addenda/standard-135-2016-bacnet-a-data-communication-protocol-for-building-automation-and-control-networks [Accessed 4 Dec. 2019].
10. "AC2000," AC2000 Access Control & Security Management | CEM Systems. [Online]. Available: https://www.cemsys.com/products/access-control-systems/ac2000/ . [Accessed: 06-Nov-2019].
11. Z. Zheng and A. Reddy, "Safeguarding Building Automation Networks: THE-Driven Anomaly Detector Based on Traffic Analysis," 2017. [Online]. Available: http://cesg.tamu.edu/wpcontent/uploads/2018/01/ICCCN-Zhiyuan-Zheng-Paper-2017-July-1.pdf