Santa Clara University
Internet of Things
RESEARCH LAB
SIOTLAB

Santa Clara University
School of Engineering

Securing Smart Homes via Software-Defined Networking and Low-Cost Traffic Classification

Holden Gordon, Christopher Batula, Bhagyashri Tushir, Behnam Dezfouli, Yuhong Liu

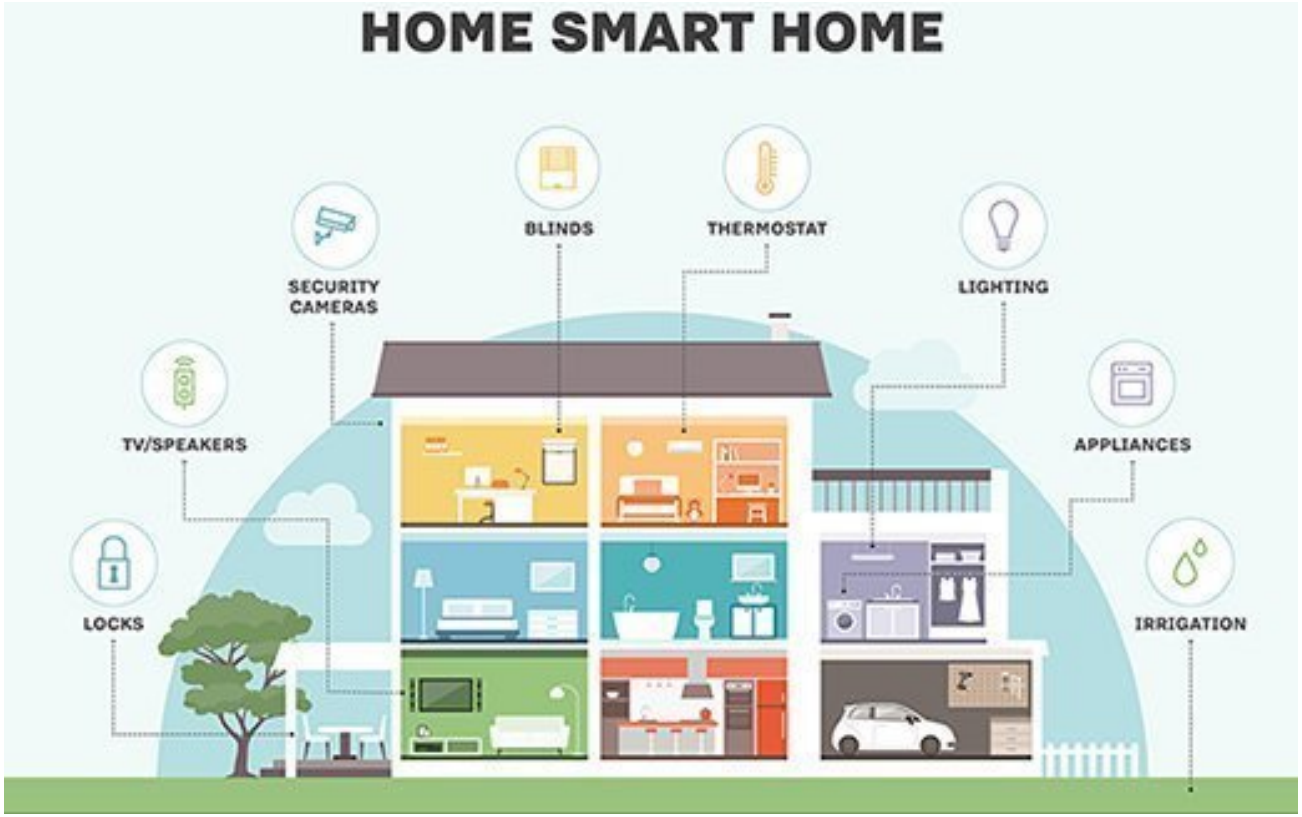
COMPSAC
July 12 — July 16
2021

Agenda

Santa Clara University
Internet of Things
RESEARCH LAB
SIOTLAB

- Smart home and smart home security
- Contribution
- Testing environment
- Result and discussions

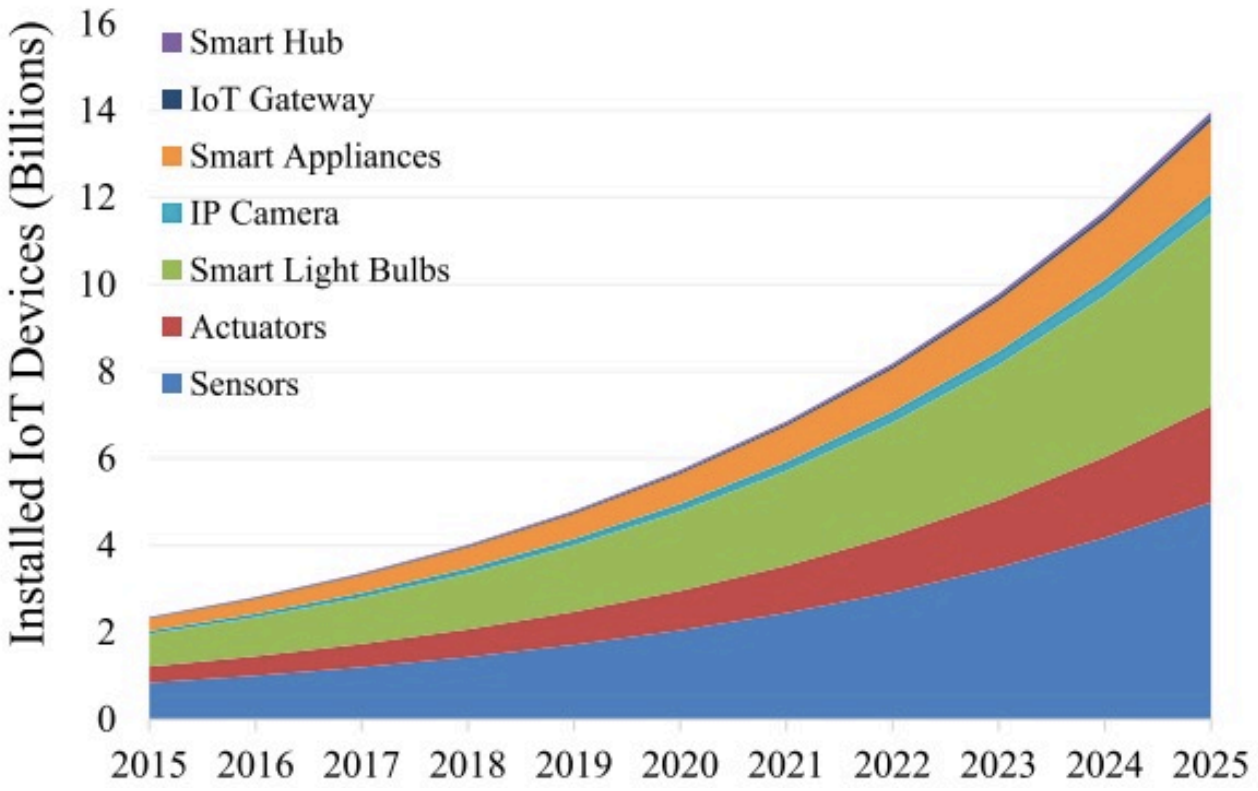
Smart Home Definition



Source: online

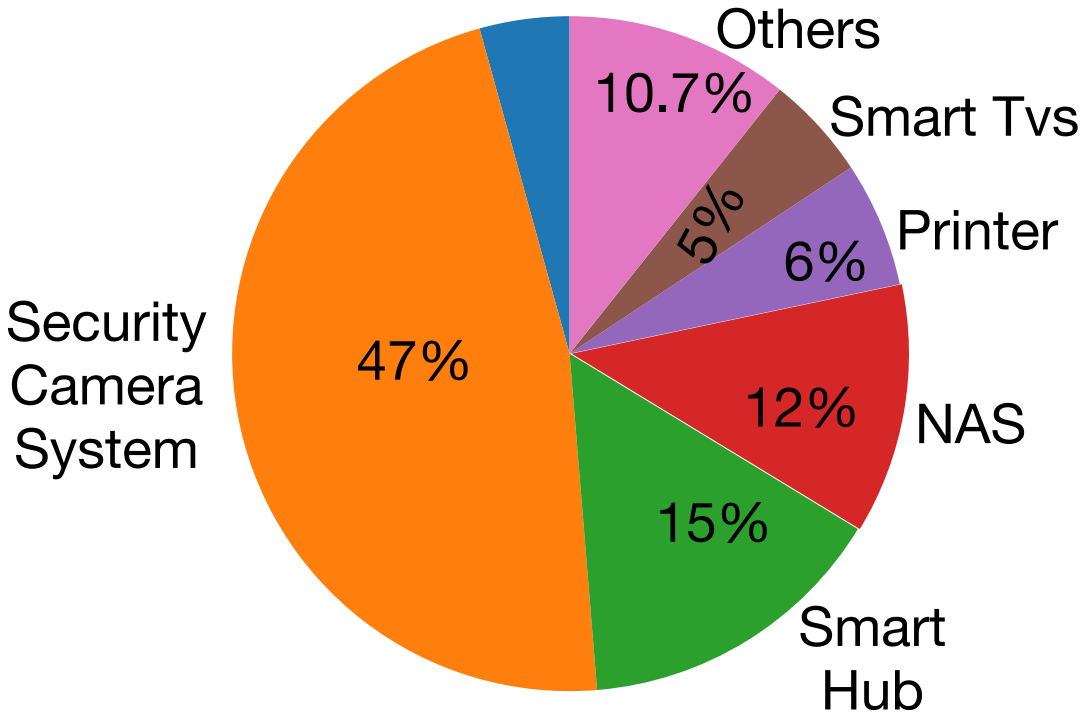
- Smart homes include wired and wireless interconnected devices
- Provide services such as voice assistants, security cameras and many more

Growth of Smart Home IoT Devices



The number of interconnected smart home devices will reach 72 billion by 2025 [1]

Smart Home Devices Security



Source: HelpNetSecurity

IoT devices are attractive to hackers

Attacks on Smart Home Devices

- IoT devices are vulnerable to various attacks:
  - **distributed Denial of Service attacks (DDoS)**
  - network scan attacks
  - SQL injections
  - zero-day attacks
  - ransomware attacks

- Reasons:
  - heterogeneity of IoT devices
  - low storage and computation resources
  - generate massive data

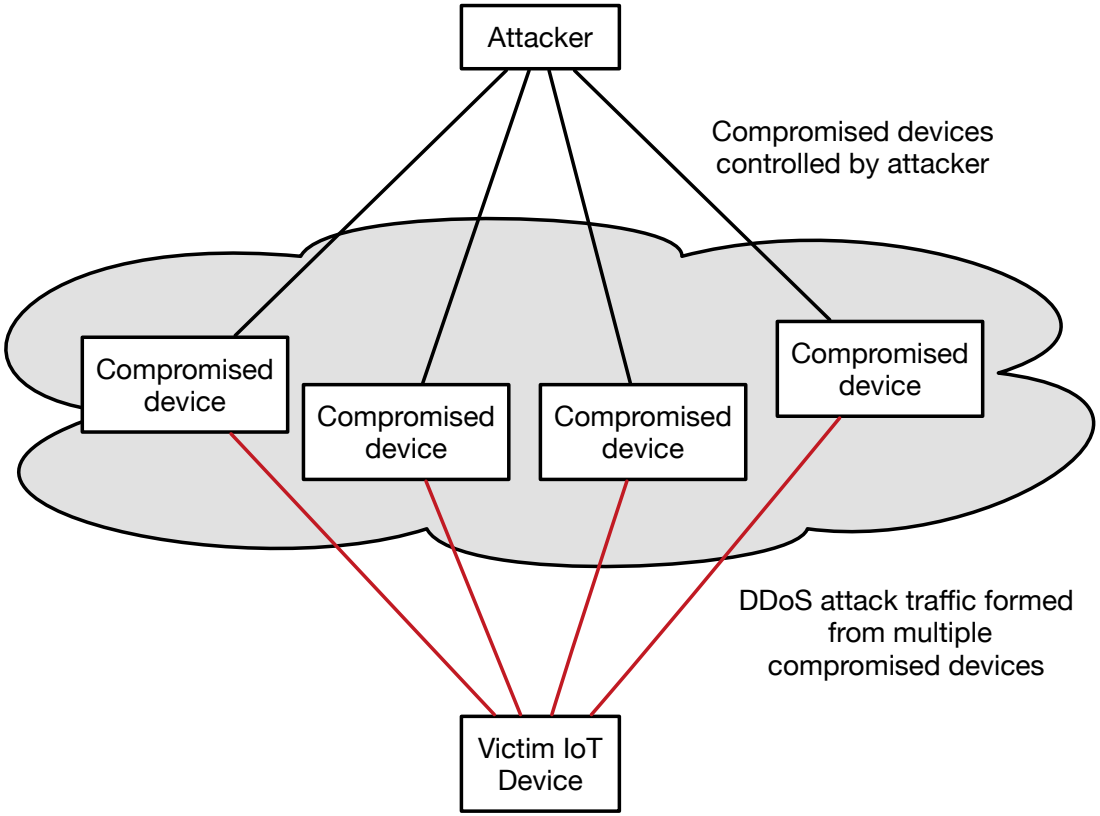
Software-Defined Networking

- Traditional systems are inadequate because:
  - IoT devices generate a high volume of data
  - a large amount of connected IoT devices

- **Solution**: Software-Defined Networking (SDN)
  - provides centralize network management
  - provides flow-based statistics
  - helps developing flexible solutions

- SDN and Machine Learning
  - helps automated decisions making
  - enhance smart home security via **IoT device classification and DDoS detection**

DDoS Attacks



- DDoS attacks disrupt the normal operation of devices
- DDoS attacks can also disconnect a device from its associated WiFi access point

Contribution

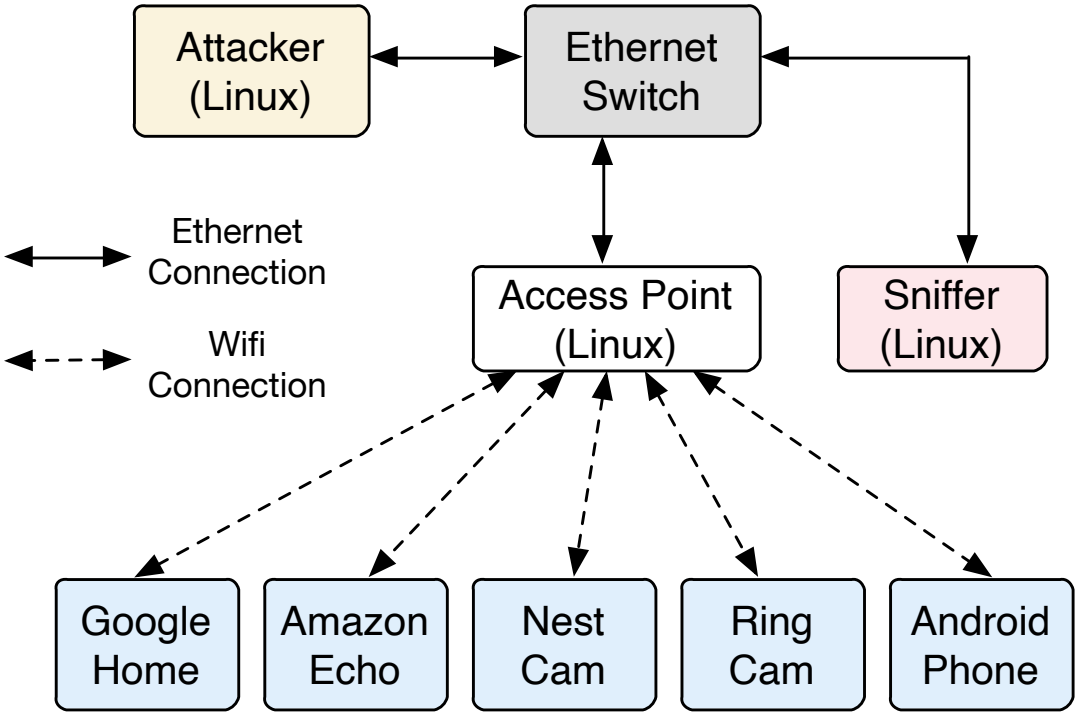
- we propose a SDN-based architecture consisting of a Faucet controller & Open vSwitch to secure smart homes
- the proposed framework is implemented on GNS3 network simulator to measure its performance
- we proposed a minimum set of flow-based features for both **IoT device classification and DDoS detection**
- we utilize non-cumulative statistics to increase framework's accuracy while reducing controller and switch communication overhead

Contribution

- to validate the efficacy and reliability of the proposed features 2 different real-world dataset are used
- the minimum dataset size to meet 95% accuracy is identified.
- the accuracy of KNN, SVM and RF are investigated in detail with varying polling interval.
- the latency for the minimum dataset is identified.

Proposed Testing Environment

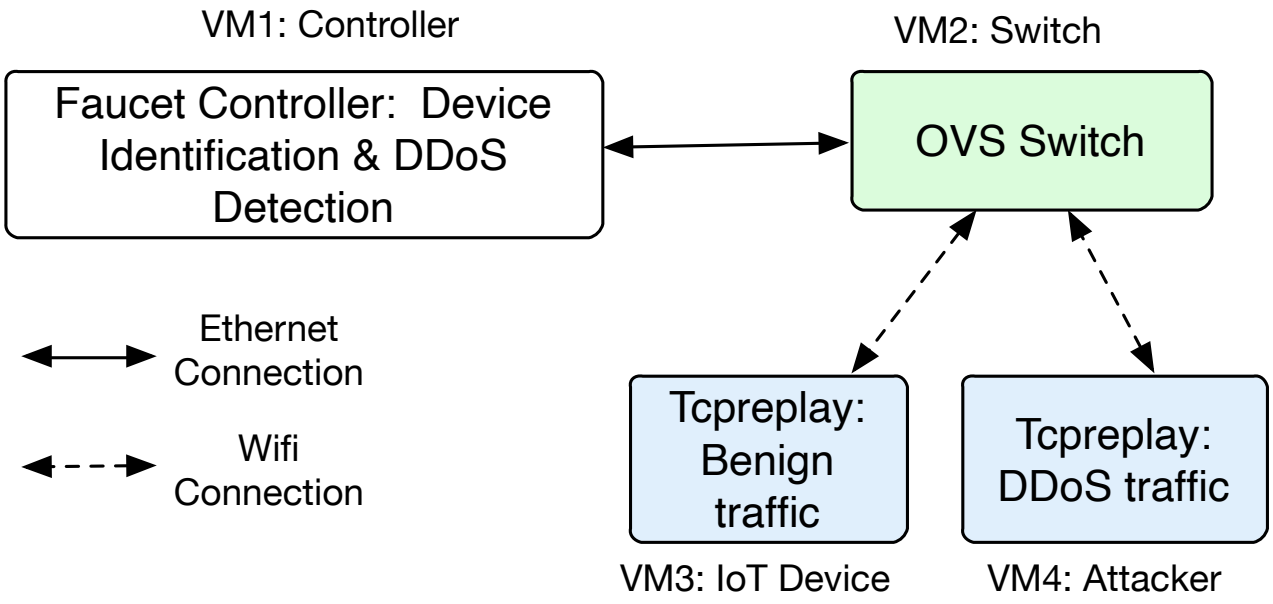
Smart Home network



Smart home network for benign and DDoS attack data collection

Proposed Testing Environment

GNS3 Environment



Virtual SDN-based smart home environment using GNS3 simulator

Proposed Testing Environment

Datasets and IoT devices used

- To confirm the robustness of machine learning models:
  - used three datasets: SIOTLAB (our lab), UNSW dataset [5], combined dataset (SIOTLAB and UNSW)

Device Category	Device
Switches/Triggers	WeMo Motion Sensor and Power Switch, TP Link Smart Plug, Chromecast
Camera Systems	Ring Camera, Nest Camera, Samsung Camera, Netatmo Camera
IoT Hub Devices	Amazon Echo, LiFX, Hue Bulb, iHome, Google Home

Proposed Flow-based Features

The proposed framework uses the following flow-based and stateless features obtained per polling interval:

- **Protocol percentage:** the percentage of ICMP, TCP, and UDP packets
- **IP diversity ratio:** calculated as the number of unique IP addresses divided by the total number of packets sent by a device
- **Packet size and count**

Proposed Testing Environment

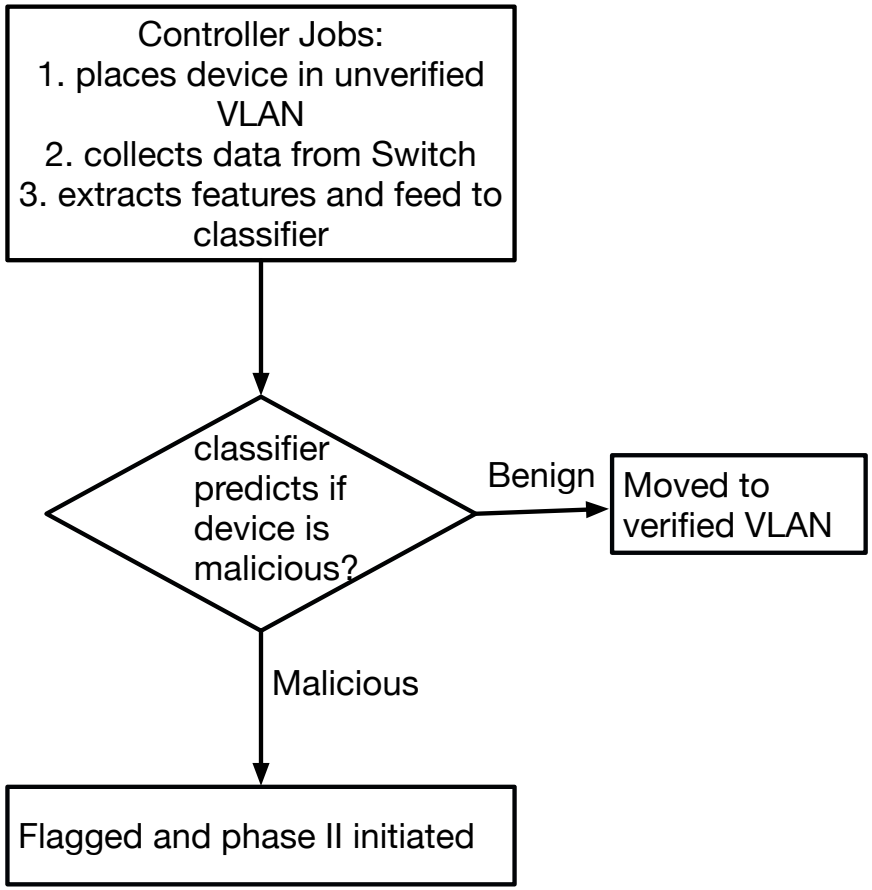
Non-cumulative statistics

- Traditional SDN networks working:
  - switches send cumulative statistics of packets sent and received to controller periodically
  - statistics are accurate at lower polling interval that increases network congestion
- we propose to use non-cumulative statistics to collect accurate statistics:
  - every poll from controller to switch resets flow gauges on controller
  - thus non-cumulative statistics are independent of events that happened before this period

Proposed Secure Home Architecture

Phase I

The proposed architecture has two phases:

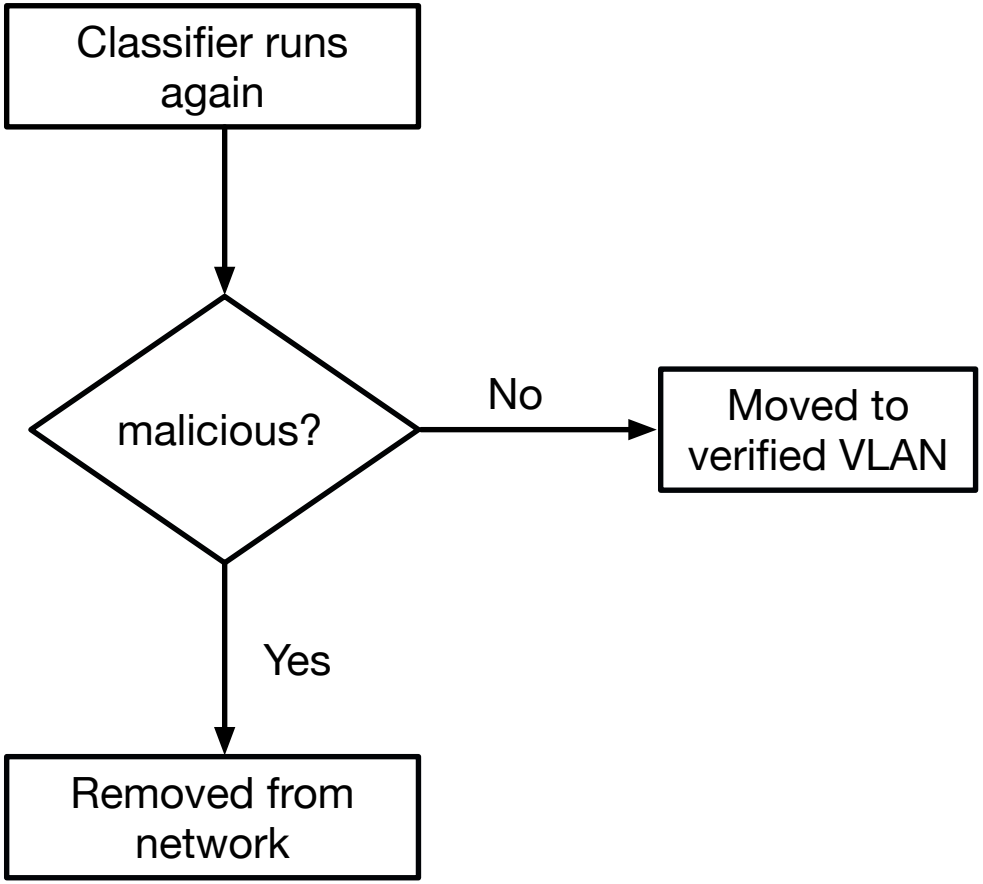


Phase I

Proposed Secure Home Architecture

Phase II

The proposed architecture has two phases:



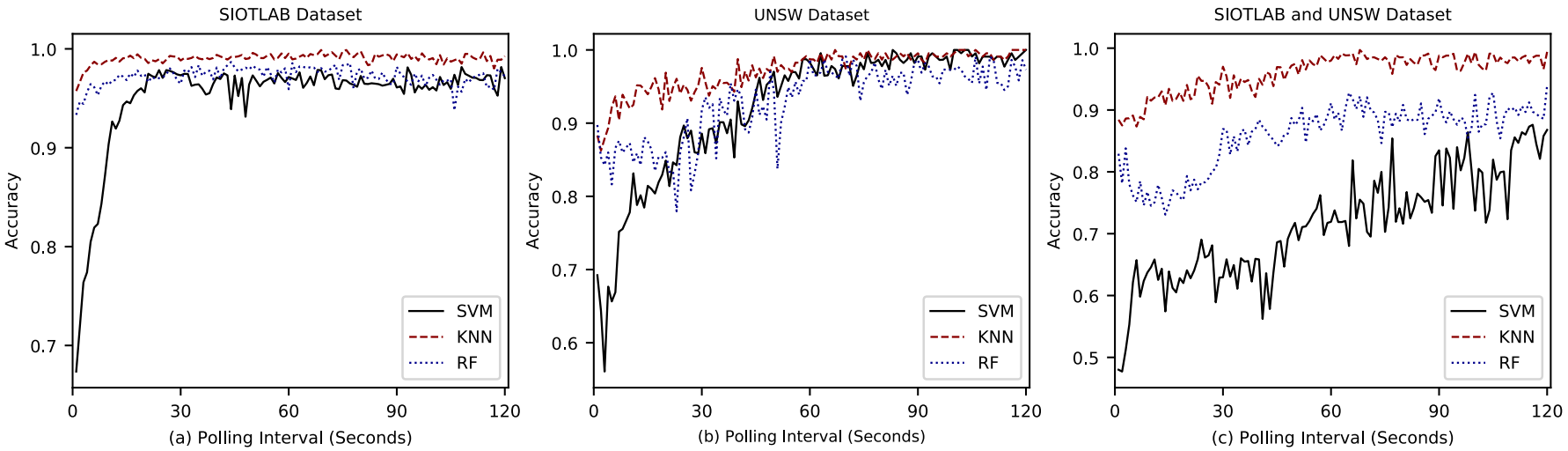
Phase II

Machine learning algorithms

- Machine Learning algorithms used:
  - K-Nearest Neighbors algorithm (KNN)
  - Support Vector Machine with linear kernel (LK-SVM)
  - Random Forest using Gini impurities (RF)
- Hyperparameters are default
- Split data into 75% train and 25% test data
- Classes are balanced

Result and Discussion

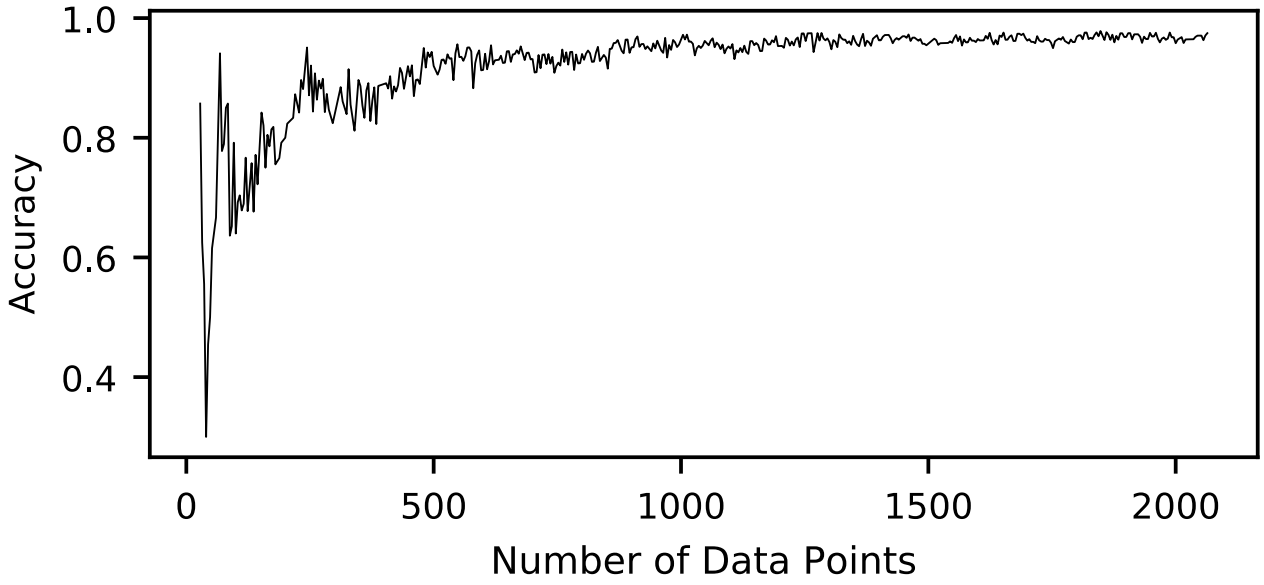
Machine learning algorithms



All models demonstrate accuracy improvement versus increasing polling interval

Result and Discussion

Minimized dataset



- for the combined dataset, KNN shows 95% accuracy for 1820 data points, meeting the low memory requirement

- for 100 trials KNN's latency on a machine with 1.4 GHz i5 processor & 4 GB RAM is 1.18 ms

Result and Discussion

Comparison with existing work

Existing work	Maximum Accuracy	Detection type	Dataset	Storgae	Latency	Feature Count
Owusu et al. [2]	92.7%	Classification	Tor	No	No	6
Reza et al. [3]	97.6%	Classification	Custom	No	No	13
Xu et al. [4]	87.8%	Classification	Custom	No	No	21
Hamza et al. [5]	97.5%	DDoS	UNSW	Yes (14.2 MB)	Yes (13 ms)	20
Doshi et al. [6]	99.8	DDoS	Custom	No	No	11
Yang et al. [7]	99.8%	DDoS	KD99	No	No	9
Proposed Solution	**99.9%**	**Classification & DDoS**	**Custom & UNSW**	**Yes (541 KB)**	**Yes (1.18 ms)**	**6**

References

[1] C. Gray, R. Ayre, K. Hinton, and L. Campbell, "'smart'is not free: Energy consumption of consumer home automation systems," *IEEE Transactions on Consumer Electronics, 2019*

[2] A. I. Owusu and A. Nayak, "An intelligent traffic classification in sdn-iot: A machine learning approach," *in 2020 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom),2020, pp. 1–6*

[3] M. Reza, M. J. Sobouti, S. Raouf, and R. Javidan, "Network traffic classification using machine learning techniques over software defined networks," *International Journal of Advanced Computer Science and Applications , vol. 8, 01 2017*

[4] J. Xu, J. Wang, Q. Qi, H. Sun, and B. He, "Deep neural networks for application awareness in sdn-based network," *in 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP) , 2018, pp. 1–6*

[5] A. Hamza, H. H. Gharakheili, T. A. Benson, and V. Sivaraman, "Detecting volumetric attacks on loT devices via sdn-based monitoring of mud activity," *in Proceedings of the 2019 ACM Symposium on SDN Research , 2019, pp. 36–48*

[6] R. Doshi, N. Apthorpe, and N. Feamster, "Machine learning ddos detection for consumer internet of things devices," in 2018 IEEE Security and Privacy Workshops (SPW). IEEE, 2018, pp. 29–35

[7] L. Yang and H. Zhao, "Ddos attack identification and defense using sdn based on machine learning method," *in 2018 15th International Symposium on Pervasive Systems, Algorithms and Networks (I-SPAN) , 2018, pp. 174–178*

Thank you!

Q & A