Towards a Conversational Agent for Threat Detection in the Internet of Things

Christopher D. McDermott, Bastien Jeanelle, John P. Isaacs

International Conference on Cyber Situational Awareness, Data Analytics and Assessment (Cyber SA 2019), 3-4 June 2019, Oxford, UK.
DDoS in IoT

Verisign DDoS Trends Report Q2 2016

256 Gbps
Peak attack size

Multi-Vector DDoS Attacks

Source: Verisign DDoS Trends Report Q4 2017

Types of DDoS Attacks

Source: Verisign DDoS Trends Report Q4 2017
Motivation

**Challenge**
- Present network information to non-technical users, for better comprehension
- Investigate the effectiveness of learning modalities in given context

**Method**
- Use of existing/familiar Smart Technologies
- Novel Conversational Agent to detect IoT Threats
- Use of Aural analysis to query a database of classified network traffic

**Outcome**
- Improve Situational Awareness of network/device activity in consumer IoT networks
- Demonstrate the effectiveness of aural analysis in detecting IoT threats
Situational Awareness Problem

Previous Work

- Participants \( n = 158 \) shown video feeds from infected/uninfected IP cameras
- Difficult to identify infected IoT device and actions
- Lack of Situational Awareness
- Full results of study published in Journal article: Exploring Awareness and Perception of Botnet Activity within Consumer IoT Networks
Research Contributions

We propose a novel conversational agent for detecting threats facing the Internet of Things.

Our contributions:

1. A scalable serverless ETL pipeline for parsing intrusion detection logs;
ETL Pipeline

**Components**

Dataset classified as *Normal* or *Unusual*.

- **Classified Dataset** – captured during previous research.
  - Local directory monitored.
  - New classified IDS logfiles extracted, transformed, uploaded to S3 bucket.

- **AWS Lambda** – execute code without provisioning server.
  - Lambda function triggered when new json file added to S3 bucket.

- **DynamoDB** – schema-less database to store daily traffic capture.
  - Function executed and data loaded into DynamoDB.

**Steps**

1. Local directory monitored.
2. New classified IDS logfiles extracted, transformed, uploaded to S3 bucket.
3. Lambda function triggered when new json file added to S3 bucket.

**Diagram**

- JSON
- Parsed Log
- AWS S3 Bucket
- AWS Lambda
- AWS DynamoDB
Conversational Agent Framework

Frontend Agent Architecture

Step 1.
- Alexa skill receives converted aural requests from the Echo device.
- Requests mapped to predefined intents which trigger specific event functionality

Step 2.
- AWS Lambda set as endpoint
- Alexa Skill invokes Lambda Function using Handler object

Step 3.
- Query request is triggered to interact with DynamoDB
- When fulfilled appropriate answer to user query is returned to Alexa skill
- Alexa skill generates an aural response from returned answer
- Echo device is invoked and communicates aural response to user

Backend Agent Architecture

NLP (Echo device) converts user input (speech) to text

IAM

AWS Lambda

Alexa Skill

DynamoDB

AWS

Amazon

Echo
**Study Design**

**Participants:**
- Consent to participate was implied
- Convenience sampling employed
- Representative sample of users used for pilot \( n=12 \) later study \( n=72 \)
Agent Evaluation

**Five Use Cases**

*uc1:* returns a summary of all activity taking place today

*uc2:* returns a summary of all activity taking place on a specified date

*uc3:* returns a summary of all activity from a specified source device on a specified date

*uc4:* returns details of the first unusual activity on a specified date

*uc5:* returns details of a specified activity ID

**Pre-Post Agreement Statements**

Strongly Disagree (1) – Strongly Agree (5)

s1: I am confident I can tell if my home network is functioning normally

s2: I am confident I can tell if an IoT device on my network has been compromised

s3: I am confident I can tell if an attack has taken place on my home network

s4: I am confident I can tell which IoT devices are using my home network

Use Case Two Example:

Echo open
Threat Detector
Hi Welcome to Threat Detector, How can I help?
give me a summary of activity for January 10 2019
Alright, on 10/01/2019 there was 74.02 percent normal activity, and 25.98 percent unusual activity.
goodbye
Agent Evaluation

<table>
<thead>
<tr>
<th>Statement</th>
<th>Pre-Test Mo</th>
<th>Post-Test Mo</th>
<th>Md</th>
<th>W</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>s1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>78</td>
<td>0.003</td>
</tr>
<tr>
<td>s2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>36</td>
<td>0.014</td>
</tr>
<tr>
<td>s3</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>28</td>
<td>0.022</td>
</tr>
<tr>
<td>s4</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>1.5</td>
<td>0.999</td>
</tr>
<tr>
<td>s5</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>0.999</td>
</tr>
</tbody>
</table>

- Statements (s1-s3) statistically significant median increase in agreement level suggesting the conversational agent had improved situational awareness in these areas.

- Statements (s4-s5) no demonstration of variance. Pre-test score indicate they were already confident they knew which devices were present and functioning correctly.
Conclusions and Future Work

Conclusions:

- Demonstrated a lack of situational awareness of threats facing consumer IoT networks.
- Demonstrated the effectiveness of aural analysis for detecting threats in consumer IoT networks.

Future Work:

- Current agent improvements:
  - Extend functionality and ability to query threat database
  - Undertake a cross-sectional study with larger sample size
- Extend the conversational agent to utilise verbal analysis:
  - Use of expressed words in written form, using a text based chatbot
  - Undertake a cross-sectional study with larger sample size
- Conduct a longitudinal study to investigate the use of multi-modal (aural, verbal, visual) analysis methods for improving situational awareness of threats in consumer IoT networks