Deriving hidden functionality from IoT Devices

Upakar Paudel

Colorado State University

Introduction

Subtitle goes here

Overview

- IoT device popularity is increasing and gaining traction
- IoT device have a large range of use case in various sectors
- IoT device contain transducers (sensors and actuators)
- IoT device are small in size and resource constrained

Problem

- Design specification and functionalities of IoT devices are abstracted away from end users
- No way to verify the device capability before purchase

Contribution

- Propose a new approach to detect context about presence of transducer from vendor materials
- Introduce the notion of composite capability
- Explore the challenge stemming from composite capability

Background

Subtitle goes here

Atomic and Composite Capability

- **Atomic capability:** Primary functionality of a transducer
- **Composite capability:** Induced functionality of a transducer by external software/dependencies, network connection and connection with other computing device

N-gram

- Sequence of N words
- Predict the occurrence of words based on previous words

BERT

- Transformer based language representation model
- Deeply bidirectional (learns information from both left and right side)
- Pretrained on large corpus of untrained text
- Two models:
  - BERT Base: 12 layers (transformer blocks), 12 attention heads,and 110 million parameters
  - BERT Large: 24 layers (trans-former blocks), 16 attention heads and, 340 million parame-ters

Affinity Propagation

- Graph based clustering algorithm
- No need to specify k initially
- Uses 4 set of matrices to cluster data points together:
  - Similarity Matrix
  - Responsibility Matrix
  - Availability Matrix
  - Criterion Matrix

Related Works

Subtitle goes here

Past works

- Many past work fingerprint IoT device by analyzing network traffic (uses machine learning)
- MUD based fingerprinting works
- Past work on context-based approach in an IoT environment for better service delivery to users

Past works contd.

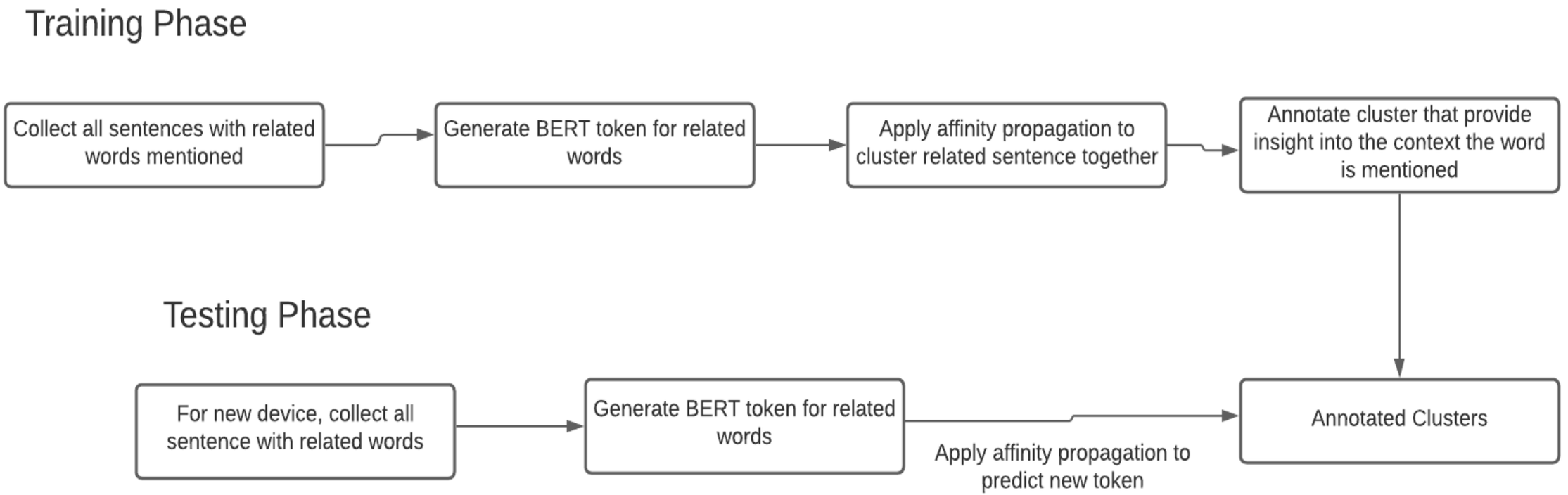
- **Key Term Set Matching:**
  - Profile device behavior proactively using vendor materials (overview page, technical specification page and manuals)
  - Suffers from high false positive rate
  - Due to ambiguity in the context in which various terms are presented in vendor materials.
    - For example, temperature is presented in multiple contexts in vendor materials (as a suitable operating temperature for a device to operate on **and** in the notion of sensing/adjusting temperature
  - Implemented context based approach to target this limitation

Methodology

Subtitle goes here

N-gram

- Extract a device specification from vendor materials (overview page, technical specification page, manuals and setup videos)
- Convert the corpus to trigram (set of 3 words)
- Analyze the trigram to justify transducer behavior
- For example, 'bright' adjacent to 'light' justify presence of led light



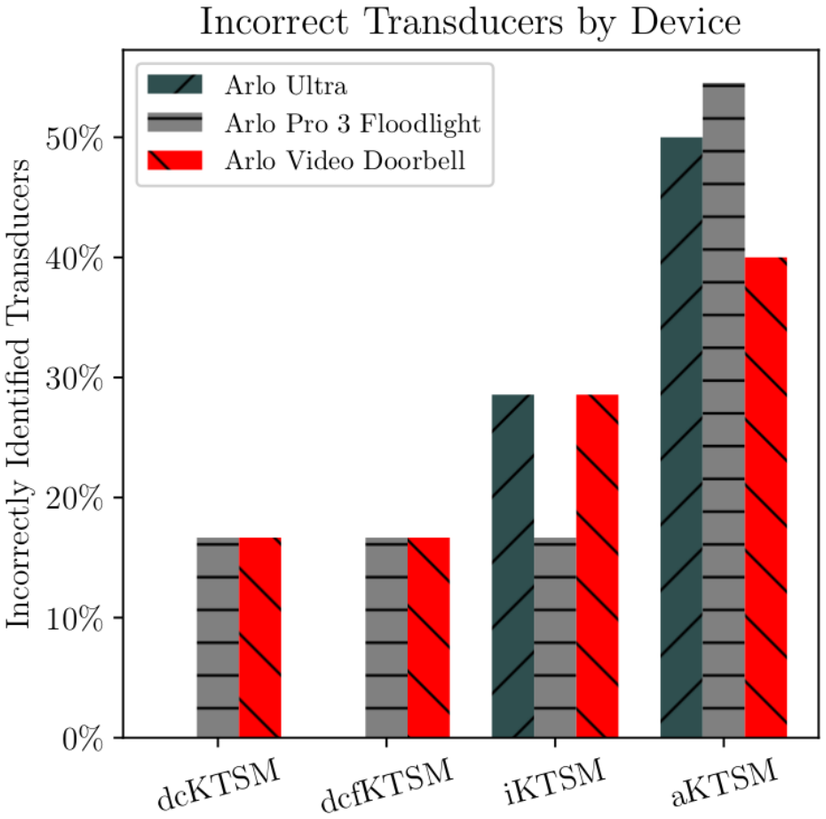
BERT and Affinity Propagation pipeline

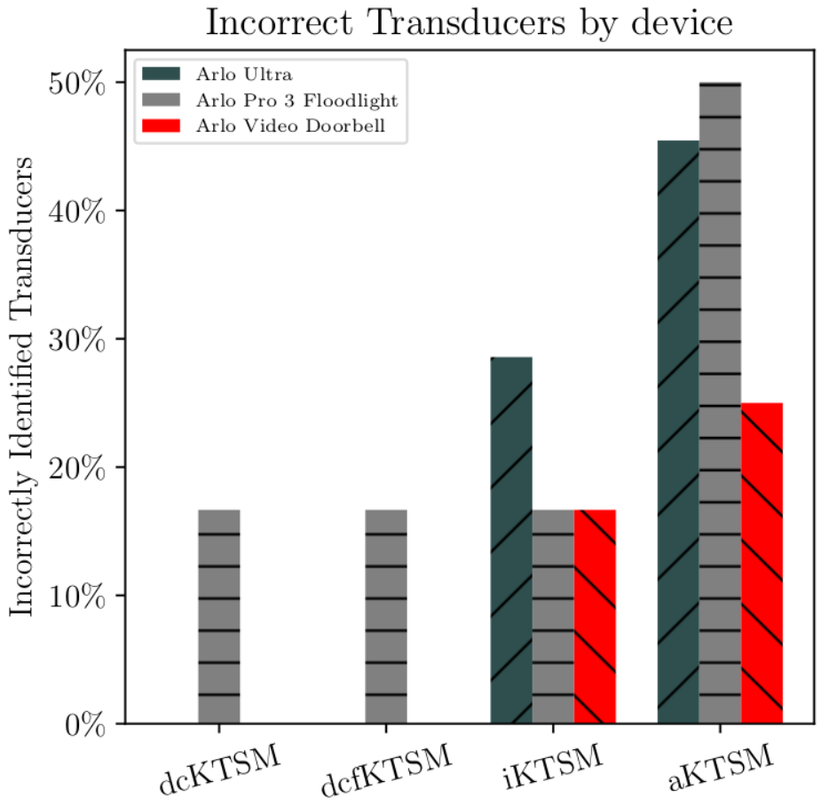
Implementation and Result

Subtitle goes here

N-gram

- LED light and temperature sensor being incorrectly identified with key term set matching approach
- Extracted trigram from corpus
- Analyzed adjacent words of key transducer we are considering
- For example, word 'bright' adjacent to 'light' justifies the presence of led light in an IoT device


Incorrect Transducers by Device
Arlo Ultra
Arlo Pro 3 Floodlight
Arlo Video Doorbell
Incorrectly Identified Transducers
0%
10%
20%
30%
40%
50%
dcKTSM
dcfKTSM
iKTSM
aKTSM


Incorrect Transducers by device
Arlo Ultra
Arlo Pro 3 Floodlight
Arlo Video Doorbell
Incorrectly Identified Transducers
0%
10%
20%
30%
40%
50%
dcKTSM
dcfKTSM
iKTSM
aKTSM

N-gram contd.

BERT and Affinity Propagation

- Applied BERT and Affinity Propagation pipeline to detect context in which word 'temperature' is being used in vendor materials
- Extracted corpus from vendor material for specific IoT device
- Filtered sentence with word 'temperature' on it
- Generated BERT token for the word 'temperature' based on context on entire sentence [Used hianxia's BERT as a service]
- Piped token to Affinity Propagation to cluster them together [used affinity propagation library from sklear.cluster]

Cluster 1
Now turn up the temperature and get comfortable
Put a Nest Temperature Sensor in any room, like the baby's room, and you can tell Nest to make that room a priority (sold separately)
The Nest Thermostat can use sensors and your phone's location to check if you've left, then sets itself to an Eco Temperature to save energy
You adjust the temperature from your phone so they'll be cozy
You may also be into Google Nest Mini From Google Nest Protect From Google Nest Temperature Sensor
The temperature they sense is warmer or cooler than home-owners feel
In a room that's used often, so Nest can read the right temperature and the homeowner can easily reach it
If your existing chime doesn't ring when someone presses your Video Doorbell, your Video Doorbell or Power Kit might not be wired correctly, or the temperature of your Arlo Video Doorbell might be too high
Check if the Arlo app is warning that your doorbell temperature is too high

Cluster 2
Operating Temperature 40°F (4 °C) to 100°F (38 °C)
Operating Temperature 32°–104°F (0°–40°C)
Operating Temperature –22° to 140°F (–30° to 60°C)
Battery temperature range: 14° to 131°F (–10° to 55°C)
Operating temperature 32° to 104°F (0° to 40°C)
Operating Temperature (F) 0°C +40°C (+32°F +104°F)
Operating Temperature -20 to 60 degree Celsius
Operating Temperature -20 to 45 degree Celsius
Operating Temperature -20 to 45 degree Celsius
The operating temperature or voltage is too low

Cluster Obtained

S.N	Training Device	Testing Device	
		Correctly Identified	**Incorrectly Identified**
1	Arlo Pro 3 Floodlight, Nest Thermostat, Nest Yale Lock, Arlo Video Doorbell, Nest Camera and Samsung Smart Cam	Nest Protect, Arlo Ultra Cam	
2	Samsung Smart Cam, Nest Thermostat, Arlo Video Doorbell, Arlo Pro 3 Floodlight, Arlo Ultra Cam, Nest Yale Lock	Nest Protect, Nest Camera	
3	Nest Yale Lock, Nest Thermostat, Samsung Smart Cam, Arlo Video Doorbell, Nest Protect, Nest Camera		Arlo Pro 3 Floodlight, Arlo Ultra Cam
4	Nest Yale Lock, Nest Thermostat, Samsung Smart Cam, Arlo Video Doorbell, Nest Protect, Arlo Ultra Cam	Arlo Pro 3 Floodlight, Nest Camera	
5	Nest Yale Lock, Nest Thermostat, Samsung Smart Cam, Nest Protect, Nest Camera, Arlo Pro 3 Floodlight	Arlo Ultra Cam	Arlo Video Doorbell
6	Nest Yale Lock, Nest Thermostat, Samsung Smart Cam, Nest Protect, Nest Camera, Arlo Ultra Cam	Arlo Pro 3 Floodlight	Arlo Video Doorbell
7	Nest Thermostat, Samsung Smart Cam, Arlo Video Doorbell, Arlo Pro 3 Floodlight, Nest Protect, Nest Camera	Nest Yale Lock, Arlo Ultra Cam	
8	Nest Thermostat, Samsung Smart Cam, Arlo Pro 3 Floodlight, Nest Protect, Nest Camera, Arlo Ultra Cam	Nest Protect	Arlo Video Doorbell

Testing Result

Composite Capability

Subtitle goes here

Communication with other computing device

- Point to point communication with other computing device
- Might give rise to additional combined capability
- Need to verify those capability in access control policy

Connectivity with communication Network

- IoT device connected to network facilitate data transfer capability
- Data transfer needs to be verified and monitored for better security

Conclusion

Subtitle goes here

Conclusion and Future work

- Context based approach to detect transducer and their capabilities
- Defined the notion of composite capability and challenge that stems from it
- **Future works:**
  - Increase the number of devices
  - Test our approach with NLP algorithm like Elmo and make comparison with current result
  - Formalize access control policy for composite capability

Thank you