Extrapolation and Prediction of User Behaviour from Wireless Home Automation Communication

Frederik Möllers
Sebastian Seitz
Andreas Hellmann
Christoph Sorge
(Wireless) Home Automation

- System performs everyday tasks
  - Locking doors, regulating heating and lighting, controlling blinds …

- Remote control and full automation

- Increasingly popular: Wireless systems
  - Benefits: Low installation effort/cost, no structural changes required
  - But: Wireless network – everyone can listen and send
  - Do state-of-the art systems use encryption / authentication?
  - Are there problems that persist?
Our Experiment

- 2 volunteers
- HomeMatic systems (default: no encryption, authentication only for door locks)
- Setup (placed inside the property)
  - Raspberry Pi
  - CC1101 USB Lite with culfw firmware
- 36 and 24 days of capturing data
- Analysis with custom software
  - 3 modules
    - Sniffer: Record data
    - Cleaner: Remove unnecessary data and organize the rest
    - Analyzer: Display data in human-readable form
Our Experiment – Methods

- **Identifying devices**
  - Apply regular expressions to messages
  - Plausibility checks (e.g. temperature values)

- **Recognizing patterns**
  - Visualization of data
    - Directed graph of connected devices
    - 2D-graphs of statuses / commands over time

- **Finding correlations**
  - Sliding window (occurrences of message pairs)

- **Identifying automation rules**
  - Commands sent at approx. the same time (almost) every day
Candidate 1

- Regular home installation
- 45,679 messages from 23 devices

- Some devices are remote controlled
  - Clear user interaction $\rightarrow$ presence / absence
Candidate 1 – Temperature / Humidity Sensors

- One temperature/humidity sensor outside the house
- Another one in living room
  - Heating controlled manually
  - Seldomly ventilated for more than 10 minutes
Candidate 1 – Tri-State Sensors

- Tri-state sensor on front door
  - Exact leaving / arrival times!

- Other results
  - Alarm function with bedroom lights
  - Automatic blind control with dawn / dusk times
Candidate 2

- 2 connected installations: Office and home
- 34,707 messages from 20 devices

- More devices are remote controlled or paired
  - Information about automation rules and user interaction
  - Remote control messages only in one location → presence / absence

- Automatic heating control
  - Heating turned off at night and on weekends, on in the morning
Encryption

• We've seen how much information is leaked
• What about encryption?
• Headers not encrypted: Nodes might still be identified
  • Communication partners
  • Frequency of communication
  • Tri-state sensors and locks only report state changes
    • State can be inferred
• Headers encrypted:
  • Amount of communication indicates presence
  • Additional power consumption
Summary / Outlook

- Current systems leak high amount of personal information
- Encryption is important
  - So is authentication – thermostats can be controlled
  - But there is more to do
- Need to hide communication
  - Create dummy traffic
  - How to determine when to send dummy messages?
  - Be energy efficient!
Questions?

frederik.moellers@uni-saarland.de
BidCos Protocol

- 868.3 MHz
- Layer 0&1: TI CC1100
- Layer 2: BidCos (no higher layers)

<table>
<thead>
<tr>
<th>Length</th>
<th>Counter</th>
<th>Control</th>
<th>Type</th>
<th>Sender</th>
<th>Receiver</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>$\text{Length}$</td>
</tr>
</tbody>
</table>
Corellation Analysis – Sliding Window

- Configurable parameters:
  - Minimum frequency
    - Number of times a single message (4) occurs
  - Minimum support
    - Number of times a message (4) is preceded by its counterpart (1/2/3)
  - Window size
Identifying Automation Rules

- **Steps**
  - Collect messages with same content
  - Discard dates (keep times)
  - Sort by time in ascending order
  - Find large number of events at roughly the same time

- **Configurable parameters**
  - Minimum frequency
    - Number of times the message occurs
  - Maximum deviation from rule
    - Time a message timestamp can deviate from the rule's time
  - Maximum time difference between messages from same rule
    - Should be $\leq \text{max. deviation} \times 2$
Candidate 2 – KeyMatic Door Locks

- KeyMatic door locks
  - Similar to a sensor on the door
  - Tells when door is locked/unlocked
  - Possible DoS targets
External Image Sources

- **Slide 2**

- **Slide 3**