IoT Unit Study Guide

**IoT** Outline

* 7.1 Network Connectivity for IoT
  + IoT network design space; GB per day, device-to-gateway range, battery life
    - What space we’ll occupy and why
* 7.2
  + How are we dealing with
    - Limited resources
    - Missing/noisy data
    - Outliers/anomalies
  + Continuous, real-time recognition
    - Detection-feature extraction-classification
  + Where is BLE low power stuff relevant
  + Strategy for collecting only data we need: Energy and Bandwidth restrict the amount of data we can collect
    - Battery life vs. quantity of data; latency vs. resolution
    - On-device processing, tradeoffs
    - WiFi/Bluetooth use MUCH less energy than 3G/LTE
    - How can data be combined as it is transmitted through the network?
    - Intermittent/low-rate radios limit what can be collected
      * Considerations
        + Continuous Monitoring vs. Alerting
        + Buffering to mitigate rate variations & dysconnectivity
    - End-to-end principle: Data should only be removed from a device when it has been delivered to permanent storage
  + IoT data is approximate: It arrives at discrete times, Data is of limited precision, Data can even be wrong
    - Interpolation; extrapolation (regression);
    - Smoothing; curve-fitting, low-pass filters
    - Alignment; auto-correlation, dynamic time warping
    - Technology usage: FunctionDB to allow users to fit continuous functions to raw data
  + Anomaly Detection: Automatically flag outliers. Identify common properties of outliers. Rank & triage outlier classes.
    - Detect anomalies via an Anomaly Detector (e.g. human labeling, robust statistics)
    - Detector feeds both an Inlier Classifier (frequent itemsets) and an Outlier classifier (frequent itemsets out of all the infrequent itemsets)
      * Feed a ranking algorithm (Mahalanobis distance)
  + Anomaly Explanation: We’ll need to tell people what the fuck is up
    - Use classifiers (e.g. decision trees, SVMs, frequent itemsets)
    - Frequent itemsets works for categorical or binned continuous variable data
    - “Given an outlier group, find a predicate over the inputs that makes the output NO LONGER an outlier”
* 7.3 Localization – nothing relevant from this module
* 8.1 Protection from Ransomware – prevent attack on the service & infrastructure
  + Security is a negative goal
  + Prevent:
    - Changes to permissions to access files fraudulently
    - Direct disk block access
    - Internet access
    - Reuse of memory
    - Accessing backups
    - Interception of network packets
    - Trojan text editors
    - Stealing disk of server(s)
    - Spoofing sysadmin – personal knowledge
  + Threat model:
    - Adversary controls some computers/networks but not all
    - Adversary controls some software on computers he controls, but not full control
    - Adversary knows some info, such as passwords or keys, but not all
    - Adversary knows about bugs in software
    - Physical attacks?
  + How we’re assuring unrealistic threat models
  + How disabling a holistic view of the system might work
    - Prevention: increase difficulty of attacks
    - Resilience: allow system to remain functional
    - Detection/Recovery: allow systems to quickly detect & recover from attacks to a fully functional state
  + Physical Unclonable Functions PUFs for enabling secure, low-cost authentication
  + Shrinking the Trusted Computing Base
  + Tamper-resistant hardware to prevent private information stored on the hardware
    - Ascend Processor for IoT to eliminate leakage over chip pins
* 8.2 Human Computer Interaction
  + Designing an interface that is User-centric
    - Use speech-to-text to avoid typing
    - Speech – understanding – dialogue – generation
  + Reasons for speech applied to information access & management:
    - Information space is broad and complex
    - Users are technically naïve
    - Device is small
* 8.3 Robotics & Autonomous Vehicles
  + Concept of “Google for the physical world”
    - Extend analogy to building a topology of professional relationships
    - What and who are where? Maintain as a function of time
  + SLAM – Simultaneous Localization And Mapping – is likely not useful here
* 9.1 Smart Buildings
  + Sensate materials: measure temperature & decibel level of meetings and work environments – automatically adjust environment to preferences
* 9.2 Smart Homes
  + Use indoor localization to identify when people who rarely come in are present; prompt fruitful connections between others
    - Multipath effect can be exploited to increase accuracy: spatial angle profile of the power reflected
    - Combine this with badges used to scan into different parts of buildings
    - Breathing & heartrate could be measured during meetings in conference rooms to determine “stress levels” when person A interacts with person B
* 9.3 Smart Cities
  + Number of connections between colleagues might help inform seating/office charts similar to mapping telephone calls
  + Local Warming type technology could be deployed to ensure users are operating at their “optimal temperature” during meetings
* 9.4 Conclusions
  + How will component technologies (location, sensors, interfaces, setup, robotics, vehicles) interact with the system design (networking, architecture, data management, security)
  + Device-Cloud-Device will be the architecture of choice since real-time data is not always required
    - This won’t be true for real-time tracking of people in the office
    - Things will be lifted to the cloud and use internet protocol & web standards; use local controller/gateway for performance