COMP3516: Data Analytics for IoT

Lecture 6: Sensing System Design

Chenshu Wu

Department of Computer Science

2025 Spring

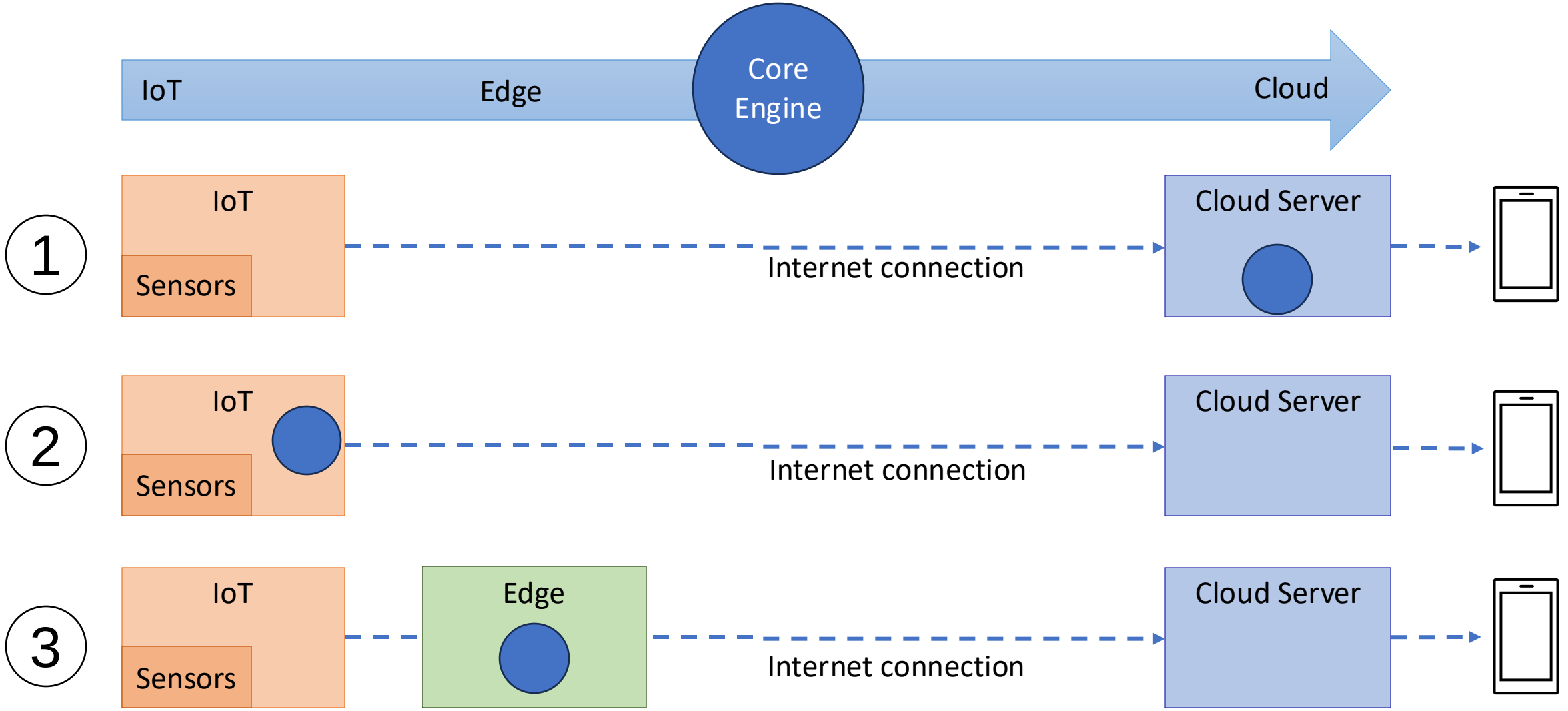
香 港 大 學
THE UNIVERSITY OF HONG KONG

HKU AIoT LAB
香港大學人工智能物联网實驗室

Contents

- A systems perspective on sensing
  - How do we put it all together and build an end-to-end sensing system?
  - What are the missing blocks for building a real system based on what we have learned thus far?
- General pipeline
- Useful techniques
  - Windowing
  - Filtering
  - Target Detection
  - Peak Finding

System Architecture

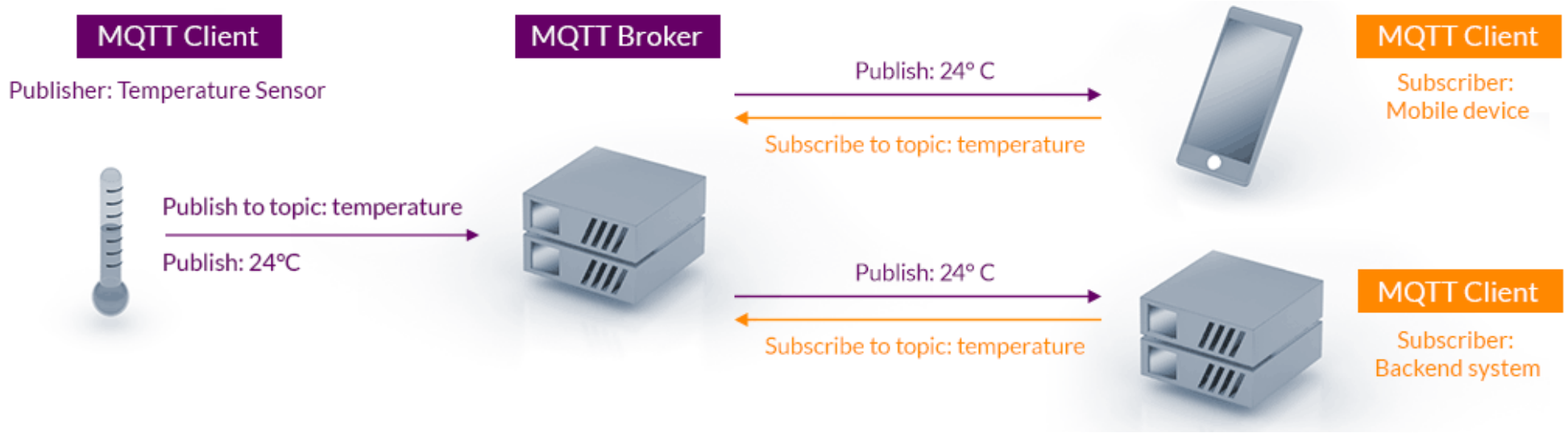


General Pipeline

- Hardware/sensor (cameras/WiFi/IMU/other sensors)
  - Standalone components? Plugged into computers?
  - Wired or wireless?
- Data Collection/Acquisition
  - How to fetch data from certain sensors from the host machine?
  - How often/how much to fetch?
  - What communication technology & network protocol?
- Preprocessing
- Core Algorithms
- User-End Interface

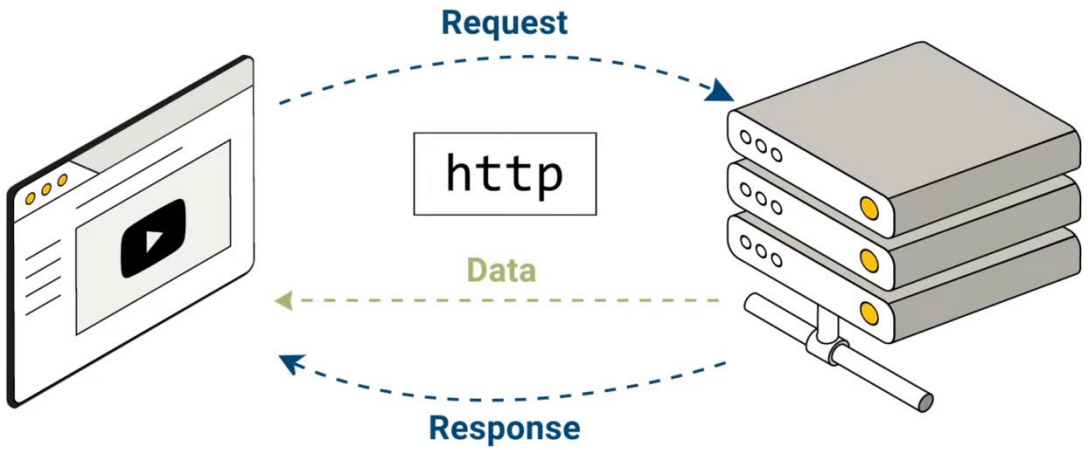
Data Transportation: MQTT

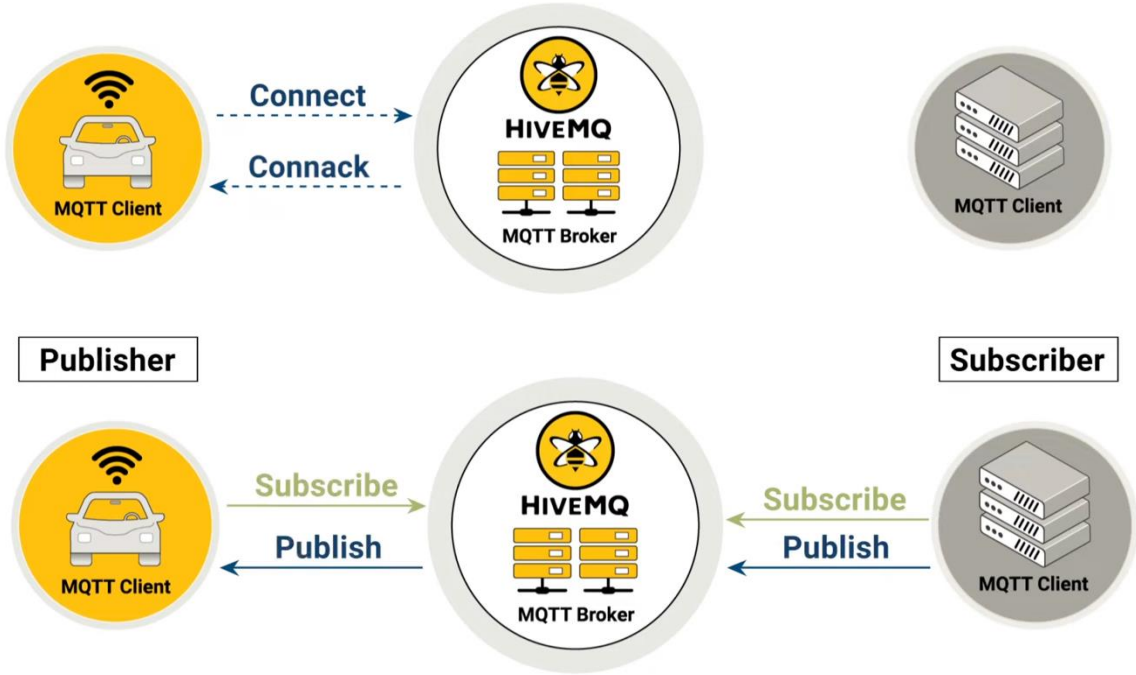
- MQTT: The Standard for IoT Messaging (source: AWS)
  - Simple implementation
  - Lightweight and efficient
  - Scalable
- MQTT (Message Queuing Telemetry Transport)



MQTT Architecture

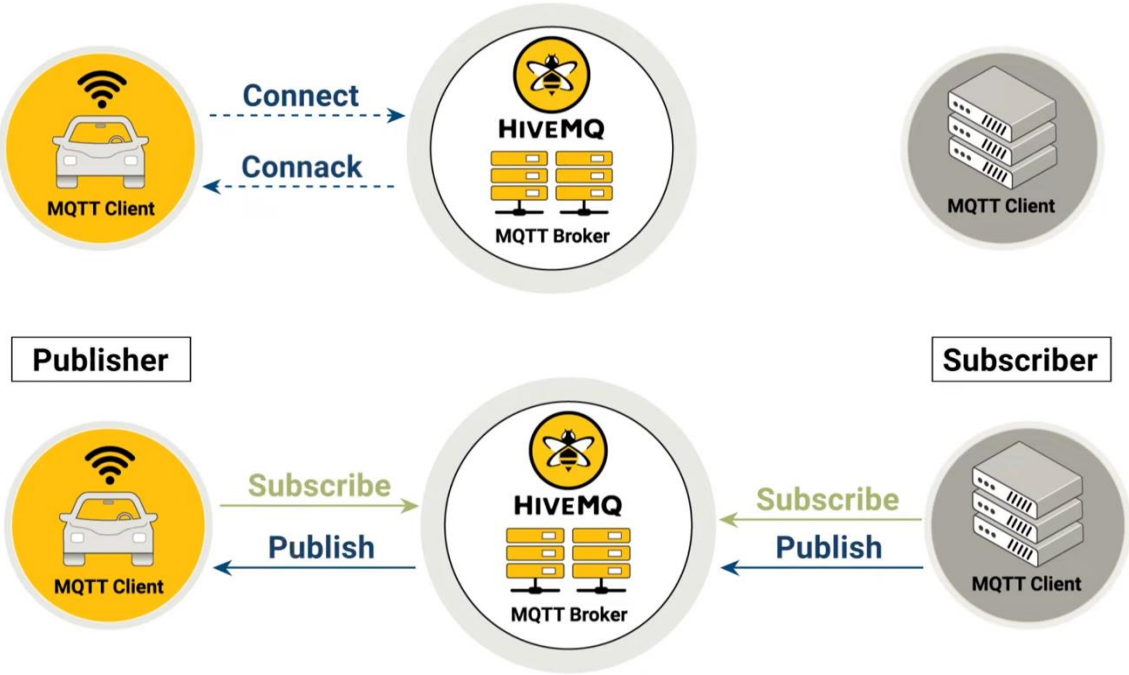
- Publish/Subscribe model (Producer/Consumer)
- Components
  - Client, Broker, Connection





MQTT Architecture

- Publish/Subscribe model (Producer/Consumer)
- Components
  - Client, Broker, Connection
- Scalability
  - Space decoupling
  - Time decoupling
  - Synchronization decoupling



MQTT Topics

Topics are structured in a hierarchy like folders and files in a file system
Allow to create a user friendly and self descriptive naming structures

myhome / groundfloor / livingroom / temperature

topic level topic level

'+' matches anything at a given tree level

myhome / groundfloor / + / temperature

- ✓ myhome / groundfloor / livingroom / temperature
- ✓ myhome / groundfloor / kitchen / temperature
- ✗ myhome / groundfloor / kitchen / brightness
- ✗ myhome / firstfloor / kitchen / temperature
- ✗ myhome / groundfloor / kitchen / fridge / temperature

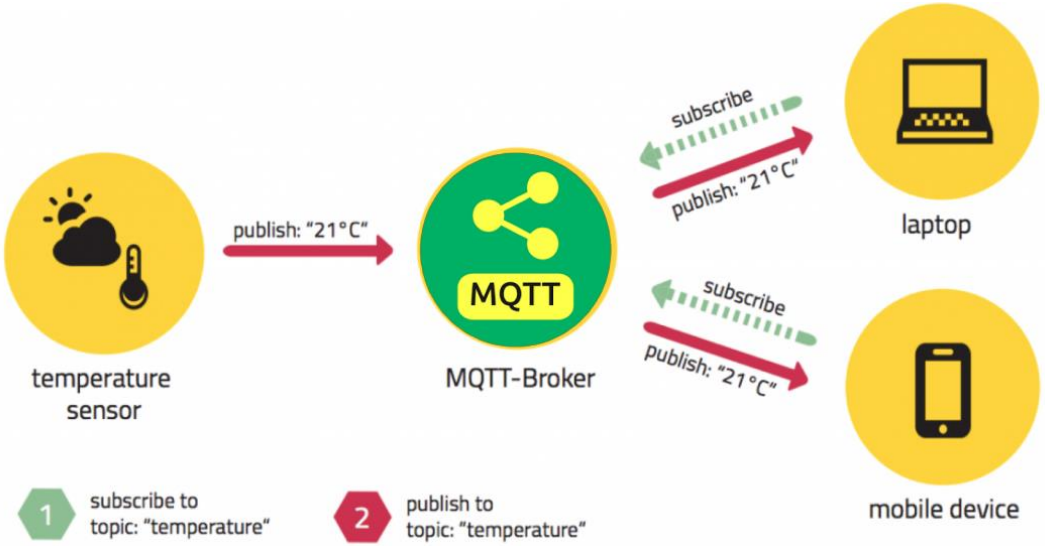
'#' matches a whole sub-tree

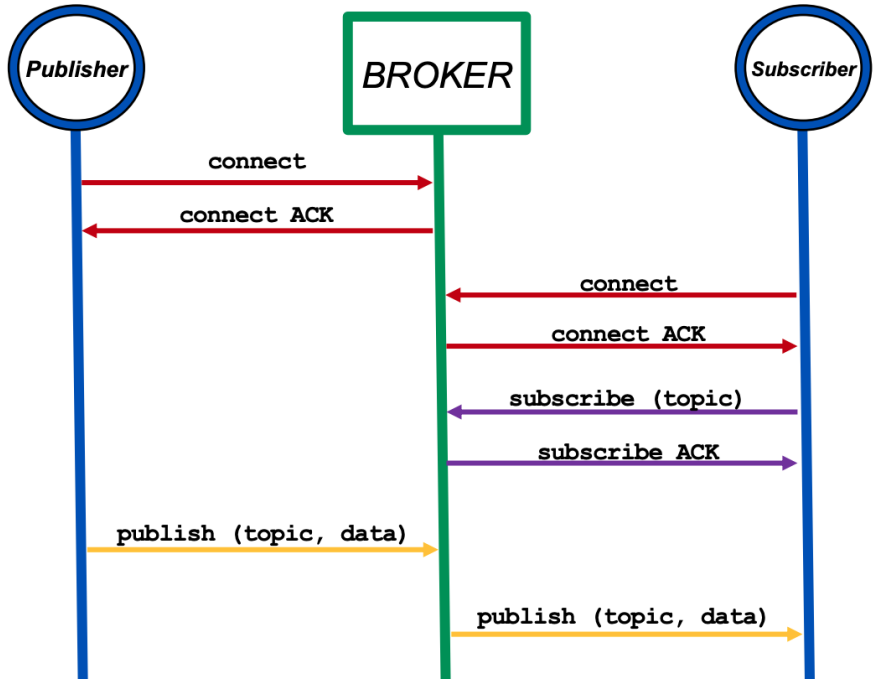
myhome / groundfloor / #

only at the end

multiple topic levels

- ✓ myhome / groundfloor / livingroom / temperature
- ✓ myhome / groundfloor / kitchen / temperature
- ✓ myhome / groundfloor / kitchen / brightness
- ✗ myhome / firstfloor / kitchen / temperature





MQTT Resources

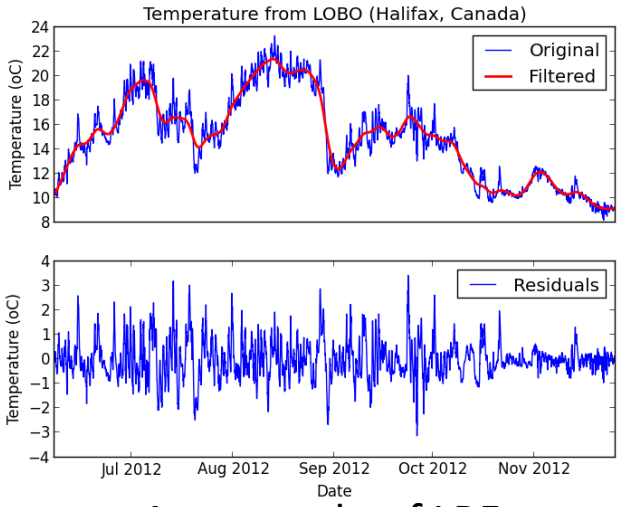
- http://mqtt.org/
- The most widely used implementations
  - http://mosquitto.org/
  - https://www.hivemq.com/
- The MQTT community wiki: https://github.com/mqtt/mqtt.github.io/wiki
- A good tutorial: http://www.hivemq.com/mqtt-essentials/

Windowing

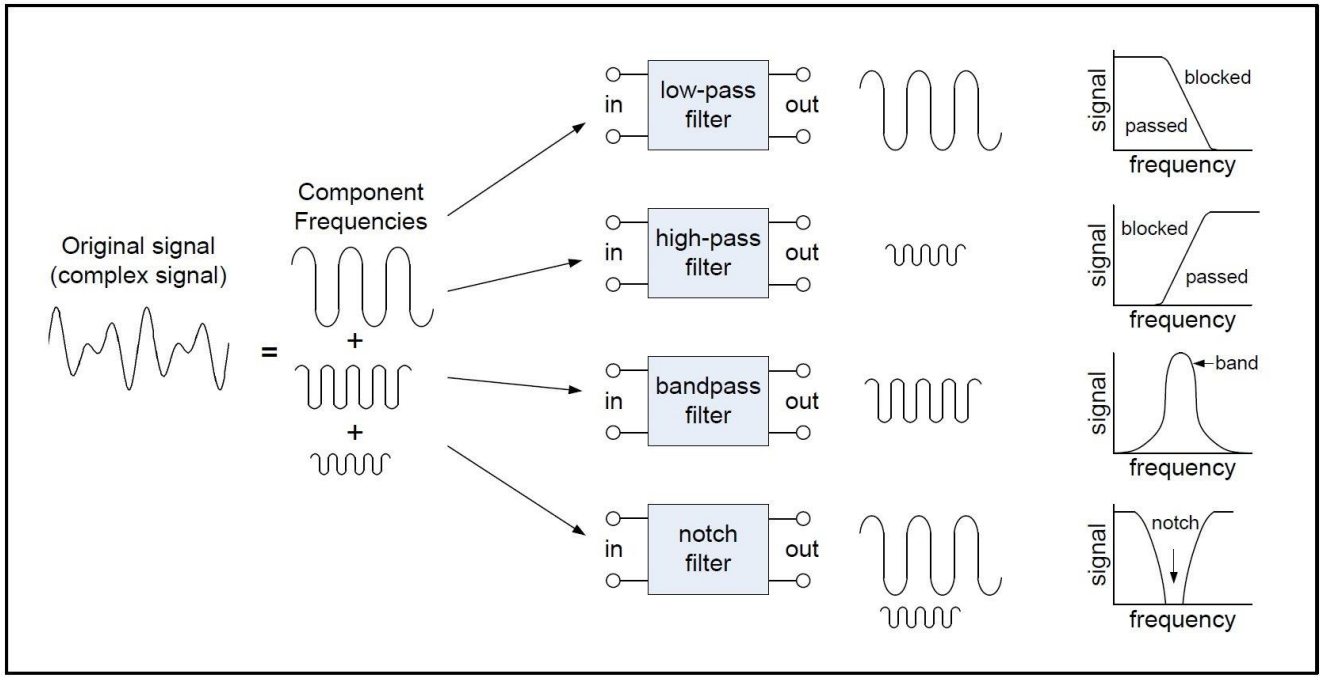
- Very often, we are processing time-series data.
- Window length
  - Any fundamental limits? E.g., FFT window requirement
  - Trade-off of accuracy vs. latency
  - Other constrains: Sampling rate, size, communication capacity, etc.
- Sliding step
  - Computation vs. responsiveness

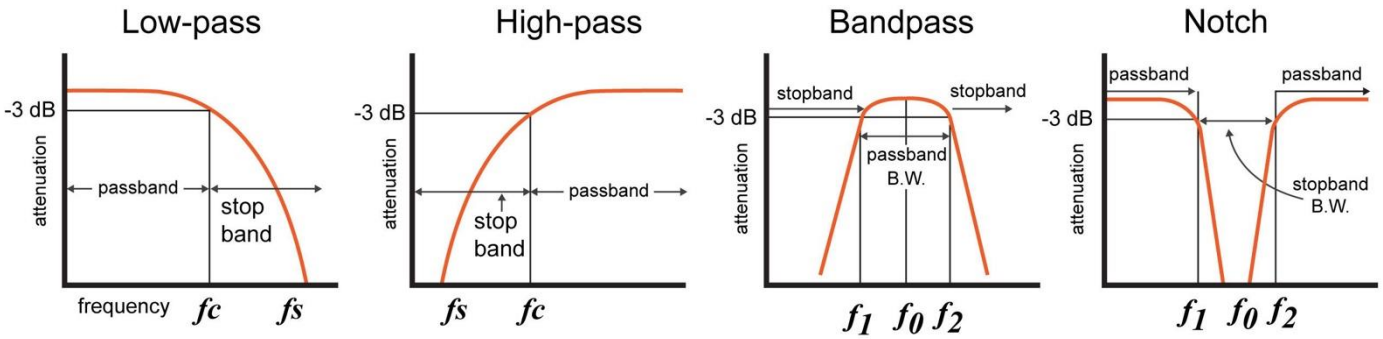
Filtering

- Low-pass filter
- High-pass filter
- Band-pass filter
- Band-stop (Notch) filter



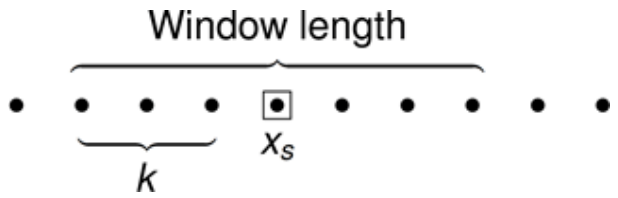
An example of LPF



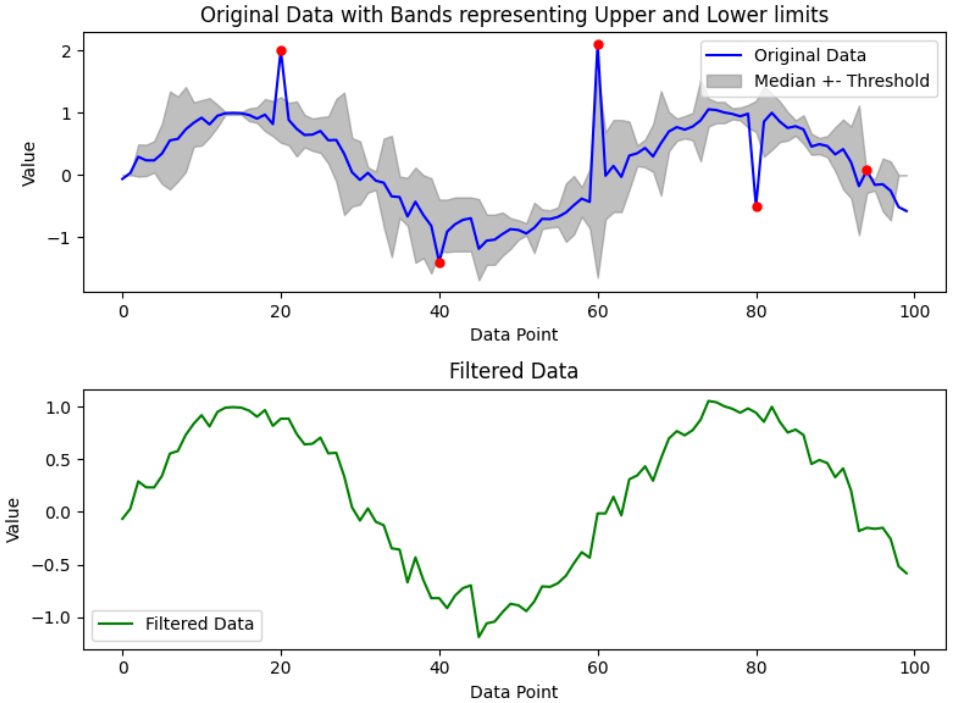


Filtering

- Hampel filter
  - A robust method for outlier detection
  - Median Absolute Deviation (MAD)



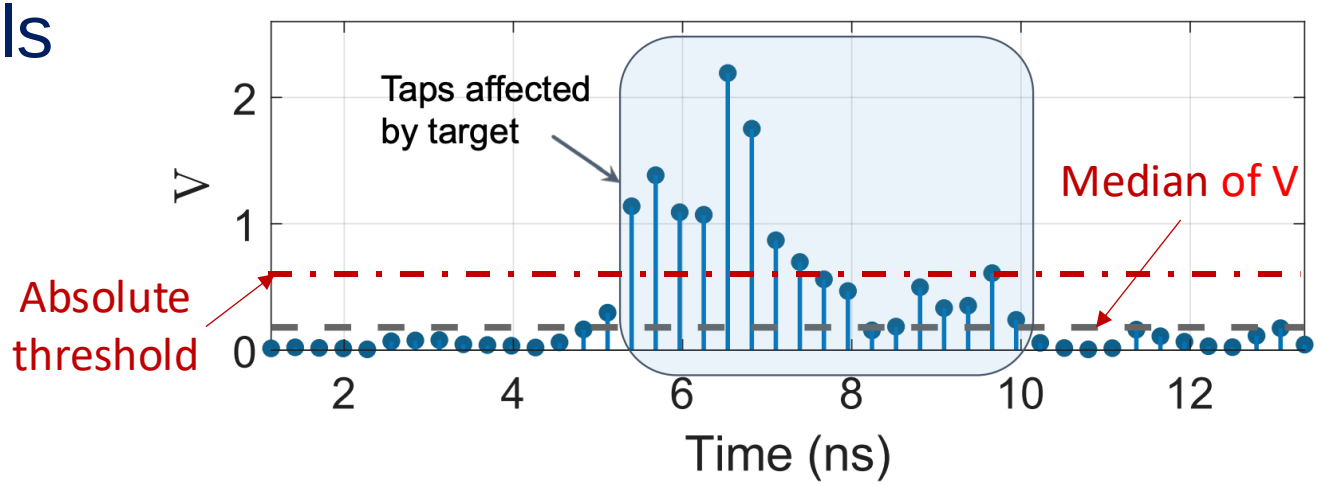
  - Set the current window (of odd length)
  - Get the local median m and SD σ
  - Outlier if $|x_s - m_i| > n\sigma$, n := threshold



https://github.com/MichaelisTrofficus/hampel_filter

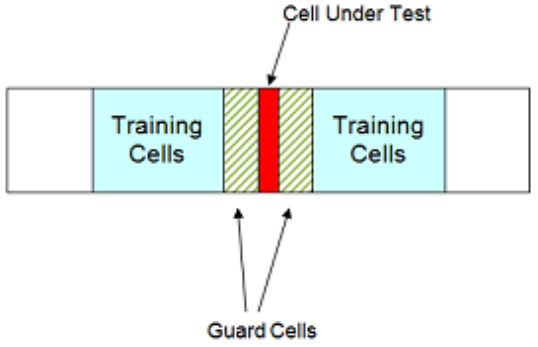
Detection: Thresholding

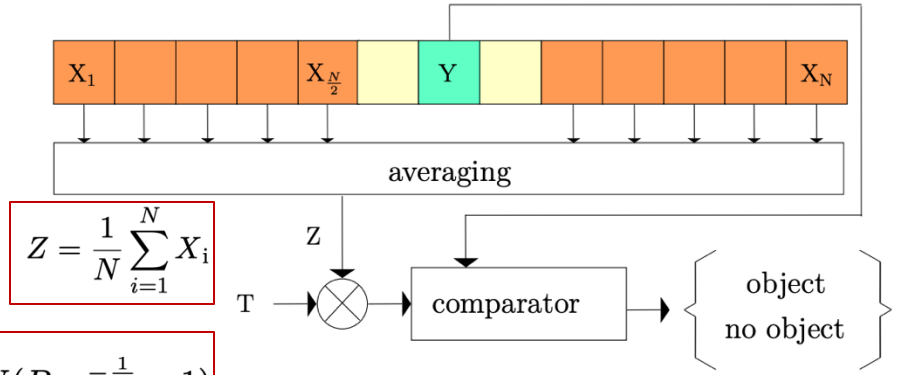
- Problem: How to detect a target/signal/event of interest?
- Hard thresholding
  - Empirical values
- Median filter for sparse signals



Detection: CFAR

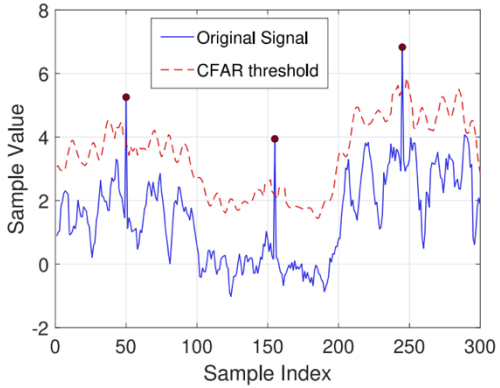
- Constant False Alarm Rate
  - Cell-Averaging CFAR (CA-CFAR)

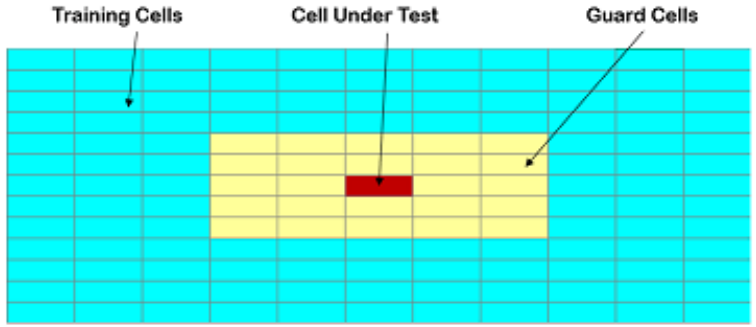


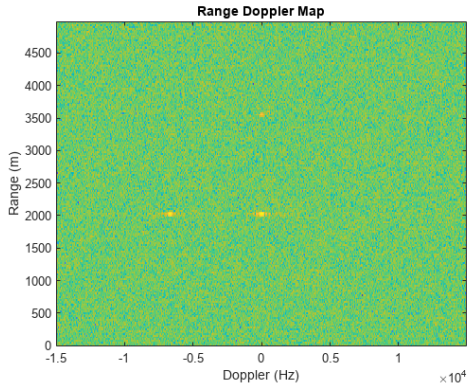


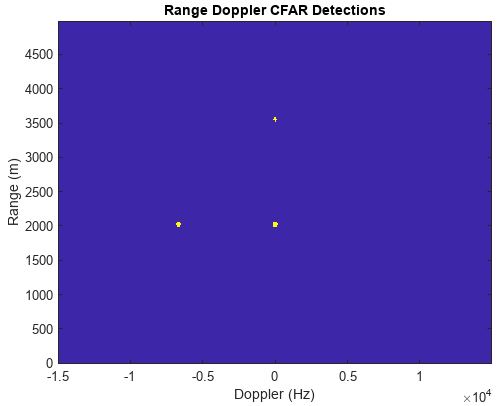
$$Z = \frac{1}{N}\sum_{i=1}^{N} X_i$$

$$T = N(P_{FA}{}^{-\frac{1}{N}} - 1)$$



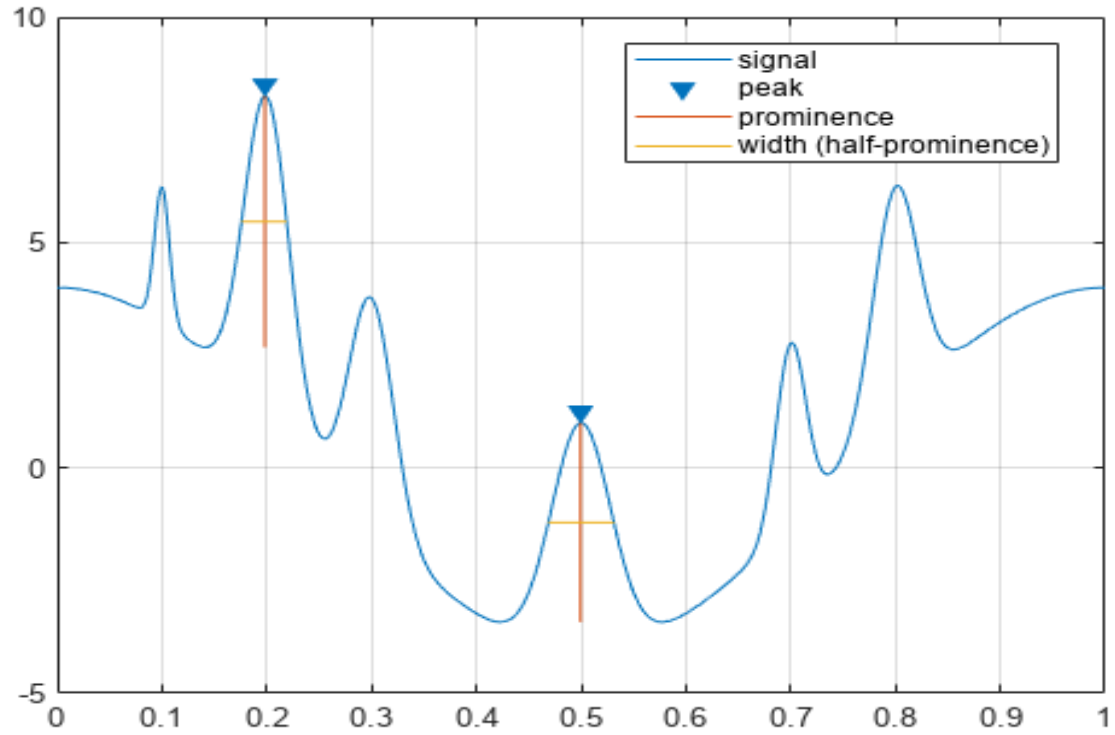




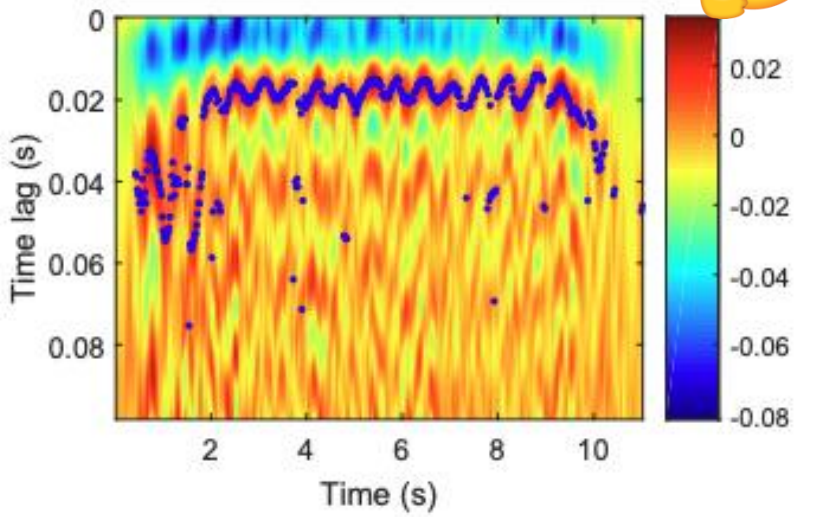


Detection: Peak

- `findpeaks(S,'MinPeakProminence',4,'Annotate','extent')`



How can you make peak finding more accurate and robust for a time series?



A Working Example

- We learn how motion detection works. How to build a system?
  - Suppose we plan to build a system using ESP32 WiFi IoT.

Questions?

- Thank you!