Toward an Easy Deployable Outdoor Parking System

Lessons from Long-term Deployment

Yu Huang, Yi-Chao Chen, **Chuang-Wen You**, Dian-Xuan Wu, Yi-Ling Chen, Kai-Lung Hua, Jane Yung-Jen Hsu

NTU IoX Center

Outline

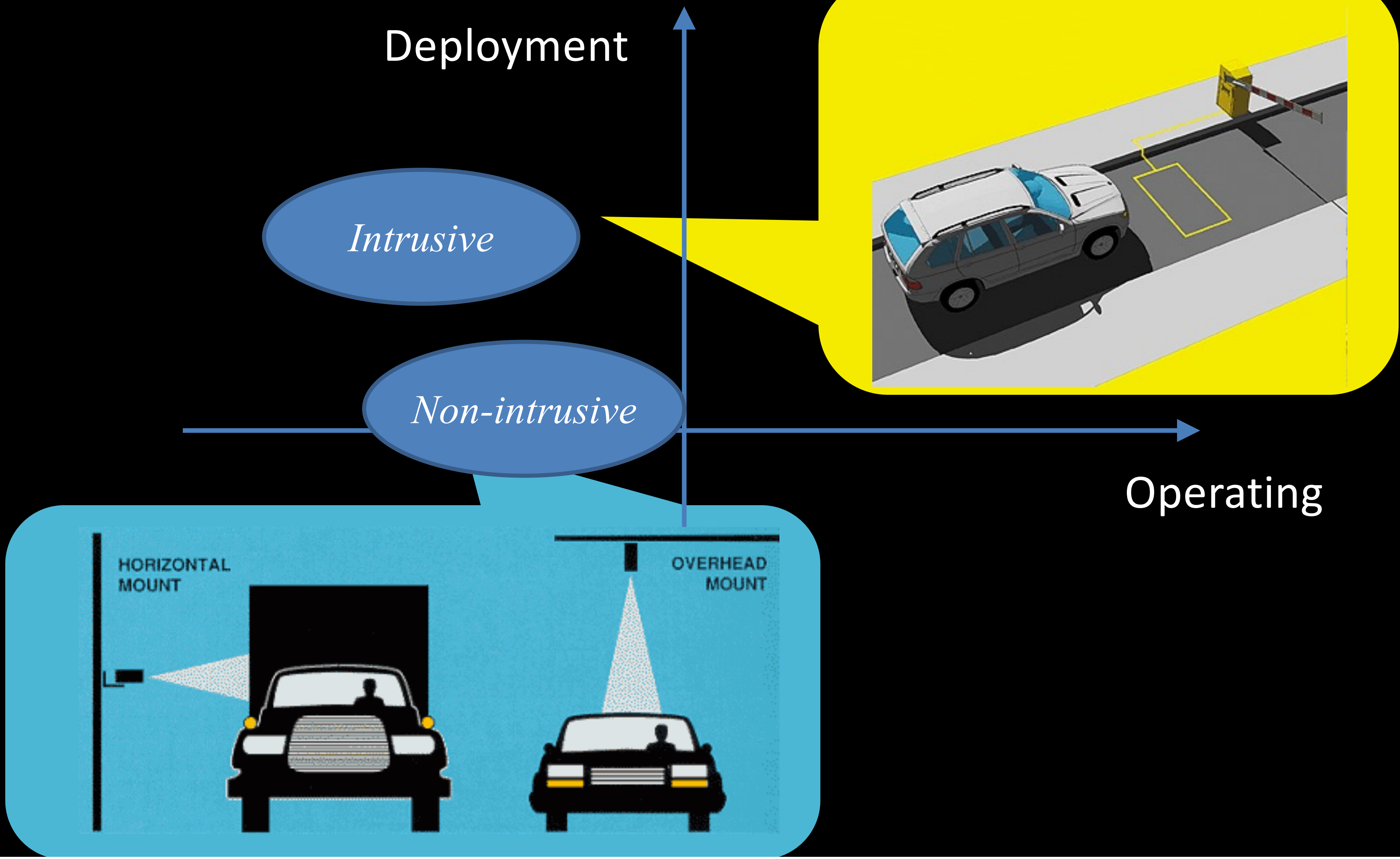
- Motivation
- Experiment
- Initial Observations
- Method
- Evaluation
- Related work
- Conclusion

Outline

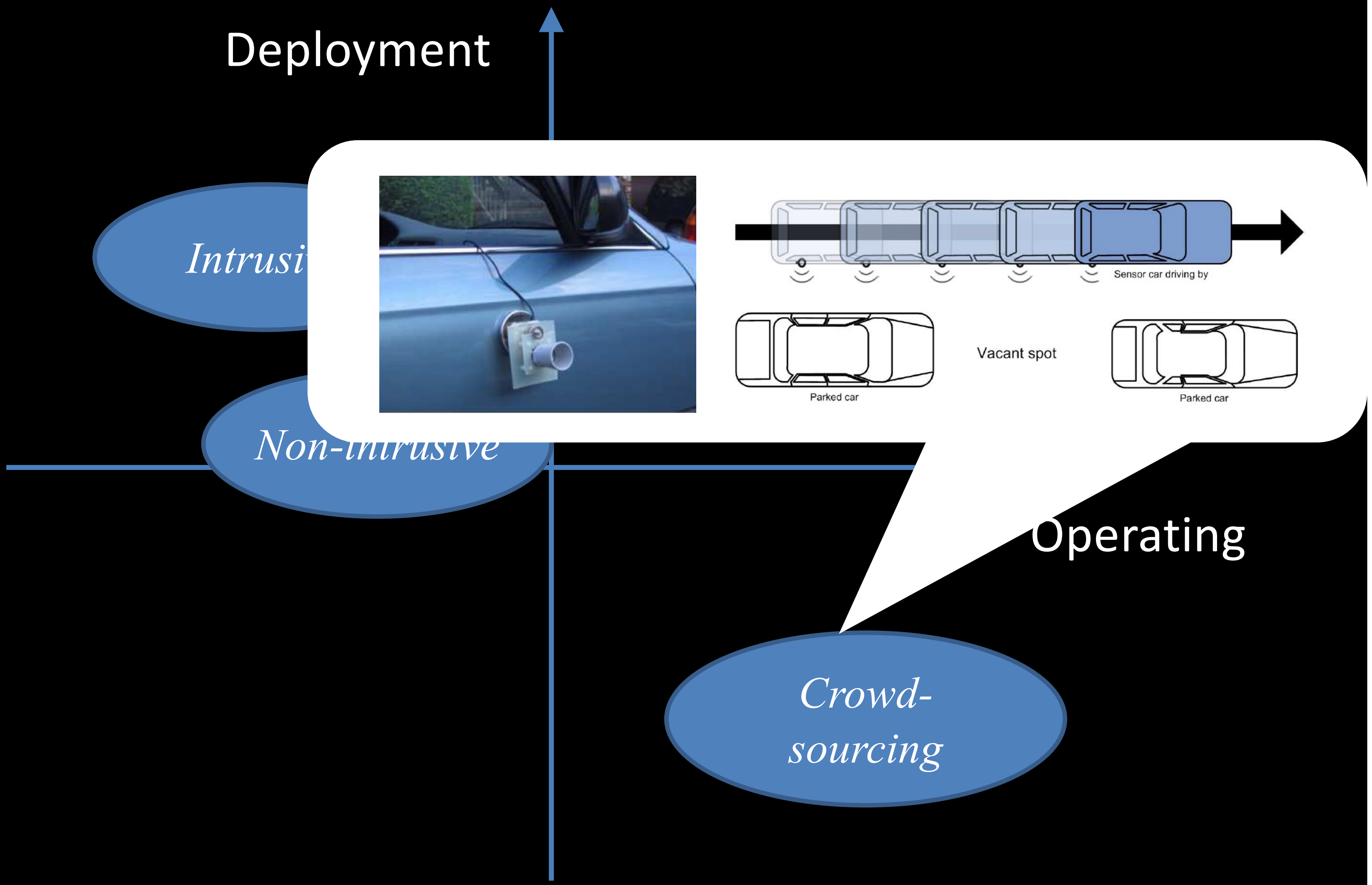
- Motivation
- Experiment
- Initial Observations
- Method
- Evaluation
- Related work
- Conclusion

The information of parking slot availability is critical to efficiently locate empty parking slots

Design considerations (1)

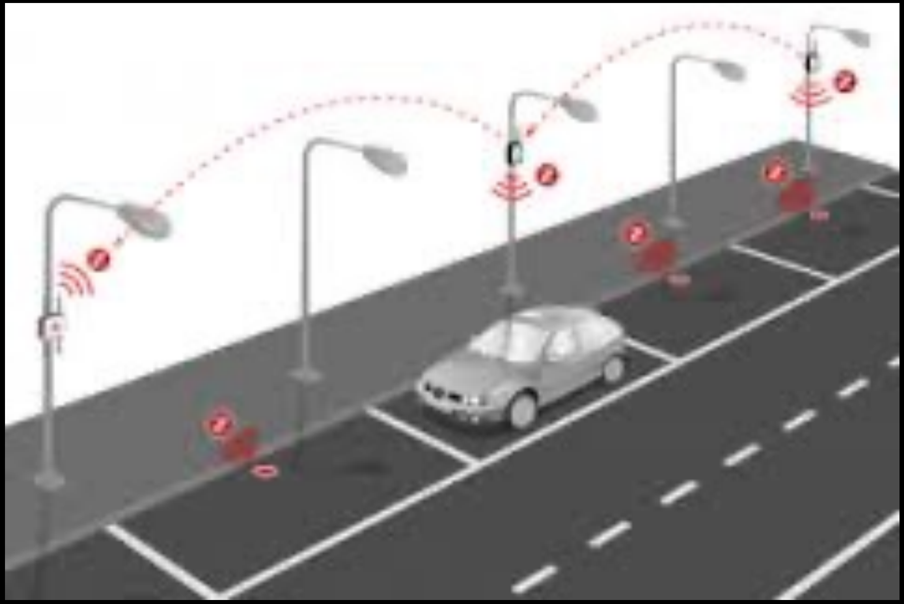


Design considerations (2)

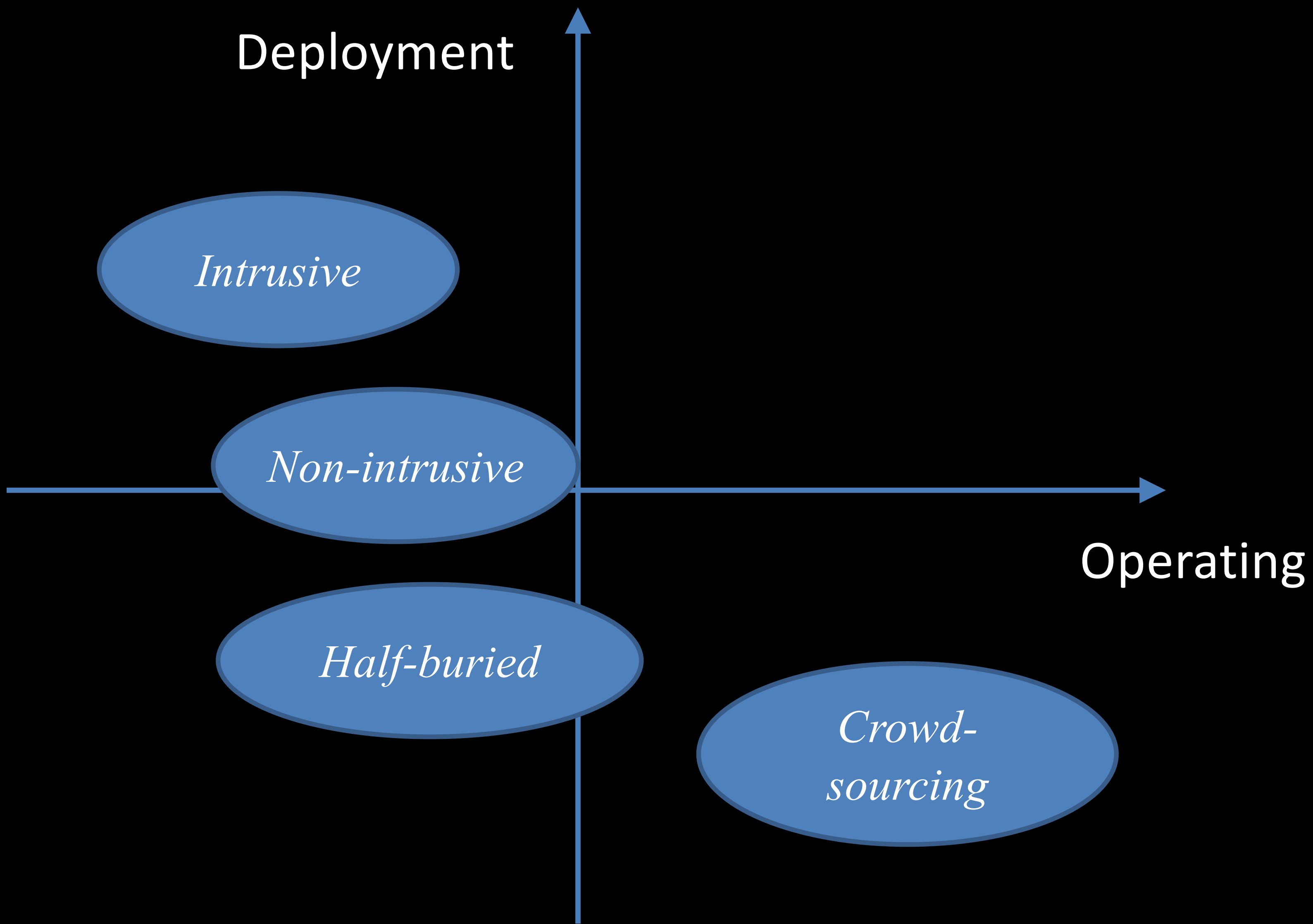


Magnetic-based sensing solutions

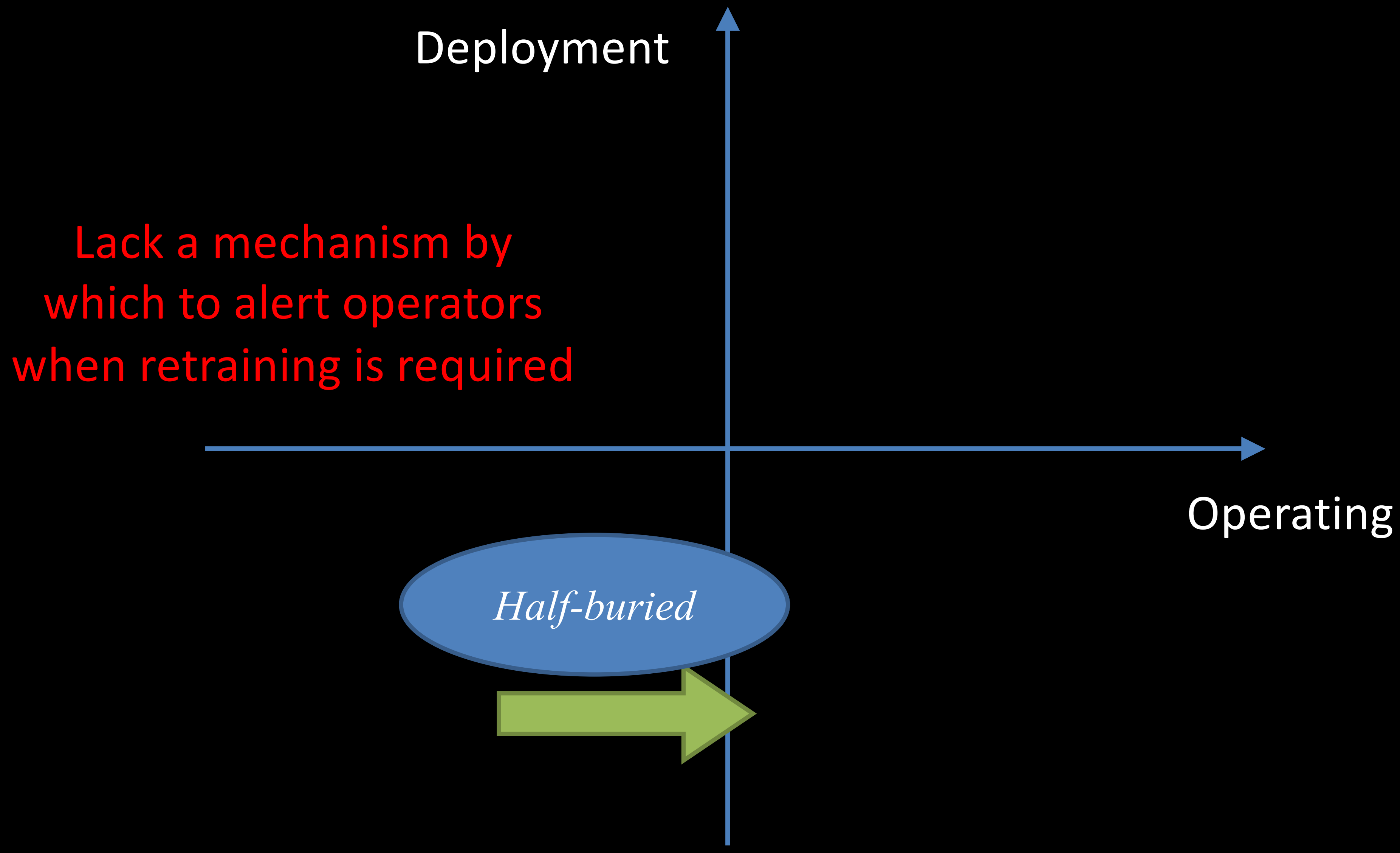
- Deploying a parking node equipped with a magnetic sensor to detect changes of magnetic fields
  - Streetline Inc., FastPrk system, etc.



Design considerations (3)



Design considerations (4)



Research problem

- Understanding time (or spatial-) dynamic issues of occupancy detection solutions
- Explore how to enable easy-deployable parking solutions

Challenges

- Easy deployable
- Easy maintainable
- Able to detect accurately

Contributions

- Easy deployable
  - Determine if the data distribution is changed quickly
- Easy maintainable
  - Designing an adaptive learning technique
- Able to detect accurately
  - Conducted a series of evaluations

Outline

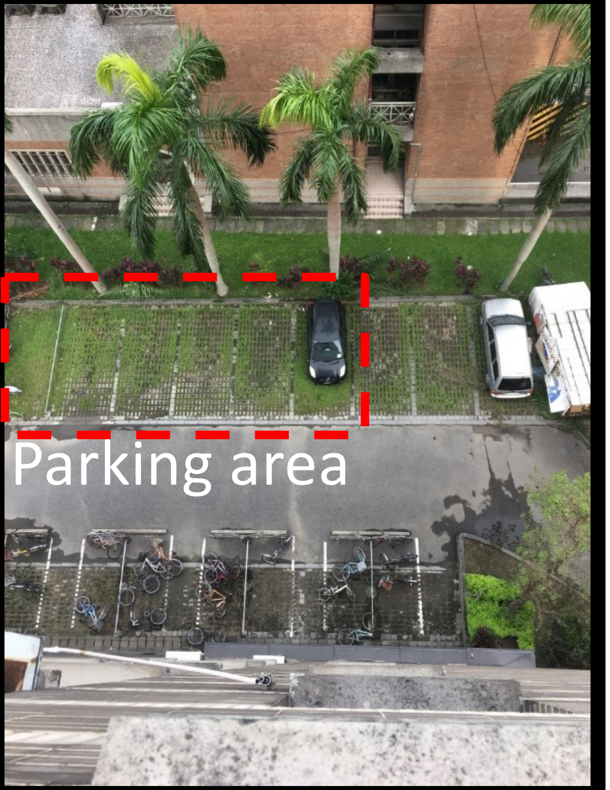
- Motivation
- Experiment
- Initial Observations
- Method
- Evaluation
- Related work
- Conclusion

Experiment: long-term data collection

- Conduct a 13 month experiment to collect real sensing data
  - Identifying the factors that affect accuracy in determining the occupancy of parking spaces

Parking area

- Four adjacent parking slots in National Taiwan University
- Easy deployment and maintenance
  - Single-point deployment of low-cost, low-power sensors
  - Placed in the center of the parking slot



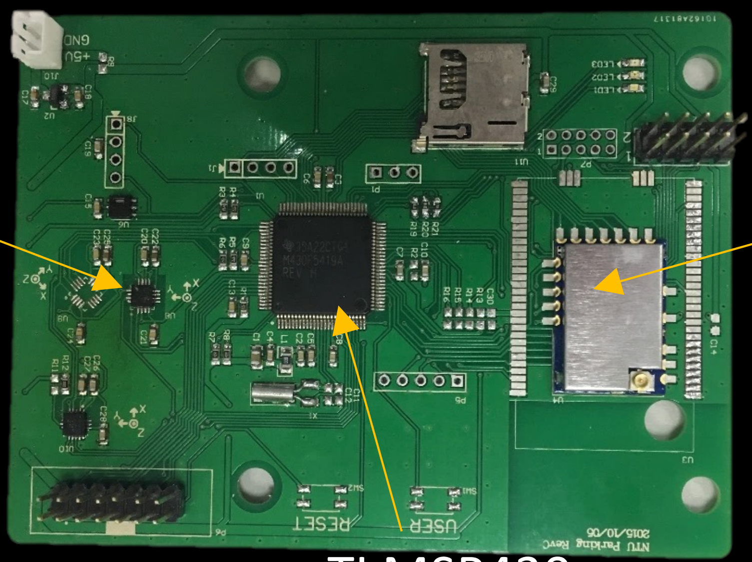
Customized parking node

- Housed a circuit board in a water-proof case
- Powered by Lithium-ion batteries
- Sensor readings
  - Recorded on an on-board micro SD card
  - Manually retrieved via a serial interface every two days

Customized circuit board (1)

- Equipped with low-cost & low-power sensors
  - TAOS TSL2561 is attached on the case

Magnetic sensor (HMC5883)



TM1276 LoRa module

TI MSP430

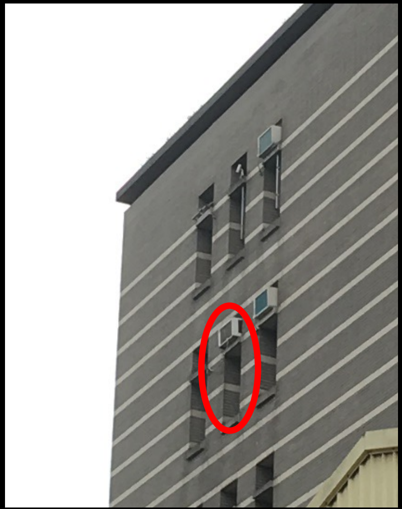
Customized circuit board (2)

- Selected the three sensing modules to demonstrate the application of a sensor selection algorithm
  - Should be applicable to switch between any combination of sensors

Module		***Average current***	***Sampling energy (time)***
System		Low-power: 22.06uA Active: 5.54mA	-
Sensing	Magnetic	2.40mA	12.68mJ (1.6ms)
	Light	3.75mA	745.63mJ (250ms)
	LoRa	Receiving: 35.88mA Sending: 76.00mA	Receiving: 1.78J (15ms) Sending: 3.16J (12.6ms)

Internet-connected gateways – LoRa senders

- Envision that a few IoT gateways in the future
  - Scattered throughout the city to enabling the exchange of messages between nodes and the backend
- Deploying three LoRa senders
  - Deployed In two buildings surrounding the area
  - Broadcasting messages every 500ms from senders to emulate message exchanges



Ground-truth for parking events

- A surveillance camera
  - Mounted on an outside wall on the 7th floor
  - Recorded footage was streamed back to a server via a wired link and stored

How to collect data

- The MSP430 microcontroller samples in a rate of 250 ms
  - Readings from magnetic and light sensors
  - RSSI values in the header of LoRa packets

The collected data

- 13 month data between 2015.4 ~ 2016.4
  - 666 arrivals and 610 departures
- A Transient event
  - Representing either a car arrival or departure
- Manually labeled each transient event

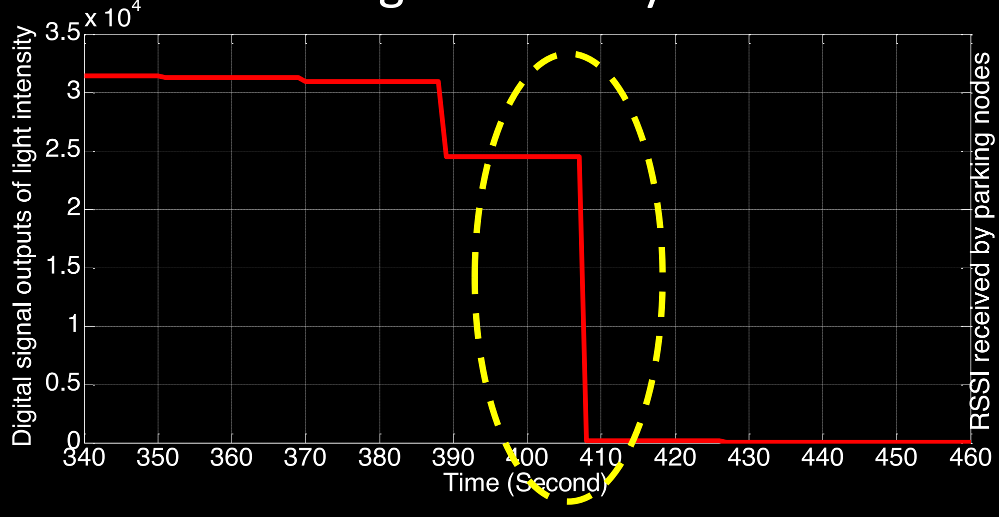
Arrival

Departure

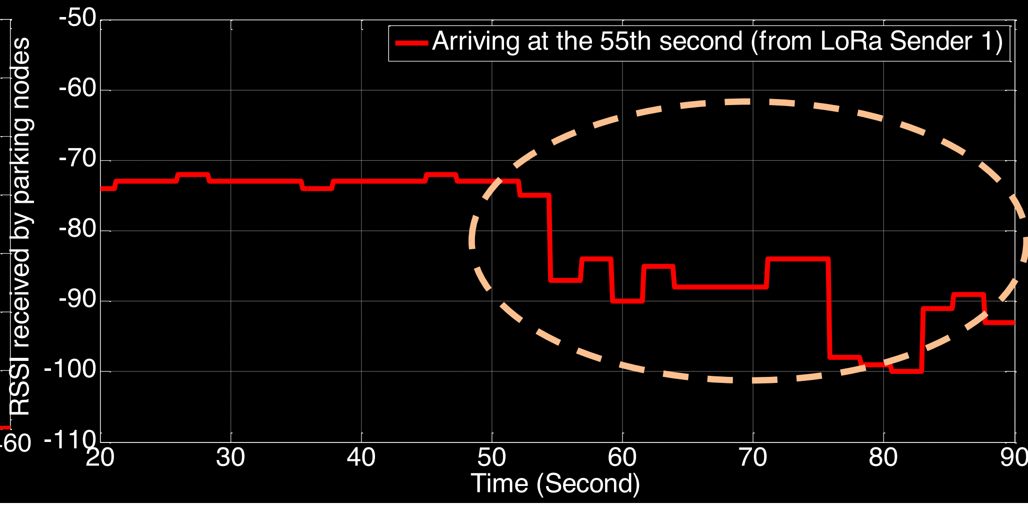
Changes in the patterns of signals

- Magnetic sensor
  - Magnetic field are easily affected by vehicles
- Light sensor
  - Light is blocked by cars entering or leaving a slot
- LoRa module
  - Car arrivals (departures) cause the attenuation (increase) of the RSSI

Light Intensity



RSSI values



Characterizing transient events

- Characterizing changes in the patterns of sensor readings
  - Preprocessing: moving average
  - Feature extraction: 15-second duration with 50% overlap
- Car presence is determined through the extraction of statistical features
  - Mean, median, variance, mode, range, etc.
  - 108 magnetic features, 30 light features, and 90 LoRa features

Outline

- Motivation
- Experiment
- Initial Observations
- Method
- Evaluation
- Related work
- Conclusion

Identifying potential inferences

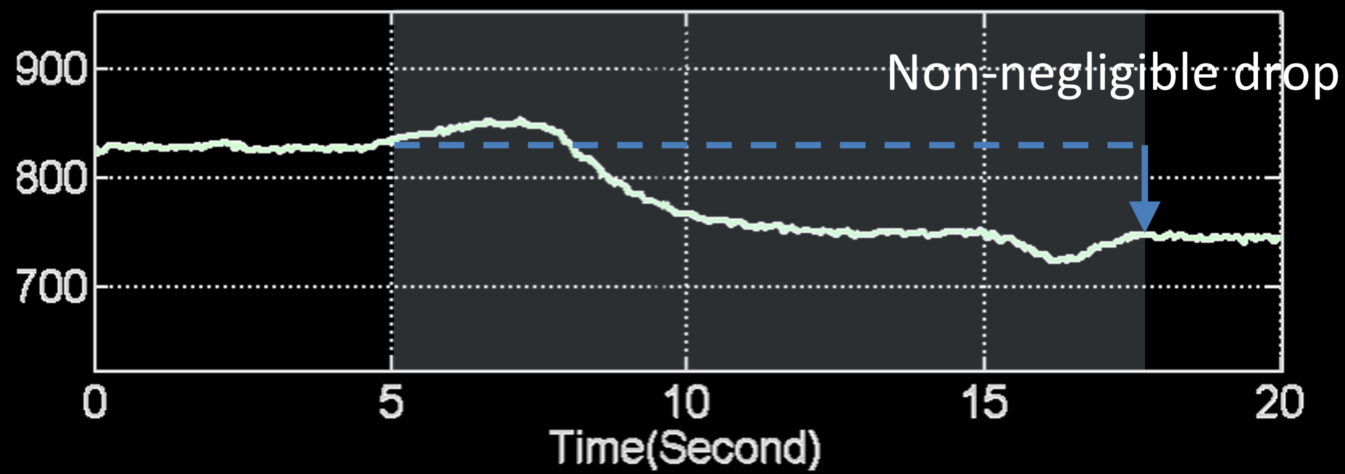
- Identify factors capable of interfering with sensor readings
  - Environmental factors
  - Deployment factors
  - Target-vehicle factors

Identifying potential inferences

- Identify factors capable of interfering with sensor readings
  - Environmental factors
  - Deployment factors
  - Target-vehicle factors

Environmental factors – magnetic sensor

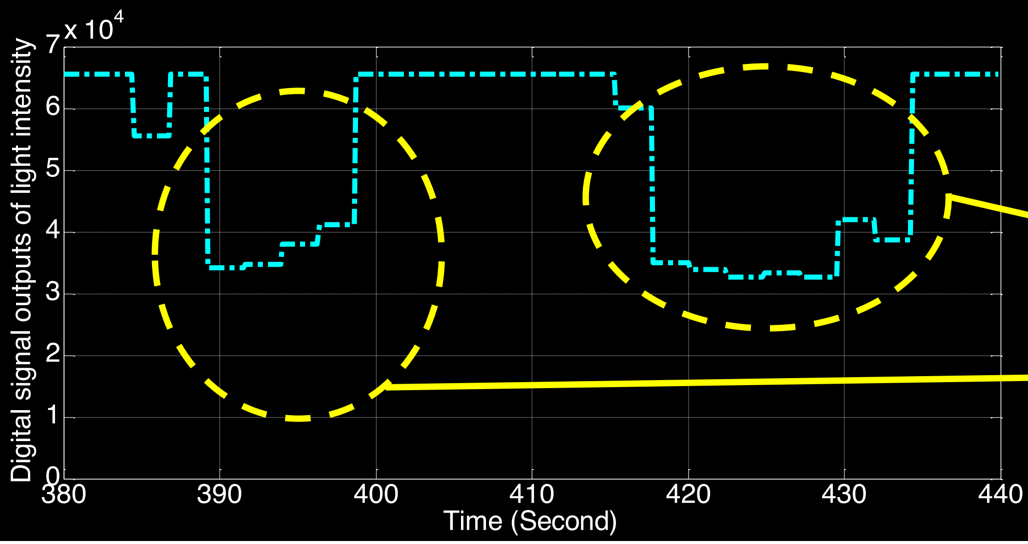
Parking in an adjacent slot during 5~18 s



- The impact on the magnetic field varies
  - The amount of ferrous metal
  - The distance from a vehicle to the parking node

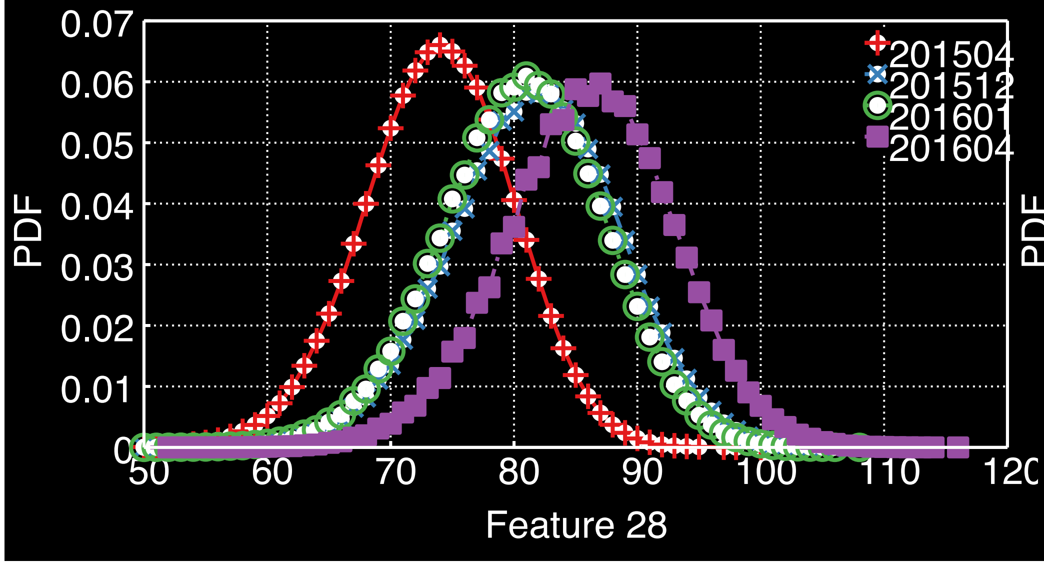
Environmental factors - light sensor

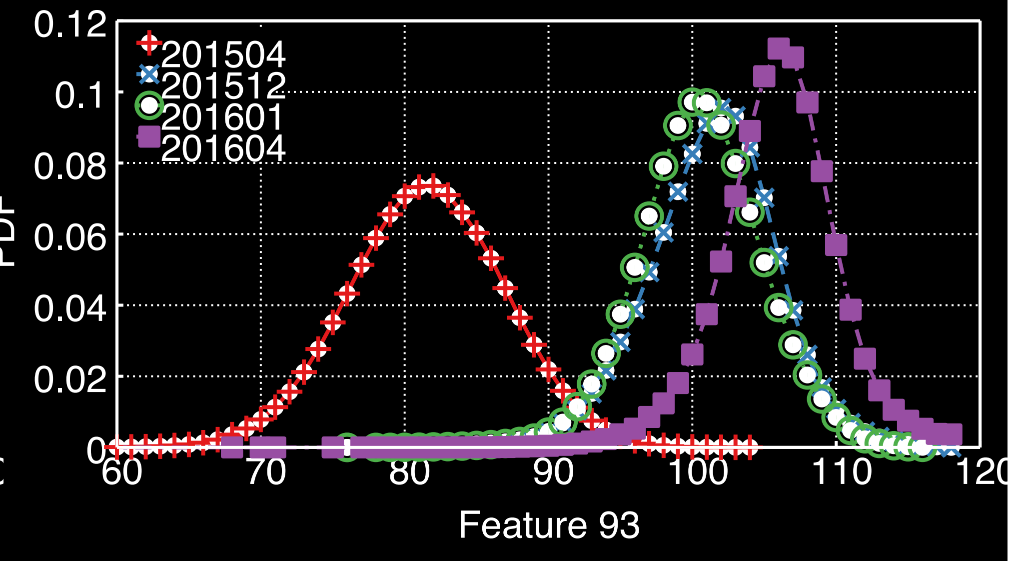
- The sun moves across the sky and sometimes be blocked by stationary objects
  - Sometimes produce patterns that do not necessarily correspond to an actual parking event



Data distribution over different months

- The values of Feature #28
  - Similarities between 2015.12 and 2016.01
  - Considerable fluctuations in the other two months
- The fluctuations in Feature #93 are also pronounced





Initial Observations from the long-term deployment

- Any difference in data distribution can confuse classifiers and thereby undermine prediction accuracy [2]
  - Increasing the size of training data set would not necessarily improve classification accuracy
- This long-term data collection
  - Guide the subsequent design of the system

[2] H. Wang, W. Fan, P. S. Yu, et al.Mining concept-drifting data streams using ensemble classifiers. In Proc.KDD '03, pages 226–235, New York, NY, USA, 2003. ACM.

Outline

- Motivation
- Experiment
- Initial Observations
- Method
- Evaluation
- Related work
- Conclusion

Proposed methods

- Three schemes to handle distribution drifts in the collection of data
  - Failure detection
    - Monitor changes in the distribution of incoming data
  - Model selection
    - Identification of the most appropriate model
  - Sensor selection
    - Switching sensors to optimize the system for accuracy and minimal power consumption

Failure detection

- Periodically conducting the statistical tests
  – Determine whether feature distributions have changed
- Trigger model selection
  – Identify another suitable trained model

Statistical tests to reveal shifts in feature distribution

- Kruskal-Wallis (KW) test
  - Measuring the distance between medians among groups
- Hellinger distance
  - Measuring similarity between two probability distributions

Quantifying the amount of drift in the distributions

- A model/feature is preferred
  - Having **feature distributions similar** to the testing data
  - **Transient and non-transient events** have distributions sufficiently **distinct**

Scoring schemes

- Defining two scoring schemes for each statistical test
  - $Score\_KW^{f}_{stable} = \frac{1}{KW(P^{f}_{T_1}, P^{f}_{T_2})}$ Similar feature distributions
  - $Score\_KW^{f}_{comb} = \frac{KW\left(P^{f}_{T_1|+}, P^{f}_{T_1|-}\right) + KW(P^{f}_{T_2|+}, P^{f}_{T_2|-})}{KW(P^{f}_{T_1}, P^{f}_{T_2})}$

    Distinct transient and non-transient events

    Similar feature distributions
- Similarly, we define the following scores based on Hellinger distance

Failure detection

- By analyzing the collected data
  - No more than 13 hours are required to collect enough data to conduct statistically meaningful tests

Model selection

- Data collected from different slots at different times may differ from new incoming data
  - Difficult to select a subset of data and features capable of maximizing detection accuracy
- Given a set of trained models or a set of training data with a number of features
  - Select the model or subset of features best suited to the classification of incoming testing data

Sensor selection

- A combination of low-cost sensors
  - Facilitate the detection of transient events
- The proposed sensor selection scheme
  - Seeking a trade-off between higher accuracy and lower power consumption
  - Switching on and off according to observed environmental factors with the aim of achieving the following objective

$$min_I \sum_{I=1}^{S} I(i) \cdot [Pow(i) - C \cdot Prob(i|F)]$$

Outline

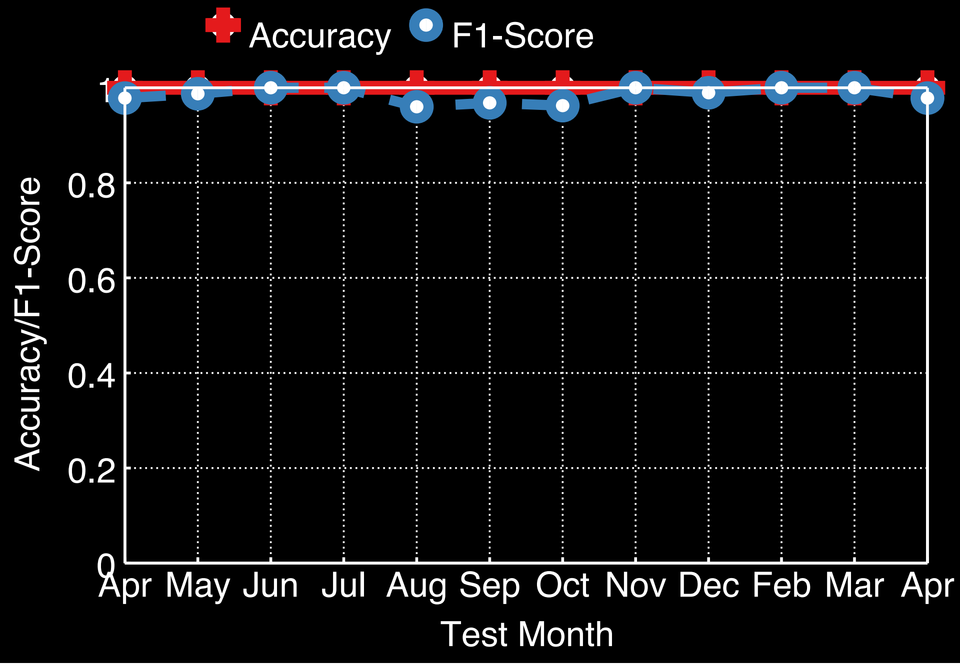
- Motivation
- Experiment
- Initial Observations
- Method
- Evaluation
- Related work
- Conclusion

Evaluation

- Evaluate the proposed schemes with the collected 13-month data
- Metrics
  - Accuracy： $\frac{(TP+TN)}{(P+N)}$
  - Precision： $\frac{TP}{(TP+FP)}$
  - Recall： $\frac{TP}{(TTP+FN)}$
  - F1-score： $2 \cdot \frac{Precision \cdot Recall}{(Precision+Recall)}$

Comparable detection performance

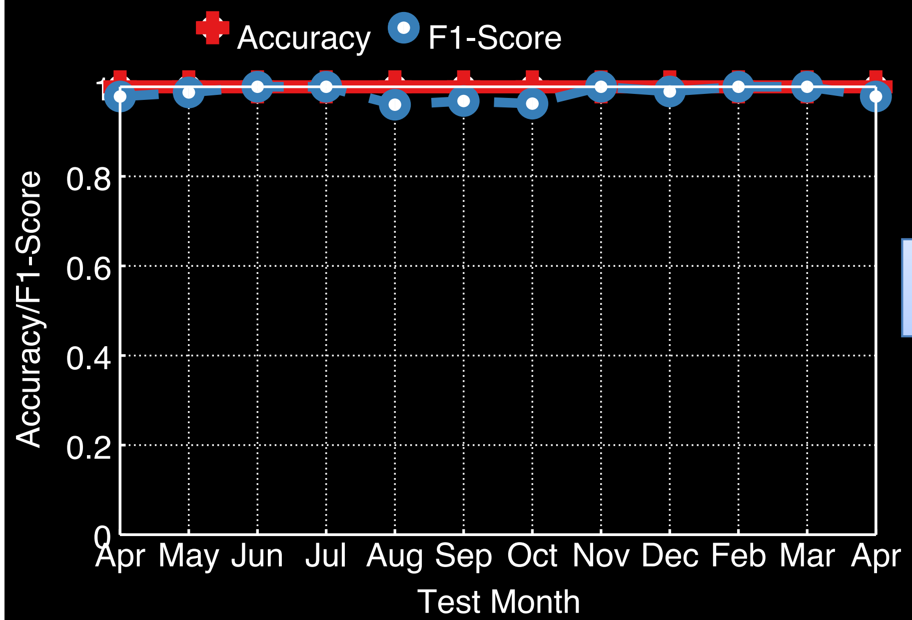
- Ten-fold cross validation to data obtained in each given month
- Comparable with the state-of-the-art commercial products [2][3]
  - Accuracy close to 1 and F1-scores ranging from 0.96 to 1

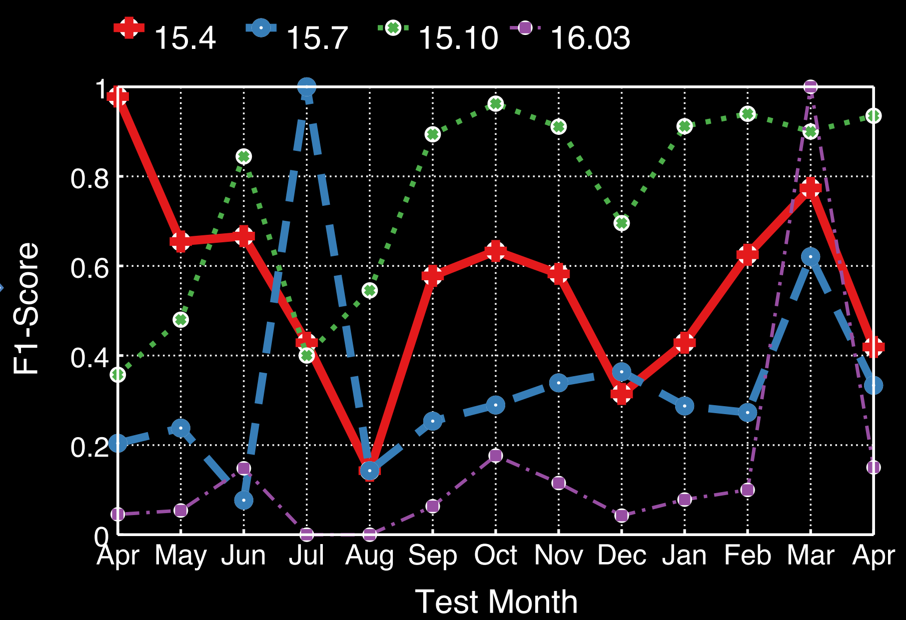


[3] ADEC Technologies. http://www.adec-technologies.ch/.
[4] Fastprk. http://www.fastprk.com/.

Impact of temporal gap between training and test data

- Selecting monthly data obtained in any of the four months (2015.04, 2015.07, 2015.10, and 2016.03) as a training set
- Testing in other months or cross-validating within the same month



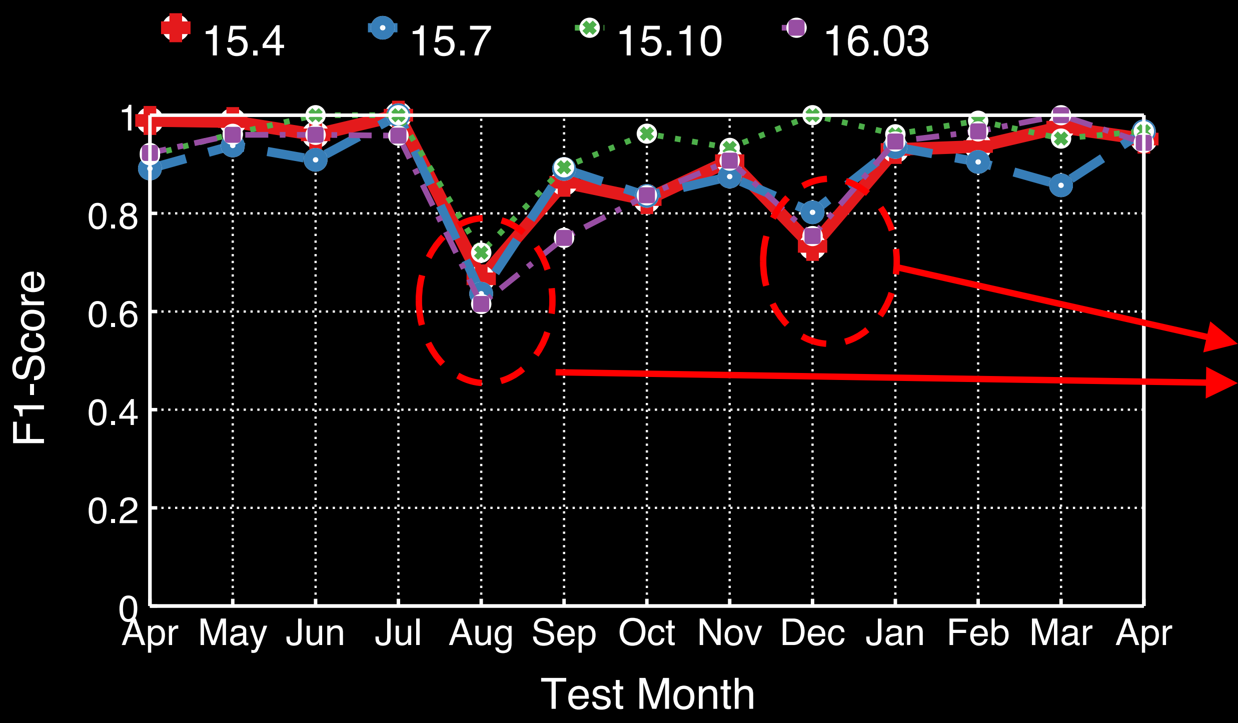


Increase the data size

- Expanding the training data from 1 month to 6 months (2016.04~2016.09)
  - Testing the data from 2016.03
- Training with a large dataset did not yield a better F1-score
  - Distribution mismatch between different months

Effectiveness of the proposed adaptive learning scheme

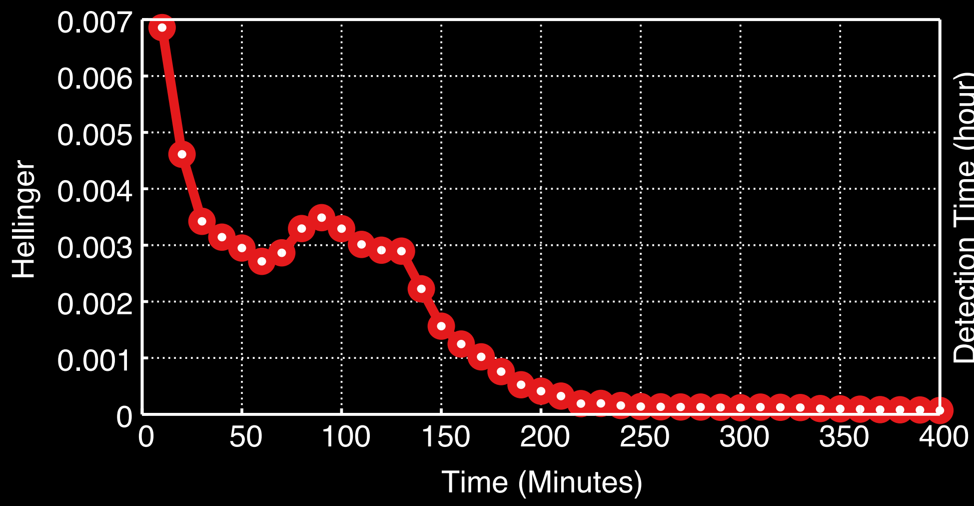
- Using features with top-40% features under the KW_{stable} scoring scheme
  - Most of the F1-Scores were more than 90%
  - F1-Scores obtained using 2015.08 or 2015.12 as training data to predict transient events in other months were lower than 90%

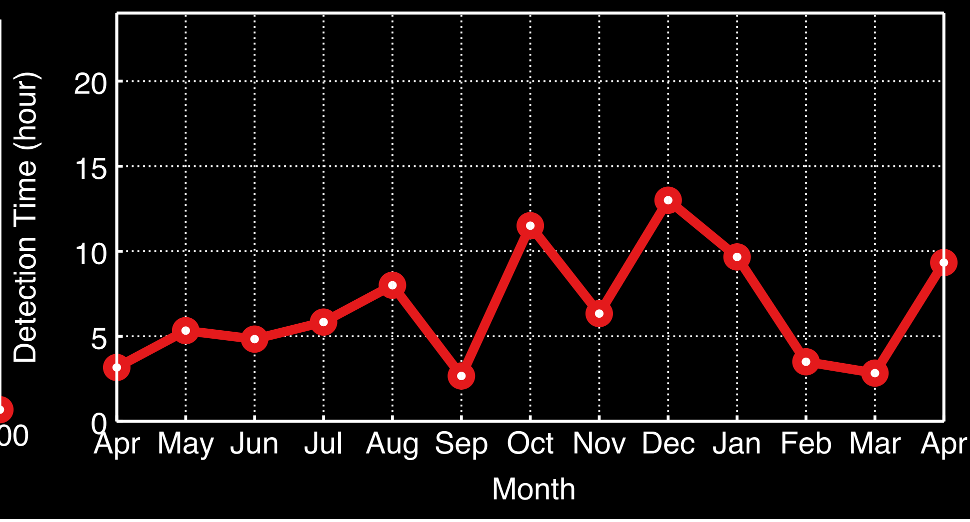


Features with significant drift in the distributions with scores ranked in the top 40% under the KW_{stable}

How long a model converges in each month

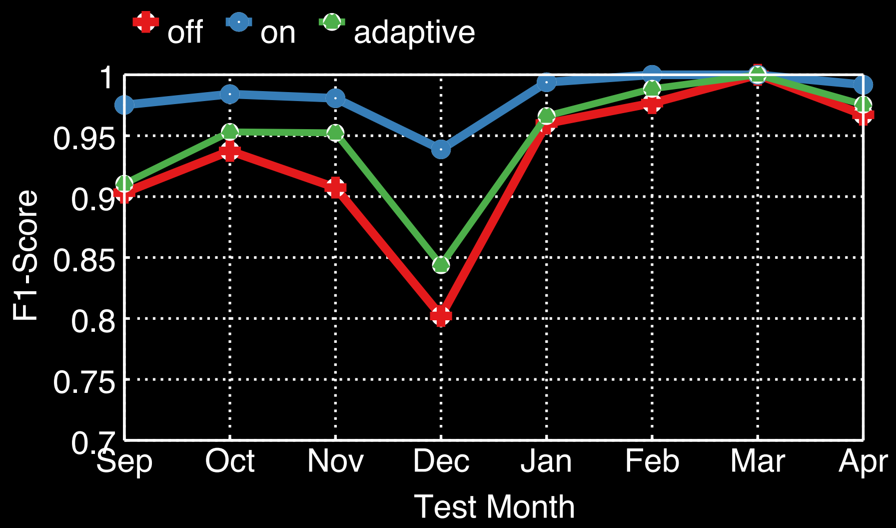
- Hellinger distance between the feature distribution
  - Taking 220 min (~= 3 hours) to determine whether incoming data presents the same distribution as data from 2015.04
- The time required for the Hellinger distance to converge (<= 0.001) in various months
  - Taking 3 ~ 13 hours

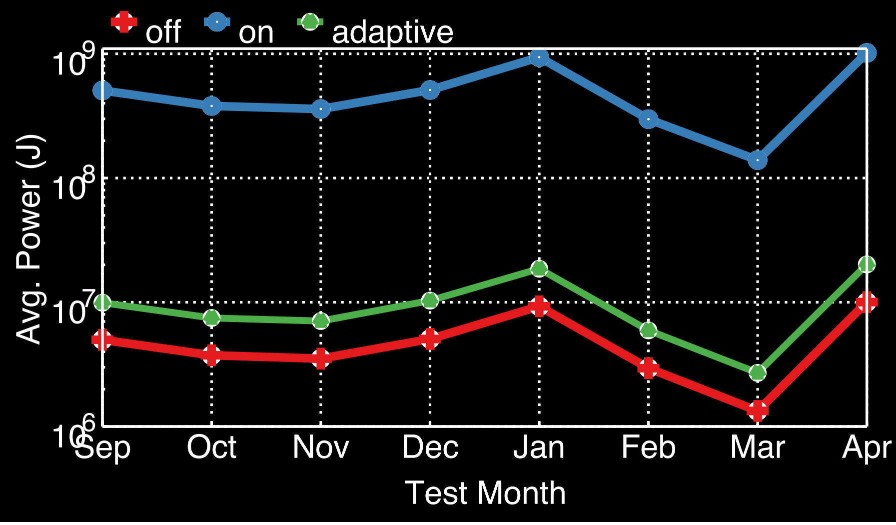




Effectiveness of sensor selection

- Additional sensors can be turned on when confidence is lower than 0.81
- The adaptive sensor selection scheme
  - Slightly Improve the F1-score
  - Maintaining power consumption at only 2% of the power consumed when all of the sensors are on





Outline

- Motivation
- Experiment
- Initial Observations
- Method
- Evaluation
- Related work
- Conclusion

Related work

Category	Solutions	Major problem
Half-buried	Streetline, Sfpark, Urbiotica, Fastprk, fastprk-2, Streetline with camera, or ADEC	- Susceptible to environmental interference - Time and spatial dynamic issues
Non-intrusive	Passive acoustic array, passive infrared sensor, passive ultrasonic sensor, RFID, microwave, or video image processing	- Expensive overhead installations and on-going maintenance - Prone to be influenced by environmental disturbances
Intrusive	Inductive loops, piezoelectric cables, or weigh-in-motion sensors	- Cutting of pavement for installation - Might install multiple detectors
Crowdsourcing	ParkNet, or ParkSense	- Requiring participation by a substantial proportion of drivers

Outline

- Motivation
- Experiment
- Initial Observations
- Method
- Evaluation
- Related work
- Conclusion

Conclusion

- Collecting data from a multi-slot parking area over a period of 13 months
  - A review of interference patterns and long-term trends in the data
    - Model selection, failure detection and sensor selection
  - A series of experiments validated the accuracy of the adaptive schemes
- Our goal
  - Highlight the need for an adaptive machine learning scheme in the design of parking occupancy detection systems

Q & A

Bing (Chuang-Wen) You
cwyou2004@gmail.com
https://sites.google.com/site/cwyou2004/