# Experience: Adopting Indoor Outdoor Detection in On-demand Food Delivery Business

Pengfei Zhou
Alibaba-NTU Joint Research Institute
Nanyang Technological University
pengfei.zhou@ntu.edu.sg

Yi Ding
Alibaba Group
University of Minnesota
dingx447@umn.edu

Yang Li
Alibaba Group
ly233039@alibaba-inc.com

Mo Li
Alibaba-NTU Joint Research Institute
Nanyang Technological University
limo@ntu.edu.sg

Guobin Shen
Alibaba Group
shen.sgb@alibaba-inc.com

Tian He
University of Minnesota
tianhe@cs.umn.edu

## ABSTRACT

This paper presents our experience in adopting recent research results of mobile phone based indoor/outdoor detection (IODetector) to support the real world business of on-demand food delivery. The real world deployment of the adopted IODetector involves three phases spanning 20 months, during which the deployment scales from a feasibility study across a few areas of interest to a citywide trial in Shanghai, and eventually to nationwide deployment over 367 cities in China. Iterative development has been performed throughout different deployment phases to excel the IODetector. Large scale evaluation and comparative A/B testing suggest key value of adopting indoor/outdoor detection in the real world business. We also present the lessons learned from the deployment experience including real world know-hows, practical limits and constraints, as well as discussions on design alternatives. We believe this paper provides insights to guide future efforts in translating research results to industry adoptions.

## CCS CONCEPTS

• **Computer systems organization → Embedded and cyber-physical systems**; • **Human-centered computing → Ubiquitous and mobile computing**; • **Networks → Location based services**; • **Applied computing → E-commerce infrastructure**.

## KEYWORDS

Indoor/outdoor Detection, Nationwide Deployment, On-demand Food Delivery, Business Adoption.

**ACM Reference Format:**
Pengfei Zhou, Yi Ding, Yang Li, Mo Li, Guobin Shen, and Tian He. 2022. Experience: Adopting Indoor Outdoor Detection in On-demand Food Delivery Business. In *The 28th Annual International Conference On Mobile Computing And Networking (ACM MobiCom '22), October 24–28, 2022, Sydney, NSW, Australia.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3495243.3517023


*ACM MobiCom '22, October 24–28, 2022, Sydney, NSW, Australia*

ACM ISBN 978-1-4503-9181-8/22/10…$15.00
https://doi.org/10.1145/3495243.3517023

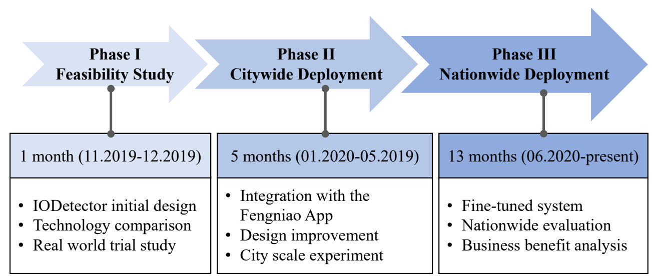



**Figure 1: Three phase deployment of the IODetector.**

## 1 INTRODUCTION

Previous studies have explored the use of various sensors on smartphones to detect the indoor/outdoor status [8, 11, 17, 27, 30, 31, 37]. Although many potential applications have been suggested in the literature including context aware sensing [16, 24, 36], adaptive localization [9], daily activity profiling [26, 29, 34], there has not been any reported experience in the commercial adoption of the technology in real world. This paper presents our experience of applying indoor/outdoor detection (IODetector) to the on-demand food delivery industry [23], which to our knowledge is the first of its kind. The adopted IODetector extends its original design as proposed in [37] and primarily utilizes four sensing resources on the food courier's mobile phone including the light sensor, the magnetometer, the cellular module, and the GNSS (Global Navigation Satellite System) module in order to derive the real time indoor/outdoor status of the courier. With accurate indoor/outdoor knowledge of its fleet of couriers, the food delive ry platform is able to make better optimized decisions in order assignment and as a result achieve improved food delivery efficiency.

We work with Ele.me [15], the second largest on-demand food delivery service platform in China with its business operating in over 300 cities, a fleet of over one million food couriers in contract, and generated revenue of 25.4 billion RMB in the fiscal year 2020 [19]. The deployment and evaluation of the IODetector have gone through three phases as illustrated in Figure 1. In Phase I, a feasibility study was performed from Nov. 2019 to Dec. 2019, when we finalized the basic architecture of the IODetector and evaluated its performance with small scale experimentation at several areas of interest in Shanghai city. In Phase II, the IODetector was integrated into the operational App of all Shanghai couriers' mobile phones for a trial study and tested in their daily business routines. The trial lasted from Jan. 2020 to May 2020, during which the IODetector was evaluated based on ~90,000 monthly active couriers in Shanghai delivering 800,000 daily food orders. A total number of 2,895 different phone models were involved in the study. In this stage, we dealt with practical challenges arising from the scale of the deployment, including the difficulty in obtaining groundtruths and the prevalence of corner cases that may impair the worst case performance. The Phase III deployment started from Jun 2020 and has lasted ever since. In this phase, the IODetector has been gradually launched for Ele.me food couriers in 367 cities. We progressively refined the parameter settings of the IODetector based on periodically updated evaluation results. The up-to-date evaluation result has been obtained from ~1 million monthly active couriers with 4,861 different phone models, which gives the average detection accuracy of 87.3% indoors and 90.6% outdoors.

A comprehensive end-to-end A/B testing has been conducted to further estimate the business benefit to the on-demand food delivery industry from the IODetector. The results show that the adoption of the IODetector shortens the average per-order delivery time by 19 seconds and reduces the late delivery rate relatively by 5.64%, which translate to yearly profit gain of ~11 million RMB to the platform and yearly cost saving of ~108 million RMB to participating food couriers and restaurants.

The lessons learned through the IODetector deployment are summarized as follows.

- We demonstrate the practical value of applying advanced mobile sensing technology to commercial business and share the real world know-hows of the application.
- We share our experiment methodology for obtaining large scale ground truth references. We discuss our thoughts and studies involved in devising the methodology.
- We show the prevalence of corner cases when a mobile sensing system scales, and share our considerations in achieving performance trade-offs between worst cases and overall performance.
- We share our design considerations when translating the research advances. In particular we comparatively study the machine learning based and rule based mobile sensing solutions when it comes to practical adoption.

The rest of this paper is organized as follows. Section 2 presents the background of the on-demand food delivery industry and the initial design of the IODetector. Section 3 introduces our three-phase deployment and re-development experience of the IODetector along with the business benefit evaluation and analysis. Section 4 discusses the lessons learned from the deployment experience and the expected future research topics. Section 5 reviews the related works and Section 6 concludes this paper.

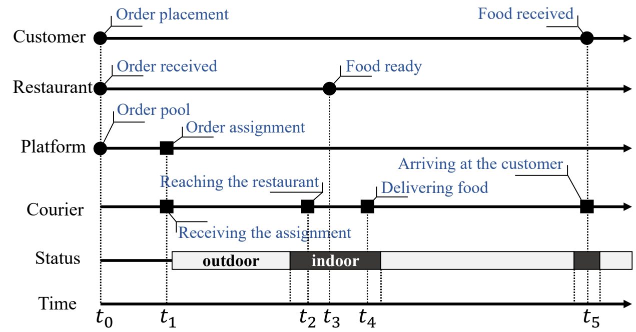



**Figure 2: The food order servicing cycle.**

## 2 IODETECTOR FOR ON-DEMAND FOOD DELIVERY SERVICE

### 2.1 On-demand Food Delivery

The on-demand food delivery business features extremely high timeliness requirements with huge order amounts unevenly distributed in time and space (over 10 million orders per day with less than 40 minutes average delivery time on Ele.me platform). A smart order dispatch strategy is necessary to best match the food orders and couriers in real time to improve delivery efficiency.

To serve the purpose of efficient order dispatch, one of the most important problems is *how to accurately track courier status?* Figure 2 illustrates the basic information flow of a food order servicing cycle. Once a customer places a food order online at $t_0$, the restaurant receives the order and starts preparing the food. The platform, at the same time, puts the order in a temporary order pool containing all the orders generated within the past 1~2 minutes. Once the pool is full or the time expires, the platform considers the current courier status, the geographic distribution of orders as well as other related information to make order assignments to each courier. When the courier arrives at the restaurant at $t_2$, if the food is ready, he/she picks up the food, leaves the restaurant at $t_4$ and reaches the customer at $t_5$ to complete the delivery. Note that in practice, a courier usually has multiple food orders on hand at a time. For example, the courier may receive new order assignments when delivering a previous order. Knowing time points $t_2$~$t_5$ is thus very valuable for the online dispatching algorithm to make subsequent assignments. The exact knowledge of $t_2$~$t_5$, however, is not possible to the platform and can only be inferred according to courier statuses. The indoor/outdoor context of the courier segments its trajectory and helps identify the key time points ($t_2$~$t_5$) within certain time ranges which helps the order dispatching platform best estimate $t_2$~$t_5$ and assign subsequent orders accordingly.

### 2.2 Initial Design

The initial design is extended from the original design of the IODetector [37], the first and the most impactful work for generic indoor/outdoor detection with smartphone sensors. The adopted IODetector makes use of three energy-efficient mobile phone sensing resources, i.e., light sensor, cellular module and magnetometer, which according to real-world measurement and experiment

**Table 1: Three phase deployment of the IODetector.**

| Phase / Metric | **Phase I: Feasibility study** (Nov. 2019.11 - Dec. 2019) | **Phase II: Citywide deployment** (Jan. 2020 - May 2020) | **Phase III: Nationwide deployment** (Jun. 2020 - present) |
|---|---|---|---|
| | 5 iPhones and 15 Android phones 2 areas of interest in Shanghai | Over 250$k$ phones comprising 2,895 phone models in Shanghai | Over 3 million phones comprising 4,861 phone models in 367 cities of China |
| Participants | 20 research assistants and couriers | 90$k$ monthly active couriers in Shanghai | 1 million monthly active couriers over 367 cities |
| Design upgrade | GNSS input integrated and cellular sensor upgrade | Improved use of GNSS as well as parameter optimization | Improved use of magnetic field signal as well as parameter refining |
| Detection accuracy | 88.5% indoors, 91.3% outdoors | 85.1% indoors, 88.9% outdoors | 87.3% indoors, 90.6% outdoors |
| Efficiency improvement | N.A. | N.A. | Avg. delivery time reduction:19 secs/order, Relative late delivery rate reduction: 5.64% |

[30, 37] exhibit different and distinct data patterns indoors or outdoors. Outdoor light, primarily the natural sunlight in the daytime, has wider spectrum and is of higher intensity in nature, while indoor light is primarily artificial light, and is of much lower intensity. The mobile phone cellular signal from the ambient cell towers may significantly increase ($> 10dB$) when the mobile phone user moves from indoor to outdoor and decrease from outdoor to indoor due to the signal blockage. The indoor Earth magnetic field signal is heavily distorted due to the existence of steel construction and electrical appliances but not when outdoors.

Figure 3 illustrates the main design rationale of the IODetector, which inherits the main framework from our previous work [37]. The indoor/outdoor context of the courier is firstly estimated separately by each of the three sensors and the final detection result is a weighted combination of the outputs of the three sensors based on their confidence levels[1]. The confidence level of the output of each sensor is a normalized value from 0 to 1, where 1 indicates the highest confidence and 0 indicates the lowest. For each detection, the IODetector considers the sensor readings in the past $w$ seconds. For the light signal, the IODetector measures the light intensity $L$ and compares it with a threshold $\alpha$ to decide indoor or outdoor status. The light sensor detection confidence is low when at night or not available when the mobile phone is identified in a pocket (with proximity sensor). For the cellular signal, the IODetector measures the signal variation $\Delta C$ and compares it with a threshold $\beta$ to decide indoor or outdoor status. The IODetector uses the magnetometer to measure the signal variation of the ambient Earth magnetic field $\Delta M$ and compare with a threshold $\gamma$ to decide indoor or outdoor status. Due to the nature of the Earth magnetic variation, the indoor status is reported with high confidence but the outdoor status is reported with low confidence. The IODetector further extends the design by implementing a fourth sensor based on mobile phone GNSS module input, which will be detailed in Section 3.1.

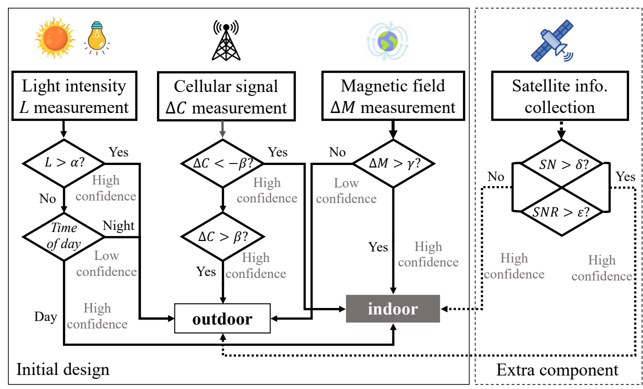



**Figure 3: The design of the adopted IODetector.**

## 3 DEPLOYMENT EXPERIENCE

Since Nov. 2019, we launched the IODetector deployment in three consecutive phases, i.e., feasibility study phase (Phase I, Nov.-Dec. 2019), city scale deployment phase (Phase II, Jan.-May 2020), and nationwide deployment phase (Phase III, Jun. 2020 - present). Table 1 summarizes key particulars during the three phase IODetector deployment. In the following, we will present the design upgrades, the scale of deployment, real-world evaluation methodology and the IODetector performance, and the efficiency improvement for food delivery.

### 3.1 Phase I Deployment

In Phase I, we deploy and evaluate the IODetector from a feasibility study perspective. We engage 20 participants including 10 research assistants and 10 Ele.me food couriers, and test the performance of the IODetector in two city plaza areas in Shanghai which contain shopping malls, office buildings, and various indoor/outdoor zones. The testing mobile phones comprise 5 iPhones and 15 Android phones of different models. Since the light sensor readings and satellite information are not available from iOS APIs, we deploy a simplified version for iPhone. The study was conducted for four weeks from Nov. to Dec. 2019.

**IODetector design upgrade.** We make the following design upgrades.

*GNSS sensor.* The Ele.me platform already employs GNSS modules of couriers' mobile phones to update location information, so we decide to leverage the GNSS input and integrate that as a fourth sensor into the IODetector. No extra energy consumption is incurred. Thanks to the deployment of GPS [5], Beidou [1], GLONASS

[1] We encourage readers to also read [37] to gain a more comprehensive understanding of the design rationale and technical considerations.

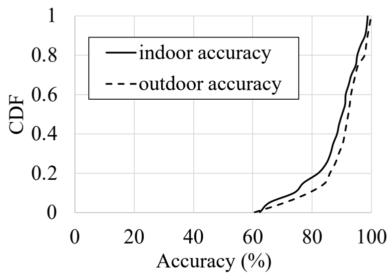



**Figure 4: Phase I accuracy.**

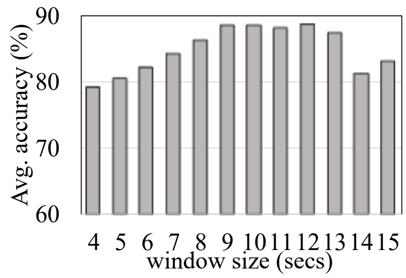



**Figure 5: Window setting.**


**Figure 6: Vision-assisted labeling.**

[4], and Galileo [3] system, over 100 positioning satellites presently orbit the Earth at the height of ~20,000$km$ and are available to provide a reference signal. We may differentiate satellites from different systems based on their pseudorandom noise (PRN) codes or space vehicle numbers (SVN). The current implementation of the IODetector only uses GPS input for phone compatibility reason, but can be easily extended to work with other GNSS inputs when needed. As Figure 3 illustrates, an extra GNSS based sensor is added. We consider both the number of visible satellites $SN$ and the received $SNR$, and threshold them with $\delta$ and $\varepsilon$, respectively. If $SN > \delta$ and $SNR > \varepsilon$, the sensor outputs outdoor with high confidence. If $SN < \delta$ and $SNR < \varepsilon$, it outputs indoor with high confidence. Otherwise, it outputs indoor or outdoor with low confidence. The output of the sensor is weighed in when aggregated with the outputs from the other three sensors to derive the final result.

*Cellular sensor.* During the feasibility study, we also observe that the effectiveness of the cellular sensor may be impaired due to unexpected disruption to the line of sight from the serving base station to the mobile, e.g., when the mobile turns around a corner. As a result, the sudden drop of signal strength (at ~15-30dB) of one certain cell tower may lead to wrong estimated outdoor-indoor transition, especially when that cell tower is highly weighted. We revise the estimation logic of the original cellular sensor and take into consideration the signal consistence across multiple cell towers. The IODetector monitors the signal variation of all visible cell towers and trusts more on the cellular sensor with consistent variations. In each detection window, $n$ (mostly 4~7 in practice) signals from neighboring cell towers are used in addition to the one that the mobile is associated with. The weight of the associated cell tower $c_a$ is set to $w_a$, and the total weight of all neighboring cell towers ($c_1$~$c_n$) is set to $1 - w_a$. For each cell tower $i$, its output indicator $e_i$ is thereafter determined by the signal variation $sv_i$:

$$e_i = \begin{cases} 1 & \text{if } sv_i > \beta, \text{ i.e., outdoor;} \\ -1 & \text{if } sv_i < -\beta, \text{ i.e., indoor;} \\ 0 & \text{otherwise, i.e., unchanged/unknown.} \end{cases} \quad (1)$$

Assembling the results from all cell towers, the output of the cellular sensor with confidence is $O_{cellular} = e_a * w_a + (1 - w_a) \sum_{i=1}^{n} \frac{e_i}{n}$, which thus has varied dependence on the main cell tower the mobile is associated with and other observable neighboring cell towers.

*IODetector for iPhone.* Due to the constrained sensor exposure from the limited iOS APIs where the detailed information of light signal, satellite signal and the neighboring cellular signal is not available, we have to adopt a much simplified algorithm for iPhones, where the variation and calibration accuracy of the magnetic field, the signal of the main associated cell tower, and the estimated localization error are used to estimate the indoor/outdoor status.

**Performance evaluation.** We evaluate the performance of the IODetector with 60 walking traces collected from the 20 participants, who are advised to manually label their movement trajectories as groundtruth. In particular, the 10 recruited couriers are advised to walk through similar paths as they normally do when delivering food orders. The overall cumulative distribution function (CDF) of the indoor/outdoor detection accuracy is plotted in Figure 4. The average detection accuracy is 88.5% for indoors and 91.3% for outdoors. The proportion of iPhones among the phones used by Ele.me couriers varies from 10% to 25% in different cities. The average detection accuracy with Android phones is 92.9%, which is much higher than that with iPhones (77.1%), mainly due to the sensing constraints we face with iPhones. The results suggest the high potential of the IODetector for large scale deployment.

## 3.2 Phase II Deployment

After the feasibility study, in Phase II we deploy the IODetector into the food delivery routine of all Ele.me couriers in Shanghai. The Phase II deployment lasted from Jan. 2020 to May 2020. We tested the IODetector in the wild with ~90$k$ monthly active couriers delivering ~800$k$ food orders in Shanghai every day. A total number of over 250$k$ mobile phones of 2,895 different models were tested in this phase.

**IODetector design upgrade.** To build the IODetector into the Ele.me food delivery routine, we integrate the IODetector in the Fenginao App [15] - a mobile phone app used by Ele.me couriers to receive order assignments, manage order operations, and update availability for assignments (e.g., *active* or *resting*). The IODetector build-in outputs the courier's indoor/outdoor status every second, and the result is uploaded to the Ele.me cloud platform periodically. Couriers are aware of the data collection [2] and for the IODetector we collect only the indoor/outdoor detection result when their working status is *active*. No raw sensor data is collected. We finalize the window size of each detection. In reality, the optimal window size varies with both the environment and the moving speed of the courier. As the detection accuracies for window sizes of 7~13 are of little difference as suggested by the experimental results shown in

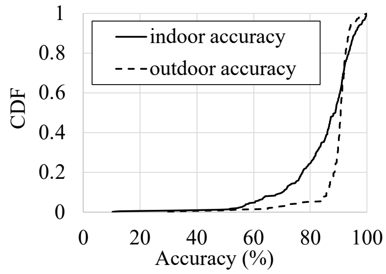
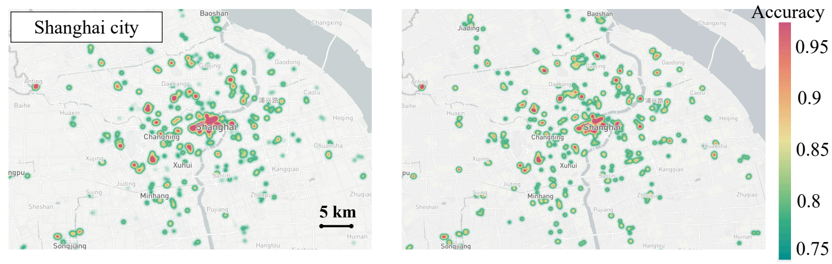



**Figure 7: Phase II accuracy.**




(a) Accuracy for indoors. (b) Accuracy for outdoors.

**Figure 8: Geographical distribution of the IODetector accuracy across Shanghai.**


(a) First month (b) Seventh month (c) Thirteenth month

**Figure 9: Nationwide deployment of the IODetector in Phase III.**

Figure 5, we use a fixed window size of $w$=9 seconds for reliability consideration.

**Performance evaluation.** We evaluate the performance of the IODetector with Ele.me food delivery routine. The ground truth is collected in two different ways.

*Vision-assisted labels.* As depicted in Figure 6, we let couriers wear cameras on the chest to record videos during their daily work. We use the HNSAT UC-30 mini camera that can rotate 180° and gives full HD 1920*1080P video. Due to the high overhead and cost, we recruit 10 Ele.me couriers for one week. They deliver food orders in 3 AoIs (areas of interest), to provide ~5,000 in-situ video clips (~420 hour video footage in total) in the experiment from which we recover precise traces of their indoor/outdoor statuses.

*Inferred labels.* The vast majority of the groundtruth labels used in the evaluation are inferred labels. We rely on the indoor BLE beacon infrastructure which has earlier been deployed by Ele.me in certain restaurants and shopping malls to audit the arrival of food couriers to those PoIs [13, 14]. There are currently over 860$k$ beacon devices in use. Receiving the beacon signal with RSSI above the threshold gives positive labels of indoor status. We also examine the moving speed of the courier and infer positive labels of outdoor status when the moving speed is continuously above the threshold because Ele.me couriers ride standard e-bikes when moving on roads.

While the *vision-assisted labels* provide holistic indoor/outdoor profiles along the couriers' movement trajectories but are limited in size, the *inferred labels* only provide sporadic reference tags mostly near indoor PoIs but enable city scale and nationwide evaluation due to the high data scale.

The overall CDF of the indoor/outdoor detection accuracy is plotted in Figure 7. The average detection accuracy is 85.1% for indoors which is 3.4% lower than that in Phase I, and 88.9% for outdoors which is 2.4% lower than that in Phase I. The decrease of detection accuracy is mainly due to the prevalence of corner cases that may yield low accuracy (the tail of the CDF in Figure 7). Our efforts to trade average accuracy for improved worst case accuracy also lead to the reduction of average accuracy. Figure 8 presents the geographical diversity of the detection accuracy in 552 AoIs across Shanghai city. Although we do not have the complete groundtruth of the entire city of Shanghai, the inferred labels from many scattered AoIs over the city allow us to derive accuracy across those regions.

## 3.3 Phase III Deployment

We started the Phase III deployment from June 2020, during which the IODetector was gradually built into the Fengniao App of all Ele.me couriers across the country. Till now, the deployment has engaged ~1 million monthly active couriers across 367 cities in China. A total number of over 3 million mobile phones of 4,861 different models have been involved. Figure 9 visualizes the nationwide deployment of the IODetector achieved at three stages - (i) Jun. 2020, the first month when we dealt with compatibility issues and deployed the tool in 30 major cities; (ii) Jan. 2021, the seventh month when we fine tuned system parameters and expanded the deployment scale to 311 cities; (iii) Jul. 2021, the thirteenth month since when we have arrived at the deployment scale to 367 cities till now.

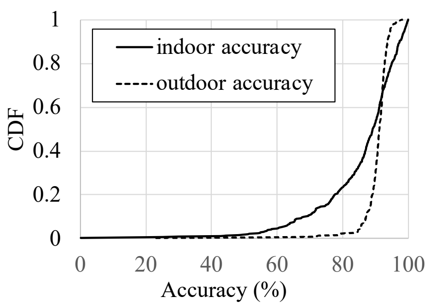



Figure 10: Phase III accuracy.

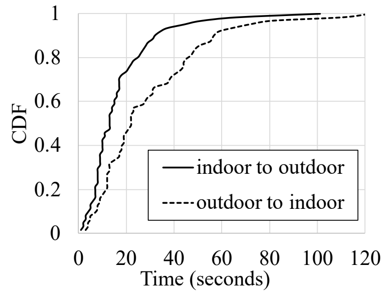



Figure 11: Detection latency in outdoor-indoor transitions.

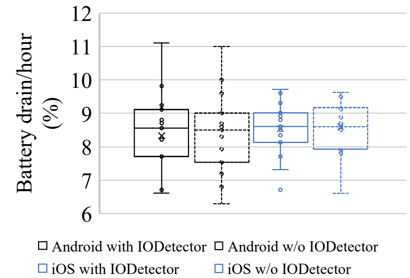



Figure 12: Energy consumption.

Table 2: System overhead for the IODetector built-in.

| Fengniao SDK (test scenario) | Size | Power consumption | CPU user time (minute, second) | CPU system time (minute, second) | WiFi / cellular transmission |
|---|---|---|---|---|---|
| w/o IODetector (static) | 170kB | 93.2 mAh | 8m 36s | 3m 26s | 2.09MB |
| w/o IODetector (mobile) | 170kB | 106.9 mAh | 7m 42s | 4m 12s | 3.33MB |
| with IODetector (static) | 194kB | 95.1 mAh | 8m 56s | 3m 45s | 2.18MB |
| with IODetector (mobile) | 194kB | 109.1 mAh | 9m 16s | 4m 56s | 3.56MB |

**IODetector design upgrade.** Since China has a vast territory and the signal conditions (in particular GNSS and cellular signals) at different cities may slightly vary, we fine-tune the system parameters across different cities to ensure reliability and efficiency. The measured signal strength of the magnetometer varies with both the environment distortion to the geomagnetic field and the orientation of the mobile phone itself. The confidence on the magnetometer sensor output degrades with the phone orientation dynamics. In the updated IODetector the overall change of the phone orientation is thus considered in addition to the signal strength to derive the confidence on the magnetometer output.

**Performance evaluation.** We evaluate the performance of the IODetector based on the data from 367 cities. The ground truth is primarily derived based on the inferred labels as described in Section 3.2. The overall CDF of detection accuracy is plotted in Figure 10. The average detection accuracy is 87.3% for indoors and 90.6% for outdoors, both slightly higher (by 2.2% for indoors and 1.7% for outdoors) than that in Phase II. We believe the accuracy increase is mainly due to the system fine-tuning and also the fact that the complicated environment is more prevalent in Shanghai. Figure 11 presents the latency in the detection during the indoor outdoor transition, which tells the responsiveness of the IODetector. The median detection latency of indoor to outdoor transition is 13 seconds and that of outdoor to indoor transition is 21 seconds. The detection latency is mainly due to the signal sensing inertia during the indoor/outdoor transition.

**System overhead.** The IODetector is a built-in component of the mobile data collection SDK in the Fengniao app. Figure 12 summarizes the energy consumption of mobile phones with and without the integration of the IODetector. We compare the phone battery drains for iPhones and Android phones, and the data is collected from 20 mobile phones during a comparative study where the fully powered phone is completely depleted. We see from the figure that the use of the IODetector has a statistically insignificant impact on the energy consumption. The differences in both the median and mean of the battery drain rate is within 0.15%.

The Battery Historian [18] is used to derive more details of the size, power consumption, communication and computation overhead of the SDK before and after integrating the IODetector. We investigate the system overhead of the two different SDK versions (with and without IODetector) with two settings, i.e., the *static* and *mobile* scenario, where the test phones are placed still or moved with Ele.me couriers for one hour. Table 2 summarizes the average testing results across all tested phone models for different settings. The CPU user time is the amount of time CPU spent in running the SDK code and the CPU system time is the amount of time CPU spent in running the kernel functions connected to the SDK. We see little differences in size, power consumption, CPU user time, CPU system time and the amount of data transmissions between the SDK versions before and after integrating the IODetector in both the static and mobile scenario. In all tests, the power used for screen-on (50%~60%) and wireless transmission (25%~30%) accounts for the majority (80%~90%) of the total power consumption.

## 3.4 Business Benefit

As previously discussed in Section 2, the IODetector's indoor outdoor outputs may contribute to estimating critical business parameters including the ETS[2] (Estimated Time of Stay) and ETA (Estimated Time of Arrival) of couriers. As Figure 2 illustrates, $ETS = t_4 - t_2$ and $ETA = t_5$. The food delivery platform incorporates the estimation of these parameters with other information including the orders in the temporary order pool, courier previous

[2]In Ele.me, there is a set of ETS parameters including ETS at malls, ETS at restaurants, etc. In this paper we use the ETS at restaurants as an example.

**Table 3: Real world A/B testing for business benefit analysis.**

| Metric \ Order sets | **A1** (w/o IODetector) | **A2** (w/o IODetector) | **Difference (A2−A1)** | **B** (with IODetector) | **Difference (B−A1)** |
|---|---|---|---|---|---|
| Order amount | 425315 | 448441 | - | 842345 | - |
| Average delivery time | 31.17 mins | 31.12 mins | -3 secs | 30.86 mins | -19 secs |
| Average time_lunch | 38.28 mins | 38.37 mins | 5 secs | 38.03 mins | -25 secs |
| Average time_1500m | 36.08 mins | 35.93 mins | -9 secs | 35.54 mins | -32 secs |
| Average time_3000m | 42.59 mins | 42.49 mins | -6 secs | 41.74 mins | -51 secs |
| Late delivery rate | 10.46% | 10.38% | -0.08% | 9.87% | -0.59% |
| Late delivery rate_lunch | 12.78% | 12.66% | -0.12% | 12.06% | -0.72% |
| Late delivery rate_1500m | 14.00% | 14.10% | 0.1% | 13.23% | -0.77% |
| Late delivery rate_3000m | 16.16% | 16.07% | -0.08% | 14.62% | -1.53% |
| Order reassignment rate | 4.55% | 4.59% | 0.04% | 4.52% | -0.03% |
| Reassignment rate_lunch | 5.67% | 5.68% | 0.01% | 5.70% | 0.03% |
| Reassignment rate_1500m | 5.25% | 5.41% | 0.16% | 5.26% | 0.01% |
| Reassignment rate_3000m | 5.87% | 6.13% | 0.26% | 6.03% | 0.16% |

records, traffic conditions, etc., in its dispatching algorithm to find a match between food orders and couriers. Better estimation of the ETS and ETA significantly impact the quality of order dispatch, and thus the business operation cost and profit [38].

**Real world A/B testing.** During Phase III deployment, we also launch a real world A/B testing to evaluate the impact of the IODetector on Ele.me food delivery efficiency. The test was conducted during 22th to 26th of July in 2020 in three major cities in China, namely Shanghai, Beijing and Hangzhou. During the experiment, a total number of ~1.7 million food orders from the three cities were examined. We randomly assign them into 2 groups, orderset A and B, respectively. Each orderset contains ~50% of the total orders. Orderset A is used as the control group, where the orders are assigned to couriers without using the IODetector output, and B is used as the test group, where the orders are assigned by the platform based on the IODetector calibrated ETS and ETA. Orderset A is further divided into subgroup A1 and A2, each of which contains ~25% of the orders, for comparison. The overall result of the test is summarized in Table 3. The key performance indicator for on-demand food delivery business includes three major metrics, i.e., the average delivery time ($T$) of each order, the late delivery rate ($LR$) of orders, and the order reassignment rate ($RR$).

**Average delivery time.** As Table 3 presents, the overall average delivery time difference between A1 and A2 is 3 seconds and the delivery time reduction from A1 to B is 19 seconds. Further analysis shows that the time reduction from A1 to B is 25 seconds for peak-hour orders at lunch time (Average time_lunch), 32 seconds for orders with delivery distance longer than 1500$m$ (Average time_1500m), and 51 seconds for orders with a delivery distance longer than 3000$m$ (Average time_3000m). For comparison, the difference between the control group A1 and A2 is consistently below 10 seconds (mostly below 6 seconds). The result indicates the statistical advantage of the IODetector in reducing the average delivery time. The benefit is higher for peak-hour orders and long-distance orders.

**Late delivery rate.** The late delivery rate measures the ratio of orders that are not delivered within the pre-estimated time duration. As Table 3 presents, the overall late delivery rate is reduced from 10.46% at orderset A1 to 9.87% at orderset B, which gives 0.59% reduction. With further analysis, we see that the late delivery rate reduction from A1 to B is 0.72% for peak-hour orders at lunch time (Late delivery rate_lunch), 0.77% for orders with a delivery distance longer than 1500$m$ (Late delivery rate_1500m), and 1.53% for orders with a delivery distance longer than 3000$m$ (Late delivery rate_3000m). Meanwhile the differences between the control group A1 and A2 remain below 0.12% for all types of orders. We perform a comparison between the late delivery rate distribution of A1 and B. The relative uplift is $\frac{LR_B - LR_A}{LR_A} = 5.64\%$ and the Z-Score difference is $\frac{LR_B - LR_A}{\sigma} = -11.08$, where $\sigma$ is the standard deviation. The result suggests 95% confidence in the IODetector advantage in reducing the late delivery rate relatively by 5.64%.

**Order reassignment rate.** The order reassignment rate indicates the ratio of reassigned orders over all previously assigned orders. The reassignment may take place due to various factors including food delay in restaurants, accidents or exceptions of couriers, traffic incidents, etc. According to Table 3, the differences of the order reassignment rates between A1, A2, and B are mostly below 0.2% for all scenarios. The results do not suggest the obvious benefit of reassignment rate reduction from the IODetector.

**Business benefit.** We estimate the business benefit of the IODetector by translating (i) the gain on the reduction of the average delivery time into the saved time for serving extra food orders and (ii) the reduced late delivery rate to the reduced penalty due to customer complaints or negative service ratings.

*Benefit to the platform.* The IODetector reduces the average delivery time of each order by 19 seconds. According to the statistics, the average delivery time per order is 31 minutes with a profit of 0.3 RMB. The number of food orders on Ele.me platform is 10 million per day. The average delivery time reduction of 19 seconds thus translates to an extra profit to the platform as $\frac{19}{31 \times 60} \times 10^7 \times 0.3 \times 365 \approx 11.2$ million RMB.

*Benefit to the courier/restaurant.* The IODetector reduces the late delivery rate by 0.59%. Assuming the penalty for a late delivery order is 3 RMB[3], the yearly penalty savings of late deliveries can be estimated as $0.59\% \times 10^7 \times 365 \times 3 \approx 64.61$ million RMB. In addition, based on the business experience, we assume that 10% of the late delivery events may lead to a negative rating or customer complaints to the service. Assuming 20 RMB penalty to every such event [4], the extra penalty savings can be estimated as $10\% \times 0.59\% \times 10^7 \times 365 \times 20 \approx 43.07$ million RMB. The total amount of savings are approximately $64.61 + 43.07 = 107.68$ million RMB.

In addition to the quantitative profit and cost saving, higher delivery efficiency would further improve the customer experience, encourage the participation of more restaurants and couriers, and establish competitive advantages for the platform in the long term business.

**Other business applications.** As a general tool for indoor outdoor trajectory segmentation, the IODetector is also beneficial to other business applications in addition to estimating ETS and ETA. The output of the IODetector can be an essential input for various knowledge discovery models in Ele.me including discovering building entrances for courier route planning, arbitration for dispute in late food preparation and delivery, and building profiling for business upgrade.

## 4 LESSONS LEARNED

We present the lessons learned throughout the IODetector deployment experience, based on which we discuss possible future research topics.

### 4.1 Real World Know-hows

Translating research results into commercial adoption involves abundant experience in addressing real world problems, through which we gain know-hows in practicing large scale mobile sensing.

**Sensor selection.** The three initially used sensors, i.e., light sensor, cellular module and magnetometer, contribute differently to the final detection result with various circumstances. During the daytime, the unobstructed light sensor is highly effective. According to our measurement results with 2,136 mobile phone models, the average intensity of outdoor light on cloudy days is 1,648 lux compared with the 587 lux of indoor light. The 5th percentile outdoor light intensity is 1,031 lux compared with the 95th percentile of indoor light intensity at 847 lux. The cellular sensor provides general indication when the mobile user transits between indoors and outdoors, which however may be compromised when the user continuously sojourns around the semi-outdoor environment. The magnetometer generally performs well with steel structured buildings with abundant electrical appliances, but may fail with indoor spaces lacking those materials. Our experimental statistics with typical shopping malls show 93.3% accuracy in detecting such environment when the mobile phone is nearby the steel or electrical appliances. Apart from the sensors currently used in the IODetector, we also considered in the design stage other smartphone embedded sensors including the WiFi interface, barometer, etc, which were not included mainly due to the low utility, e.g., nowadays WiFi APs have been widely deployed indoors and outdoors leading to indistinguishable WiFi signals.

**GNSS module.** The GNSS based sensor in the IODetector significantly improves the detection accuracy as compared with the initial design of the IODetector. The rationale is that due to the signal blockage of walls, the indoor SNR and visible satellite number SN are much lower. According to the comparative evaluation, the inclusion of the GNSS module improves the detection accuracy by 13%. On the other hand, the GNSS module introduces an extra signal sensing delay of 5~20 seconds during indoor/outdoor transitions, in particular for outdoor-to-indoor transitions. For applications with strong low latency requirements, the GNSS based sensor may have to be switched off.

**Diversity.** The diversity in phone models, usage styles, and environment is high, and contributes to the deviation of the end performance. Based on our analysis of the data from over 3 million phones of 4,861 models, the IODetector detection accuracy may vary from 10.2% to 99.9% with an average of 87.0% and a standard deviation of 13.4%. The IODetector's performance from 1,398 couriers over one month (without changing their phones) shows that the standard deviation of the detection accuracy from the same courier may vary from 2.7% to 6.5%, mainly due to the varied phone placement across time. The IODetector's accuracies across the three major cities in China, namely Beijing, Shanghai, and Shenzhen are 86.5%, 87.3%, and 88.4%, which shows the impact of geographical and environmental diversity. In our deployment, we use adaptive parameter settings across phone models and city environment to mitigate the deviations.

**Groundtruth labeling.** As mentioned in Section 3.2, evaluating the IODetector in real world faces the challenge of collecting large scale groundtruth. Throughout the deployment, we planned a variety of practical methods in groundtruth labeling, assessed the effectiveness of those methods, and adopted the effective and efficient ones among them. We detail our considerations and experience in Section 4.2.

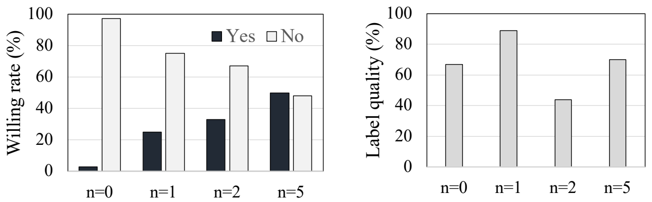
**Corner cases.** The IODetector performance is subject to varied operating parameters and real world usage diversity. In the deployment, we spotted prevalent corner cases where the detection accuracy is much lower than the expected which limit the IODetector performance, especially among the low performance cases. We detail their impact on the application goal as well as some of our efforts in combating them in Section 4.3.

### 4.2 Groundtruth Labeling

**Labeling methodology.** Other than manual groundtruth collection, we leveraged two different ways to obtain groundtruth labels during the deployment. As Section 3.2 describes, the *vision-assisted labeling* provides full recordings of courier movement trajectories and derives very accurate groundtruth labels from them. The cost of *vision-assisted labeling* however is too high for large scale evaluation. In addition to the hardware cost and overhead of mounting the hardware in the daily activities, we engaged external contractors to annotate the collected ~420 hour video footage, which costed over 2 weeks and 100,000 RMB. It is hard to apply to massive evaluation. The *inferred labels*, despite of being not as accurate as the vision-assisted labels or not complete in terms of the coverage in

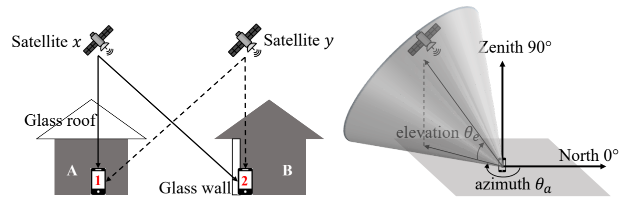
[3]In practice, the penalty depends on the actual delay in the event of late delivery. We use a median value of 3 RMB based on statistics.

[4]The penalty for customer complaints is normally higher than that of negative ratings, but we treat them equally for simplicity.




(a) Willing rate v.s. incentives. (b) Incentives v.s. label quality.

**Figure 13: Study of crowdsourcing the groundtruth labels.**




(a) Corner case examples. (b) Satellite direction.

**Figure 14: Improving GNSS performance in corner cases.**

space, provides us large-scale observation, with which we were able to evaluate the IODetector's performance in a vast number of areas of interest. The BLE beacon infrastructure used to infer indoor labels was deployed during Jan. 2018 to Apr. 2020. It covers ~34% of the registered restaurants in Shanghai and much lower in other cities. The deployment cost of each beacon is ~9.5 US dollars, including 8 dollars for hardware, 1 dollar for transportation and 0.5 dollars for deployment. About 160 million outdoor labels from ~53,000 couriers and ∼10 million indoor labels from ~10,000 couriers in Shanghai can be inferred throughout a day. The vast majority of the groundtruth labels for nationwide performance evaluation are inferred labels.

The label inference methodology is not built into the IODetector design itself because the IODetector is supposed to be solely based on the food couriers' mobile phones for general application. The indoor label inference depends on the availability of the BLE infrastructure and is thus not generally applicable. The outdoor label inference is based on moving speed which is only available when a courier rides the e-bike and is thus not complete.

**Crowdsourcing labels.** We also examined the possibility of crowdsourcing groundtruth labels from the couriers themselves. We interviewed 82 couriers to understand their willingness in contributing to labeling their own indoor/outdoor statuses. Figure 13(a) summarizes the willing rate with regard the incentives (n RMB). It is shown that the willingness increases with the provided incentives, and the survey projects to a monthly cost of at least 150,000 RMB to gain 10~20% labels from the test data. We launched a mini trial with 4 recruited food couriers to label the data with different incentives (n = 0, 1, 2, and 5 RMB). Figure 13(b) plots the quality of the crowdsourced labels when verified with the *vision-assisted labels*. Interestingly, we see no correlation between the label quality and the amount of incentives paid. The highest quality (89%) is obtained from the labels from a courier receiving 1 RMB incentive per label, and the lowest quality (44%) is obtained from the courier with 2 RMB incentive per label. The overall label quality is not adequate to reflect the groundtruth. Their feedback is not accurate/reliable, and also incurs extra manpower overhead from couriers, which is not preferred in Ele.me business. According to the cost and quality analysis, we finally concluded that crowdsourcing groundtruth labels from the couriers is not feasible and dropped the plan.

**Future research.** We expect future studies on how traditional crowdsourcing solutions may be complemented with unsupervised or self-supervised learning technology [35] to deal with the scale of unlabeled data that we face.

## 4.3 Corner Cases

The prevalent corner cases, mainly due to real world diversity as discussed in Section 4.1, lead to performance degradation. As Figure 7 and Figure 10 in Section 3 suggest, the low-performance cases of the IODetector constitute a long tail of its overall accuracy. The 10th percentile accuracies in Phase II and III drop to 73.4% and 79.2%, respectively. We make the following attempts to improve the corner case performance.

We tackle the corner cases with miss interpreted GNSS signals. Figure 14(a) illustrates two types of such corner cases when the building has a glass roof (Building A) or glass wall (Building B). Although the phone stays indoors, it may still receive high SNR satellite signals (e.g., satellite $x$). To address such corner cases, we develop a solution that takes into consideration the directions of satellite signals. As Figure 14(a) illustrates, the mobile phone may receive high SNR signals from one direction (e.g, satellite $x$), but due to concrete blockage only low SNR signals from other directions (e.g., satellite $y$). We estimate whether a majority of satellite signals come from similar directions, e.g., within a cone as suggested in Figure 14(b), which highly suggests indoor phone placement but with signal penetration from glasses or similar circumstances. As Figure 14(b) illustrates, the azimuth $\theta_a$ and elevation $\theta_e$ are used to estimate the signal direction. We consider the maximum and minimum value of the azimuth $\theta_a$ and elevation $\theta_e$ of satellite signals. If $\theta_a^{max} - \theta_a^{min} > \theta_1$ and $\theta_e^{max} - \theta_e^{min} > \theta_2$, the satellite module outputs outdoor with higher confidence and less otherwise, where $\theta_1$ and $\theta_2$ are empirically set to 150° and 60°, respectively in current IODetector design. After upgrading the GNSS sensor design, the 10th percentile detection accuracy is increased by 3.1% while the overall average is reduced by 0.1%, which we believe is a worthwhile trade-off because the low-performance cases are a major limit to the IODetector utility in business.

We also tackle the corner cases when diverse and unreliable sensor readings across phone models are experienced. After the wide system deployment, we observe prevalent diversity in light sensor readings across different phone models. The distributions of light intensity measurements from different mobile phone brands differ. As an example, Figure 15 plots the probability distribution function (PDF) of the measured outdoor light intensity from three different phone brands (with over 500 specific phone models involved in our test) during one day operation, i.e., HONOR (with 122 phone models), HUAWEI (with 232 phone models) and OPPO (with 149 phone models), which are the top-3 phone choices of Ele.me couriers. The

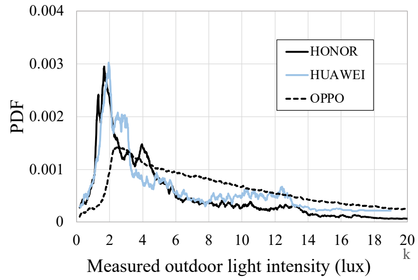
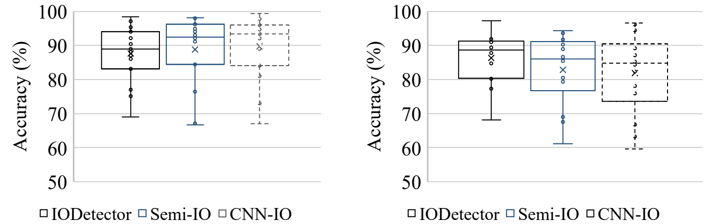



**Figure 15: Outdoor light intensity distributions measured with different phone brands.**




(a) Test with the same 3 AoIs. (b) Test with 15 different AoIs.

**Figure 16: Comparing ML-based and rule-based solutions.**

measured light intensity distribution of HONOR phones is similar to that of HUAWEI phones, but both differ from that of OPPO phones. Such diversity is prevalent, and lead to underperforming cases for certain phone models. Based on the statistics, we re-assign differently fine tuned light sensor thresholds for the top 5 mobile phone brands, which account for over 90% of the phone models used by Ele.me couriers. We also identify a few phone models which give unreliable light sensor readings or experience frequent sensor failures, and tune the IODetector design to trust more on the cellular and GNSS sensors for those phone models. After the system upgrade, the achieved 10th percentile accuracy and the overall average improve by 2.9% and 0.2%, respectively.

**Future research.** We expect future studies that deal with the trade-off between the worst cases and overall performance for better technology adoption in business.

## 4.4 Rule-based v.s. ML-based

Machine learning (ML)-based sensing techniques [6, 20, 22] have been popular for a wide range of applications in recent years. There are rule-based solutions [8, 33, 37] as well as ML-based methods [30, 31] proposed for indoor/outdoor detection. We carefully investigated the two technology threads in our feasibility study. We implemented and compared three solutions, i.e., the initial IODetector, Semi-IO which is an ML-based design reproduced from [30], and CNN-IO built with CNN-based indoor/outdoor classification. Semi-IO and CNN-IO are trained with 80% of the vision-based labels collected from 3 AoIs in Shanghai (see Section 3.2). Our major observation is that while ML-based solutions may arrive at the high average performance if sufficiently trained, their performance deviation can be much higher and they may be very sensitive to the difference between the training and testing environments.

Figure 16(a) depicts the detection accuracies of the three solutions when tested with the same 3 AoIs for training the ML-based solutions. We see CNN-IO achieves the highest average accuracy of 91.6%, while the IODetector achieves the lowest average accuracy of 88.2%. Nevertheless, the worst cases of both CNN-IO and Semi-IO are lower than the IODetector. Figure 16(b) further depicts the results when the three solutions are tested with 15 different AoIs in Shanghai. While the IODetector's accuracy maintains at 86.3%, the accuracies of Semi-IO and CNN-IO drop to 83.8% and 81.4%, respectively. The 10th percentile accuracies of Semi-IO and CNN-IO significantly drop to 68.9% and 66.8%. We see rule-based solutions achieve better reliability and generality, which are the major considerations for large scale deployment. Thus the rule-based design is employed in our final adoption. The underperformance of ML-based solutions we believe is primarily due to the limitation in obtaining sufficient high-quality training data to cope with the environmental diversity.

**Future research.** We expect future studies on how to balance the reliability, generalizability and accuracy when designing ML-based solutions for general or large scale mobile sensing applications.

# 5 RELATED WORK

**Real world mobile system experience.** Many recent research works [7, 10, 21, 25, 28, 32] report the experience of large scale system deployment. Ding *et al.* [14] introduce the experience of deploying a virtual arrival detection system for on-demand food delivery. Li *et al.* [25] report the lessons of developing and using MobileInsight for mobile network analytics. Sevilla *et al.* [32] present the experience of deploying CoLTE for small-scale, community owned and operated LTE networks. Boateng *et al.* [10] describe the experience of implementing a wrist-worn computing device for mobile health apps. Optasia [28] is a data flow system for processing queries on video feeds from many cameras. MONROE [7] is an open access platform for experimenting on operational mobile networks. Kim *et al.* [21] introduce the experience of deploying smart market applications in an urban large-scale market. To the best of our knowledge, this paper is unique in its focus on ambient mobile sensing and its deployment scale that concerns over 3 million phones across 367 cities in a country.

**Indoor/outdoor detection.** There are two categories of indoor outdoor detection solutions in the literature, i.e., general purpose approaches [30, 37] for all scenarios and dedicated approaches [8, 11, 17, 27, 33] for specific applications or with extra hardware/resource in addition to COTS mobile phones. The first and most impactful work on indoor/outdoor detection is the IODetector [37], based on which we design the basic architecture of our tool used in the deployment. Subsequent research works propose alternatives to IODetector. Radu *et al.* [30] make indoor/outdoor detection with semi-supervised learning for adaptive parameter selection. Liu *et al.* [27] propose to combine light sensor and cellular module for indoor/outdoor detection. Sung *et al.* [33] propose sound-based indoor/outdoor detection to achieve seamless handover of positioning systems. Canovas *et al.* [11] use Wi-Fi signal with AdaBoost to infer the indoor/outdoor condition. SenseIO [8] infers the environment type (e.g., rural and urban) leveraging the sensor-rich smartphones. Esmaeili Kelishomi *et al.* [17] detect the indoor/outdoor environment according to different user activities. To our knowledge, none of above works report results from large scale experimentation. In this paper, we compare different design alternatives and choose to adopt a rule-based solution for practicability and generalizability considerations.

**On-demand food delivery.** While the on-demand food delivery has been a booming business that has mainly been driven in the industry, there are also many research works to understand its characteristics and improve its efficiency. Li *et al.* [23] present a technology review on positive and negative impacts of the on-demand food delivery. Zhu *et al.* [38] propose a DNN-based model to predict critical time points in the order servicing cycle. Ding *et al.* [12] develop a delivery scope generation framework to draw suitable delivery scopes for millions of restaurant partners.

## 6 CONCLUSION

This paper presents our experience in adopting indoor outdoor detection to support a real world business of on-demand food delivery. We share our deployment experience, lessons learned on real world know-hows and practical considerations of the technology adoption, based on which we also discuss possible future research topics.

## ACKNOWLEDGMENTS

We sincerely thank the anonymous shepherd and reviewers for the insightful comments and suggestions to improve this paper. This research is supported by the National Research Foundation Singapore under its Industry Alignment Fund – Pre-positioning (IAF-PP) Funding Initiative, Alibaba Group through Alibaba Innovative Research (AIR) Program and Alibaba-NTU Singapore Joint Research Institute (JRI), and Ministry of Education Singapore MOE AcRF Tier 2 MOE-T2EP20220-0004. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of National Research Foundation Singapore and other funding agencies.

## REFERENCES


[1] 2021. Beidou Navigation Satellite System. http://en.beidou.gov.cn/. [Online; accessed 19-Aug-2021].

[2] 2021. Ele.me courier privacy policy. https://logisticsapp.ele.me/static/zb-h5/dist/protocol.html#/?type=secret&opt_type=0. [Online; accessed 19-Aug-2021].

[3] 2021. Galileo Positioning System. https://www.gsc-europa.eu/. [Online; accessed 19-Aug-2021].

[4] 2021. GLObal NAvigation Satellite System. https://www.glonass-iac.ru/en/. [Online; accessed 19-Aug-2021].

[5] 2021. The Global Positioning System. https://www.gps.gov/systems/gps/. [Online; accessed 19-Aug-2021].

[6] Mohamed S. Abdelfattah, undefinedukasz Dudziak, Thomas Chau, Royson Lee, Hyeji Kim, and Nicholas D. Lane. 2020. Best of Both Worlds: AutoML Codesign of a CNN and Its Hardware Accelerator. In *Proceedings of the 57th ACM/EDAC/IEEE Design Automation Conference (DAC '20)*. Article 192, 6 pages.

[7] Özgü Alay, Andra Lutu, Miguel Peón-Quirós, Vincenzo Mancuso, Thomas Hirsch, Kristian Evensen, Audun Hansen, Stefan Alfredsson, Jonas Karlsson, Anna Brunstrom, Ali Safari Khatouni, Marco Mellia, and Marco Ajmone Marsan. 2017. Experience: An Open Platform for Experimentation with Commercial Mobile Broadband Networks. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking (MobiCom '17)*. 70–78.

[8] Mohsen Ali, Tamer ElBatt, and Moustafa Youssef. 2018. SenseIO: Realistic Ubiquitous Indoor Outdoor Detection System Using Smartphones. *IEEE Sensors Journal* 18, 9 (2018), 3684–3693.

[9] Roshan Ayyalasomayajula, Aditya Arun, Chenfeng Wu, Sanatan Sharma, Abhishek Rajkumar Sethi, Deepak Vasisht, and Dinesh Bharadia. 2020. *Deep Learning Based Wireless Localization for Indoor Navigation*.

[10] George Boateng, Vivian Genaro Motti, Varun Mishra, John A. Batsis, Josiah Hester, and David Kotz. 2019. Experience: Design, Development and Evaluation of a Wearable Device for MHealth Applications. In *The 25th Annual International Conference on Mobile Computing and Networking (MobiCom '19)*. Article 31, 14 pages.

[11] Oscar Canovas, Pedro E. Lopez-de Teruel, and Alberto Ruiz. 2017. Detecting Indoor/Outdoor Places Using WiFi Signals and AdaBoost. *IEEE Sensors Journal* 17, 5 (2017), 1443–1453.

[12] Xuetao Ding, Runfeng Zhang, Zhen Mao, Ke Xing, Fangxiao Du, Xingyu Liu, Guoxing Wei, Feifan Yin, Renqing He, and Zhizhao Sun. 2020. Delivery Scope: A New Way of Restaurant Retrieval for On-Demand Food Delivery Service. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery Data Mining (KDD '20)*. 3026–3034.

[13] Yi Ding, Ling Liu, Yu Yang, Yunhuai Liu, Desheng Zhang, and Tian He. 2021. From Conception to Retirement: a Lifetime Story of a 3-Year-Old Wireless Beacon System in the Wild. In *18th USENIX Symposium on Networked Systems Design and Implementation (NSDI 21)*. 859–872.

[14] Yi Ding, Yu Yang, Wenchao Jiang, Yunhuai Liu, Tian He, and Desheng Zhang. 2021. Nationwide Deployment and Operation of a Virtual Arrival Detection System in the Wild. In *SIGCOMM '21: Proceedings of the 2021 Annual conference of the ACM Special Interest Group on Data Communication on the applications, technologies, architectures, and protocols for computer communication (Sigcomm '21)*.

[15] Ele.me. 2021. Fengniao delivery. https://fengniao.ele.me/. [Online; accessed 19-Aug-2021].

[16] Salma Elmalaki, Lucas Wanner, and Mani Srivastava. 2015. CAreDroid: Adaptation Framework for Android Context-Aware Applications. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*. 386–399.

[17] Aghil Esmaeili Kelishomi, AHS Garmabaki, Mahdi Bahaghighat, and Jianmin Dong. 2019. Mobile User Indoor-Outdoor Detection Through Physical Daily Activities. *Sensors (Basel)* 19 (2019).

[18] Google. 2021. Profile battery usage with Batterystats and Battery Historian. https://developer.android.com/topic/performance/power/setup-battery-historian. [Online; accessed 19-Nov-2021].

[19] Alibaba Group. 2021. Fiscal Year 2020 Annual Report. https://doc.irasia.com/listco/hk/alibabagroup/annual/2020/ar2020.pdf. [Online; accessed 19-Aug-2021].

[20] Sojeong Ha and Seungjin Choi. 2016. Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors. In *2016 International Joint Conference on Neural Networks (IJCNN)*. 381–388.

[21] Hyung-Sin Kim, JeongGil Ko, and Saewoong Bahk. 2017. Smarter Markets for Smarter Life: Applications, Challenges, and Deployment Experiences. *IEEE Communications Magazine* 55, 5 (2017), 34–41.

[22] Jack Kosaian, Amar Phanishayee, Matthai Philipose, Debadeepta Dey, and Rashmi Vinayak. 2021. Boosting the Throughput and Accelerator Utilization of Specialized CNN Inference Beyond Increasing Batch Size. In *2021 International Conference on Machine Learning (ICML 2021)*.

[23] Charlene Li, Miranda Mirosa, and Phil Bremer. 2020. Review of Online Food Delivery Platforms and their Impacts on Sustainability. *Sustainability* 12, 14 (2020).

[24] Mo Li and Yunhao Liu. 2009. Underground Coal Mine Monitoring with Wireless Sensor Networks. *ACM Trans. Sen. Netw.* 5, 2, Article 10 (April 2009), 29 pages.

[25] Yuanjie Li, Chunyi Peng, Zhehui Zhang, Zhaowei Tan, Haotian Deng, Jinghao Zhao, Qianru Li, Yunqi Guo, Kai Ling, Boyan Ding, Hewu Li, and Songwu Lu. 2021. Experience: A Five-Year Retrospective of MobileInsight. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking (MobiCom '21)*. 28–41.

[26] Yang Liu, Zhenjiang Li, Zhidan Liu, and Kaishun Wu. 2019. Real-Time Arm Skeleton Tracking and Gesture Inference Tolerant to Missing Wearable Sensors. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '19)*. 287–299.

[27] Zhongwei Liu, HyunCheol Park, Zili Chen, and Hosik Cho. 2015. An Energy-Efficient and Robust Indoor-Outdoor Detection Method Based on Cell Identity Map. *Procedia Computer Science* 56 (2015), 189–195.

[28] Yao Lu, Aakanksha Chowdhery, and Srikanth Kandula. 2016. Optasia: A Relational Platform for Efficient Large-Scale Video Analytics. In *Proceedings of the Seventh ACM Symposium on Cloud Computing (SoCC '16)*. 57–70.

[29] Xiaomin Ouyang, Zhiyuan Xie, Jiayu Zhou, Jianwei Huang, and Guoliang Xing. 2021. ClusterFL: A Similarity-Aware Federated Learning System for Human

Activity Recognition. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '21)*. 54–66.

[30] Valentin Radu, Panagiota Katsikouli, Rik Sarkar, and Mahesh K. Marina. 2014. A Semi-Supervised Learning Approach for Robust Indoor-Outdoor Detection with Smartphones. In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems (SenSys '14)*. 280–294.

[31] Illyyne Saffar, Marie Line Alberi Morel, Kamal Deep Singh, and Cesar Viho. 2019. Semi-Supervised Deep Learning-Based Methods for Indoor Outdoor Detection. In *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*. 1–7.

[32] Spencer Sevilla, Matthew Johnson, Pat Kosakanchit, Jenny Liang, and Kurtis Heimerl. 2019. Experiences: Design, Implementation, and Deployment of CoLTE, a Community LTE Solution. In *The 25th Annual International Conference on Mobile Computing and Networking (MobiCom '19)*. Article 45, 16 pages.

[33] Rakmin Sung, Suk-hoon Jung, and Dongsoo Han. 2015. Sound based indoor and outdoor environment detection for seamless positioning handover. *ICT Express* 1 (2015), 106–109. Issue 3.

[34] Yanwen Wang, Jiaxing Shen, and Yuanqing Zheng. 2020. Push the Limit of Acoustic Gesture Recognition. In *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*. 566–575.

[35] Huatao Xu, Pengfei Zhou, Rui Tan, Mo Li, and Guobin Shen. 2021. LIMU-BERT: Unleashing the Potential of Unlabeled Data for IMU Sensing Applications. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems (SenSys '21)*. 220–233.

[36] Pengfei Zhou, Yuanqing Zheng, and Mo Li. 2012. How Long to Wait? Predicting Bus Arrival Time with Mobile Phone Based Participatory Sensing. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services* (Low Wood Bay, Lake District, UK) *(MobiSys '12)*. 379–392.

[37] Pengfei Zhou, Yuanqing Zheng, Zhenjiang Li, Mo Li, and Guobin Shen. 2012. IODetector: A Generic Service for Indoor Outdoor Detection. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems (SenSys '12)*. 113–126.

[38] Lin Zhu, Wei Yu, Kairong Zhou, Xing Wang, Wenxing Feng, Pengyu Wang, Ning Chen, and Pei Lee. 2020. Order Fulfillment Cycle Time Estimation for On-Demand Food Delivery. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery  Data Mining (KDD '20)*. 2571–2580.