From Conception to Retirement: a Lifetime Story of a 3-Year-Old Wireless Beacon System in the Wild

NSDI 21
Problem Statement

• Design and deploy a system for an Instant delivery platform that is capable to infer its couriers’ indoor status.

• Examples of Instant delivery platform: Doordash, uberEats, Alibaba Local Services.

Most annoying behavior of delivery apps?

Basic Flow

1) Order
2) Notify
3) Assign
4) Pick up
How to prevent apps from breaking promised ETA?

• Better ETA and tracking. Of course!
• Cost associated with overdue – Typically $1. If customer has insurance -> 200% x (order value) 😃
• Efficient scheduling algorithms that can assign the right courier to the task.
• In places like Shanghai, merchants are located in multistoried malls. C can be 33% of the total time. BIG CHUNK.
• How to find C?
Arrival and Departure detection

• Most basic problem ever. Extensive research exists in this domain
• BUT, it’s usually done in a controlled environment.
• Small scale or private environment. Not “in the wild”.
• Existing solutions: Manual, WiFi, RFID, LED fixtures.

1. Continuous scanning required – Power consumption
2. Access points are expensive for merchants still in the stone age

Additional equipment at both ends – Price issues.

Doesn’t scale as hardware modification is required.

Unreliable – Intentional, unintentional QR code scanning.
BLE Beacon (Preferred Solution)

• BLE – Bluetooth Low Energy
• Easy to **deploy** (still an issue though)
• Acceptable cost ($10)
• Transparent to couriers

• Continuous scanning – 2% extra power consumptions – way less than wifi
• Only at merchant end unlike RFID
• Battery powered small devices that are portable unlike LED fixtures
aBeacon System

• Commissioned by Alibaba. Experiment took place in Shanghai.
• Uses customized BLE devices.
• Some numbers:
  • Experiment length – 3 years.
  • Total number of merchants - 12109
  • Total number of couriers - 109K
  • 64 million delivery orders for 7.3 million customers!!
  • $600 million in order values.
aBeacon Architecture

Server uses data to crunch numbers – ex: run scheduling algorithms
Justifying the cost of the experiment

- A metric based approach is used for this; to understand the cost-performance tradeoff
- Metrics –
  - Cost: Device cost and deployment cost
  - Lifetime: Battery – Estimate (2yrs)
  - Reliability: Failure rates
  - Utility: Reduction in overdue delivery rate and hence overdue cost
Justifying the cost of the experiment

- **Gain Equation:**

\[
G_T = \sum_{t=1}^{T} \sum_{i=1}^{N_t} B_t^i - C_T
\]

\[
C_T = N_T \cdot C_{Dev}
\]

\[
B_t^i = F_i(P_{Life}, t, t_0)
\]

\[
F_1(O_t^i, P_{Reli}, P_{Util}, C_{Dev})
\]

- **The real benefit:** How much cost saved.

- **Costs and variables:***
  
  - \(C_{Dev}\): cost of a device, i.e., hardware & deployment
  - \(C_{Over}\): cost of overdue penalty per order, e.g., $1.
  - \(P_{Life}^i\): lifetime of a device \(i\)
  - \(P_{Reli}^i\): reliability of \(i\)
  - \(P_{Util}^i\): utility of \(i\)
  - \(t_0^i\): day of \(i\) was deployed
  - \(T\): # of days since aBeacon deployed
  - \(N_t\): # of deployed devices until the \(t\)th day
  - \(O_t^i\): # of orders at \(t\)th day in the merchant with \(i\)

- **Calculation:**

\[y = P_{Life}^i - (t-t_0) > 0? 1:0;\]
Phase 1 – Pilot
Conception stage:
3 months
18 merchants, 3 bands = 54 devices
Average reliability - 98%

Cost looks high. Should try to reduce it. Each dollar counts in a huge experiment like this.
**Deployment Timeline**

- **Phase 1** – Last 5 months
  - BLE module with low power consumption
  - Alkaline battery proffered over lithium since less mAh/$
  - Casing for future proof – Outdoors
  - A/B testing with custom and commodity sensor
  - 200 units in 200 stores. Similar reliability and lower cost 😊

**Customized beacon devices:**
- Less cost ($8 each)
- Longer lifetime (≥2 years)
Deployment Timeline

- **Stage 1:** Conception
- **Stage 2:** Customization
- **Stage 3:** Deployment
- **Stage 4:** Operation

**2017**
- Server Issues

**2018**
- Spring Festival

**2019**
- Battery Runout (Retirement)

**2020**
- Spring Festival & COVID-19

**Spatial Heatmap**

- 2018/01: 75 km
  - SHA
  - PVG
  - (a) 2 Weeks into Deployment

- 2018/03: SHA
  - PVG
  - (b) Fully Operational

- 2019/09: SHA
  - PVG
  - (c) 2 Months Before Battery Runout

- 2020/03: SHA
  - PVG
  - (d) Retired (Non-Operational)
Impact of area:

Densely populated areas have more utility.

Lesson:

More orders might not imply more utility, it’s the uncertainty of courier behaviors that lets aBeacon shine => more utility.

Impact of Floor

Observation: aBeacon is more beneficial in higher floors and basements

Impact of area: Densely populated areas have more utility.

Lesson: More orders might not imply more utility, it’s the uncertainty of courier behaviors that lets aBeacon shine => more utility.
Performance: Reliability
(how many arrival events can be detected among all events?)

Impact of Staying Duration

Impact of Smartphone Hardware

<table>
<thead>
<tr>
<th>Android Reliability</th>
<th>iOS Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>75%</td>
<td>85%</td>
</tr>
</tbody>
</table>
Impact of device placement

Performance: Reliability

*(how many arrival events can be detected among all events?)*
Performance: Reliability
(how many arrival events can be detected among all events?)

Lesson Learned: Reliability in the Wild
Even for arrival detection, the reliability is far from guaranteed in the wild due to multiple factors.
Performance: Lifetime (the lifetime of each device)

Lesson Learned: Lifetime in the Wild
- Battery may NOT be the major constraint for mobile/wireless devices, since 40% devices survive longer than the environment (i.e., shops).
# Implication for Building Industrial Systems

<table>
<thead>
<tr>
<th>Lessons</th>
<th>System Evolution</th>
<th>Reliability</th>
<th>Lifetime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Physical devices fail earlier than expected.</td>
<td>Wireless beacon devices are NOT reliable (for regulation).</td>
<td>Device lifetime are significantly affected by the environment.</td>
</tr>
</tbody>
</table>
Additional Application of aBeacons

• Order delivery time estimation
• Merchant location correction
• Anomaly detection
Next Generation of aBeacon: aBeacon+

Alibaba Cloud

(0) Deploying & Binding

(1) Advertising

ID Tuple

(2) Uploading

(3) Mapping

Merchant

Shop

Couriers

Courier

- No hardware or deployment cost.
- No lifetime worries.
- No battery worries.
- Hybrid solution.
Thank You
- Metric based evaluation of networked systems to determine
gain is a takeaway

- Why does adacomp require batch rollouts?