Container Throughput Estimation Leveraging Ship GPS Traces and Open Data

Longbiao CHEN, Daqing ZHANG, Gang PAN, Leye WANG, Xiaojuan MA, Chao CHEN, Shijian LI

Zhejiang University, China
Institut Mines-TELECOM/TELECOM SudParis, France
Huawei Noah’s Ark Lab, Hong Kong, China

UbiComp 2014 09-17-2014 Seattle, WA
Outline

I. Introduction
II. Related Work

III. Berthing Event Identification
IV. Berthing Event Container Throughput Estimation

V. Evaluation
VI. Conclusion
Outline

I. Introduction
II. Related Work
III. Berthing Event Identification
IV. Berthing Event Container Throughput Estimation
V. Evaluation
VI. Conclusion
Background

- **Container throughput of a sea port (e.g. Seattle)**
  - number of containers handled at berth
  - include: input + output

- **A crucial measurement**
  - port performance
  - regional economy
I. INTRODUCTION

- Container throughput of a sea port (e.g. Seattle) is
  - number of containers handled at terminal
  - include: input + output

- A crucial measurement of
  - port performance
  - regional economy

![Container throughput chart](chart.png)
**Motivation**

- **Real-time accurate** container throughput data is the key to the success of
  - port planning
  - liner optimization

- However, existing methods are either
  - **Delayed** (1-2 months, port authority statistics)
  - **Inaccurate** (10% error, history based estimations)
Motivation (cont.)

- Ubiquitous Maritime Data
  - AIS (Automatic Identification System) data: ship traces in port
    - e.g. ship ID, GPS position, heading direction, etc.
  - Maritime open data: ship and port attributes
    - e.g. ship dimensions, port facility parameters, etc.
Objectives and Challenges

- We attempt to leverage ship traces and open data to enable real-time accurate container throughput estimation.

1. How many ships arrive at the container berth?
   - Identify from ship traces

2. How many containers are handled for each ship?
   - Estimate using ship and port open data
Contributions

1. First work
2. Two phase approach
   - Berthing Event Identification
   - Berthing Event Container Throughput Estimation
3. Evaluation with real-world data
   - Hong Kong and Singapore, 2011
   - 1% error rate
Outline

I. Introduction
II. Related Work
III. Berthing Event Identification
IV. Berthing Event Container Throughput Estimation
V. Evaluation
VI. Conclusion
Related Work

- Event Identification from GPS traces
  - Human mobility \((Amini, 2012)\)
  - Taxi operation \((Zhang, 2011)\)

- Container Throughput Estimation
  - Historical throughput data based \((Peng, 2009)\)
  - Economic factors (e.g. GDP) based \((Seabrooke, 2003)\)

- No known work on throughput estimation leveraging GPS traces and open data
Outline

I. Introduction
II. Related Work

III. Phase 1: Berthing Event Identification
IV. Phase 2: Berthing Event Container Throughput Estimation

V. Evaluation
VI. Conclusion
We attempt to identify **berthing event** (Event B) from ship traces.

- (A – anchoring event; C – temporary stop event)
We attempt to identify berthing event (Event B) from ship traces.

- (A – anchoring event; C – temporary stop event)
Berthing Event Identification Method

1. Extract stop events: sliding-window
2. Find densely-clustered events: DBSCAN
Outline

I. Introduction
II. Related Work

III. Phase 1: Berthing Event Identification
IV. Phase 2: Berthing Event Container Throughput Estimation

V. Evaluation
VI. Conclusion
Berthing Event Throughput Estimation

- We attempt to estimate how many containers are handled for each berthing event.
- Containers are handled using Quay Cranes.
Berthing Event Throughput Estimation

- We attempt to estimate how many containers are handled for each berthing event.
- Containers are handled using Quay Cranes.
Berthing Event Throughput Estimation

1. A ship may be assigned several quay cranes.
2. A quay crane transships containers in row.
Berthing Event Throughput Estimation

Three factors affecting the throughput ($\pi_i$):

1. Quay Crane Number ($N_i$): related to ship length
2. Quay Crane Efficiency ($E_i$): related to ship width
3. Container Handling Time ($T_i$): unproductive time

\[
\pi_i = N_i \cdot E_i \cdot T_i
\]
\[
= \left[ \frac{l_i}{A} \right] \cdot \frac{V}{W + b_i} \cdot (T_i - \Delta T_1 - \Delta T_2)
\]
Outline

I. Introduction
II. Related Work

III. Berthing Event Identification
IV. Container Throughput Estimation

V. Evaluation
VI. Conclusion
Datasets Description

1. **Ship GPS traces** collected by AIS stations
   - Two ports: Hong Kong and Singapore
   - One year: 2011
   - Sampling interval: 3min
   - 50,000 ship traces

2. **Container ship database** from marinetrack.com
   - 4,896 container ships
   - Ship ID, length, breadth, capacity, etc.

3. **Port statistics** from port authority websites
   - Two ports: Hong Kong and Singapore
   - Number of ship arrival at berth (monthly, 2011)
   - Container throughput data (monthly, 2011)
#1. Performance of Berthing Event Identification

- Ground truth: published number of ship arrivals
- Month by month comparison

Graphs showing the number of ships for Hong Kong and Singapore, comparing ground truth and identified data for each month from January to December.
#2. Performance of Container Throughput Estimation

- Ground truth: published port throughput data
- Baseline: Berthing-time-based method
  - Assume: throughput proportional to berthing time
    \[ \pi_i = \frac{T_i}{T_{\text{max}}} \cdot 2C_i \]
- Evaluation metric: Mean Absolute Percentage Error (MAPE)

\[
MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|
\]
Evaluating Port Container Throughput

- Evaluation Results
  - Hong Kong and Singapore, 6 months, 2011

<table>
<thead>
<tr>
<th></th>
<th>Hong Kong</th>
<th>Singapore</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>1.0%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Proposed</td>
<td>8.2%</td>
<td>7.1%</td>
</tr>
</tbody>
</table>

V. EVALUATION
Outline

I. Introduction
II. Related Work
III. Berthing Event Identification
IV. Container Throughput Estimation
V. Evaluation
VI. Conclusion
Conclusions

- We proposed a real-time accurate container throughput estimation method, leveraging ship AIS traces and maritime open data.

- We proposed a two-phase approach
  - First identify berthing events from ship traces
  - Then estimate container throughput of each berthing event using ship and port open data

- We evaluated our approach using real-world data from two ports, and achieve accurate results.
Thank you!
And Questions 😊
Answers to Questions
Details of Berthing Event Identification

- Stationary and Stable
  - Position is fixed
  - Heading direction keeps unchanged
- BEs are densely located near the berth

\[ \begin{align*}
\delta_p &= \ldots \\
p_1 &= \ldots \\
p_2 &= \ldots \\
p_3 &= \ldots \\
p_4 &= \ldots \\
p_5 &= \ldots \\
p_6 &= \ldots \\
p_7 &= \ldots 
\end{align*} \]
Why not use digital berth maps?

- This is considered in our future work
Quay Crane Number Estimation

- QC number is related to **ship length**: the longer a ship is, the more QCs can be assigned to it. *(Steenken, 2004)*

\[ N = \left\lfloor \frac{l}{A} \right\rfloor \]

- \(L_i\): Ship length
- \(A\): average safe distance between QCs
Quay Crane Efficiency Estimation

- Number of containers moved per hour by the QC
  - The wider the ship, the longer distance a trolley has to travel, thus the longer time it takes to move a container on average.

\[ t = \frac{W + b}{V} \]

\[ E = \frac{1}{t} = \frac{V}{W + b} \]

- \( W \): QC span width
- \( b \): ship breadth
- \( V \): trolley speed
The overall berthing time is $T'$.

There are two kinds of unproductive time (Steenken, 2004):

- QC Ready Time ($\Delta T_1$): drop down, lift up, etc.
- QC Shift Time ($\Delta T_2$): shift between container rows

We obtain the container handling time $T$ as:

$$T = T' - \Delta T_1 - \Delta T_2$$
Historical Throughput Data based Method

- Hong Kong, 2011
- Classical decomposition model (*Peng, 2009*)

**Methods**

<table>
<thead>
<tr>
<th>Methods</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historical</td>
<td>6.01%</td>
</tr>
<tr>
<td>Berthing time</td>
<td>4.65%</td>
</tr>
<tr>
<td>Our method</td>
<td>1.76%</td>
</tr>
</tbody>
</table>
Generalization to Other Ports

- Two issues in generalization
  - Data quality: not enough coverage, data missing
  - Terminal facility parameters, e.g. trolley speed