Container Port Performance Measurement and Comparison Leveraging Ship GPS Traces and Maritime Open Data

Longbiao Chen, Daqing Zhang, Xiaojuan Ma, Leye Wang, Shijian Li, Zhaohui Wu, Member, IEEE, Gang Pan

Abstract—Container ports are generally measured and compared using performance indicators such as container throughput and facility productivity. Being able to measure the performance of container ports quantitatively is of great importance for researchers to design models for port operation and container logistics. Instead of relying on the manually collected statistical information from different port authorities and shipping companies, we propose to leverage the pervasive ship GPS traces and maritime open data to derive port performance indicators, including ship traffic, container throughput, berth utilization, and terminal productivity. These performance indicators are found to be directly related to the number of container ships arriving at the terminals and the number of containers handled at each ship. Therefore, we propose a framework that takes the ships’ container handling events at terminals as the basis for port performance measurement. With the inferred port performance indicators, we further compare the strengths and weaknesses of different container ports at terminal level, port level, and region level, which can potentially benefit terminal productivity improvement, liner schedule optimization, and regional economic development planning. In order to evaluate the proposed framework, we conduct extensive studies on large-scale, real-world GPS traces of container ships collected from major container ports worldwide through the year, as well as various maritime open data sources concerning ships and ports. Evaluation results confirm that the proposed framework can not only accurately estimate various port performance indicators, but also effectively produces port comparison results such as port performance ranking and port region comparison.

Index Terms—container port; GPS trace; open data; Intelligent Transportation System (ITS); urban computing

I. INTRODUCTION

CONTAINER ports have become logistics centers of international trade [1]. The competency and competitiveness of a container port is generally characterized by its operational performance, such as container throughput and facility productivity. Therefore, being able to measure and compare the port performance quantitatively is of great importance for researchers to design and optimize port operation and container logistics, such as berth allocation [2], quay crane assignment [3], ship scheduling [4], and shipping line optimization [5].

Existing work on container port performance measurement and comparison mainly rely on statistical data published by port authorities and shipping companies [5]–[8]. Researchers have to collect and process the data manually, which often involves huge human effort, and sometimes leads to sub-optimal, incomplete results due to the delay and absence of key performance indicators required by their models [9], [10]. Moreover, in the absence of industry standards of performance measurement [11], different port authorities and shipping companies often use different, proprietary measures and timescales to produce their statistical data, making it difficult to conduct comprehensive performance analysis and comparison on a global scale [12].

In this work, we introduce an automatic, low-cost, and more accessible approach to derive the key performance indicators of container ports, including ship traffic, container throughput, berth utilization, and terminal productivity [13], [14]. According to existing literature [6], [7], [15], the key performance indicators of container ports are directly related to the container handling events at terminals, especially the number of container ships arrived at terminals and the number of containers handled by terminal facilities. Therefore, we take container handling events at terminals as the basis for port performance study, and propose a framework to derive key performance indicators of container ports. More specifically, we first identify the container handling events of a ship from its GPS traces extracted from the AIS (Automatic Identification System [16]) database, including when the ship arrives at the terminal and how long it stays there. Then, we estimate the number of containers handled by terminal facilities during each container handling event, leveraging the relevant information extracted from various maritime open data sources, including data about the corresponding ship (e.g. length and capacity) and terminal facilities (e.g. berth and quay crane parameters). Finally, we select a set of commonly used key performance indicators, and aggregate all the identified container handling events into the corresponding terminals and ports in chronological order to estimate these indicators. The resulting figures allow researchers to compare the strength and weakness of different container ports in terminal level, port level, and region level, and thus enabling applications on terminal productivity improvement, liner schedule optimization,
and regional economic development planning.

In designing the framework, there are several research issues that have to be addressed:

1) It is not trivial to identify container handling events from ship GPS traces. A ship may be engaged in various kinds of events in a container port during a port call, from a temporary stay near the terminal to a long-term repair at the shipyard. These events may share similar mobility patterns with real container handling events and make it difficult to separate them accurately. Therefore, in order to effectively identify container handling events, we need to extract a set of container-handling-event-specific features from ship GPS traces, such as staying location, moving speed and heading direction.

2) It is not straightforward to estimate the number of containers handled during a container handling event. For instance, we can not simply estimate this figure using a ship's capacity. Although a ship's capacity indicates the maximum number of containers a ship can carry, the ship may not be fully loaded and it is likely that not all containers are unloaded during a container handling event. Therefore, we need to estimate the actual number of containers handled based on other operational evidence during a container handling event, such as the ship's berthing duration, the number of terminal facilities used to handle containers, and the handling speed of these facilities.

3) It is not easy to select a set of key performance indicators for systematic comparison across different ports. Since container ports vary significantly in scale and follow different operation regulations, there are no industrial standards for what to measure, how to measure, and with what metric these ports can be compared in an informative and consistent manner [11]. In this work, we select four key performance indicators commonly used in the literature, including ship traffic, container throughput, berth utilization and terminal productivity, and accurately estimate their values based on container handling events.

In summary, the main contributions of this paper include:

1) To the best of our knowledge, this is the first work on container port performance measurement and comparison leveraging ship GPS traces and maritime open data.

2) We propose a container handling event-based framework to study container port performance. First, we identify container handling events from ship GPS traces. Then, we estimate the number of containers handled during a container handling event by leveraging information about ship and terminal facilities from various maritime open data sources. Finally, we aggregate identified container handling events into terminal and port levels to estimate four key performance indicators to enable performance comparison in terminal level, port level, and region level.

3) We evaluate our framework using large-scale, real-world ship GPS traces and various maritime open data about ships and terminals. The results show that the proposed framework can not only accurately estimate various key port performance indicators, but can also enable systematic container port comparison on a global scale.

The rest of this paper is organized as follows. We begin by reviewing related work in Section II. After introducing the preliminary knowledge about container ports in Section III, we present an overview of the proposed framework in Section IV. In the three following sections, we detail the three phases of the framework one by one, i.e., container handling event identification in Section V, container handling event quantity estimation in Section VI, and key performance indicator calculation in Section VII. Extensive evaluation results are reported in Section VIII to verify the accuracy and effectiveness of the proposed framework. Finally, we conclude the paper and chart the future directions in Section IX.

II. RELATED WORK

In this section, we review the related work in three groups. The first group consists of the work on container port performance analysis, the second group focuses on exploiting GPS traces for purposes other than container port performance study, and the third group reviews the existing literature on open data applications.

A. Container Port Performance Analysis

Some researchers and organizations have proposed to measure and compare container port performance from different perspectives. For instance, UNCTAD [17] suggested two categories of port performance indicators: macro performance indicators quantifying aggregate port economic impacts on economic activity, and micro performance indicators evaluating port operation efficiency. Kemmer et al. [15] grouped most frequently used container port performance indicators into two categories, i.e. design indicators affecting the resulting design of a container port in terms of equipment choice and capacities, and service indicators providing figures about the fulfillment of demands from port customers. Esmer et al. [11] reviewed four categories of port performance indicators, i.e. production, productivity, utilization and service measurement.

Among the above-mentioned previous studies, the following key performance indicators are commonly used. (1) Ship traffic, which measures the number of ships arrived at a port during a period of time. This indicator is employed by many port authorities, including Hong Kong [18] and Singapore [19], to evaluate the regional competitiveness of the port. (2) Container throughput, which is the number of containers handled in a port during a period of time. This indicator is considered one of the most important measurement of port competency and determinants of the future development, and thus has been intensively studied [20]–[22]. (3) Facility utilization, which measures how intensively the terminal facilities are being occupied, especially for berth occupancy. This indicator is widely used to evaluate berth allocation strategies and assist berth planning [2], [23]. (4) Operation productivity, which evaluates the efficiency of terminal facilities, especially quay cranes that are used to handle containers at terminals [24]. This indicator is often a major concern of both port operators.
and shipping companies [15], [24] in port facility scheduling and shipping route planning.

Researchers have proposed various approaches to calculate the above-mentioned indicators. For instance, Peng et al. [21] estimated future container throughput from historical throughput data, whereas Seabrooke et al. [25] leveraged macro-economic conditions and regional competition to model port throughput growth. Shabayek et al. [26] proposed a framework to estimate the berth utilization of Hong Kong container terminals using statistical data and facility parameters, while Kia et al. [27] investigated facility efficiency using a predictive model with existing operation data. The JOC Group evaluates berth productivity based on the data about container handling quantity for each ship provided by shipping companies, and ranks world ports based on their berth productivity [14]. Nevertheless, most of these studies are based on statistical data published by port authorities or shipping companies, which are usually delayed, heterogeneous, and sometimes inconsistent across different ports. Therefore, the results of these studies are usually port-specific, and thus can not be extended to enable comparison across global ports. In this paper, we propose to estimate container port performance indicators in a different way, i.e., we leverage ship GPS traces and maritime open data instead of port statistics to model container handling events at terminals, and derive a set of key performance indicators that allow researchers to analyze and compare different container ports on a global scale.

B. GPS Trace Mining

In the literature, there have been many studies on mining various kinds of GPS traces for different application scenarios. On human mobility study, Zheng et al. [28] proposed a framework to mine interesting locations and travel sequences from human GPS trajectories, while Alvarens et al. [29] proposed a model to enrich human GPS trajectories with semantic geographical meanings. On taxi operation study, Zhang et al. [30] detected anomalous passenger delivery trips from taxi GPS traces, Chen et al. [31] explored citywide night bus planning issues leveraging taxi GPS traces, and Zhang et al. [32] identified taxi refueling behaviors from GPS trajectories to estimate citywide petrol consumption and analyze gas station efficiency.

As for ship GPS trace mining, Ristic et al. [33] proposed a framework to detect ship motion anomaly by extracting motion patterns from historic ship AIS data, while Miola et al. [34] presented a methodology to evaluate maritime air pollution in port by modeling ship traffic patterns from ship AIS data. Our previous work [20] proposed a method to estimate container throughput by leveraging ship GPS traces and open data. This paper is based on it but differs in the following respects. First, we propose a novel approach to accurately identify container handling events by leveraging terminal location as well as ship mobility pattern. As a result, fake events sharing similar mobility patterns to real container handling events, such as repair at shipyard, can be effectively eliminated. Second, this paper targets different objectives from [20]. With the proposed framework in this paper, we not only estimate container throughput of the port, but also calculate a set of key performance indicators. In this way, we are able to enable comprehensive analysis and comparison of different container ports. Last but not least, we conduct two case studies to evaluate the effectiveness of the proposed framework using real-world data from major world container ports.

C. Open Data Applications

Using open data available over the Internet has been a common practice in the literature [20], [35]–[37]. For instance, Bichard et al. [38] proposed the Great British Public Toilet Map, a web-based service derived from public-sector open data to help improve accessibility of public toilets, while Garbett et al. [39] presented Fearsquare, an application that meshes geo-located crime statistics with personal check-ins in Foursquare to give users a realistic representation of the crime levels of the places they frequent or visit. Sadilek et al. [40] also presented a system to automatically identify restaurants posing public health risks leveraging Twitter user posts on food-borne illness.

However, considering maritime open data, few studies have been conducted in the literature. Nevertheless, various maritime open data have been available in the industry. For instance, the Marine Cadastre Project1 of U.S. National Ocean Service have released numerous data on container ports (such as terminal layout and navigation courses), while the Marine-Traffic API2 provides comprehensive container ship information (such as ship capacity and length), as well as worldwide container port information. In this paper, we explore the potential to incorporate these heterogeneous maritime data with AIS trace data to study the performance of container terminals and ports.

III. PRELIMINARY

In this section, we provide preliminary knowledge of container ports and define several basic terms used in this paper.

A. Preliminaries of Container Port and Container Ship

As shown in Fig.1, a container port usually consists of several container terminals, as well as other marine service infrastructures, such as anchorage and ship repair yards. Usually,

---

1http://marinecadastre.gov/data/
2http://www.marinetrack.com/en/p/api-services
a container ship is loaded and unloaded at container terminals. The stone or metal platform along the waterside of a terminal for loading/unloading ships is called quay, and a ship’s allotted place at a quay is called berth.

A container ship may visit a port for various reasons, such as container handling, maintenance, repair, and taking on supplies or fuel. When a ship arrives at a port for container handling (as shown in Fig.1), it first requests a berth from the port authority, and then stays at the anchorage (Place A) waiting for the assignment. When the requested berth becomes available, the ship maneuvers to the terminal areas and moors at the berth alongside the quay (Place B). During the container handling process, in general some containers are unloaded from the ship to the shore, followed by some other containers loaded to the ship from the shore. When all is finished, the ship leaves the terminal. The GPS trace of the ship is depicted in Fig.1.

B. Automatic Identification System (AIS)

AIS is an on-board marine tracking system that sends a ship’s mobility information to nearby ships and to port authorities automatically [16]. Such information includes (1) static ship-related information such as ship name, type, and dimension; (2) dynamic motion-related information such as the ship’s GPS position and heading direction, together with a unique identifier and a time-stamp; and (3) voyage-related information such as cargo type, destination, and estimated time of arrival. Most dynamic information is automatically updated through the AIS-connected sensors and thus is reliable, while the static and voyage-related information is usually entered manually and may suffer from human errors [16]. To obtain reliable results, in this paper, we only use the dynamic information from AIS data, and retrieve static ship-related information (e.g. ship size and capacity) from credible maritime data sources.

In practice, AIS base stations deployed in port areas collect AIS data sent by nearby ships, and upload it to a central database in the cloud. We consider an AIS record can be denoted as a data point, and define an AIS point as

\[ p = (\text{ship}_id, \text{time}\_\text{stamp}, \text{position}, \text{heading}) \] (1)

For a specific container ship, its AIS trace is then defined as a sequence of AIS points in chronological order. We represent a ship’s AIS trace as

\[ s : p_1 \rightarrow p_2 \rightarrow \cdots \rightarrow p_n \] (2)

where \( p_1, p_2, \ldots, p_n \) share the same \( \text{ship}_id \).

C. Container Handling Event

A container handling event is defined as the process of a container ship loading and unloading containers at a terminal [15]. Usually, containers are handled by quay cranes (QCs) [13] installed at terminals. QCs are equipped with specialized handling tools called spreaders, which can be driven to locate a container and transport it to the land side or to the ship side. QCs normally transport a single container at once. To increase the speed of container handling, usually several QCs work on a ship concurrently. The total number of containers handled in a port during a period of time is calculated as the container throughput and is often measured in thousand TEUs (Twenty-foot Equivalent Units) [26].
IV. FRAMEWORK OVERVIEW

Having defined the necessary preliminaries, we present the overview of the proposed framework in this section. As shown in Fig. 2, the framework consists of three phases, i.e., container handling event identification, container throughput estimation, and performance measurement and comparison. We briefly elaborate on the whole process as follows.

1) Container Handling Event Identification: We first detect a set of stationary events as candidates from ship GPS traces. We then apply a supervised model to select real container handling events from these candidate events, leveraging features based on terminal location and ship mobility patterns.

2) Container Handling Quantity Estimation: We estimate the number of containers handled in each container handling event by taking into consideration the number of QCs assigned to the ship, the QC speed, and the container handling time. Such information can be inferred based on the attributes of the corresponding ship (e.g., length and capacity) and terminal facilities (e.g., berth and QC parameters).

3) Port Performance Measurement and Comparison: We aggregate identified container handling events into terminal and port levels in chronological order, and derive a set of key performance indicators, including ship traffic, container throughput, berth utilization, and terminal productivity. With these indicators, we can analyze and compare the performance of different terminals and ports on a global scale.

In the following three sections, we elaborate on the design of the three phases of the proposed framework.

V. CONTAINER HANDLING EVENT IDENTIFICATION

In this section, our goal is to identify container handling events from ship AIS traces. It seems intuitive that the AIS points in the berths correspond to container handling events. However, there are several challenges in practice: (1) GPS readings are noisy. In general, AIS devices have an error of about 10 meters in reporting GPS positions [16], making it difficult to locate the actual position of a container handling event on the map; (2) the area of berths occupied by ships of different dimensions may vary significantly. Therefore, simply applying a distance threshold to extract AIS points may lead to unreliable results; (3) ships may be engaged in other events in berthing areas, such as temporary stops, and thus causing errors in identifying container handling events.

To address the above challenges, we propose a two-step process to accurately identify container handling events. First, we detect stationary events from ship AIS traces using an adaptive sliding-window based method. Then, we use a supervised model to separate container handling events from other events based on various spatial-temporal features extracted from terminal locations and ship mobility patterns.

A. Stationary Event Detection

Due to GPS errors, the readings of a ship’s position drift around even when the ship actually stays stationary. Therefore, we extract stationary events as clusters of AIS points from ship AIS traces. First, we apply an adaptive sliding-window based method to extract trace segments consisting of geographic-clustered AIS points. Then, we map each trace segment to a stationary event. Specifically, for a ship trace \( p_1 \rightarrow p_2 \rightarrow \cdots \rightarrow p_n \), we extract every trace segment \( p_m \rightarrow p_{m+1} \rightarrow \cdots \rightarrow p_{m+k} (1 \leq m < n, 1 \leq k \leq n - m) \) in which the distance \( (\text{dist}) \) between each pair of adjacent points is less than a threshold \( \delta_p \). i.e.,

\[
\forall m \leq i < m + k, \text{dist}(p_i, p_{i+1}) < \delta_p
\]  

(3)

We use a sliding-window with adaptive size along the trace to find such trace segments. We extend the window size by adding new points until the newly-formed segment violate requirement (3). More specifically, we elaborate on the process using an example shown in Fig. 3. For a ship trace \( p_1 \rightarrow p_2 \rightarrow \cdots \rightarrow p_7 \), we start by creating a window consisting of the first two points \( (p_1, p_2) \) in this case, and check whether the distance between \( p_1 \) and \( p_2 \) is less than \( \delta_p \). Since \( \text{dist}(p_1, p_2) > \delta_p \), we discard this window, and slide the window to start over from the end point \( (p_2, p_2) \), and create a new window \( (p_2, p_3) \). We see \( \text{dist}(p_2, p_3) < \delta_p \) so the window is kept; since \( \text{dist}(p_3, p_4) < \delta_p \), we extend the window by adding \( p_4 \), and repeat this procedure for the next adjacent points until the distance constraint is violated. In this way, we obtain a window containing a set of consecutive points \( p_2 \rightarrow \cdots \rightarrow p_6 \).

We map each extracted trace segment to a stationary event \( (\text{ship}_id, \text{position}, \text{duration}) \), where \( \text{ship}_id \) corresponds to the unique ship ID, \( \text{position} \) is estimated as the center position of all the AIS points in the segment, and \( \text{duration} \) is calculated as the time interval between the last and first AIS points in the segment.
B. Container Handling Event Selection

The detected stationary event in the previous step may correspond to various types of ship events. For instance, as shown in Fig. 4, a ship may stay at anchorage for several hours waiting for an available berth, or stop temporarily near the terminal for pilots boarding or course changing. A ship may also come to a port for regular maintenance or emergency repair, and thus it can stay at a shipyard for days or even months. These events share similar spatial-temporal patterns with real container handling events, and need to be filtered out. In this step, we apply a supervised model to select real container handling events from the candidate events, using the following features extracted from terminal location and ship mobility patterns:

1) **Ship-Terminal Distance.** This feature indicates whether the position of a stationary event is in a berth, and is designed to eliminate stationary events detected in areas such as anchorage or shipyard. Note that due to GPS errors, the position of a stationary event is usually not accurate, therefore we cannot simply use a binary threshold to determine whether the ship is in a berth or not. Instead, we leverage the distance between the ship and the terminal quay to characterize the ship’s position, as illustrated in Fig. 5. More specifically, we first separate the quay into a consecutive sequence of straight quay segments \( q_1, q_2, \ldots, q_n \). We then draw a bounding box for each berth alongside the quay segment \( b_1, b_2, \ldots, b_n \), respectively. Note that, since the breadth of each container ship varies significantly (ranging from 11m to 59m according to the ship dataset), we adaptively set the width of the bounding box to match the breadth of the corresponding ship. Finally, we calculate the distances between the position of the stationary event and the centers of the bounding boxes, and choose the nearest neighbor to determine the ship-terminal distance.

2) **Heading Direction Variation.** During a container handling event, the ship is fixed by mooring ropes and keeps a constant heading direction. In contrast, when a ship is fixed by anchors, for instance, staying at anchorage, its heading direction usually keeps changing (Fig. 4 (c)). Therefore, we consider the variation of ship heading direction as an effective feature to distinguish a container handling event from others.

3) **Berthing Duration.** A container handling event typically takes several hours [27]. In contrast, a temporary stop usually takes much less time, while ship maintenance at shipyard might take days to months. Therefore, we consider the duration of a stationary event as a feature from the temporal perspective.

With the above three features extracted, we train an SVM classifier [41] with a manually labeled dataset of stationary events, and use the trained model to separate real container handling events from other events. Finally, we obtain a container handling event with the corresponding ship ID, terminal ID, port ID, and berthing duration, i.e.,

\[
b = (\text{ship}_\text{id}, \text{terminal}_\text{id}, \text{port}_\text{id}, \text{duration}) \tag{4}
\]

VI. CONTAINER HANDLING QUANTITY ESTIMATION

In this phase, our goal is to calculate the number of containers handled during each container handling event. This is not trivial since there is no directly available data from the AIS dataset or the maritime open data sources. Therefore, we resort to publicly available information about ships and terminal facilities to help estimate the number of containers handled. We explored several strategies.

First, we attempt to leverage the *draft variation* of the ship during a container handling event to infer the number of containers handled. The draft value indicates the ship’s displacement in the water and can be retrieved from the AIS records [16]. The basic idea is that when containers are unloaded from the ship to shore, the draft of the ship will decrease, and vice versa. Together with the ship’s dimension and the average container weight, we can estimate the number of containers handled in real-time. However, this approach often fails to work. On the one hand, crew members may not enter or update the draft value in the AIS transponder in time during container handling [16]. On the other hand, the ship may dynamically adjust its ballast water [42] during the container handling process to keep its height align with the quay. In summary, the draft values may not be proportional to the number of containers on board.

The second approach leverages the ship’s *capacity*, which can be retrieved from the ship open data. Intuitively, if a ship is always fully loaded and unloaded during each container handling event, the number of containers handled can be estimated as twice the ship’s capacity. However, this assumption does not hold in most cases. According to [13], [43], the number of containers loaded and unloaded at each container...
terminal is demand-driven and highly dynamic. For instance, a container ship with capacity up to 8,000 TEUs might only load and unload about 2,000 TEUs during a container handling event [13]. Therefore, leveraging ship capacity to estimate the number of containers handled might be inaccurate.

In this section, we propose a QC-operation-based method to estimate the number of containers handled. Based on [13], [44], the number of containers handled is mainly affected by the following three factors: 1) the number of QCs assigned to the ship, 2) the container handling speed of a single QC, and 3) the working time of QCs. However, such information is not directly available. Instead, we need to leverage data we have at hand, i.e. the corresponding ship ID, terminal ID, and berthing duration of a container handling event. More specifically, we notice that:

1) The number of QCs assigned to a ship is related to the ship’s length. The longer the ship, the more QCs can be assigned to work on it simultaneously.

2) The container handling speed of a single QC is related to the average distance to transfer a container from ship to shore (and vice versa), which is mainly affected by the ship’s breadth. The wider the ship, the farther a container has to be transferred, thus decreasing the container handling speed of the QC.

3) The working time of QCs is determined by the duration of the container handling event and the unproductive time of QCs, such as the time for QCs to move to specific container rows.

Based on these observations, we propose an approach to estimate the values of the three factors from ship dimensions and terminal facility parameters, and then estimate the number of containers handled using these factors. We elaborate on the details of the approach in the following subsections.

A. QC Number Estimation

In general, a container terminal is equipped with several QCs, which are mounted on rails alongside the terminal quay [15]. During the container handling process, a set of QCs are assigned to the container ship. For instance, as illustrated in Fig.6, ship A is assigned with three QCs, while ship B has only one appointed QC. In this paper, we assume that during the transshipment process, the number of assigned QCs does not vary. In practice, the number of QCs that can be assigned to a container ship $N_q$ is constrained by the ship length as well as the number of available QCs in the terminal. Consequently, we present a method to estimate this number as follows.

First, we estimate the maximum number of QCs that a ship of length $L_s$ can take. According to [3], QCs should operate at a safety margin $D_q$ between each other (see Fig.6). Hence, the maximum QC number $N_{q,max}$ can be estimated as follows:

$$N_{q,max} = \left\lfloor \frac{L_s}{D_q} \right\rfloor$$

We determine the value of $D_q$ based on the QC parameters published by manufacturers, and verify the actual settings for each terminal by manually measuring the average distance between working QCs on satellite maps.

Then, we consider the constraints of QC availability, i.e. the actual number of QCs assigned to a ship can not exceed the number of available QCs in the terminal. For instance, in Fig.6, since there are only four QCs in the terminal and ship A has been assigned with three QCs, ship B can only be assigned with one. We denote this number as $N_{q,avail}$.

Altogether, when a ship arrives at a terminal with $N_{q,avail}$ available QCs, we estimate the number of QCs assigned to the ship $N_q$ as follows:

$$N_q = \min\{N_{q,avail}, N_{q,max}\}$$

Meanwhile, we update the available QC number by $N_q^*_{avail} = N_{q,avail} - N_q$. When the ship leaves the terminal, we return the assigned QCs to the available QC pool and update $N_{q,avail}$ again. Initially, we set $N_{q,avail}$ to be the total number of QCs in the terminal $N_{q,t}$. We retrieve the value of $N_{q,t}$ for each terminal from the port open data sources, and verify the values by manually counting on satellite maps.

B. QC Handling Speed Estimation

QC handling speed can be measured by the number of containers handled per hour [13]. In order to estimate this figure, we investigate the time spent in a container handling operation (Fig.7). When unloading containers from a ship, the QC trolley first locates a container in the row, and then picks up the container via the QC spreader, transfers it to the terminal side along the QC boom, and puts it on a yard truck. Yard trucks serve in parallel tracks under the QC to transport containers from terminal to yards. The container loading operation is a reverse of the above steps. In this paper, we assume that yard trucks arrive in different tracks equally often and transport containers without delay or congestion. According to the existing studies on QC handling speed [45], [46], the time spent on handling a container mainly consists of the following two parts:

1) Time for the spreader to locate and pick up/drop off the container on the ship or the yard truck ($T_p$). $T_p$ can be estimated as $\frac{H_q}{v_p}$, where $H_q$ is the height of the

Fig. 7. Container handing by QC at terminal.
QC boom and $V_s$ is the hoisting/lowering speed of the spreader, respectively. Both parameters can be found in the terminal facility open data. In general, it takes about 30 seconds to pick up and drop off a container for a typical container handling operation [47].

2) Time for the trolley to travel between the ship and the terminal ($T_1$). $T_1$ is related to the ship’s breadth. As illustrated in Fig.7, the average distance for the trolley to transfer a container can be approximated by $2(D_1 + D_2)$ (including a return trip without container payload), where $D_1 = \frac{W_s}{2}$ is half the QC’s rail span [45], and $D_2 = \frac{W_q}{2}$ corresponds to half of the ship’s breadth. Suppose the trolley travel speed along the boom is $V_t$, then the trolley travel time can be estimated as $\frac{2(D_1 + D_2)}{V_t}$. Altogether, the average time spent to handle a container is $T = T_p + T_1$. We measure the above distances in meters and speeds in meters per hour, and estimate the the QC handling speed $V_q$ as the number of containers handled per hour, i.e.,

$$V_q = \frac{\frac{T_q}{V_s} + \frac{W_s + W_r}{V_t}}{q}$$

\[ (7) \]

C. Container Handling Time Estimation

It is worth noting that during a container handling event, QCs are not handling containers all the time. In fact, there are two main kinds of unproductive time of QCs:

1) QC Preparation Time. The QC booms have to raise up when a ship is arriving at or leaving the terminal to avoid collision with the ship [13]. Each action usually takes up to 5 minutes [45]. We denote this time as $\Delta T_p$, and approximate it by a constant value for each container handling event.

2) QC Shift Time. During the container handling process, QCs may need to shift horizontally among container rows to located target containers [44]. We denote this time as $\Delta T_s$, and also approximate it by a constant value for each container handling event.

Hence, we deduce the real QC container handling time as:

$$T_q = T - \Delta T_p - \Delta T_s$$

\[ (8) \]

where $T$ is the duration of the container handling event.

D. Container Handling Quantity Estimation

We estimate the number of containers handled $\pi$ in a container handling event as the product of the assigned QC number $N_q$, the single QC handling speed $V_q$, and the QC container handling time $T_q$, i.e.,

$$\pi = N_q \cdot V_q \cdot T_q$$

$$= \min\{N_{q_{\text{avail}}}, \left\lfloor \frac{L_q}{D_q} \right\rfloor \} \cdot \frac{T - \Delta T_p - \Delta T_s}{V_s} \frac{W_s + W_r}{V_t}$$

\[ (9) \]

VII. PORT PERFORMANCE MEASUREMENT

With all the container handling events identified and their corresponding container handling quantity estimated, we are now ready to assess the performance of different ports in different time scales (e.g. month or year) in a quantitative way. In this section, we first aggregate container handling events into different terminals and ports in chronological order, and then calculate a set of key performance indicators at both terminal and port levels. In particular, we select the following four key performance indicators:

1) Ship Traffic, defined as the number of ships arriving at a terminal (or a port) for container handling during a period of time. This indicator is used by many port authorities to measure the port’s regional competitiveness and connectivity to other ports [18], [19].

2) Container Throughput, defined as the number of containers handled at a terminal (or a port) during a period of time. It is generally considered as one of the most important measurements of port productivity and regional trade growth [1], [48].

3) Berth Utilization, defined as the percentage of time that the berths of a terminal (or a port) are occupied by ships during a period of time [11]. Understanding berth utilization is critical to port management [49], since it not only indicates how intensively the berths are being used, but also reflects the potential ship waiting time for berths [49]. For instance, 90% berth utilization may indicate that the berths are almost fully used, but with the potential of congestion and long waiting time.

4) Terminal Productivity, defined as the number of containers handled per hour at a terminal over a period of time [50]. Shipping companies often use this indicator to evaluate a terminal, since it directly measures the speed at which container ships are loaded and unloaded [14]. Moreover, as most container terminals are charged by time, this indicator also affects the charges of container handling at terminals and ports [15].

A. Container Handling Event Aggregation

We aggregate identified container handling events based on their locations and occurring time, as illustrated in Fig.8. At terminal level, each container handling event occupies a series of consecutive time slots based on their location, beginning time, and duration. Note that these container handling events may overlap with each other since there may be more than one ship berthing at the same terminal. At port level, we put together terminals in the same port to obtain a table of
global container ports, and map the container handling events in different terminals to the corresponding ports.

**B. Key Performance Indicator Calculation**

A wide variety of performance indicators can be calculated based on container handling events. In this section, we select four key performance indicators widely used by researchers in port performance evaluation [7], [11], [15].

1) **Ship Traffic:** We calculate this indicator by summing up the number of container handling events identified in a terminal or a port during a given period of time. We further analyze the detailed traffic volumes of different types of ships, and the common traffic origins and destinations. More specifically, we focus on the ship traffic of different ship types. We divide ships into two categories, i.e. deep-sea ships and feeders, according to [15]. Deep-sea ships are large mother ships connecting major container ports all over the world, while feeders are smaller ships collecting and distributing containers within a continent or region [15].

2) **Container Throughput:** We calculate this indicator by summing up the number of containers handled in all of the container handling events in a terminal or a port during a given period of time. We also analyze the detailed types of container throughput. In particular, we focus on the transshipment throughput, which is the number of containers unloaded from one ship at the terminal and, after temporary storage in the stack, transferred to another ship to reach their destinations [44]. Transshipment business is driven by re-export and trans-ocean trades, and has become increasingly important in major container ports [51]. Since transshipment containers are either collected or distributed by feeders [15], we estimate the transshipment volume as twice the overall feeder throughput in a terminal, assuming that most of these containers will be handled again by other ships and counted again. We note that due to the potential losses of containers caused by other means of transshipment (e.g., road transport), the actual transshipment volume might be smaller than our estimation.

3) **Berth Utilization:** We calculate berth utilization on a terminal basis instead of per berth, since berths do not have clear boundaries in most terminals, and a large ship may occupy more than one berths. More specifically, for a specific terminal, we first sum up the durations of all the container handling events during a period of time, and then divide the value by the number of berths to estimate the average berth utilization time. Finally, we calculate the berth utilization as the ratio of the average berth utilization time to the duration of the given period. We perform a similar procedure to calculate port level berth utilization.

4) **Terminal Productivity:** We calculate terminal productivity by averaging the total number of containers handled at a terminal over a period of time. This indicator is usually measured in moves per hour [47]. We average the productivity of all the terminals in a port to obtain the port level productivity.

**VIII. Evaluation**

In this section, we evaluate the performance of our framework based on a large-scale, real-world AIS dataset and various maritime open data sources. First, we describe the datasets we use. Then, we evaluate the performance of container handling event identification and container handling quantity estimation. Next, we evaluate the key performance indicator measurement, and compare the performance on terminal level, port level, and region level.

**A. Dataset Description**

We use the following three datasets in our evaluation.

1) **Container Ship AIS Trace Dataset:** This dataset is provided by our partner shipfinder.com\(^4\), one of the leading AIS tracking service provider. After a data cleansing process that removes records containing invalid ship IDs or positions, we obtain the AIS traces of 4,881 container ships collected from the world’s major container ports in the year 2011 and 2012. The sampling interval ranges from 2 seconds for fast maneuvering ships to 3 minutes for stationary ships.

2) **Container Ship Information Dataset:** This dataset is also provided by shipfinder.com. It contains information about the 4,881 container ships, where each record includes the ship’s type (either deep-sea or feeder ship), length, breadth, capacity in TEUs, and its unique ship ID known as MMSI [16].

3) **Port Information Dataset:** For each container port used in our evaluation, we retrieve its information from the Marine Cadastre Portal and the corresponding port authority portals (e.g. [18], [19]). More specifically, we obtain (1) the port digital map with detailed terminal layout, and (2) the numbers and parameters of port facilities, especially QCs.

**B. Evaluation on Container Handling Event Identification**

We first determine the parameters and features used in the container handling event identification method by leveraging a human-labeled dataset of stationary events. We then validate the effectiveness of the method on the Port of Hong Kong and the Port of Singapore, respectively, and further apply the method to world-wide container ports.

1) **Parameter and Feature Setting:** We extract 500 ship traces from the Port of Hong Kong, and manually label the stationary events by plotting the traces on digital maps. For each stationary event in the terminal area, we check whether it is a container handling event by looking up the Hong Kong Marine Department Arrivals and Departures Register\(^5\), and obtain the exact information about the ship ID, terminal ID, and berthing duration for verified container handling event.

We use this manual-labeled dataset as the ground truth for the following experiments. In the first step (stationary event detection), we select the distance threshold \(\delta_p = 1m\), which yields closest results to the manual labels over a series of repeated experiments. In fact, the selection of the distance threshold \(\delta_p\) is determined by several factors, and sampling rate is exactly one of them. We notice that when the sampling rate is very low, setting a small \(\delta_p\) might omit some stay events due to insufficient GPS points within \(\delta_p\).

---

\(^3\)A sample dataset is available at: http://longbiaochen.com/dataset/shipdata.html
\(^4\)http://www.shipfinder.com/
spends 7.1 hours at terminal for container handling. In average, a container ship arrive at terminal for container handling events in both ports during 2011. The statistical data are retrieved from the ship arrival at terminal for container handling sections from [18] and [19]. The comparison between our identification results and the published statistics indicates that our method accurately identifies container handling events in both ports. We further apply the identification method on world-wide container ports (Dataset 1), and present the statistics of identification results in TABLE III. In general, we identify 257,181 container handling events generated by the 4,881 container ships in 2011. In average, a container ship spends 7.1 hours at terminal for container handling.

<table>
<thead>
<tr>
<th>Features</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>0.671</td>
<td>0.721</td>
</tr>
<tr>
<td>Distance + Heading</td>
<td>0.831</td>
<td>0.928</td>
</tr>
<tr>
<td>Distance + Heading + Duration</td>
<td>0.906</td>
<td>0.942</td>
</tr>
</tbody>
</table>

TABLE II
IDENTIFIED AND PUBLISHED CONTAINER HANDLING EVENTS IN HONG KONG AND SINGAPORE DURING 2011

<table>
<thead>
<tr>
<th>Port</th>
<th>Identified</th>
<th>Published</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hong Kong</td>
<td>13,211</td>
<td>13,347</td>
<td>1.02%</td>
</tr>
<tr>
<td>Singapore</td>
<td>18,973</td>
<td>19,290</td>
<td>1.64%</td>
</tr>
</tbody>
</table>

TABLE III
STATISTICS OF CONTAINER HANDLING EVENT IDENTIFICATION IN WORLDWIDE PORTS IN 2011

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Container Ships</td>
<td>4,881</td>
</tr>
<tr>
<td>Number of Identified Container Handling Events</td>
<td>257,181</td>
</tr>
<tr>
<td>Average Duration at Terminal</td>
<td>7.1 hour</td>
</tr>
</tbody>
</table>

C. Evaluation on Container Handling Quantity Estimation

We evaluate the container handling quantity estimation on the Port of Hong Kong and the Port of Singapore. Since we do not have the number of containers handled in each container handling event, we perform our evaluation on a port level.

The port authorities of Hong Kong and Singapore both provide monthly container throughput statistical data [18], [19]. We use the first 6 months of 2011 for parameter selection, and the next 6 months of 2011 for evaluation. More specifically, we first use the data of the first half year to determine the port-specific QC parameters in Eq. 9, i.e., \( D_q \), \( H_q \), \( W_q \), \( V_1 \), \( V_t \), \( \Delta T_p \), \( \Delta T_s \). We find the range for each parameter from the port information dataset (Dataset 3), and select the optimal values by minimizing the differences between the estimated monthly throughput and the published statistical data. The selected parameters are shown in TABLE IV. We then estimate the number of containers handled in the second half year in both ports using the selected parameters. The results are shown in the left part of Fig.9. The Mean Absolute Percentage Error (MAPE) [52] between our estimation and the published data across the second half year of 2011 are 1.0% for Hong Kong and 0.4% for Singapore, respectively. We further predict the monthly container handling number in the first 6 months of 2012, assuming that the selected parameters do not change significantly for the year. We compare our prediction with the published figures of port authorities, as shown in the right part of Fig.9. We observe that our prediction is consistent with the published figures, and the published number is slightly larger than the predicted value. The prediction error is increasing over months, probably due to improvement of operation efficiency and upgrade of terminal equipments, suggesting that the estimation parameters need to be trained and updated on a regular basis, e.g., every year or every two years.

We also check the validity of the estimated QC number and QC speed in Hong Kong. The results show that (1) the average QC number assigned to a ship is 1.7 for feeders, and 4.2 for deep-sea ships. This results is consistent with the common crane split practice [15]. (2) The average QC speed is 37 containers per hour, which is close to the estimation of [53].

D. Port Performance Measurement

We measure port performance using the four key performance indicators. Fig.10 shows our port performance visualization interface. The interface consists of a port list (the left part of Fig.10) from which we can select a port, and the content...
Fig. 9. Estimated and published number of containers handled in Hong Kong (HK) and Singapore (SG) during the second half year of 2011 and the first half year of 2012, respectively.

panel where different performance indicators are visualized in charts and figures (the right part of Fig.10). The period of time can be adjusted by selecting different start and end dates.

More specifically, for each terminal in the port, we use bar charts to visualize the accumulative ship traffic and container throughput volume over a period of time, and use terminal layout maps to visualize the average berth utilization and terminal productivity over a period of time. We further separate different types of container ships (i.e. deep-sea and feeder) and container throughput (i.e. transshipment and other) in each terminal. For an overview of the port, we use pie charts to present the percentage of different categories ship traffic and container throughput, and plot the monthly trends of berth utilization and terminal productivity over a period of time.

E. Performance Comparison At Terminal Level

We conduct a case study on the Port of Hong Kong to compare the performance of different terminals. The Port of Hong Kong contains 9 container terminals run by 5 terminal operators [53]. Each terminal has different number of berths and QCs, as detailed in TABLE V.

We present the key performance indicators at terminal level in 2011, as shown in Fig.11. We can see that the ship traffic and container throughput volume vary significantly among different terminals due to different terminal sizes and number of facilities, whilst the average berth utilization and terminal productivity are similar among these terminals. In particular, we observe that (1) the average berth utilization is about 63.28% (Fig.11(c)), which is higher than the empirical optimal value of 60% [49], and might indicate ship traffic congestion in port waiting for berths. (2) the average terminal productivity is 67 moves per hour (Fig.11(d)), which is consistent with the estimation of the JOC Port Productivity Report [14], and exceeds the world’s average value of about 60 moves per hour, indicating that Hong Kong is one of the most efficient container ports in the world [9].

In order to better understand the roles of each operator, we also group terminals by operators and calculate the corresponding aggregated key performance indicators, as shown in Fig.12. We take HIT, the largest terminal operator in Hong Kong [6], as an example. HIT has observed the largest ship traffic and container throughput as well as high facility utilization and efficiency. In particular, HIT has the highest percentage of feeder traffic and transshipment throughput. This might be the result of its business strategy that focuses on the transshipment business in Southeast Asia rather than the export business from Mainland China [48]. In recent decades, export business has shrunk significantly in Hong Kong due to competition from nearby Chinese ports. However, transshipment business has become an advantage of Hong Kong since the Chinese maritime cabotage law forbids foreign ships from handling domestic cargo at Chinese ports, but from which Hong Kong is excluded [9]. Consequently, the burgeoning transshipment demand from China in the past decades has brought large revenue to HIT. In summary, HIT not only handles large volumes of container traffic efficiently, but also focuses on the fast-growing transshipment business to benefit from the economies of scale in China [6].
F. Performance Comparison At Port Level

For port-level comparison, we conduct the following two studies. First, we compare the performance of the Port of Hong Kong and the Port of Singapore, two of the largest container ports in the world. Second, we rank container ports based on their annual container throughput, and compare our list with the publicly announced one.

1) Pair-wise Comparison: We present the performance comparison results in Fig.13. We observe that Singapore has a much higher percentage of feeder traffic, and thus higher transshipment throughput than Hong Kong (Fig.13(a) and (b)). In fact, Singapore is one of the largest transshipment hubs in the world [51]. According to [9], this can be explained by Singapore’s strategic location in the Malacca straits and its relative small home market. In contrast, Hong Kong is generally considered a gateway port with transshipment functions [9], thanks to the large volume of export cargo from its hinterland, the Pearl River Delta in South China [54]. The berth utilization of Hong Kong in different months varies more significantly than that of Singapore (Fig.13(c)). In particular, we observe two significant decreases of berth utilization in February and September in Hong Kong, which might correspond to the Chinese New Year holidays and the typhoon seasons in Hong Kong [53]. The comparison results also indicate that Singapore has slightly higher terminal productivity than Hong Kong (Fig.13(d)). This is probably due to the fact that in Singapore, most container terminals are operated by one company (i.e. PSA International), therefore global scheduling strategies of ships and terminals can be applied to improve terminal productivity [55].

2) Port Ranking: To compare container ports on a global scale, we rank ports based on their container throughput in 2011. We compare our list with The JOC Top 50 World Container Ports 2011 [50], which is based on aggregated port container throughput statistical data. The top ten ports of the two lists are shown in TABLE VI. We can see that in both lists, Asian ports hold nine of the ten top port rankings, indicating that Asia has become the world’s main seaborne trade region [1]. We also observe two differences: (1) In our list, Hong Kong ranks 5 while it ranks 3 in the JOC list, which indicates that we underestimated its container throughput. This is probably caused by Hong Kong’s unique mid-stream operation [56] that handles containers at anchorages or mooring buoys other than at container terminals. Our estimation only counts container handling events at terminals, whilst the JOC list counts all. (2) The rankings of Dubai and Qingdao in our list are the reverse of the JOC list. This is probably due to the very close container throughput values of both ports (Qingdao at 13.02 kTEU and Dubai at 13.00 kTEU in the JOC list).

G. Performance Comparison at Region Level

Container port performance is largely influenced by its regional economic development [57]. For instance, the container throughput of the Port of Hong Kong grows significantly in the past decades thanks to the fast economic development in
Southeast Asia countries and China. Moreover, the economic development of different regions are strongly connected via container shipping. In order to compare container ports in different regions, we construct a container port network based on the performance indicators, and analyze the connections of container ports in different regions.

More specifically, we use the ship traffic and container throughput data in 2011 to construct a network for the top-100 ports from the above-mentioned rank list. We treat each port as a vertex, and the ship traffic between ports as link weights. The size of each vertex corresponds to the container throughput of the port. Fig.14a shows the results of the port network, where darker links indicate higher ship traffic intensity, and vertices with larger disks correspond to larger container throughput of ports. In this network, we can identify two major types of connections between different regions, as mentioned in [5]. The first type is the East-West connections which link through the three poles of the global economy, i.e. the East Asia region, the West Europe region, and the North America region. The second type is the North-South connections which extend the container shipping network to the Africa, the Latin America, and the Australia regions [5]. Most hub ports of these connections are located at the interconnect points of regions, such as the Malacca straits, the Suez Canal, and the Panama Canal.

We then perform a community detection on the container port network. We define communities as groups of ports with many links within the groups but few links between different groups. A community may contain ports from several regions, and thus reveals the connections of these regions. We detect these communities with a modularity optimization method proposed by [58]. The results are shown in Fig.14b, where each community is represented by a color. More specifically, we detect the following four communities in the network:

1) Asia-Pacific and Trans-Pacific Regions (blue). This community includes large ports such as Shanghai, Busan, and Los Angeles, indicating strong economic connections between the East Asia and North America regions.

2) Southeast Asia Regions (yellow). This community indicates the frequent trade between China and Southeast Asia. Interestingly, Australian ports are also strongly connected to the region, probably due to Australia’s large-scale raw material export to the Asia market.

3) Trans-Atlantic Regions (red). These are the traditional international trade regions of the western world, connecting countries in Europe and the Americas. Large hubs, such as Rotterdam, Antwerp, and Felixstowe, can be found in this community.

4) Middle East and South Asia (green). This community includes several large ports in the Middle East region (e.g. Dubai) and the South Asia region (e.g. Tanjung Pelepas). It’s worth noting that this community also includes the Port of Cape Town, one of the largest ports of South Africa, probably indicating the distribution of manufactured cargoes from South Asia to Africa.

IX. Conclusion and Future Work

The pervasiveness of digital traces and open data provides us with an unprecedented opportunity to understand the dynamics in container shipping and international trade. In this paper, we introduce a cheap, automatic, and more accessible approach to measure and compare container port performance on a world scale. This work is motivated by the fundamental need of researchers to understand container port performance, and could benefit a variety of applications such as terminal productivity improvement, liner schedule optimization, and regional economic development. We propose a three-phase framework to derive key performance indicators of container ports by leveraging ship GPS traces and maritime open data. First, we identified container handling events from ship GPS traces. Then, we estimate the number of containers handled by leveraging information about ship and terminal facilities from various maritime open data sources. Finally, we aggregate container handling events into terminal and port levels to derive a set of key performance indicators, including ship traffic, container throughput, berth utilization, and terminal productivity. By leveraging these indicators, we can compare container port performance at terminal level, port level, and region level. We evaluate our framework with large-scale, real-world ship GPS traces and maritime open data. The results show that our framework not only accurately estimates various port performance indicators, but also effectively produce port comparison results.

In the future, we plan to broaden and deepen this work in two directions. First, we plan to incorporate more maritime open data sources, such as berth pricing policies of terminals, to study the revenue and economic strategies of container ports. Second, we plan to study the long-term variation of container port performance and container port network structure by leveraging more ship AIS traces, and explore their impact on regional economy.
X. ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the editors for their helpful comments and suggestions. The authors would like to thank the Maritime and Port Authority of Singapore for providing valuable statistical data, Dr. Paul Gibson for proofreading, and Dr. Chao Chen and Dr. Dingqi Yang for their inputs and comments. This research is supported by the National 973 Program (2013CB329504), Program for New Century Excellent Talents in University (NCET-13-0521), and Zhejiang Provincial Natural Science Foundation of China (LR15F020001).

REFERENCES

[18] Hong Kong Marine Department, “Port of Hong Kong Statistical Tables,” Hong Kong, 2014.


Longbiao Chen received the B.Sc. degrees in computer science from Zhejiang University, Hangzhou, China, in 2010. He is currently pursuing the Ph.D. degree in Department of Computer Science, Zhejiang University, and visiting Institut Mines-TELECOM / TELECOM SudParis in France. His research interests are mainly in urban computing, big data applications, and ubiquitous computing.

Daqing Zhang is a Full Professor at Telecom SudParis, Institut Mines-Telecom, France. His research interests include context-aware computing, urban computing, mobile computing, big data analytics, pervasive elderly care, etc. Dr. Zhang has published more than 200 technical papers in leading conferences and journals, where his work on context model is widely accepted by the pervasive computing, mobile computing and service-oriented computing communities. He served as the general or program chair for more than 10 international conferences, giving keynote talks at more than 16 international conferences. He is the associate editor for 4 journals including ACM Transactions on Intelligent Systems and Technology, IEEE Trans. On Big Data, etc.. He is the winner of the Ten-years CoMoRea impact paper award at IEEE PerCom 2013, the Honorable Mention award of ACM Ubicomp 2015, the Best Paper award at IEEE UIC 2015 and 2012, and the Best Paper Runner Up award at MobiQuitous 2011. Daqing Zhang obtained his Ph.D. from University of Rome La Sapienza in 1996.

Xiaojuan Ma is an assistant professor of Human-Computer Interaction (HCI) at the Department of Computer Science and Engineering (CSE), Hong Kong University of Science and Technology (HKUST). She received the Ph.D. degree in Computer Science at Princeton University. She was a post-doctoral researcher at the Human-Computer Interaction Institute (HCII) of Carnegie Mellon University (CMU), and before that a research fellow in the National University of Singapore (NUS) in the Information Systems department. Before joining HKUST, she was a researcher of Human-Computer Interaction at Noah’s Ark Lab, Huawei Tech. Investment Co., Ltd. in Hong Kong.

Leye Wang is currently a Ph.D. student at Institut Mines-TELECOM / TELECOM SudParis and Universite Pierre et Marie Curie, France. He received his M.Sc. and B.Sc. in Computer Science from Peking University. His research interests include mobile crowdsensing and ubiquitous computing.
Shijian Li  Shijian Li received his PhD from the College of Computer Science and Technology, Zhejiang University, in 2006. He worked in the Institute Telecom SudParis, France, as a visiting scholar in 2010. His research interests include sensor networks, ubiquitous computing, and social computing. He has published over 40 papers in the above domains. He works as an editor of the International Journal of Distributed Sensor Networks, and as reviewer or PC member of more than 10 conferences.

Zhaohui Wu  Zhaohui Wu received the B.Sc. and Ph.D. degrees in computer science from Zhejiang University, Hangzhou, China, in 1988 and 1993, respectively. He is currently a Professor with the Department of Computer Science, Zhejiang University. His research interests include distributed artificial intelligence, semantic grid, and pervasive computing. Dr. Wu is a Standing Council Member of the China Computer Federation.

Gang Pan  received the B.Sc. and Ph.D. degrees in computer science from Zhejiang University, Hangzhou, China, in 1998 and 2004, respectively. He is currently a Professor with the College of Computer Science and Technology, Zhejiang University. He has published more than 90 refereed papers. He visited the University of California, Los Angeles, Los Angeles, during 2007-2008. His research interests include pervasive computing, computer vision, and pattern recognition. Dr. Pan has served as a Program Committee Member for more than ten prestigious international conferences, such as IEEE International Conference on Computer Vision and IEEE Computer Society Conference on Computer Vision and Pattern Recognition.