IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 26, 2018, accepted November 30, 2018, date of publication December 5, 2018, date of current version January 7, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2885081

Planning City-Wide Package Distribution Schemes Using Crowdsourced Public Transportation Systems

GEYAO CHENG[1], DEKE GUO[1,2], (Member, IEEE), JIANMAI SHI[1], AND YUDONG QIN[1]
[1]Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha 410073, China
[2]College of Intelligence and Computing, Tianjin University, Tianjin 300350, China

Corresponding authors: Deke Guo (guodeke@gmail.com) and Jianmai Shi (jianmaishi@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61772544, in part by the Hunan Provincial Natural Science Fund for Distinguished Young Scholars under Grant 2016JJ1002, and in part by the Guangxi Cooperative Innovation Center of Cloud Computing and Big Data under Grant YD16507 and Grant YD17X11.

ABSTRACT Due to the rapid development of online retailers, there is a great demand for package express shipping services, which causes traffic congestion, resource consumption, and environmental pollution (e.g., carbon emission). However, there is still a large amount of under-utilized capacity in the public transportation systems during off-peak hours. In this paper, we investigate the same-day package distribution using crowdsourced public transportation systems (CPTSs). Specifically, given a number of packages and the timetable of available CPTSs trips, we optimize the schemes of delivering the packages using the under-utilized capacity of the CPTS trips, without impacting the quality of passenger experience. To estimate the amount of under-utilized capacity of each trip across any two adjacent stations, we propose the passenger transit model based on the history data. To assign the under-utilized capacity of each trip to the package deliveries, we develop the minimum limitation delivery (MLD) method, which only utilizes the minimum amount of under-utilized capacity of the whole trip to deliver packages. However, the available capacity is not fully utilized at most stations by MLD. Therefore, we further propose the adaptive limitation delivery (ALD) method, which loads as many packages as possible, until the volume of loaded packages reaches the available capacity in theory. The experimental results and theoretical analysis show that both MLD and ALD could distribute packages efficiently. Moreover, given a set of packages, scheduling of ALD only consumes about 67% time compared to the scheduling of MLD, with a little higher risk of impacting passengers.

INDEX TERMS Package distributions, crowdsourced, public transportation systems, quality of passenger experience.

I. INTRODUCTION

The pervasive use of transportation vehicles facilitates the round-the-clock package distributions, which urges the sustainable development of the logistics industry [1]. Specifically, online ordered products have generated over one billion package distributions in 2013, and this number is predicted to grow by 28.8% in 2018 [2]. Despite the significant growth of logistics, companies still face many challenges in the successful fulfilment of package deliveries. One of the main challenges is to provide a convenient *same-day delivery* service [3] in a cost-efficient way. Additionally, the consumed manpower and other resources increase continuously due to the large number of packages to be delivered [4]. Data shows that the postal enterprises in China have built more than 1,000 warehousing and distribution centers, and newly opened 153 trunk postal routes in 2017 [5]. Furthermore, the dedicated urban vehicles for logistics have a serious impact on air pollution and traffic congestion, especially when the total traffic volume is huge [6]. Therefore, cities are looking for instruments and policies to guarantee an efficient and effective urban transmission for both passengers and packages [7].

Many efforts aim to reduce the unnecessary environmental pollution and resource consumption caused by dedicated transportation systems in package distributions. The first kind of efforts focuses on the mixed logistics [1], [8]–[10]. In this situation, several kinds of vehicles would be successively utilized to realize the package deliveries, such as trucks & city freighters [11]. The second kind of efforts is based on Crowdsourced Delivery, where the package delivery tasks are outsourced to a non-specific Crowdsourced network in a free and voluntary manner [12], [13]. Some efforts aim at having packages take hitchhiking rides, such as taxis [6] and private cars [13], while others suggest drivers to accept several package delivery requests in one trip [14]. In the third kind of efforts, a connected Personal and Freight Rapid Transit System is established [15], where the Personal Rapid Transit and the Freight Rapid Transit serve the transportation requests together [16].

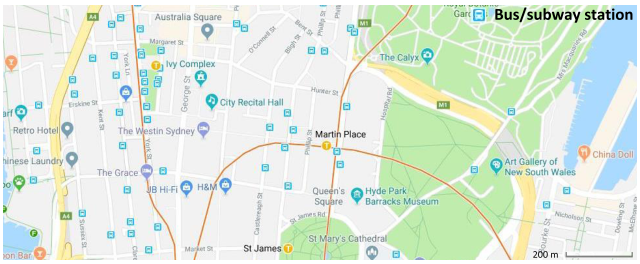

FIGURE 1. The bus/subway stations in Sydney, Australia.

While various efforts have been made to realize the economic and green logistics, it is still urgent to find a way to tackle the package distributions from a global perspective and achieve the overall advantage. Public Transportation Systems (PTS) [17], as an urban infrastructure, would be competent to accomplish the package delivery tasks. The PTS is stable, time-scheduled, economically friendly, and widely covered [17]. Moreover, PTS exhibits a huge amount of under-utilized capacity in most cases. As Figure 1 shows, the bus/subway stations have covered over 75% blocks ($200 \times 200\ m^2$) of central Sydney, Australia. However, the adequate resources are not utilized efficiently, because there always exists free seats on bus/subway in off-peak hours [18]. Inspired by such observations, we propose a novel idea of the Same-day Package Distribution using Crowdsourced Public Transportation Systems, named the SPDCP problem. The Crowdsourced Public Transportation Systems (CPTS) allocate the resources of PTS in a free and voluntary manner. Thus, the capacity utilization of the CPTS could be significantly enhanced. Moreover, the package deliveries would cause less resource consumption and environmental pollution, due to the share of the under-utilized capacity in CPTS vehicles. Therefore, the business model of SPDCP is really feasible in the future.

The business model of CPDCP is really feasible in the future, due to its abundant idle capacity, effectiveness, and environmental friendliness. Actually, some other transportation systems have already been adopted as part of new logistic methods in the commercial sector called last-mile delivery, such as taxis [6] and drones [19]. The package is traceable in the whole delivery process, which ensures the package security. The CDC, where the package is pretreated, has an ID; the station, which is chosen as the target destination of the package, has an ID; the trip, where the package is loaded, has an ID; and the freighter, which delivers the package to its final destination, has an ID. This ensures the package security and avoids package delivery mistakes or losses. In addition, our SPDCP mode is effective for delivering a huge number of packages every day. As the postal company said, the number of packages handled and unloaded could reach up to 1.22 million at peak seasons [20]. In our experiments, the number of daily delivered packages can reach about 90 thousand when the SPDCP mechanism only involves 10 Crowdsourced routes. This number would increase to 2-3 times if the SPDCP mode involves all of the bus routes in the city. Thus, our SPDCP mode could refresh about 1/5 pressure of package delivery, which is really helpful for the *same-day delivery service*.

The package distribution consists of three stages in the SPDCP problem. In the first stage, packages would experience pretreatment at CDC. Specifically, for each package, the nearest station to their final destination would be selected as its target station. All packages with the same target station would be tagged the same category and clustered together. The clustered packages would be assigned to the departure station of a route that passes through their target stations. In the second stage, packages would wait to be transmitted at the departure station of the chosen route. Once the upcoming bus/subway trip is predicted to have sufficient under-utilized capacity, the packages would be loaded in a determined sequence. These packages would be unloaded when the trip arrives at the target stations of the packages. In the third stage, a number of local freighters would collect the unloaded packages and distribute them to their final destinations.

In this paper, we mainly focus on solving the issues involved in the second stage. That is, given a set of packages to be delivered, we plan a scheme using the minimum number of continuous bus/subway trips such that their under-utilized capacities are sufficient for delivering such packages, without impacting the quality of passenger experience (QoPE) [21]. Note that minimizing the number of utilized trips would not only decrease the resource consumption, but also ensure the *same-day delivery* service. To realize this design goal, we address the following two research challenges: 1) estimating the amount of under-utilized capacity of each trip across any two adjacent stations; 2) assigning the unoccupied capacity of each trip to the package deliveries, without influencing QoPE.

To solve the first challenge, we present the Passenger Transit Model to estimate the number of passengers at each station of each trip. Thus, the space of the unoccupied seats at each station of each trip, which reflects the corresponding under-utilized capacity, could be determined. To solve the second challenge, we first propose the Minimum Limitation Delivery Method (MLD), which only utilizes the minimum amount of under-utilized capacity of the whole trip to accommodate packages, so as to ensure QoPE. However, there is still a considerable amount of under-utilized capacity that has not been fully exploited at most stations. Inspired by this fact, we further propose the Adaptive Limitation Delivery Method (ALD). It prefers to load as many packages as possible in each trip, until the volume of loaded packages reaches the available under-utilized capacity.

One important metric of our methods is the impact rate of QoPE. The impact appears when the volume of loaded packages exceeds the realistic amount of under-utilized capacity. The root cause is that the package loading scheme is made based on the expected number of passengers at each station in a trip, while the estimated value may mismatch with the value in reality. To guarantee the availability of our schemes, we further give a theoretical analysis of the impact rate of MLD and ALD.

We conduct extensive evaluations to verify the effectiveness and impact of our schemes of the SPDCP problem. Results show that more than 60% packages can be successfully delivered when our SPDPT mode involves 10 routes and the distance constraint is set as 500 m. The average delivering time is always less than 1 hour in off-peak hours, which satisfies the requirement of the *same-day delivery* service greatly. Given a set of packages to be delivered, the scheduling of ALD utilizes 2/3 of trips to complete the task, compared to that of MLD, at the cost of incurring a higher risk of impacting the quality of passenger experience. Overall, the impact on passenger experience is not severe and is acceptable in the nowadays public transportation system.

II. RELATED WORK

Various efforts aim to reduce the unnecessary resources consumption and environmental pollution caused by the dedicated transportation of packages. The efforts can be divided into three types:

A. MIXED LOGISTICS

Some efforts focused on mixed logistics, where more than one kind of vehicles may be utilized to realize the successive package deliveries. For example, literature [11] introduced a heterogeneous delivery team of two cooperating vehicles: a truck carries a shipment of packages to the street blocks, and a micro aerial vehicle carries individual packages from the truck to the specific delivery points in the region. There were still other kinds of vehicles used in mixed logistics, such as buses and city freighters [1], [8], close-open mixed two-echelon [9], and trucks and city freighters [10].

B. CROWDSOURCED DELIVERY

The second kind of efforts is based on Crowdsourced Delivery [22]–[24], where the tasks of package deliveries are outsourced to a non-specific mass Crowdsourced network in a free and voluntary manner. Some efforts aimed at having packages take hitchhiking rides, such as taxis [6] and private cars [13]. For example, Chen *et al.* [6] suggested that voluntary taxi drivers would first collect the packages before picking up the passengers, and then leave the packages at the appointed locations after dropping off the passengers. There were some other efforts optimizing the assignment of package delivery tasks to drivers. Setzke *et al.* [12] optimized the assignment of packages to drivers, subject to transportation routes and time constraints. Arslan *et al.* [13] proposed a peer-to-peer platform to realize the matches between packages and drivers. Additionally, drivers were suggested to accept several requests in one trip for less time and resource consumption [14], compared to that of direct deliveries, where one trip only accepts one request.

TABLE 1. Summary of the related works.

Literature	Efforts types	Features and solved problems	Stages
[1] [8]	Mixed logistics	Successive package deliveries with buses and city freighters	②③
[9]		Successive package deliveries with close-open mixed two-echelon	②③
[10]		Successive package deliveries with trucks and city freighters	②③
[11]		Successive package deliveries with trunks and micro aerial vehicles	②③
[6]	Crowdsourced delivery	Crowdsourced Deliveries using taxis	②
[12]		Optimization of the assignment from packages to drivers	①②
[13]		Realization on the matches between packages and private car drivers	①②
[14]		Optimization of accepting several requests in one trip	①②
[15]	Shared passengers and goods	Optimization of the joint use of PRT/FRT vehicles	②
[16]		Determination of the set of routes (a fleet of PRT/FRT vehicles) with the minimum transportation cost	②

C. SHARED PASSENGERS AND GOODS [15]

In the third kind of work, the potential of integrating *a shared package and passenger transit system* [15] is investigated. The Personal Rapid Transit (PRT) and the Freight Rapid Transit (FRT) serve the transportation requests together. Fatnassi *et al.* [16] modeled an asymmetric distance constrained vehicle routing problem. The objective was to find the best set of routes using a fleet of PRT/FRT vehicles, which satisfied all trip requests for a typical period at the minimum transportation cost. Some other works optimized the joint use of PRT/FRT vehicles between passenger and goods flows [15].

There is an increasing attention solving the package delivery problem using diverse transportation tools, such as private cars [13], taxis [6], and Unmanned Aerial Vehicles (UAV) [19]. They aim to utilize the under-utilized capacity of different transportation tools to reduce the resource consumption and release traffic congestion. However, there is still a lack of works on delivering packages utilizing the Crowdsourced public transportation systems. The few works on this topic are [1] and [8], which mainly focused on the conjunction with the second and the third stage in the SPDCP problem proposed in this paper. Masson *et al.* [1] optimized the handshaking process between the buses and the city freighters. The primary objective of Trentiniet *et al.* [8] was to minimize the number of city freighters used and the total time traveled by these vehicles. In this paper, we mainly focus on solving the issues involved in the second stage of the SPDCP problem. Specifically, given a set of packages to be delivered, we plan a scheme using the minimum number of continuous trips such that their under-utilized capacities are sufficient for delivering such packages, without impacting the quality of passenger experience.

TABLE 2. Abbreviations for our SPDCP mode.

Abbreviation	Description
SPDCP	the Same-day Package Distribution using Crowdsourced Public Transportation Systems;
CPTS	the Crowdsourced Public Transportation Systems;
PTS	the Public Transportation Systems;
MLD	the Minimum Limitation Delivery method;
ALD	the Adaptive Limitation Delivery method;
CDC	the Consolidation and Distribution Centers;
QoPE	the quality of passenger experience;
nip	the number of impacted passengers;
ris	the ratio of impacted stations;
tnr	the ratio of utilized trips between MLD and ALD;

III. FRAMEWORK OF SPDCP

In this section, we provide definitions of some basic concepts, and then formally state the framework of the SPDCP problem. Thereafter, we provide a problem statement of the SPDCP problem.

A. BASIC CONCEPTS

We first propose several important definitions used in the remainder of this paper.

Definition 1 (Route and Trip): We define a **route** to be the travel sequence of a bus/subway, linked by all involved stations alongside. A **trip** is actually a specific route starting at a scheduled time [25].

Definition 2 (Bus/subway Schedule): For a given route, the daily schedule at the station i is a nearly deterministic arrival (and departure) process: $\Re_i = \{t_i^r : r = 1, 2, \ldots, R\}$. Let t_i^r represent the time when the r^{th} trip arrives at station i, and R represent the total number of trips of the route.

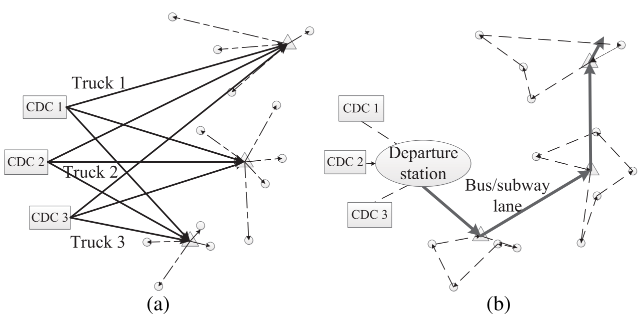

FIGURE 2. The comparison between the traditional logistics distribution mode and our proposed SPDCP mode [1]. (a) Conventional logistics distribution mode. (b) Our proposed SPDCP mode.

B. FRAMEWORK OF SPDCP

In the conventional logistics distribution mode, as shown in Figure 2(a), each logistics company would distribute packages from their subordinate CDC to customers independently and separately, using their own vehicles (both trucks and local freighters). Therefore, it consumes a large amount of labor power as well as material resources, due to the long total distance. However, our proposed SPDCP mode, which is shown in Figure 2(b), can resolve the problem greatly.

The main idea of SPDCP is to deliver packages from the CDCs to their final destinations economically and ecologically, utilizing the considerable under-utilized capacity of CPTS vehicles. The process is usually composed of three stages. In the first stage, packages experience pretreatment at the CDCs of their logistics companies. Specifically, each package would be tagged with a target station, which has the shortest distance to its real final destination. Obviously, the chosen stations must be in the Crowdsourced routes, which have the cooperation with the logistics companies. Thereafter, the packages with the same target station would be grouped together. Finally, the truck would deliver the grouped packages to the departure station of the route that passes through their target stations. Note that, there might exist several routes passing through that target station. We choose the route that has the shortest distance from the CDC to its departure station. When there is no free space in the following trips or the selected route is canceled, the scheduling personals would select another route that has the shortest distance from the CDC to the departure station. This ensures the robustness of the SPDCP mode. In the second stage, packages from different CDCs would be collected at the departure station of the chosen route and wait to be transmitted. When an upcoming bus/subway is predicted to offer under-utilized capacity, the packages would be loaded in a determined sequence. The packages would be unloaded when the trip arrives at their target stations. In the third stage, the given number of local freighters would collect all of the unloaded packages and distribute them to their final destinations. Our proposed SPDCP contains the following advantages, against the existing conventional logistics distribution mode maintaining dedicated shipping crews.

1) EFFECTIVENESS

The bus/subway stations are usually close to the densely populated areas. The Flint Hills Area Transportation Agency report states that 75% of off-campus students and 35% of employees live within five minutes of the bus city-wide routes [26]. Moreover, the stations of CPTS are scattered throughout our cities with a high coverage rate. For example, the bus layer of Great Britain covers the largest fraction of the island [27]. Thus, most of the unloaded packages can be easily transmitted to their destinations by city freighters, this facilitates the delivery tasks in the last kilometer [28].

2) CONSIDERABLE AMOUNT OF UNDER-UTILIZED CAPACITY

PTS exhibits a huge amount of under-utilized capacity in off-peak hours, which offers a great chance for sharing capacity between passengers and packages. Moreover, the bus/subway is operated stably with a time schedule, whereas existed services are less efficient during weekends, holidays, and days of bad weather [29]. Thus, the packages would be delivered to their target destinations in the operational time using the considerable amount of under-utilized capacity, with high feasibility.

3) ECONOMIC AND ENVIRONMENTAL FRIENDLINESS

Since our solution only leverages the under-utilized capacity of the CPTS vehicles for package distributions, it induces little extra air and noise pollution [4]. Moreover, we cluster packages from the CDCs of different logistics companies and then deliver them to their destinations from a global perspective. This would reduce the total travel distance, resulting in less resource consumption, traffic congestion, and environmental pollution.

Moreover, the SPDCP incurs a low impact on the quality of passenger experience. Note that, only the under-utilized capacity of the CPTS vehicles can be used for logistics. This avoids the occupation on the seats for passengers. In addition, bus/subway stops at every station originally. This facilitates unloading packages, without causing too much additional waiting time. Additionally, the package is traceable in the whole delivery process, with the help of IoT technologies. This ensures the package security and avoids the delivery mistakes or package losses.

Although the existing conventional logistics distribution mode aggravates traffic congestion and environmental pollution, the loaded amount of each operating trunk is much larger than that of the SPDCP mode. This enhances transportation efficiency. What's more, the separate transportation for passengers and packages would not incur chaos in the capacity utilization, thus ensures the quality of passenger experience. The SPDCP mode also has some disadvantages. For example, the number of passengers entering and exiting a station, which is estimated from passenger surveys or bus/subway data, may be inaccurate, compared with the real situation. Moreover, the data of passengers taking on or off the subway is easy to get, because the subway system has stored the information when people travel by subway. However, the data may be incomplete for bus routes, because the exact traveling record for some passengers paying by coins is difficult to collect.

C. PROBLEM STATEMENT

In this paper, we mainly focus on the second stage of the SPDCP mode and optimize the package delivering schemes utilizing the under-utilized capacity of the CPTS trips. To this end, we first design two design rationales in the process of package distribution. Thereafter, we present a model formulation.

1) RATIONALE 1: NO PACKAGE COULD BE UNLOADED UNLESS IT ARRIVES AT ITS TARGET STATION

The most efficient way of delivering packages in SPDCP is to fully utilize the under-utilized capacity of all stations in a trip. However, in this situation, when the actual number of passengers exceeds the expected number, some packages should be unloaded to ensure QoPE. This would cause an additional amount of labor power as well as resources consumption. Besides, the frequent loading up and down the packages among the involved stations of the chosen trip might cause delivery mistakes and even package losses. Therefore, we assume that no package could be unloaded unless it reaches its target station.

2) RATIONALE 2: THE PACKAGES OF THE SAME TARGET STATION SHOULD BE UNLOADED SUCCESSIVELY IN ADJACENT TRIPS

We should note that, it requires more manpower input and longer working hours at each station when the packages of the same target station are allocated into separate trips discontinuously. The reason is that, the workers at each station of the route will wait for package arrivals until their freighters are fully loaded, then they begin to distribute the packages to the customers. Therefore, we hope the packages of the same target station can be unloaded successively in adjacent trips. To do this, we generate a Package Flow that ranks the packages based on target stations. When any available trip begins, the packages would be loaded in the sequence of the Package Flow, until the loaded volume reaches the trip's available capacity.

TABLE 3. Notations of model formulation.

Notations	
Indices:	
i/j	index of stations;
k	index of packages;
r	index of trips;
Sets:	
N	the set of stations;
R	the set of available trips for package distributions;
K	the set of packages ranked by their target stations;
Parameters:	
Q_j^r	the under-utilized capacity for package delivery between station j and station $j+1$ in trip r;
v_k	the volume of package k;
a_{ki}	1 if package k should be delivered to station i, else 0;
Decision variables:	
x_{kr}	binary variable that denotes whether package k is delivered to its target station i_k in trip r;
y_r	binary variable that denotes whether trip r is utilized for package delivery;

3) MODEL FORMULATION

The notations used in the model description are presented in Table 3. Accordingly, the model can be developed as follows.

$$\min \sum\nolimits_{r\in R} y_r \tag{1}$$

$$s.t. \sum\nolimits_{r\in R} x_{kr} = 1 \quad \forall k \in K, \tag{2}$$

$$\sum\nolimits_{k\in K} \sum\nolimits_{i\in N \& i>j} a_{ki} x_{kr} v_k \le y_r Q_j^r \quad \forall r \in R, \tag{3}$$

$$y_{r_2} \le y_{r_1} \quad \forall r_1,\ r_2 \in R \;\&\; r_1 \le r_2, \tag{4}$$

$$x_{k_2 r_2} \le x_{k_1 r_1}$$
$$\forall r_1 \in R, \exists r_2 \in R \;\&\; r_1 \le r_2;\ \forall k_1,\ k_2 \in K \;\&\; k_1 \le k_2, \tag{5}$$

$$y_r \in \{0,1\} \quad \forall r \in R, \tag{6}$$

$$a_{ki} \in \{0,1\} \quad \forall k \in K,\ i \in N, \tag{7}$$

$$x_{kr} \in \{0,1\} \quad \forall k \in K,\ r \in R. \tag{8}$$

The optimization objective (1) minimizes the number of trips utilized for package delivery. Constraint (2) ensures that all packages are delivered to their target stations and each package can only be delivered by one trip. Constraint (3) ensures that the volume of loaded packages between any two adjacent stations of a trip is no more than the corresponding under-utilized capacity. Constraint (4) ensures that all available trips should be utilized in a time sequence. This constraint minimizes the time consumption of the task completion, which is important for *same-day delivery* services. Constraint (5) ensures that all packages at the departure station are loaded on the trips in a sequence of the Package Flow, which is described in Rationale 1. Specifically, the package with a farther target station should be assigned to a later trip. Finally, constraints (6)-(8) define the domains of the variables.

IV. ESTIMATION OF THE UNDER-UTILIZED CAPACITY IN A TRIP

To utilize CPTS for package distributions, the first important challenge is to estimate the amount of under-utilized capacity between any two adjacent stations in a trip. The amount of under-utilized capacity is defined as the space of unoccupied seats. It is under a direct relationship with the number of passengers getting on/off at each station in a given trip. Therefore, we set Passenger Transit Model in Section IV-A, to compute the number of passengers between each two adjacent stations in a trip. With that, the under-utilized capacity of a CPTS trip can be calculated, which is presented in Section IV-B.

A. PASSENGER TRANSIT MODEL

Given a route with N stations, we firstly define a passenger vector, which is represented by $S := [S_1, S_2, \ldots, S_N]^T$. The element S_i denotes the number of passengers who travel from station i in a whole day. To describe the destinations of such passengers, we set a passenger partition matrix. The matrix is represented as H, with N rows and N columns. Each element $H_{i,j}$ in H refers to the fraction of passenger who travel from station i to station j, where $0 \leq H_{i,j} \leq 1$ and $\sum_{j=i+1}^{N} H_{i,j} = 1$. The population partition matrix of any route can be represented as follows.

$$H := \begin{bmatrix} 0 & H_{1,2} & H_{1,3} & \ldots & H_{1,N} \\ 0 & 0 & H_{2,3} & \ldots & H_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & H_{N-1,N} \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

With the vector S and the matrix H, we can compute the number of passengers who travel along the route starting from station i and ending at station j daily, that is $S_i H_{i,j}$. Note that, $H_{i,j} = 0$ in any route when $i \geq j$, because only the passengers travel on the route in the forwarding direction are considered.

Besides the number of passengers travelling on the route from station i to station j daily, we need to know approximately when they get on and off the CPTS vehicles. Therefore, we define a departure distribution matrix V, which represents the probability density function of the departure time distribution for a given passenger partition matrix H. The departure distribution matrix V can be given by:

$$V := \begin{bmatrix} 0 & V_{1,2}(t) & V_{1,3}(t) & \ldots & V_{1,N}(t) \\ 0 & 0 & V_{2,3}(t) & \ldots & V_{2,N}(t) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & V_{N-1,N}(t) \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix},$$

where $V_{i,j}(t)$ represents the probability density function of the departure time distribution of passengers travelling from station i to station j at time t. Note that the distributions should be converted to continuous probability density functions for our analysis.

The values in S, H, and V can be estimated from passenger surveys or bus/subway data (such as the number of passengers entering and exiting a station). Note that, the data of passengers taking on or off the subway is easy to get, because the subway system has stored the information when people travel by subway. However, the data may be incomplete for bus routes, because the exact traveling record for some passengers paying by coins is sometimes difficult to collect. A feasible way to address this difficulty is to introduce facial recognition systems in the bus, which can help collect detailer passenger traveling data in the route. This issue is out of the scope of this paper.

In any time period, there is an increasing number of passengers at station i, who wait for the arrival of a new trip of the given route. We let $\hat{P}^r_{i,j}$ denote the expected number of passengers who wait at station i and aim at station j, in a given route r. Then, $\hat{P}^r_{i,j}$ can be given by:

$$\hat{P}^r_{i,j} = S_i H_{i,j} \int_{t_i^{r-1}}^{t_i^r} V_{i,j}(t)dt. \tag{9}$$

where $1 \leq i < j \leq N$. Therefore, the number of accumulative passengers waiting at station i between the $r-1^{th}$ trip and the r^{th} trip is denoted as $\hat{P}^r_i$, and can be calculated as:

$$\hat{P}^r_i = \sum_{j=i+1}^{N} \hat{P}^r_{i,j} = \sum_{j=i+1}^{N} S_i H_{i,j} \int_{t_i^{r-1}}^{t_i^r} V_{i,j}(t)dt. \tag{10}$$

Similarly, the expected number of passengers who get off the route at station j in the r^{th} trip ($\check{P}^r_j$) is estimated as follows.

$$\check{P}^r_j = \sum_{i=1}^{j-1} \hat{P}^r_{i,j} = \sum_{i=1}^{j-1} S_i H_{i,j} \int_{t_i^{r-1}}^{t_i^r} V_{i,j}(t)dt. \tag{11}$$

With Equation (10) and Equation (11), we can obtain the expected number of passengers between stations i and $i+1$ on the r^{th} trip. It is denoted as $\bar{P}^r_j$, and can be given by:

$$\bar{P}^r_i = \sum_{j=1}^{i} (\hat{P}^r_j - \check{P}^r_j). \tag{12}$$

So far, with the above Passenger Transit Model, we can obtain the expected number of passengers between any two adjacent stations in a specific trip by Equation (9) $\sim$ Equation (12).

B. CALCULATION OF THE UNDER-UTILIZED CAPACITY BETWEEN EACH PAIR OF ADJACENT STATIONS IN A TRIP

After predicting the number of passengers between each two adjacent stations of a trip, we can compute the corresponding under-utilized capacity that can be utilized for package distributions. We consider the space of the unoccupied seat as the under-utilized capacity. Therefore, the delivery task cannot be implemented in a trip, where all seats are occupied for passengers. The capacity of each seat is denoted as β in a CPTS vehicle. We assume that there are κ seats in that vehicle. Therefore, the total capacity in a CPTS vehicle is $C = \beta \times \kappa$. The total capacity of the CPTS vehicle is made up of the capacity for passengers and the under-utilized capacity for packages. Let $\bar{P} = \{\bar{P}_1, \bar{P}_2, \ldots, \bar{P}_{N-1}\}$ denote the number of passengers between each two adjacent stations in a trip, and $Q = \{Q_1, Q_2, \ldots, Q_{N-1}\}$ denote the corresponding under-utilized capacity. Then, the under-utilized capacity between each two adjacent stations in a trip can be given by:

$$Q_i = \begin{cases} \beta \times (\kappa - \bar{P}_i), & \bar{P}_i \leq \kappa, \\ 0, & \bar{P}_i > \kappa. \end{cases} \tag{13}$$

where $1 \leq i \leq N-1$. As Figure 3 shows, the big rectangle can be considered as the total capacity of a CPTS vehicle, which consists of two parts of capacity. The capacity dedicated to passengers is represented as the orange striped blocks in the figure, and the under-utilized capacity available for packages is denoted as the blue vertical striped blocks. Note that, the packages can only be distributed by utilizing the under-utilized capacity of a trip to avoid the impact on the quality of passenger experience.

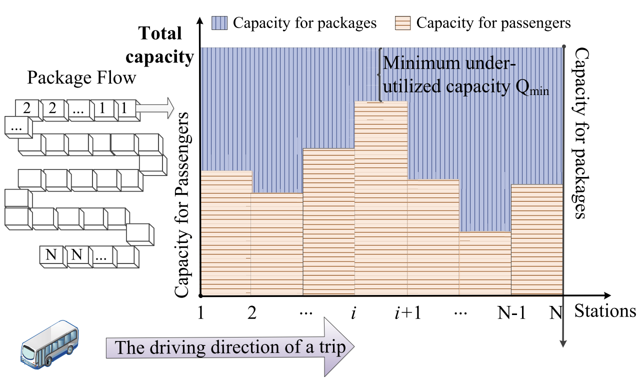

FIGURE 3. The illustrative examples of the capacity for passengers and under-utilized capacity for packages.

V. PACKAGE LOADING SCHEDULES

With the calculated under-utilized capacity in each trip described in Section. IV, we develop two package loading methods in this section.

A. MLD: THE MINIMUM LIMITATION DELIVERY METHOD

For fear of the situation when the volume of packages exceeds the amount of under-utilized capacity, MLD propose a particular case of the model shown in Section. III-C. Specifically, MLD uses the minimum amount of under-utilized capacity of all involved stations in a trip as the maximum capacity to be actually utilized for package distributions. Consequently, MLD can avoid negative impact on QoPE effectively. For convenience, we let Q_{min} denote the minimum amount of under-utilized capacity of a trip. As Figure 3 shows, when the trip travels from station i to $i+1$, the vehicle suffers from the busiest time and the amount of under-utilized capacity is Q_{min}. Our MLD method only loads packages with a volume no more than Q_{min} in that trip at the departure station.

To complete the package delivery task, we should estimate the minimum under-utilized capacity (Q_{min}) of the following upcoming trips. Let Q^r_{min} represent the under-utilized capacity of the r^{th} upcoming available trip of the chosen route. Thus, the only different constraint is Constraints (3), compared to the model in our problem statement. Constraint (3) could be changed as follows in our MLD:

$$\sum_{k \in K} \sum_{i \in N} x_{kr} a_{ki} v_k \leq y_r Q^r_{min}, \quad \forall r \in R. \tag{14}$$

This constraint ensures that the total volume of packages to be delivered via trip r is no more than the minimum under-utilized capacity of this trip (Q^r_{min}). This makes our MLD method not impact on the quality of passenger experience.

B. ALD: THE ADAPTIVE LIMITATION DELIVERY METHOD

Although MLD can avoid the negative impact on QoPE greatly, there exists considerable under-utilized capacity that has not been fully exploited at most stations. For this reason, we propose ALD, which load as many packages as possible in each trip, until the volume of loaded packages reaches the available under-utilized capacity. We account for the basic idea with the example in Figure 3. The total volume of loaded packages at station 1 is at most Q_1. After passengers getting on and off, as well as packages being unloaded, the remained volume of the loaded packages could not exceed Q_2 at station 2. The constraints at the rest stations can be satisfied in the same manner.

For an upcoming trip r, the expected under-utilized capacity at each station is denoted as $Q^r = \{Q^r_1, Q^r_2, \ldots, Q^r_{N-1}\}$. Thus, the minimum number of utilized trips for package delivery could be calculated through the model proposed in Section. III-C, with Equation (1)~(8).

VI. THEORETICAL PERFORMANCE ANALYSIS OF MLD AND ALD METHODS

One important performance matric of the methods we proposed is the impact rate. This kind of negative impact appears when the quality of passenger experience is influenced, because the pre-determined volume of loaded packages exceeds the realistic amount of under-utilized capacity. The impact rate refers to the rate when the negative impact appears at any station in a trip. Note that the package loading schemes are made based on the expected number of passengers at each station in each trip. However, the estimated value may violate the real value. In this section, we first calculate the probability distribution of the realistic number of passengers in Section VI-A. Then, we analyze the conflict between the capacity for packages and the capacity for passengers by calculating the theoretical impact rate in Section VI-B.

A. PROBABILITY DISTRIBUTION OF THE NUMBER OF PASSENGERS

For a specific trip in a route, we let P_i denote the actual number of passengers when the trip travels from station i to $i+1$. Let P^u_j and P^d_j denote the actual number of passengers who get on or off the bus/subway when the trip stops at station j, respectively. Therefore, P_i can be given by:

$$P_i = \sum_{j=1}^{i}(P^u_j - P^d_j). \tag{15}$$

where we have $P^u_j, P^d_j \geqslant 0$ for any station. The passenger flows taking on or off a bus/subway can be modeled as the Poisson point process, which is widely used in customer-server queue systems [30]. In our context, the process may not be stationary as the arrival rate of passengers changes over time. Nevertheless, for any sub-process in the time period from t^r_i to t^{r+1}_i, the arrival rate can be thought to be constant. This assumption would not bring an effect on the result. Thus, a Poisson process is applicable in the probability distribution of the number of passengers in our SPDCP problem [31]. Therefore, the probability that there are exactly m passengers getting on the bus at station j can be given by:

$$P_{P^u_j}(m) = (\hat{P}_j)^m e^{-\hat{P}_j}/m!. \tag{16}$$

where $0 \le m \le S_j$, and $\hat{P}_j$ refers to the expected value of P^u_j. Similarly, the probability of $P^d_j = m$ can be given by:

$$P_{P^d_j}(m) = (\check{P}_j)^m e^{-\check{P}_j}/m!. \tag{17}$$

Therefore, $P_{P_i}(m)$, the probability that there are exactly m passengers on the bus/subway when the trip travels from station i to station $i+1$, can be derived from recursion with P_{i-1}, P^u_i and P^d_i. Specifically, the probability of $P_i = m$ $(1 \le i \le N)$ can be given by:

$$\begin{aligned} P_{P_1}(m) &= P_{P^u_1}(m) = (\hat{P}_1)^m e^{-\hat{P}_1}/m! \\ P_{P_i}(m) &= \Sigma_{m_1=0}^{S_1+\ldots+S_{i-1}} \Sigma_{m_2=m-m_1}^{S_i} P_{P_{i-1}}(m_1) P_{P^u_i}(m_2) \\ &\quad \times P_{P^d_i}(m_1+m_2-m). \end{aligned} \tag{18}$$

B. THEORETICAL COMPARISON OF THE IMPACT RATE

Given a set of packages, the volume of loaded packages with different target stations in each trip should be constrained using the proposed methods. For convenience, we let x^M_{kr} and x^A_{kr} represent the binary variables, which denote whether the package k is assigned to be loaded on the r^{th} trip using MLD and ALD methods, separately. The package loading schemes can avoid negative impact on the quality of passenger experience in theory. However, when the trip travels, the realistic number of passengers may mismatch the estimated value. This may arise from the capacity conflict between passengers and packages. For example, obeying the determined package loading schemes of MLD, when the r^{th} trip leaves from station i, the volume of packages loaded on the CPTS vehicle is $\sum_{k\in K}\sum_{j=i+1}^{N} x^M_{kr} a_{kj} v_k$. Accordingly, as long as the capacity allocated to passengers does not exceed $C - \sum_{k\in K}\sum_{j=i+1}^{N} x^M_{kr} a_{kj} v_k$ at any station i in the trip r, the capacity of the trip is sufficient for transmitting both passengers and packages. Otherwise, the negative impact on the quality of passenger experience would happen. The impact rate refers to the rate when the predetermined volume of loaded packages exceeds the realistic amount of under-utilized capacity at any station in a trip. Let O^r_M and O^r_A denote the impact rate in the r^{th} trip of MLD and ALD, respectively. Then, O^r_M and O^r_A can be given by:

$$O^r_M = 1 - \Pi_{i=1}^{N-1} \sum_{m=0}^{C-\sum_{k\in K}\sum_{j=i+1}^{N} x^M_{kr} a_{kj} v_k} P_i(m), \tag{19}$$

$$O^r_A = 1 - \Pi_{i=1}^{N-1} \sum_{m=0}^{C-\sum_{k\in K}\sum_{j=i+1}^{N} x^A_{kr} a_{kj} v_k} P_i(m). \tag{20}$$

where N denotes the number of stations of the trip and m is always set as an integer. We should note that $\sum_{k\in K} x^A_{kr}$ is always larger than $\sum_{k\in K} x^M_{kr}$ in any trip r, because the ALD utilizes the under-utilized capacity more efficiently. Therefore, $O^r_A \ge O^r_M$.

We further consider the worst situation, where all packages aim at the last station of the route. In other words, $\sum_{k\in K}\sum_{i=0}^{N} x_{kr} a_{ki} = \sum_{k\in K} x_{kr} a_{kN}$. In this situation, the capacity suffers from an extreme occupation. Note that, when packages are unloaded at their target stations along the trip, the volume of residual loaded packages would decrease. This would reduce space occupation and avoid capacity conflict. However, in the worst situation, the packages would not be unloaded unless the trip arrives at the last station. Thus, the corresponding under-utilized capacity would be occupied from the first station to the last station. For the MLD method, the volume of loaded packages at any station in trip r is Q^r_{min} in the worst situation. Therefore, if the realistic capacity for passengers exceeds $C - Q^r_{min}$ at any station, the capacity would suffer a conflict between passengers and packages, and some passengers would be influenced. Interestingly, the impact rate of the ALD method is as same as that of the MLD method in the worst situation. It is because that, the minimum under-utilized capacity of a trip limits the volume of loaded packages at the departure station in the ALD method. We can make a formulation explanation as follows. When all packages end at the last station, Equation (3) can be specified as follows:

$$\sum_{k\in K}\sum_{i\in N \& i>j} a_{ki} x_{kr} v_k = \sum_{k\in K} x_{kr} a_{kN} v_k \le y_r Q^r_j,$$
$$\forall r \in R,\ \forall i, j \in N \tag{21}$$

Note that station j in Equation (21) can be any station in the trip, including the station with the minimum amount of under-utilized capacity. Thus, in the worst situation, the maximum volume of loaded packages of the ALD method in trip r is also Q^r_{min}. Let $\widetilde{O^r_M}$ and $\widetilde{O^r_A}$ denote the impact rate in the worst situation in trip r under our MLD and ALD method, separately. Then the impact rates can be given by:

$$\widetilde{O^r_A} = \widetilde{O^r_M} = 1 - \Pi_{i=1}^{N-1} \sum_{m=0}^{C-Q^r_{min}} P_i(m). \tag{22}$$

VII. PERFORMANCE EVALUATION

In this section, we evaluate the performance of our methods of the SPDCP problem. We start with the experimental setting, and then conduct a performance comparison between MLD and ALD, in terms of the effectiveness and the impact.

A. EXPERIMENTAL SETTING

1) DATASET

We use the real-world datasets for the evaluation, i.e., the road network data and the traffic data in Changsha, China. The data sets do not provide any information about package deliveries and the real-time passenger flows. Therefore, we apply different mechanisms to emulate the two kinds of information. To emulate a package delivery task, we randomly generate a set of packages (K) and their target stations (i_k) in the road network. To emulate the passenger flows, we first randomly partition the passengers between each pair of stations and thus construct the passenger partition matrix H, using a uniform distribution; We then randomly set the number of passengers getting on the route at each station daily (S_i) as a value ranging from 200 to 800. For a CPTS vehicle, we assume the number of seats is 25. Each passenger occupies one seat. The volume of packages (v_k) is generated with a uniform distribution, whose average value is 1/4 seat.

2) PERFORMANCE METRICS

We adopt the following two metrics, effectiveness and impact, to evaluate the methods of the SPDCP problem. **Effectiveness.** Given a set of packages to be delivered, we adopt four terms to measure the effectiveness of our methods.

- The ratio of deliverable packages. The deliverable packages refer to the packages whose final destinations are within the distance constraint from the target stations in the given road network. Note that, the distance constraint should be set as a small value to facilitate the last mile delivery and satisfy *same-day delivery* service.
- The average delivering time of a package, which starts from the package being loaded at the departure station and ends at the package being unloaded at the target station. Although MLD loads fewer packages in a given trip compared to ALD, the average delivering time is the same as that of ALD at the same time slot, due to the same delivering path and traffic condition.
- The number of utilized trips for the delivery task. The package delivery task means delivering the given set of packages to their final destinations.
- The ratio in the number of utilized trips between MLD and ALD method, given a package delivery task. We call the ratio as the Trip Number Ratio, which is abbreviated as tnr, and derived from:

$$tnr = \frac{number\ of\ utilized\ trips\ of\ MLD}{number\ of\ utilized\ trips\ of\ ALD}. \tag{23}$$

Impact. We use three terms to measure the negative impact on the quality of passenger experience.

- The impact rate in a package delivery task, where

$$impact\ rate := \frac{number\ of\ impacted\ trips}{number\ of\ total\ trips}. \tag{24}$$

- The Number of Impacted Passengers at each station in a trip, which is abbreviated as *nip*
- The Ratio of Impacted Stations, which is abbreviated as *ris*, and derived from:

$$ris := \frac{number\ of\ impacted\ stations}{number\ of\ total\ stations}. \tag{25}$$

Our evaluation methodology is represented as follows. First, we generate parameter values as described in Section. VII-A1, and then build the traffic (S, H and $\Re$) and passenger (V) matrices. Thereafter, given a set of packages, we would plan package loading schemes using our MLD and ALD method, separately. With these schemes, we compare our methods from two aspects, i.e., effectiveness and impact. In the following subsections, we conduct large-scale experiments to evaluate our MLD and ALD methods. All of the experimental results are determined from 200 repeated experiments.

B. THE EFFECTIVENESS OF OUR PROPOSED METHODS

To begin with, we compare the performance of the methods we proposed in terms of effectiveness.

1) RATIO OF DELIVERABLE PACKAGES w.r.t NUMBER OF ROUTES

Figure 4(a) plots the ratio of deliverable packages in our SPDPT mode under a varied number of involved routes, with different distance constraints (300m, 500 m, 700 m). With the number of involved routes increases, more packages could be successfully delivered to their final destinations via our SPDPT mode. The root cause is that more involved Crowdsourced routes would enlarge the coverage ratio of the SPDCP system. Thus, for any package, the distance between the target station and the final destination would be reduced. Moreover, the ratio of deliverable packages grows up with the increase of the distance constraint. It is because that the larger distance constraint would cover more final destinations for a given target station.

2) AVERAGE DELIVERING TIME w.r.t TIME SLOT

Figure 4(b) plots the average delivering time for a set of packages appeared in different time slots in a day, with 6 routes being involved in the SPDPT mode. The results show

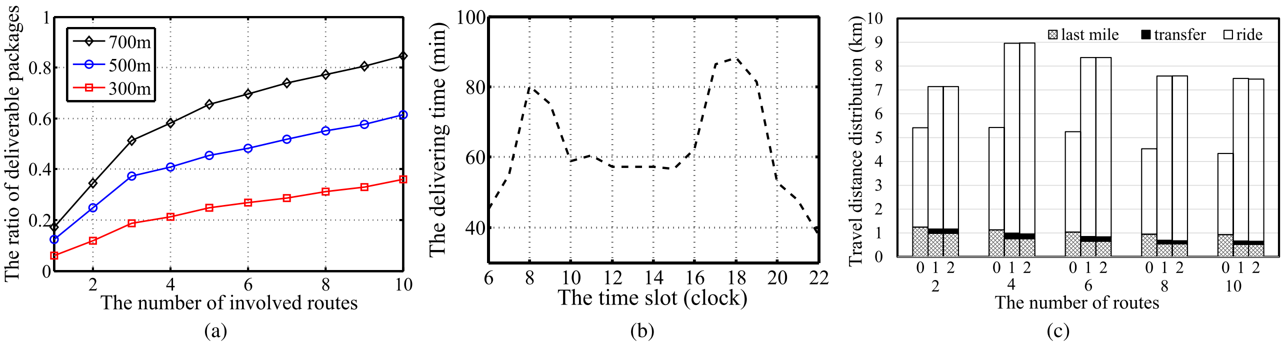

FIGURE 4. **The effectiveness of our proposed SPDPT, with a given number of packages to be delivered. (a) The ratio of deliverable packages under varied number of routes, with different distance constraints. (b) The average delivering time for a package in different time slots in a day (06:00-22:00). (c) The travel distance distribution under varied number of route, with different transfer times.**

that packages could be delivered within 1 hour in off-peak working hours (10:00-16:00), which can satisfy the *same-day delivery* service greatly.

3) TRAVEL DISTANCE DISTRIBUTION w.r.t TRANSFER TIMES

Figure 4(c) shows the travel distance distributions with different permitted times of package transfer. The travel distance of any package consists of three parts: 1) ride, the distance when the package on any bus; 2) transfer, the distance between the transfer station of ROUTE#1 to the transfer station of ROUTE#2; 3) last mile, the distance of a package being transmitted from its target station to its final destination. Note that, to decrease the distance of transferring, which may cause a huge amount of manpower resource, we select the transfer station of each pair of connected routes as the one that has the shortest distance to the other route. The figure shows that the last mile is the longest, when the package transfer is not permitted. It is because that one package could only utilize one specific route in this mechanism, which limits the coverage scale. Additionally, transfer-permitted mechanism consumes more resource in the riding process. It is because that the package transfer only appears between the dedicated transfer stations, which requires more riding distance for arriving the transfer stations. The total travel distance from the departure station to the final destination is the smallest when the number of routes is two. The root cause is that the packages can be transferred only once in this situation, which limits the distance of riding. What's more, the total travel distance becomes smaller with more involved routes (from 4 to 10). It is because that more available resources of Crowdsourced routes would trigger more efficient transmission strategy.

4) NUMBER OF UTILIZED TRIPS w.r.t NUMBER OF PACKAGES

Figure 5(a) plots the number of utilized trips in MLD and ALD, with respect to the number of packages, from 10:00 to 16:00. As the number of packages increases, the number of utilized trips grows up nearly linear. It is because that, when the package delivery task contains more packages, a larger

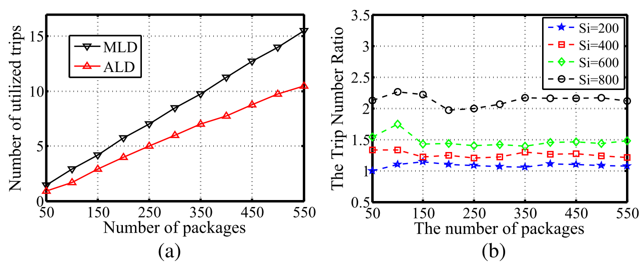

FIGURE 5. **The effectiveness of our proposed SPDPT, with a given number of packages to be delivered. (a) The number of utilized trips under varied number of packages in off-peak working hours (10:00-16:00). (b) The *tnr* under varied number of packages in off-peak working hours, with different number of passengers.**

amount of under-utilized capacity is required, resulting in more utilized trips. Moreover, the number of utilized trips in ALD is always larger than that in MLD. The reason is that, the scheduling in MLD inefficiently utilizes the under-utilized capacity of each trip compared to that of ALD. Thus, our MLD method utilizes more trips to accomplish the delivery task.

5) TRIP NUMBER RATIO w.r.t NUMBER OF PACKAGES

Figure 5(b) reports the *tnr* with different numbers of packages to be delivered in four cases, where the numbers of passengers at each station daily are set as $Si = 200, 400, 600, 800$, respectively. As the number of packages increases, the *tnr* of the four cases stay stable. It is because that, the number of utilized trips in both MLD and ALD grows up linearly with the number of packages. Thus, their ratio would not change whatever the given number of packages is. Moreover, *tnr* of the four cases are always larger than 1, and the case with a larger Si triggers a bigger *tnr*. The reason is that, as the increase of Si, the under-utilized capacity of MLD decreases much more quickly, resulting in a larger quotient value of the number of utilized trips.

C. THE NEGATIVE IMPACT ON THE QUALITY OF PASSENGER EXPERIENCE

Although our SPDCP mode could be utilized at any time when the public transportation systems operate, the efficiency

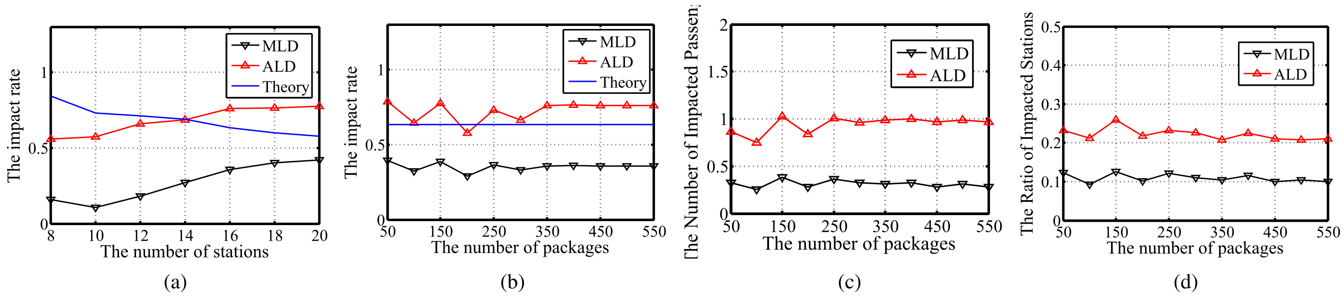

FIGURE 6. **The impact of our proposed SPDPT, with a given number of packages to be delivered. (a) The impact rate under varied number of stations in a route. (b) The impact rate under varied number of packages. (c) The changing trend of Number of Impacted Passengers with the number of delivered packages increases. (d) The changing trend of Ratio of Impacted Stations with the number of delivered packages increases.**

and feasibility are higher in off-peak hours. Therefore, we only analyze the impact of the SPDCP problem, i.e., the impact rate, *nip*, and *ris*, from 10:00 to 16:00 for practicality.

1) IMPACT RATE w.r.t NUMBER OF STATIONS IN A ROUTE

Figure 6(a) shows the theoretical results about the impact rate in the worst situation and the experimental results about the impact rate under MLD and ALD, with respect to the number of involved stations of a route, given a delivery task containing 550 packages. The theoretical impact rates are calculated using Equation (19) and Equation (22), while the experimental impact rates are derived from the simulations of passenger arrivals and departures. We can observe that the theoretical impact rate goes through a slow decrease as the route contains more stations, which can be explained with Equation (19). The experimental results of both MLD and ALD increase gradually with the increase of the involved stations. It is because that the number of passengers could be a large value at some stations in experiments, resulting in a small product of $\sum_{m=0}^{(C-Q_i)/\beta} P(m)$ and then a large impact rate. What's more, the experimental impact rate of ALD exceeds the theoretical impact rate when the number of stations exceeds 14.

2) IMPACT RATE w.r.t NUMBER OF PACKAGES

Figure 6(b) reports the theoretical results about the impact rate in the worst situation and the experimental results about the impact rate under MLD and ALD, with respect to the number of packages, in a route containing 16 stations. As the number of packages increases, the impact rates all fluctuate and end with stable values. The reason is that, each trip is a sample for computing the impact rate. Therefore, with more samples, the impact rate tends to be a more stable value. Moreover, the theoretical result is larger than the experimental result of MLD and smaller than that of ALD. The reason is that, the target stations of all packages in experiments of MLD are chosen randomly, which offers more under-utilized capacity for package deliveries, compared to the theoretical result where all packages aim at the last station. When the route involves a large number of stations, the number of passengers could be a large value at some stations in experiments, resulting in a small product of $\sum_{m=0}^{(C-Q_i)/\beta} P(m)$ and then a large impact rate.

3) NUMBER OF IMPACTED PASSENGERS & RATIO OF IMPACTED STATIONS w.r.t NUMBER OF PACKAGES

In Figure 6(c) and Figure 6(d), we plot the *nip* and *ris* with respect to the number of packages, respectively. As the increase of the number of packages to be delivered, the *nip* of both MLD and ALD fluctuate and end with stable values. The reason is that, when the delivery task contains more packages, more samples for computing the average *nip* are offered. Thus, the results tend to be more stable values. The changing trend of *ris* is similar to that of *nip*. The essential reason is that the impacted station appears when there exist passengers been influenced at that station. Moreover, the *nip* and *ris* of ALD are always larger than that of MLD, because the scheduling of MLD utilizes the under-utilized capacity of PTS vehicles more efficiently.

In summary, more than 60% packages are deliverable, when the SPDCP mode contains 10 routes and the distance constraint is set as 500 m. The average delivering time is always less than 1 hour in off-peak hours, which satisfies the *same-day delivery* service greatly. Given a set of packages to be delivered in off-peak working hours, the scheduling of ALD utilizes two-thirds of the trips to complete the task, compared to that of MLD. However, ALD may cause a higher risk of impacting the quality of passenger experience. Specifically, using the scheduling of MLD, at most 0.48 passengers lose their seats due to the package deliveries at any station, and at most 14.7% stations are impacted. In the scheduling of ALD, there are at most 1.03 passengers and 20.4% stations being impacted, separately. Such results indicate that the impact of inaccurate passenger estimation is not severe and is acceptable in the nowadays public transportation system.

VIII. CONCLUSION AND FUTURE WORK

In this paper, we present a novel framework called the SPDCP problem. The framework proposes to exploit the under-utilized capacity of CPTS vehicles to deliver packages to their final destinations, without degrading QoPE. We first propose a Passenger Transit Model, through which we can estimate the amount of under-utilized capacity in each trip. Thereafter, we develop two package loading methods, MLD and ALD, which utilize the minimum amount of under-utilized capacity and all available amount of under-utilized capacity in a trip, separately. We further give a theoretical analysis of the caused impact rate. Finally, we evaluate the performance of our methods and find that compared to MLD, ALD is more efficient and effective for package distributions, at the loss of a little higher risk of impacting QoPE.

It is the first attempt to investigate the package loading problem under the SPDCP framework, and this topic can be further extended in several directions. First, some design rationales for package loading would be revisited. For instance, we assume that package cannot be unloaded unless it arrives at its target station. This may not gain the highest efficiency in accomplishing the package delivery tasks. Thus, more flexible and efficient package loading rationales need to be studied. Second, it is a meaningful work for real-world case study and benchmark instance design, which can be utilized to test different models and methods in this field. For the difficulty of gaining the real-world datasets of passengers on the bus routes, a feasible way to overcome is to introduce facial recognition systems in the bus, which can help collect detailer passenger traveling data in the route. Finally, the first and third stage of the SPDCP mode should also be taken into consideration. Specifically, for the first stage, the SPDCP mode should provide a schedule to deliver the clustered packages from the Consolidation and Distribution Center to the departure stations, with the smallest time consumption. For the third stage, designing the transmitting path of each city freighter is meaningful, so that the transmission task can be completed within the limited time constraint, with the minimum number of city freighters.

REFERENCES

[1] R. Masson, A. Trentini, F. Lehuédé, N. Malhéné, O. Péton, and H. Tlahig, "Optimization of a city logistics transportation system with mixed passengers and goods," *Euro J. Transp., Logistics*, vol. 6, no. 1, pp. 81–109, 2017.

[2] R. Lowe and M. Rigby, "The last mile exploring the online purchasing and delivery journey," Barclays, London, U.K., Tech. Rep., 2014.

[3] S. A. Voccia, A. M. Campbell, and B. W. Thomas, "The same-day delivery problem for online purchases," *Transp. Sci.*, 2017.

[4] N. Geroliminis and C. F. Daganzo, "A review of green logistics schemes used in cities around the world," UC Berkeley, Berkeley, CA, USA, Tech. Rep., 2005.

[5] J. Jin. (2017). *China's Express Delivery comes to the Era of 100 Million a Day*. [Online]. Available: http://china.zjol.com.cn/ktx/201801/t20180108_6278004.shtml

[6] C. Chen *et al.*, "crowddeliver: Planning city-wide package delivery paths leveraging the crowd of taxis," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 6, pp. 1478–1496, Jun. 2017.

[7] J. S. G. Cheng, D. Guo, and Y. Qin, "When packages ride a bus: Towards efficient city-wide package distribution," in *Proc. Int. Conf. Parallel Distrib. Syst.*, 2018, p. 1.

[8] A. Trentini, R. Masson, F. Lehuédé, N. Malhéné, O. Péton, and H. Tlahig, "A shared 'passengers & goods' city logistics system," in *Proc. Int. Conf. Inf. Syst., Logistics Supply Chain*, 2012, p. 10.

[9] Z. Zeng, X. U. Weisheng, X. U. Zhiyu, and W. Shao, "Open-close mixed two-echelon vehicle routing problem in city logistics," *Inf. Control*, vol. 43, no. 6, pp. 744–749, 2014.

[10] T. G. Crainic, N. Ricciardi, and G. Storchi, "Models for evaluating and planning city logistics systems," *Transp. Sci.*, vol. 43, no. 4, pp. 432–454, 2009.

[11] N. Mathew, S. L. Smith, and S. L. Waslander, *Optimal Path Planning in Cooperative Heterogeneous Multi-Robot Delivery Systems*. New York, NY, USA: Springer, 2015.

[12] D. S. Setzke, C. Pflügler, M. Schreieck, S. Fröhlich, M. Wiesche, and H. Krcmar, "Matching drivers and transportation requests in crowdsourced delivery systems," in *Proc. Amer. Conf. Inf. Syst.*, 2017, pp. 10–12.

[13] A. Arslan, N. Agatz, L. Kroon, and R. Zuidwijk, *Crowdsourced Delivery: A Dynamic Pickup and Delivery Problem With Ad Hoc Drivers*. New York, NY, USA: Social Science Electronic Publishing, 2016.

[14] V. Paskalathis and S. N. Azhari, "Ant colony optimization on crowdsourced delivery trip consolidation," *Indonesian J. Comput. Cybern. Syst.*, vol. 11, no. 2, p. 109, 2017.

[15] A. Trentini, A. Campi, F. Boscacci, and N. Malhene, "Shared passengers and goods urban transport solutions," *Territorio*, vol. 7, pp. 38–44, 2011.

[16] E. Fatnassi, J. Chaouachi, and W. Klibi, "Planning and operating a shared goods and passengers on-demand rapid transit system for sustainable city-logistics," *Transp. Res. B, Methodol.*, vol. 81, pp. 440–460, Nov. 2015.

[17] Y. H. Cheng and S. Y. Chen, "Perceived accessibility, mobility, and connectivity of public transportation systems," *Transp. Res. A, Policy Pract.*, vol. 77, pp. 386–403, Jul. 2015.

[18] *Transportation in Sydney Trains*. Accessed: 2017. [Online]. Available: http://www.sydneytrains.info/

[19] Q. M. Ha, Y. Deville, Q. D. Pham, and M. H. Hà, "On the min-cost traveling salesman problem with drone," *Transp. Res. C, Emerg. Technol.*, vol. 86, pp. 597–621, Jan. 2015.

[20] *Central Bureau of Changsha Post District: Sorting More Than 1.2 Million Express Parcels a Day*. Accessed: 2018. [Online]. Available: http://hn.rednet.cn/c/2018/02/11/4554471.htm

[21] D. A. Hensher, C. Mulley, and N. Yahya, "Passenger experience with quality-enhanced bus service: The Tyne and wear 'superoute' services," *Transportation*, vol. 37, no. 2, pp. 239–256, 2010.

[22] Y. Chen, P. Lv, D. Guo, T. Zhou, and M. Xu, "Trajectory segment selection with limited budget in mobile crowd sensing," *Pervasive Mobile Comput.*, vol. 40, pp. 123–138, Sep. 2017.

[23] Y. Y. Chen, P. Lv, D. K. Guo, T. Q. Zhou, and M. Xu, "A survey on task and participant matching in mobile crowd sensing," *J. Comput. Sci., Technol.*, vol. 33, no. 4, pp. 768–791, 2018.

[24] Y. Chen, P. Lv, D. Guo, T. Zhou, and M. Xu, "ProSC+: Profit-driven participant selection in compressive mobile crowdsensing," in *Proc. IEEE/ACM Int. Symp. Qual. Service*, Jun. 2018, pp. 1–6.

[25] C. Shahabi, M. R. Kolahdouzan, and M. Sharifzadeh, *A Road Network Embedding Technique for K-Nearest Neighbor Search in Moving Object Databases*. Norwell, MA, USA: Kluwer, 2003.

[26] Flint Hills Area Transportation Agency, Inc. (2014). *ATA Annual Report*. [Online]. Available: http://www.rileycountyks.gov/DocumentCenter/View/11595

[27] R. Gallotti and M. Barthelemy, "The multilayer temporal network of public transport in great Britain," *Sci Data*, vol. 2, p. 140056, Jan. 2015.

[28] Z. Guo *et al.*, "'Last kilometer' distribution strategy research of the county logistics," *Open J. Social Sci.*, vol. 5, no. 5, pp. 251–262, 2017.

[29] X. Tan, Y. Shu, X. Lu, P. Cheng, and J. Chen, "Characterizing and modeling package dynamics in express shipping service network," in *Proc. IEEE Int. Congr. Big Data*, Jun. 2014, pp. 144–151.

[30] O. Kella and W. Stadje, "A clearing system with impatient passengers: Asymptotics and estimation in a bus stop model," *Queueing Syst.*, vol. 80, nos. 1–2, pp. 1–14, 2015.

[31] S. Siddappa, "Statistical modeling approach to airline revenue management with overbooking," *Ind. Manuf. Eng.*, 2007.

GEYAO CHENG received the B.S. degree in management science and engineering from the National University of Defense Technology, Changsha, China, in 2016, where she is currently pursuing the M.S. degree with the College of Systems Engineering. Her research interests include mobile computing and crowdsensing.

DEKE GUO (S'06–M'08) received the B.S. degree in industry engineering from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 2001, and the Ph.D. degree in management science and engineering from the National University of Defense Technology, Changsha, China, in 2008. He is currently a Professor with the College of Systems Engineering, National University of Defense Technology. His research interests include distributed systems, software-defined networking, data center networking, wireless and mobile systems, and interconnection networks. He is a member of ACM.

JIANMAI SHI received the Ph.D. degree in management science from the National University of Defense Technology (NUDT), Changsha, China, in 2011. He is currently an Associate Professor with the College of System Engineering, NUDT. His main research interests include heuristic algorithm design, reverse logistic, vehicle routing problems, and supply chain management.

YUDONG QIN received the B.S. degree in management science and engineering from the National University of Defense Technology, Changsha, China, in 2016, where he is currently pursuing the M.S. degree with the College of Systems Engineering. His research interests include data centers and software-defined networks.