

When Packages Ride a Bus: Towards Efficient City-wide Package Distribution

Geyao Cheng, Deke Guo, Jianmai Shi, Yudong Qin

Science and Technology on Information Systems Engineering Laboratory

National University of Defense Technology, Changsha Hunan 410073, P.R. China

Email: chenggeyaoyao13@nudt.edu.cn; guodeke@gmail.com; jshi1980@163.com; qinyudong12@nudt.edu.cn

Abstract—The demand for package delivery services is extremely huge every day, because of the rapid development on online retailers. This leads to huge traffic congestion, resource consumption and environmental pollution (e.g. carbon emission). In this paper, we investigate the City-wide Package Distribution problem using Crowdsourced Public Transportation Systems (CPDCP). That is packages are delivered to their final destinations within the city limits by utilizing the considerable amount of idle capacity of the Crowdsourced public transportation systems (CPTS). Specifically, given a number of packages and the timetable of available CPTS trips, we optimize the package delivering scheme by determining the four identified states of any package at any time slot (waiting, riding, re-waiting and being unloaded). The delivering scheme can be modeled as an instance of the multi-commodity flow problem, and formulated by the Mixed Integer Linear Programming techniques. We then propose an efficient heuristic solution for this NP-hard problem. Finally, our work is validated via comprehensive emulations of a city environment with a real bus transportation network.

Index Terms—package distribution; crowdsourced; public transportation system; idle capacity; package states.

I. INTRODUCTION

The rapid development on online retailers leads to the huge demand of package delivery services, which urges the sustainable development of logistics industry [1]. Specifically, Chinese postal enterprises have built more than 1,000 warehousing and distribution centers, and newly opened 153 trunk postal routes; the number of new energy vehicles has increased to 7,158 in 2017 [2]. Although the logistics grow significantly, these companies still face many challenges in the package distribution problem. One of the main challenges is providing convenient *same-day delivery* services [3] in an efficient way. However, the large number of packages to be delivered gives rise to the continuous increase of the consumed manpower and other resources [4]. Data shows that, online ordered products have generated over one billion package distributions in 2013, and this number is predicted to grow by 28.8% in 2018 [5]. Furthermore, the dedicated urban vehicles for logistics would incur serious air pollution and traffic congestion, especially when the traffic volume is huge [6]. Therefore, cities are looking for instruments and policies to ensure an efficient and effective urban transmission for both passengers and packages.

The package delivery problem has a trend to be solved by utilizing different transportation tools, such as unmanned aerial vehicles (UAV) [7], taxis [6], and private cars [8]. These researches aim at reducing the resource consumption and

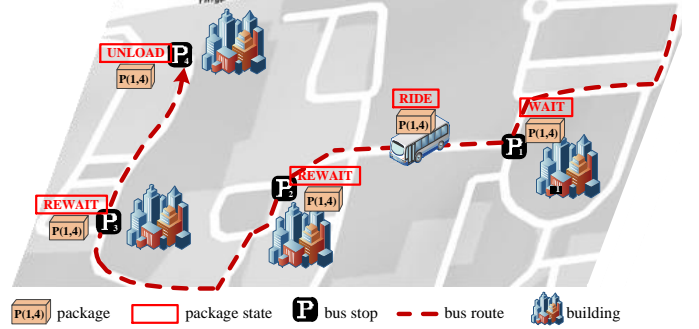


Fig. 1. An illustrative example of the package delivering state.

releasing traffic congestion. The first kind of efforts focuses on the mixed logistics [1] [9] [10] [11], where different kinds of vehicles would be successively utilized to realize the package deliveries, such as trucks & city freighters [12]. The second kind of efforts is based on Crowdsourced Delivery, where the package delivery tasks are outsourced to a non-specific mass network in a free and voluntary manner [8] [13].

Although many kinds of researches have been made to realize the efficient and effective urban transmission, we urgently need a way to solve the package distribution problem from a global perspective and achieve the overall advantage. Public Transportation Systems (PTS) [14] would be competent to accomplish the package delivery tasks. As a urban infrastructure, the PTS is stable, time-scheduled, economically friendly, and widely covered [14]. Moreover, it exhibits a huge amount of under-utilized capacity at off-peak hours, which gives probability to the capacity share between passengers and packages. Inspired by such observations, we propose a novel idea of realizing the City-wide Package Distribution using Crowdsourced Public Transportation Systems, named the CPDCP problem. It is the first attempt to deliver packages to their final destinations economically and ecologically, utilizing the under-utilized capacity of Crowdsourced Public Transportation Systems (CPTS) vehicles. Thus, passengers and packages would share the capacity of CPTS vehicles together, and the capacity utilization of the CPTS could be significantly enhanced.

The process of CPDCP problem usually consists of three stages. In the first stage, each package is allocated to a bus route which has the shortest sum distance to the package's original location and final destination. Obviously, the selected route must be in the CPTS, which have the cooperation with

the logistics companies. In the second stage, packages wait at the start stops of their selected routes and then ride a trip to their target stops. Specifically, when under-utilized trips come, the packages would be loaded in a determined sequence and finally be unloaded when arriving at their target stops. Note that in the delivering process, the packages could be unloaded at the intermediate bus stops so as to vacate capacity for ensuring passenger experience in the busy time. In the third stage, the local freighters would collect the unloaded packages and distribute them to their final destinations.

In this paper, we mainly focus on solving issues involved in the second stage. Specifically, given a set of packages to be delivered from a set of bus stops, we plan a scheme using the minimum number of continuous under-utilized trips such that their idle capacities are sufficient for delivering such packages. Any package may have four states in the delivering process. As Fig. 1 shows, a package, whose start stop is P_1 and target stop is P_4 , would *wait* at its start stop P_1 for a under-utilized trip, and then *ride* a bus trip from P_1 to P_2 . When the package is unloaded at intermediate stop (P_2 and P_3), its state becomes *re-wait* until the following under-utilized trip comes; finally, the state of the package would remain as *unloaded* after arriving at P_4 . We derive a mathematical model for packages being delivering to their final destinations by determining their states at any time slot. Our model takes into account the temporal/spatial availability of the amount of idle capacity in the given bus trips in the Crowdsourced platform. Thus, the experience of passengers would not be impacted. We model the scheduling task of a given set of packages as an instance of the multi-commodity flow problem, and formulated it using Mixed Integer Linear Programming techniques (MILP). therefore, propose an efficient heuristic solution for this NP-hard problem.

We conduct extensive evaluations to verify the efficiency and effectiveness of our solution for the CPDCP problem. Results show that more than 60% packages could be delivered to their target stops within 500m to their final destinations, when the Crowdsourced platform involves 10 routes. Considering 1000 packages on the Crowdsourced platforms of 10 routs, our heuristic method only occupies 7.5 trips in their off-peak working hours, and just consumes less than 4 hours to accomplish the delivering task.

The rest of this paper is organized as follows. Section II introduces the related work. Section III demonstrate the framework of our proposed CPDCP. In Section IV, we discuss the system model, the MILP-formulation, and the heuristic solution. Section V reports our experimental results. Finally, Section VI concludes this paper.

II. RELATED WORK

There is an increasing attention of researches aim to reduce the unnecessary resources consumption and air pollution caused by the dedicated transmission of packages. The researches can be divided as follows:

Mixed logistics. The first kind of efforts focuses on mixed logistics, where more than one kind of vehicles would be

successively utilized to realize the 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 then city freighters carry individual packages from the blocks to the specific delivery points in the region. There are still other kinds of vehicles used in mixed logistics, such as buses and city freighters [1] [9], close-open mixed two-echelon [10], and trucks and Micro Aerial Vehicles [12].

Crowdsourced delivery. The second kind of efforts focuses on Crowdsourced Delivery, where the tasks of package deliveries are outsourced to a non-specific mass network in a free and voluntary manner. Some efforts aimed at having packages take hitchhiking rides, such as taxis [6] and private cars [8]. For example, Chen et al. [6] suggested that voluntary taxi drivers would pick up the passengers in cooperating with collecting packages. There were some other efforts optimizing the assignment of package delivery tasks to drivers. Setzke et al. [13] optimized the assignment of packages to drivers, subject to transportation routes and time constraints. Arslan et al. [8] proposed a peer-to-peer platform to realize the matching between packages and drivers. Additionally, drivers were suggested to accept several requests in one trip [15] for less time and resource consumption, compared to those direct deliveries accepting one request in one trip.

There is an trend to utilize different kinds of transportation tools, such as unmanned aerial vehicles (UAV) [7] and taxies [6], to realize the package deliveries. These researches aim at reducing the resource consumption and releasing traffic congestion. However, it remains open to deliver packages utilizing the under-utilized capacity of the Crowdsourced public transportation system. A few proposes are literature [1] and [9], which mainly focused on the third stage in our proposed CPDCP problem in this paper. Masson et al. [1] optimized the handshaking process between the buses and the city freighters. The primary objective of literature [9] 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 issues involved in the second stage. Specifically, given a set of packages to be delivered, we plan a scheme using the minimum number of continuous under-utilized trips such that their idle capacities are sufficient for delivering such packages.

III. FRAMEWORK OF THE CROWDSOURCED PUBLIC TRANSPORTATION SYSTEMS

In this section, we start with definitions of some basic concepts, and then formally state the framework of CPDCP.

A. Basic concepts

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

Definition 1 (Route/Trip): We define a *route* as the travel sequence of a bus, linked by all involved bus stops alongside. A *trip* is a specific route starting at a scheduled time [16].

Definition 2 (Original location/Final destination): We define the place where a package locates before delivery as

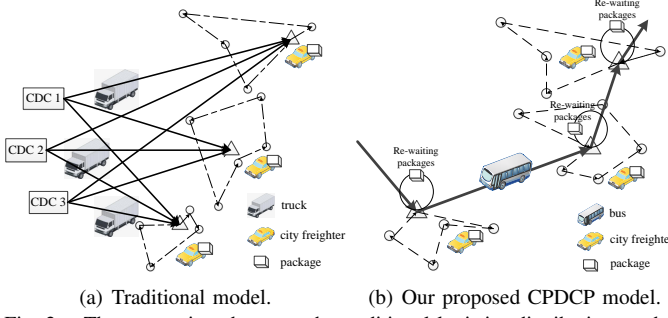


Fig. 2. The comparison between the traditional logistics distribution model and our proposed CPDCP model.

Original location, and the place where a package aims at as *Final destination*.

Definition 3 (Start stop/Target stop): With a selected route, the bus stop where the package is loaded on firstly is defined as *Start stop*, and the bus stop which the package is unloaded finally is defined as *Target stop*. Note that, *Start stop/Target stop* is selected as the one having the shortest distance to the package's *Original location/Final destination*, separately.

B. Framework of CPDCP

In the conventional logistic distribution model, as shown in Fig. 2(a), each logistic company would distribute packages from their subordinate CDC to their customers independently and separately, using their own vehicles (both trucks and local freighters). Therefore, it incurs serious traffic congestion and air pollution due to the large number of operating vehicles. However, our proposed CPDCP, as shown in Fig. 2(b), can solve such serious problem greatly.

The main idea of CPDCP is delivering packages from their *original locations* to their *final destinations* economically and ecologically, utilizing the considerable idle capacity of CPTS vehicles (from *start stop* to *target stop*). The process usually consists of three stages. In the first stage, packages are delivered to the start stops of their selected bus routes. For city-wide packages, couriers pick up the packages at each street/block, and then transmit the packages to nearby bus stops. For country-wide packages, the logistic companies sort them at Consolidation and Departure Center (CDC), and deliver them to the departure stop of the dedicated bus route. Obviously, the selected Crowdsourced bus routes must have cooperation with the logistics companies. Thereafter, the sorted packages would be transmitted to the nearby bus stops of their dedicated routes. In the second stage, packages ride the bus trips to their target stops. Specifically, when under-utilized trips of the selected route come, the packages would be loaded with a determined sequence, and finally, be unloaded when the trips arrive at their target stops. In the delivering process, the packages would be unloaded at the intermediate stops at a busy time, for ensuring the experience of passengers. As Fig. 2(b) shows, after the packages are unloaded at the intermediate stops, they would re-wait at these stops for the following upcoming trips. In the third stage, the local freighters would collect the unloaded packages and distribute them to their final destinations.

Compared to the existed logistics distribution model that maintains dedicated shipping crews, our CPDCP model has the following advantages. 1) Effectiveness. The bus stops 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 [17]. Moreover, the stops of CPTS have pervaded our cities with a high coverage rate. For example, the bus layer of Great Britain covers the largest fraction of the island [18]. Thus, most of the unloaded packages have a high probability to be transmitted to their destinations by city freighters. This facilitates the delivery tasks in the last kilometer [19]. 2) Sustained and stable operation. The bus is operated stably with a time schedule, whereas current services are less efficient during weekends, holidays, and days of bad weather [20]. As long as in the operational time, the packages would be loaded on the trips continuously and delivered to their target destinations, with a great feasibility. 3) Economic and environmental friendliness. Our proposed solution only leverages the under-utilized space of the CPTS vehicles for package distributions and induces little extra air and noise pollution [4]. Moreover, clustering packages from the CDCs of different logistic companies and delivering them to their destinations from a global perspective would reduce the total travel distance. This further results in less resource consumption, traffic consumption, and air pollution.

Additionally, the CPDCP incurs a low impact on the quality of passenger experience. Note that, we only utilize the under-utilized capacity of the CPTS vehicles for logistics. Thus, the seats for passengers would not be occupied. In addition, the original stopping time at each bus stop facilitates loading and unloading packages, causing less additional waiting time for passengers. Additionally, couriers, selected bus routes, selected bus trips, and the freighters all have an ID. Thus, packages are traceable in the whole delivery process, which guarantees the package security and avoids delivery mistakes or losses.

IV. SOLUTIONS OF THE CPDCP PROBLEM

In this section, we first define the system model and its main objectives for assigning each package to the specific trips of the selected route (Section. IV-A). Then we address the problem via Mixed Integer Linear Programming (MILP) in Section. IV-B. Finally, we provide an efficient heuristic solution (Section. IV-C).

A. CPDCP Problem Statement

Given a bus, the daily schedule at bus stop s is a nearly deterministic arrival (and departure) process. We consider a time slotted model $\xi_s = \{t_s^k : k=1,2,\dots\}$, where t_s^k represents the time slot when the k^{th} trip arrives at bus stop s . The length of each slot can be determined by history data about bus arrival and departure. The scenario can be defined as a tuple $\langle S, P, TP, BR \rangle$, where:

- $S = \{s_1, s_2, \dots, s_{|S|}\}$ is the set of bus stops that can be served by the Crowdsourced bus routes.

TABLE I
NOTATIONS OF MODEL FORMULATION

<i>Indices:</i>	
$s/o/d$	index of bus stops;
k	index of trips;
i	index of packages;
<i>Sets:</i>	
S	the set of bus stops;
P	the set of packages;
TP	the set of under-utilized trips for package distributions;
BR	the set of bus routes in the Crowdsourced platform;
<i>Parameters:</i>	
Q_s^k	the idle capacity for package delivery between bus stop s and bus stop $s+1$ in trip k ;
vol_{p_i}	the volume of package p_i ;
<i>Decision variables:</i>	
$x_{p_i}^{t_s^k}$	binary variable that denotes whether package p_i waits at its start stop at time t_s^k ;
$y_{p_i}^{t_s^k}$	binary variable that denotes whether package p_i rides a bus at time t_s^k ;
$z_{p_i}^{t_s^k}$	binary variable that denotes whether package p_i re-waits at an intermediate bus stop at time t_s^k ;
$v_{p_i}^{t_s^k}$	binary variable that denotes whether package p_i has been unloaded at its target stop at time t_s^k ;
w_s^k	binary variable that denotes whether the delivery tasks have been finished at time t_s^k ;

- $P = \{p_1, p_2, \dots, p_{|P|}\}$ represents the set of packages of a delivering task. Specifically, the information of each package can be presented as a tuple $p_i = p_{i,r,o,d}\{s,k,x,y,z,v\}$. Let o and d represent the corresponding start stop and target stop of the selected route r , separately; s represent the current located bus stop; k represent the upcoming trip of route r . Additionally, x, y, z, v represents four states of the delivering package, i.e., waiting, riding, re-waiting, and being unloaded, separately.
- $TP = \{k_1, k_2, \dots, k_{|TP|}\}$ represents the set of trips defined by the bus stops and arrival/departure time for a given bus route, which operates continuously in a time schedule.
- $BR = \{r_1, r_2, \dots, r_{|BR|}\}$ represents the set of bus routes. Note that, all of the bus routes in BR are willing to accept the package delivering tasks in a Crowdsourced platform.

Based on the previous definition, we assume that any package can be delivered to their final destinations using the bus route in the Crowdsourced platform. In the second stage of our proposed CPDCP problem, when the k^{th} upcoming trip of the selected route arrives at its s^{th} bus stop, i.e., at time slot t_s^k , each package p_i is in one of the following four states:

- waiting, i.e., staying at the start stop. We introduce the following binary state variable $x_{p_i}^{t_s^k} \in \{0, 1\} \forall p_i \in P, t_s^k \in \xi$, such that:

$$x_{p_i}^{t_s^k} = \begin{cases} 1 & \text{if } p_i \text{ stays at its start stop at time } t_s^k \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

- riding, i.e., riding a bus and aiming at its target stop. We introduce the following binary state variable $y_{p_i}^{t_s^k} \in$

$\{0, 1\} \forall p_i \in P, t_s^k \in \xi$, such that:

$$y_{p_i}^{t_s^k} = \begin{cases} 1 & \text{if package } p_i \text{ rides a bus at time } t_s^k \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

- re-waiting, i.e., being unloaded at the intermediate bus stop because of the space limitation, and then ride another trip to reach its target stop. We introduce the following binary state variable $z_{p_i}^{t_s^k} \in \{0, 1\} \forall p_i \in P, t_s^k \in \xi$, such that:

$$z_{p_i}^{t_s^k} = \begin{cases} 1 & \text{if } p_i \text{ stays at an intermediate bus stop at time } t_s^k \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

- being unloaded, i.e., having being unloaded at its target stop. We introduce the following binary state variable $v_{p_i}^{t_s^k} \in \{0, 1\} \forall p_i \in P, t_s^k \in \xi$, such that:

$$v_{p_i}^{t_s^k} = \begin{cases} 1 & \text{if } p_i \text{ has arrived at its target stop at time } t_s^k \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Let $T(p_i, t_s^k)$ be the consumed time of delivering package p_i at time t_s^k . Based on the state variables defined above, for any package p_i , $T(p_i, t_s^k)$ is updated at each $t_s^k \in \xi$ as follows:

$$\begin{aligned} T(p_i, t_s^k) = & (x_{p_i}^{t_s^{k-1}} + z_{p_i}^{t_s^{k-1}}) \cdot [T(p_i, t_s^{k-1}) + t_s^k - t_s^{k-1}] \\ & + y_{p_i}^{t_s^{k-1}} \cdot [T(p_i, t_s^{k-1}) + t_s^k - t_s^{k-1}] \\ & + v_{p_i}^{t_s^{k-1}} \cdot T(p_i, t_s^{k-1}) + v_{p_i}^{t_s^{k-1}} \cdot T(p_i, t_s^{k-1}) \end{aligned} \quad (5)$$

Based on the value of $T(p_i, t_s^k)$, we introduce an additional state-transfer-off decision variable $w_s^{t_s^k} \in \{0, 1\} \forall p_i \in P, t_s^k \in \xi$ such that:

$$w_s^{t_s^k} = \begin{cases} 1 & \text{if } v_{p_i}^{t_s^{k-1}} = 1 \text{ for any } p_i \in P \text{ at time } t_s^k \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

The delivering time DT is defined as the minimum time slot such that $DT > w_s^{t_s^k} \times t_s^k$. Informally, the CPDCP problem can be defined as: determine the variable statues $x_{p_i}^{t_s^k}, y_{p_i}^{t_s^k}, z_{p_i}^{t_s^k}$, and $v_{p_i}^{t_s^k}$ for any package p_i and any time slot t_s^k , so that DT is minimized. The constraint is that all packages are delivered to their target stops by riding the dedicated bus routes. In the following, we prove that this model is a special instance of the multi-commodity flow model, which can be formulated via MILP techniques.

B. MILP-based formulation of the CPDCP problem

We devise a multi-period directed multi-graph $G(V, A)$ in order to model the movement of packages across the start stop, target stop, and the intermediate stops. For a given bus route, the set of nodes V is composed by the couple (s, k) , $\forall s \in S, \forall k \in TP$. We denote the generic travel hop as an arc e_s represented by $[(s, k), (s', k')]_l$, where (s, k) and (s', k') are the tail and the head of the arc, respectively. Each arc also has an associated weight $W(e_s)$ and a subscript l that denotes one of the four possibilities:

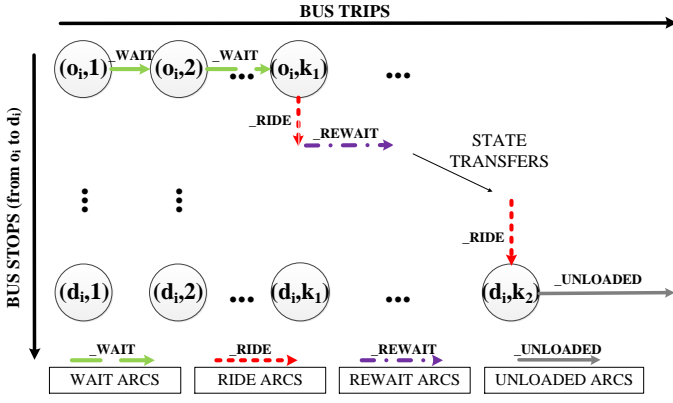


Fig. 3. An illustrative example of the state transfers when a package rides a bus from its start stop to its target stop.

- Case 1. *wait arc* $[(s, k), (s, k+1)]_{WAIT}$, where s is the start stop of a given package. This kind of arcs refers to the behavior of a package, which waits at its start stop when the k^{th} trip of the selected route comes. Here, the weight is given by the utilized time caused by waiting for the next trip, i.e., $W(e_s) = t_s^{k+1} - t_s^k$;
- Case 2. *ride arc* $[(s, k), (s+1, k)]_{RIDE}$. This kind of arcs models the behavior of a package riding a bus trip to its next stop. Here, the weight is given by the consumed time by riding the selected trip between two adjacent bus stops, i.e., $W(e_s) = t_{s+1}^k - t_s^k$;
- Case 3. *re-wait arc* $[(s, k), (s, k+1)]_{REWAIT}$, where s is an intermediate bus stop in package delivering. This kind of arcs models the behavior of a package being unloaded at bus stop s at busy time of trip k , and then re-waiting at that stop for the next upcoming trip $k+1$. Here, the weight is given by the utilized time caused by re-waiting for the next trip, i.e., $W(e_s) = t_s^{k+1} - t_s^k$;
- Case 4. *unloaded arc* $[(s, k), (s+1, k)]_{UNLOAD}$, where s is the target stop in the package delivering. This kind of arcs models the behavior of a package having been unloaded at its target stop, and would be further transmitted by city freighters. Here, the package delivering process has been terminated i.e., $W(e_s) = 0$.

Fig. 3 shows an example of a multi-period directed multi-graph built according to the rules above. Specifically, package p_i aims to be delivered from its start stop o_i to its target stop d_i . The package initially waits at bus stop o_i , when an under-utilized trip k_1 comes, it rides that trip to the following bus stops. If the trip suffers from peak hours (space limitation), the package would be unloaded at intermediate bus stops and re-wait for the following trips, for fear of impacting passengers. Finally, the package is unloaded at bus stop d_i from k_2 , where $k_2 \geq k_1$. Note that, a delivering path of a package can be presented as a continuous fold line and each sub-fold line directs right or down, because of the time sequence of the bus trips and the direction of the selected routes.

We model the action selection process of each package p_i through the following non-splittable flow variables:

$$\varphi_{[(s,k),(s',k')]_l}^{p_i} = \begin{cases} 1 & \text{if } p_i \text{ uses arc } [(s,k),(s',k')]_l \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

Here, we want to determine the flow variables $[(s,k),(s',k')]_l$ for all packages of any Crowdsourced route to minimize the delivering time $DT \in T$ so that $DT > t_s^k \cdot w_s^k$. Additionally, the volume of packages which ride a specific trip k between any two adjacent bus stops (s and $s+1$) is limited by the corresponding amount of idle capacity (Q_s^k). Thus, the following constraints must met:

$$\min t_s^k \cdot w_{s,k} \quad (8)$$

s.t.

$$\sum_{[(o_i,k),(o_i,k+1)]_l \in A} \varphi_{[(o_i,k),(o_i,k+1)]_{WAIT}}^{p_i} = x_{p_i,o_i,k} \quad \forall p_i \in P, o_i \in S, k \in TP \quad (9)$$

$$\sum_{[(s,k),(s+1,k)]_l \in A} \varphi_{[(s,k),(s+1,k)]_{RIDE}}^{p_i} = y_{p_i,s,k} \quad \forall p_i \in P, s \in S, k \in TP \quad (10)$$

$$\sum_{[(s,k),(s,k+1)]_l \in A} \varphi_{[(s,k),(s,k+1)]_{REWAIT}}^{p_i} = z_{p_i,s,k} \quad \forall p_i \in P, s \in S \& s \neq o_i, k \in TP \quad (11)$$

$$\sum_{[(d_i,k),(d_i,k')]_l \in A} \varphi_{[(d_i,k),(d_i,k')]_{UNLOAD}}^{p_i} = v_{p_i,d_i,k} \quad \forall p_i \in P, d_i \in S, k, k' \in TP \& k < k' \quad (12)$$

$$\sum_{k \in TP} y_{p_i,s,k} = 1 \quad \forall p_i \in P, s \in S \& s \in [o_i, d_i) \quad (13)$$

$$\sum_{s \in S, k \in TP} x_{p_i,s,k} + y_{p_i,s,k} + z_{p_i,s,k} + v_{p_i,s,k} \leq 1 \quad \forall p_i \in P \quad (14)$$

$$x_{p_i,o_i,k} \leq \sum_{k' \in TP \& k < k'} y_{p_i,o_i,k'} \quad \forall p_i \in P, o_i \in S, k \in TP \quad (15)$$

$$y_{p_i,s,k} \leq y_{p_i,s+1,k} + z_{p_i,s+1,k}$$

$$\forall p_i \in P, s \in S \& s \in [o_i, d_i - 1), k \in TP \quad (16)$$

$$z_{p_i,s,k} \leq y_{p_i,s,k+1} + z_{p_i,s,k+1}$$

$$\forall p_i \in P, s \in S \& s \in (o_i, d_i), k \in TP \quad (17)$$

$$y_{p_i,d_i-1,k} \leq y_{p_i,d_i,k'} \quad \forall p_i \in P, d_i \in S, k, k' \in TP \& k < k' \quad (18)$$

$$\sum_{p_i \in P} y_{p_i,s,k} \cdot vol_{p_i} \leq Q_s^k \quad \forall s \in S, k \in TP \quad (19)$$

$$v_{p_i,s,k} \leq v_{p_i,s+1,k}, v_{p_i,s,k} \leq v_{p_i,s,k+1}$$

$$\forall p_i \in P, s \in S, k \in TP \quad (20)$$

$$w_{s,k} \leq v_{p_i,s,k} \quad \forall p_i \in P, s \in S, k \in TP \quad (21)$$

$$w_{s,k} \leq w_{s+1,k}, w_{s,k} \leq w_{s,k+1} \quad \forall p_i \in P, s \in S, k \in TP \quad (22)$$

$$x_{p_i,s,k}, y_{p_i,s,k}, z_{p_i,s,k}, v_{p_i,s,k}, w_{s,k} \in \{0, 1\}$$

$$\forall p_i \in P, s \in S, k \in TP \quad (23)$$

The objective function (8) aims to minimize the consumed time for delivering the packages, i.e., starting from the first package is loaded on a trip and ending at all packages are unloaded at their target stops. The constraints could be explained as follows: Constraints (9)-(12) correspond the arcs in the multi-graph with the four kinds of states, separately; Constraint (13) ensures that each package p_i is delivered between any two adjacent stops by exactly one bus trip $k \in TP$. Constraint (14) says that at any time slot $t_s^k \in \xi$, package p_i only exhibits at most one state. Constraints (15)-(18) declare

the conversion relationships between the four package states: after *waiting* at the start stop o , package would ride an under-utilized trip k' to the following bus stops; arriving stop $s+1$ by *riding* trip k , packages would still ride that trip or be unloaded; *rewaiting* at stop s when trip $k+1$ arrives, packages would still *re-wait*, or *ride* that trip; after *riding* a trip to its target stop, the state of the package would be fixed as *being unloaded*; Constraint (19) confines the volume of loaded packages when any trip k travels from stop s to stop $s+1$, considering the limited idle capacity. Constraint (20) states that after being unloaded at time t_s^k , the state of that package would not change at its next time slot, i.e., t_{s+1}^k or t_s^{k+1} . Constraint (21) states that variable $w_{s,k}$ is only active when all packages are unloaded; (22) ensures that when the delivery task has been accomplished, the states of any packages would remain unchanged. Finally, constraint (23) defines the domains of the variables.

C. An efficient heuristic solution of the MILP-based model

The package delivering problem represents a complicated variant of an unsplittable multi-period multicommodity network design problem with side constraints, which is known to be NP-hard [21]. In order to further reduce the computational complexity, we propose an efficient heuristic approach that determines a suboptimal solution in polynomial time. Specifically, the heuristic solution can be addressed as a path planning problem using our constructed multi-period multicommodity network. Each package should pick a continuous fold line which starts from its start stop and ends at its target stop. The following designed rationales would facilitate us understanding the process of determining the fold lines for all packages.

Rationale 1: When an under-utilized trip arrives, packages would be loaded in a determined sequence until the loaded amount reaches the upper bound. In this situation, the idle capacity of any trip could be fully exploited and the package delivering system would have a great scalability in the number of packages.

Rationale 2: Re-waiting stage should be avoided to the greatest extent. The re-waiting stage incurs more load and unload, which increases the input of labor resources and the risk of package loss. Therefore, we propose a rational that re-waiting stage should be avoided for work convenience and package security.

With such rationales, we design our solution as Algorithm. 1 shows. Firstly, the traffic condition should be collected and evaluated before the delivering task begins (line 1 - 7). Then the packages would select their delivering bus routes, and be transmitted to the start stops (line 8 - 11). The under-utilized trip picks up the packages along the bus route, and drops off packages when 1) its capacity must be reserved for passengers; 2) the packages arrive at their target stops. All packages are loaded with a priority, which could affect the *delivering time*. The priority that incurs the least *delivering time* could be found through the Genetic algorithm [22] (line 13). In the package delivering process, all ranked packages wait at the start stops until their *selectedtrip* arrives. The *selectedtrip* of package p_i

Algorithm 1 Our efficient heuristic solution of the MILP-based model

```

1: for each arc  $[(s,k), (s',k')]$  in  $G(V,A)$  do
2:   initialize the arc weight
3: end for
4: for each arc  $[(s,k), (s+1,k)]$  in  $G(V,A)$  do
5:   initialize the idle capacity  $Q_s^k$ 
6:   initialize the loaded package volume  $V(s,k) = 0$ 
7: end for
8: for all packages in the delivering task do
9:   initialize the package volume
10:  determine the corresponding selected route
11: end for
12: while  $\Delta results > \sigma$  do
13:  determine package priorities by the Genetic algorithm
14:  for  $i = 1$  to  $|P|$  do
15:     $selectedtrip = 1$ 
16:    while any  $j$  in  $(o_i \text{ to } d_i)$  so that  $V(j, selectedtrip) + v_{p_i} \geq Q_j^{selectedtrip}$  do
17:       $selectedtrip++ = 1$ 
18:    end while
19:    update the loaded package volume  $V(s,k)$ 
20:  end for
21:   $result = \max\{selectedtrip\}$  of the  $|P|$  packages
22: end while
23: return solution

```

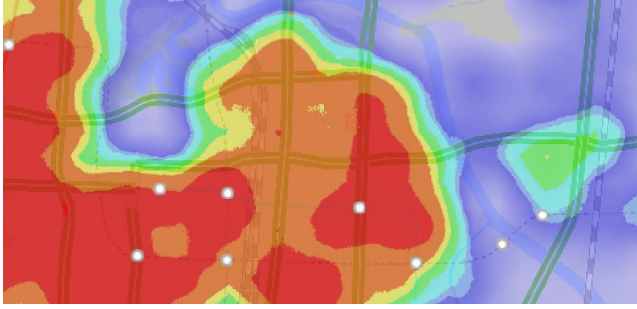
is determined as the one that still has idle capacity from its start stop o_i to its target stop d_i . Thereafter, the loaded volume of packages and idle space would be updated for each stop and each trip. In this way, the suboptimal solution of our CPDCP problem is found (line 12 - 23). The solution can also be extended into the traffic network with a great number of bus routes, considering transfers. The transfer stop of any two routes is fixed as the one that has the shortest distance to the other route. The time consumption of transfer consists of 1) transferring between the transfer stops of the two routes, 2) re-waiting at the transfer stop of the second route and re-waiting for an upcoming trip of that route.

V. PERFORMANCE EMULATION

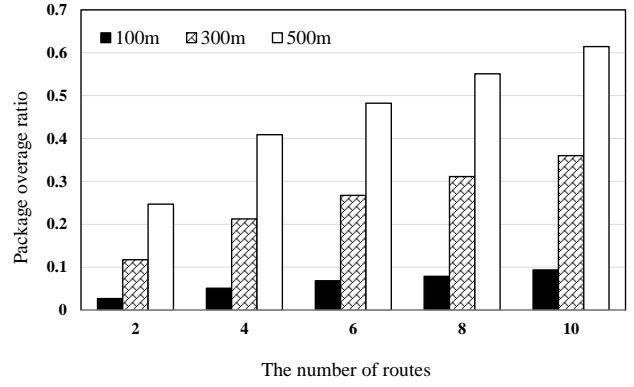
In this section, we empirically evaluate the performance of CPDCP. At first, we introduce the experimental settings used for demonstration in Section. V-A, including the experimental data, comparison algorithms, and evaluation metrics. Then we conduct performance emulations in Section V-B.

A. Experimental settings

1) *Experimental Data:* We use the real-world datasets for the evaluation, i.e., the road network data and the traffic data in the city of Changsha, China. This data is readily accessible using the published API from most city-run transportation systems. Since the datasets do not contain information about package deliveries and the real-time passenger flow, we apply different mechanisms to emulate these two kinds of information. To emulate a package delivery task, we randomly



(a) The population heat map at 10:00 pm in Furong District, Changsha, China.



(b) The package coverage ratio under varied number of route, with different visible distance (100m, 300m, 500m).

Fig. 4. The preliminary experimental studies of package coverage ratio and travel distance distribution.

generate a set of packages (K), with their original locations and final destinations being determined based on the population heat map (see Fig. 4(a)). To emulate the passenger flow, we randomly generate the number of waiting passengers at each evolved bus stop using the Poisson distribution. In a CPTS vehicle, we assume the number of seats is 25. Each passenger occupies 1 seat. The volume of packages (v_k) is generated with a uniform distribution, whose average value is $1/4$ seat. Δ , the system parameters related to the Genetic algorithm, is set as $\Delta = 1$. When the $\Delta results < 1$, the loop iteration terminates.

2) *Comparison Algorithms and Evaluation Metrics*: We complemented and compared the following algorithms:

No-rewaiting method (NRW). In this algorithm, no re-waiting stage is permitted. The priorities of packages are determined by the Genetic Algorithm.

Start-from-departure-stop method (SDT). In this algorithm, all of the collected citywide packages would firstly be delivered to the departure stop of the selected bus routes, for better global organization and dispatch.

Rewait-permitted methods with several priority policies. In this algorithm, the packages could be unloaded at the intermediate bus stops along the trip. The priority policies are:

- Genetic algorithm policy: packages are loaded in a priority determined by the Genetic Algorithm.
- Pri-distance policy: packages with short delivering distance have higher loading priority.
- Pri-start-stop policy: packages with a closer start stop to the departure stop have higher loading priority.

We rigorously analyze the above methods using the following performance metrics.

Delivering Time (DT): the utilized time starting from the first package being loaded from its start stop and ending at the last package being unloaded to its target stop, in a given package delivering task. In this paper, we convert DT into the number of utilized trips.

Package State Distribution (PSD): the average state distribution of the packages in the delivering task. The states of the delivering process include three parts: waiting at the original bus stop, re-waiting for upcoming under-utilized trips, and riding the corresponding trip.

The evaluation methodology is as follows. First, we generate

parameters as described in Section. V-A1 and then build the traffic (ξ) and passenger (Q) matrices. Given a set of packages, we plan package loading schemes based on our comparison algorithms, separately, and then emulate the performance from two aspects, i.e., DT and PSD.

B. Performance analysis

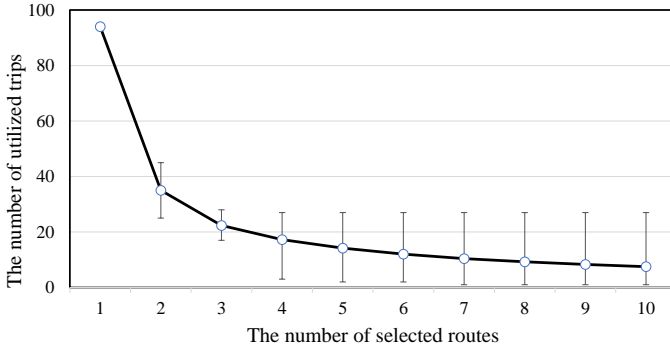
In this subsection, we conduct large-scale experiments to evaluate our CPDCP methods. We only extract traffic data at 12:00-12:30 am, since the bus trips at off-peak hours offer more idle capacity for package delivering. What's more, all of the experimental values are the average of 1000 repeated experiments.

1) *Preliminary experimental studies*: Firstly, we do some preliminary experimental studies in terms of package coverage ratio and travel distance distribution. Note that, the population heat map at 10:00 pm (see Fig. 4(a)) can represent the population distribution. Therefore, the packages' original locations and final destinations can be generated. The blocks/streets with more population would trigger more packages.

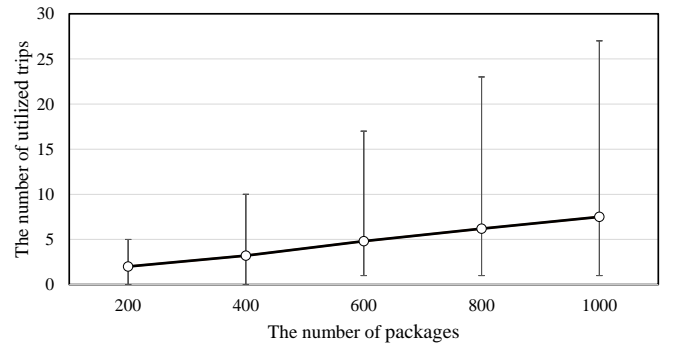
Package coverage ratio w.r.t Different visible distance: In our tests, the packages waiting at the bus stops are collected from the nearby streets and blocks within a visible distance. The visible distance is defined as the covering range of any bus stop. The packages can be delivered in our CPDCP mode when its original location is within the coverage range from a bus stop. Fig. 4(b) shows that the package coverage ratio grows up with the increase of visible distance. Moreover, when the number of utilized routes increases, more packages could be delivered by the public transportation system.

2) *DT: the number of utilized trips*: Then, we emulate the performance in terms of the number of utilized trips.

Number of utilized trips w.r.t Number of routes. Fig. 5(a) plots the number of utilized trips (average, max, and min), with respect to the number of involved routes. Note that, there are hundreds of bus routes in Changsha, but only part of them could be involved in our CPDCP model, considering the feasibility and effectiveness. This figure plots the routes with the highest coverage rate. Given the package delivering task with 1000 randomly generated packages, fewer trips would be utilized for each route on average, as the number of involved routes increases. It is because that the packages would

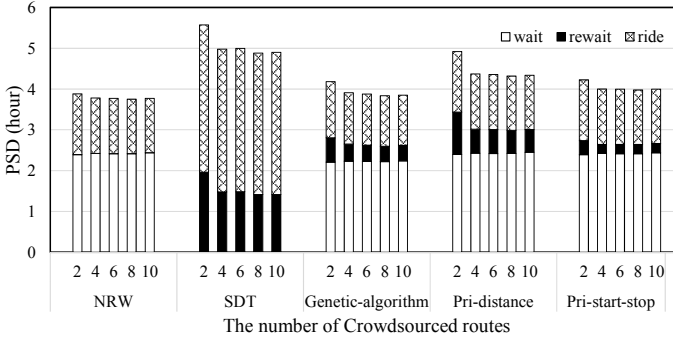


(a) Number of utilized trips under varied number of Crowdsourced routes (1000 packages).

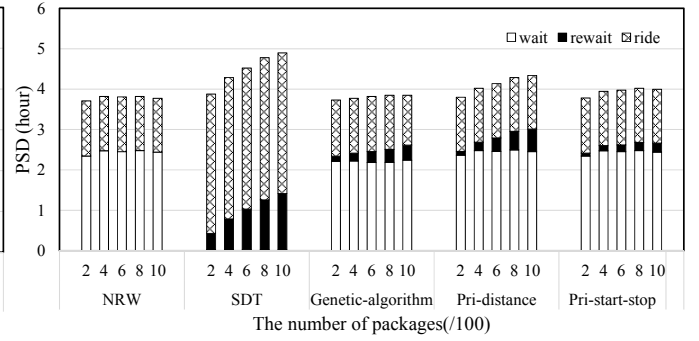


(b) Number of utilized trips of each route when there are 10 Crowdsourced routes in total (1000 packages).

Fig. 5. The performance of the number of utilized trips.



(a) The PSD under varied number of Crowdsourced routes.



(b) The PSD under varied number of packages.

Fig. 6. The package state distribution of different methods under varied number of Crowdsourced routes and varied number of packages.

be allocated to various routes. This decreases the resource requirement and workload for each evolved route.

Number of utilized trips w.r.t Number of packages. Fig. 5(b) plots the number of utilized trips (average, max, and min) of the 10 selected routes, with respect to the number of packages in the delivering task. When the delivering task contains more packages, the number of utilized trips grows linearly. The reason is that, more package deliveries demand more capacity resource in the utilized routes, leading to a larger number of utilized trips.

3) *PSD: package state distribution*: Finally, we compare the performance in terms of state distribution.

Package state distribution w.r.t Number of selected routes. Fig. 6(a) plots the PSD of several proposed methods, with respect to the number of involved routes ranging from 2 to 10. With the number of selected routes increases, the average utilized time for delivering a package decreases gradually. It is because the package delivering task is allocated to more routes, thus the workload of each route is relieved. In the SDT method, because all packages are delivered to the departure stop of the corresponding route firstly, 1) the time at wait stage is extremely short, because the packages can be loaded when a trip starts to operate; 2) the total utilized time increases a lot, due to the longer delivering distance. For the rewait-permitted methods with different priority schemes, the total utilized time is longer than that of NRW. It is because that in rewait-permitted methods, more packages would be loaded for

each trip because that the upper bound of the loaded amount is only confined by the idle capacity of the start stop. However, these packages have a high probability to be unloaded at the following stops for ensuring passenger experience. Therefore, the re-waiting stage would be extremely long waiting for the next trips.

Package state distribution w.r.t Number of packages. Fig. 6(b) plots PSD of several proposed methods with respect to the number of packages ranging from 200 to 1000. With the number of packages increases, the average utilized time for delivering a package increases, due to more demand for capacity resources. Moreover, for SDT method and rewait-permitted policies, the re-wait stage occupies more time when the delivering task contains more packages, because of the capacity resource contention.

In summary, the package delivering task can be accomplished satisfactorily by our proposed CPDCP model. More than 60% packages could be delivered to their final destinations within 500m around of their dedicated routes, when the CPTS involves 10 routes. Given 1000 packages to be delivered in off-peak working hours, our heuristic method only occupies 7.5 trips on average and utilizes less than 4 hours to accomplish the delivering task, while other methods consume more resources and time consumption.

VI. CONCLUSION AND FUTURE WORK

In this paper, we present a novel framework called CPDCP. The framework exploits the idle capacity of CPTS vehicles

to deliver packages to their destinations. We first identify the four states of a package, i.e., waiting at a bus stop, riding a bus to its destination, re-waiting at an intermediate bus stop to vacate space for passengers, and being unloaded when arriving at its final destination. Thereafter, we carefully plan the delivering scheme by determining the state of any package, to minimize the utilized time for delivering all packages. We mathematically formulate this problem using Mixed Integer Linear Programming techniques and propose an efficient heuristic solution for this NP-hard problem. In the evaluation, we compare our proposed methods and find that NRW method could accomplish the package delivering task in a more effective and efficient way.

This paper is the first attempt to investigate the package loading problem under the CPDCP framework. The research topic can be further extended in several directions. First, some design rationales for package delivering would be revisited. For instance, we assume that package should avoid re-waiting for less labor resource input and package loss in the heuristic solution, but this assumption may not gain the highest efficiency in accomplishing the package delivery tasks. Second, the proposed work just focus on package distribution on the fixed Crowdsourced route. It is an important extension to investigate the package delivery problem with alterable Crowdsourced routes in a volunteer manner. Third, our emulation is based on the traffic data in off-peak hours. It is meaningful to update the live traffic in further studies, which is much closer to the practical application. Finally, it is a meaningful work for real-world case study and benchmark instance design, which can be used to test different models and methods in this field.

ACKNOWLEDGMENT

This work is partially supported by the National Natural Science Foundation for Outstanding Excellent young scholars of China under Grant No.61422214, National Natural Science Foundation of China under Grant No.61772544, National Basic Research Program (973 program) under Grant No.2014CB347800, the Hunan Provincial Natural Science Fund for Distinguished Young Scholars under Grant No.2016JJ1002, and the Guangxi Cooperative Innovation Center of cloud computing and Big Data under Grant Nos.YD16507 and YD17X11.

REFERENCES

- [1] R. Masson, A. Trentini, F. Lehud, N. Malhn, O. Pton, and H. Tlahig, "Optimization of a city logistics transportation system with mixed passengers and goods," *Euro Journal on Transportation & Logistics*, vol. 6, no. 1, pp. 81–109, 2017.
- [2] J. Jin, "China's express delivery comes to the era of 100 million a day," http://china.zjol.com.cn/ktx/201801/t20180108_6278004.shtml, 2017.
- [3] S. A. Voccia, A. M. Campbell, and B. W. Thomas, "The same-day delivery problem for online purchases," 2015.
- [4] N. Geroliminis and C. F. Daganzo, "A review of green logistics schemes used in cities around the world," vol. 116, no. 11, pp. 2086–95, 2005.
- [5] R. Lowe and M. Rigby, "The last mile: exploring the online purchasing and delivery journey," *Technical report, Barclays*, 2014.
- [6] C. Chen, D. Zhang, X. Ma, B. Guo, L. Wang, Y. Wang, and E. Sha, "crowddeliver: Planning city-wide package delivery paths leveraging the crowd of taxis," *IEEE Transactions on Intelligent Transportation Systems*, vol. PP, no. 99, pp. 1–19, 2017.

- [7] Q. M. Ha, Y. Deville, Q. D. Pham, and M. H. H, "On the min-cost traveling salesman problem with drone," *Transportation Research Part C Emerging Technologies*, vol. 86, pp. 597–621, 2015.
- [8] A. Arslan, N. Agatz, L. Kroon, and R. Zuidwijk, "Crowdsourced delivery: A dynamic pickup and delivery problem with ad-hoc drivers," *Social Science Electronic Publishing*, 2016.
- [9] A. Trentini, R. Masson, F. Lehud, N. Malhn, O. Pton, and H. Tlahig, "A shared " passengers & goods " city logistics system," 2013.
- [10] Z. Zeng, X. U. Weisheng, X. U. Zhiyu, and W. Shao, "Open-close mixed two-echelon vehicle routing problem in city logistics," *Information & Control*, 2014.
- [11] T. G. Crainic, N. Ricciardi, and G. Storchi, "Models for evaluating and planning city logistics systems," *Transportation Science*, vol. 43, no. 4, pp. 432–454, 2009.
- [12] S. L. W. Neil Mathew, Stephen L Smith, *Optimal Path Planning in Cooperative Heterogeneous Multi-robot Delivery Systems*. Springer International Publishing, 2015.
- [13] D. S. Setzke, C. Pflügler, M. Schreieck, S. Fröhlich, M. Wiesche, and H. Krcmar, "Matching drivers and transportation requests in crowd-sourced delivery systems," in *Proc.of Americas Conference on Information Systems*, 2017.
- [14] Y. H. Cheng and S. Y. Chen, "Perceived accessibility, mobility, and connectivity of public transportation systems," *Transportation Research Part A*, vol. 77, pp. 386–403, 2015.
- [15] V. Paskalathis and S. N. Azhari, "Ant colony optimization on crowd-sourced delivery trip consolidation," vol. 11, no. 2, p. 109, 2017.
- [16] C. Shahabi, M. R. Kolahdouzan, and M. Sharifzadeh, *A Road Network Embedding Technique for K-Nearest Neighbor Search in Moving Object Databases*. Kluwer Academic Publishers, 2003.
- [17] "Flint hills area transportation agency, inc. (2014). ata annual report," <http://www.rileycountyks.gov/DocumentCenter/View/11595>.
- [18] R. Gallotti and M. Barthelemy, "The multilayer temporal network of public transport in great britain," *Sci Data*, vol. 2, 2015.
- [19] Z. Guo, P. Guo, W. Zhao, J. Wu, H. Xiao, Q. Liang, and Q. Zou, "last kilometer distribution strategy research of the county logistics," *Open Journal of Social Sciences*, vol. 05, no. 5, pp. 251–262, 2017.
- [20] X. Tan, Y. Shu, X. Lu, P. Cheng, and J. Chen, "Characterizing and modeling package dynamics in express shipping service network," in *Proc.of IEEE International Congress on Big Data*, 2014, pp. 144–151.
- [21] F. D'Andreagiovanni, J. Krolikowski, and J. Pulaj, "A fast hybrid primal heuristic for multiband robust capacitated network design with multiple time periods," *Applied Soft Computing Journal*, vol. 26, pp. 497–507, 2015.
- [22] J. Yang and V. Honavar, "Feature subset selection using a genetic algorithm," *IEEE Intelligent Systems & Their Applications*, vol. 13, no. 2, pp. 44–49, 2002.