ON FUNDAMENTAL PRINCIPLES OF THE OPTIMAL NUMBER AND LOCATION OF LOADING BAYS IN URBAN AREAS

Tomislav LETNIK¹, Iztok PERUŠ², Stane BOŽIČNIK³, Matej MENCINGER⁴*

¹, ², ³, ⁴Faculty of Civil Engineering, Transportation Engineering and Architecture; University of Maribor, Maribor, Slovenia
⁴Center of Applied Mathematics and Theoretical Physics, Maribor, Slovenia
⁴Institute of Mathematics, Physics and Mechanics; Ljubljana, Slovenia

Received 13 July 2018; revised 22 October 2018; accepted 17 December 2018

Abstract. The paper is dealing with the problem of finding the optimal number and location of Loading Bays (LBs) for efficient urban last mile deliveries. To solve the problem a multi-parametric model of the idealized urban area is introduced and applied to various instances of a rectangular urban grid structured zones. Multi-parametric approach is used to assess statistically the most relevant number and location of LBs. Computational and graphical results of the idealized model exhibit geometric patterns showing that the optimal Number of LBs (#LB) naturally tends to perfect squares. Moreover, even in case of generalized instances, at a selected number of LBs their distribution is not random but follows specific laws. The optimality is closely related to the prefixed (maximal) walking distance $d_{\text{max}}$ from the LB to the customer. Based on various simulations the existence and robustness of a descending convex dependence $d_{\text{max}} = (#\text{LB})$ is proven. The results might serve as a decision-making tool to determine the optimal number and location of LBs for any real-life city centre.

Keywords: freight transportation, last mile delivery, facility location, fuzzy clustering, decision-making, loading bay.

Introduction

Cities are irrepressibly becoming larger, which leads to various problems related to increasing traffic in urban areas. Competition for urban street space between cars, vans, delivery vehicles, public transport, pedestrians and cyclists is becoming more and more intense. One of the major problems is related to delivery vehicles, which are contributing considerably to congestion, pollution and emissions (Lindholm 2013; Tozzi et al. 2014; Zou et al. 2016).

Cities are implementing various restrictive measures (e.g. ban for heavy-duty vehicles, low emission zones, time windows, congestion charging, etc.) to mitigate negative effects of urban freight deliveries (Buldeo Rai et al. 2017; Fu, Jenelius 2018; Holguín-Veras, Sánchez-Díaz 2016; Marcucci et al. 2017; Quak, De Koster 2009; Russo, Comi 2011). Access to urban areas (especially to historical city centres or some specific parts of the city) is often prohibited or allowed only at certain time of the day and/ or only with small and environmentally friendly delivery vehicles (freight bikes, light commercial vehicles, electric delivery vehicles). When direct delivery to customers is not possible a Loading Bay (LB) is needed for transhipment of goods and parking of delivery vehicles during the final (last mile) delivery operations (Alho et al. 2014; Guastaroba et al. 2016; Letnik et al. 2018).

In general, three main issues related to LBs can be identified in the literature (Malik et al. 2017): lack of LBs; inappropriate location of LBs and the illegal use of LBs. In all these cases, delivery vehicles are not able to use LBs and they are consequently forced to double-park or to circulate around the city to find the next unoccupied LB. In European cities, this occurs in 70…80% of deliveries (Lopez et al. 2016), this is time consuming and contributes significantly to the reduction of the available road capacity and to the increase of urban traffic congestion (Russo, Comi 2012; De Abreu e Silva, Alho 2017). Last part of the delivery, from the LB to the customer, is mostly performed on foot or by trolley, therefore number and location of LBs should be carefully planned to keep the walking distances reasonably low.

*Corresponding author. E-mail: matej.mencinger@um.si

Copyright © 2019 The Author(s). Published by VGTU Press

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
The problem of finding optimal number and location of LBs should be, according to Alho et al. (2014), treated as a distance-based Facility Location-allocation Problem (FLP). The facility FLP was originally introduced by Cooper (1963) with the main aim to locate multiple facilities in an Euclidian space and allocate customers to a facility while minimizing transportation costs. Since then, many researchers have studied discrete and continuous versions of Multiple FLP (MFLP) extensively and many variations are nowadays available. The \( p \)-median problem is to place \( p \) facilities so that the total distance from all customers to their corresponding closest facilities is minimized (Colménar et al. 2018; Jánosi Ková et al. 2017; Wang et al. 2018). The \( p \)-center problem is to place \( p \) facilities in the network on such a way that the maximum distance from each customer to its closest facility is minimized (Callaghan et al. 2017; Kikuno et al. 1980; Tüzün Aksu, Ocak 2012). The maximal coverage location problem is to place one or more facilities in the network on such a way that all customers are covered (served) within a predefined radius from at least one facility (Berman et al. 2016; Blanquero et al. 2016; Farahani et al. 2012). The clustering problem is, where customers correspond to data points and where facility locations are determined as cluster centres. Clustering criteria is to minimize the sum of distances from each data point to the corresponding cluster centre (Iyigun, Ben-Israel 2010; Zhao et al. 2019). The multi-objective (or multiple-criteria) facility location problem, where different objectives: minimizing total or average weighted time/distance travelled, maximizing the number of customers receiving service before a critical time threshold and minimizing the maximum weighted time/distance travelled, are used in combination (Farahani et al. 2010; Karatas, Yakıcı 2018).

The above mentioned solution methods play an important role in solving location problems in many different areas, such as: press delivery networks, locating post boxes, telecommunication systems, fuel/gasoline service stations, retail outlets, and also city logistics terminals (Hajipour et al. 2016), which are studied in this paper. Several researchers have tried to determine the optimal number and location of LBs using different approaches and methods. Aiura and Taniguchi (2005) developed a methodology for determining optimal locations of on-street LBs to evaluate the efficiency of enforcement for controlling illegal parking of passenger cars at loading-unloading spaces. The methodology has been applied to test road network (a single road section) and was able to determine a configuration of parking spaces that achieved a cost reduction of approximately 16%. Dezi et al. (2010) studied the optimal number and position of LBs for the city centre of Bologna, Italy. They developed a methodology for placement of LBs taking into consideration expected urban freight delivery demand, specific characteristics of particular commercial entities (type, size, etc.) and characteristics of the urban road network (distances, pedestrian crossings, gradients and architectural barriers). In the case when all needed detailed data is not available (and this is often the case), they proposed to use the simplified process of defining circles of the particular radius of influence (50 m) based only on knowledge of buildings and road network. Their final aim was then to minimize overlapping areas of adjacent circles. Delaitre and Routhier (2010) studied the possibilities to combine two different tools, one dealing with the generation of vehicles movements (Freturb), another dealing with the simulation of delivery areas (Dalsim) to define the best possible number and location of LBs. The model has been applied on several characteristic streets of La Rochelle city centre in France to identify the best possible micro locations of LBs for the particular street. They found out that the number and position of LBs depends mainly on characteristics of the road network (length of streets) and availability of space, which can be dedicated to LBs. Alho et al. (2014) critically analysed state-of-the-art of the freight modelling methodologies to optimize the configuration of LBs and proposed alternative modelling framework, integrating simulation models and optimization strategies that take into account also double-parking of delivery vehicles. Their model is based on the identification of customers’ locations and their demand, again along the street segment and from the micro location point of view.

As seen from the literature review, studies on urban LBs have been so far limited only to specific cases and to selected cities (with particular characteristics). General rules and instructions for setting optimal number and location of LBs, that would fit into different urban areas, is missing in the literature and in practice (Alho, De Abreu e Silva 2014; Comi et al. 2017; Malik et al. 2017). In addition to that, most of the solutions in the existing literature have some restrictions. Either are limited to micro locations, or to the Euclidean (aerial) distances, or consider the freight delivery demand as a static criterion, implying that any change of demand (in terms of quantity and/or the location of customer) requires new calculation of optimal location. A good city logistics tactical planning should therefore include also the uncertainty of demand (Crainic et al. 2016). For solving the problem of setting the optimal number and location of LBs and overcome before mentioned restrictions, we introduced the new multi-parametric model, which is combining fuzzy clustering and probability densities of cluster centres to evaluate the efficiency of placing a different number of LBs in a typical urban grid cell of 1 km², considering uncertainties related to delivery demand. The aim of this paper is two-fold:

1) to present and introduce a multi-parametric model;
2) to provide some general recommendations for decision-making related to setting the optimal number and locations of LBs in different urban area contexts.

1. Multi-parametric LB location model

The problem of finding the optimal number and location of LBs is a very complex, since we have to take into consideration various characteristic parameters of a city and the volume and dynamics of delivery demand. For this kind of problems, a multi-parametric models and princi-
ple areas with a population density of 1000 or 1500 per inhabitant and number of inhabitants for a typical urban area. Assuming, on average, 0.1 delivery/pick-ups per inhabitant and number of inhabitants for a typical urban area per day. This may help to assess the number of deliveries if we know the population density of a certain urban area. Assuming, on average, 0.1 delivery/pick-ups per inhabitant and number of inhabitants for a typical urban area cell of 1 km² ranging from 1000 and up to 10000 people/km², the following parameters are to be studied: 100, 300, 500, 700 and 1000 deliveries/day/km².

1.1. Main parameters of the model

The optimal number and location of LBs mainly depend on the size and shape of an urban area, the characteristics of the road network, the location, distribution and density of customers and on the demand characteristics. Urban areas considerably differ from each other; therefore, we first need to consider and define some common characteristics when modelling urban freight deliveries. According to the Statistical Office of the European Union (EuroStat), an urban centre is defined as contiguous grid cells of 1 km² with a population density of at least 1500 inhabitants/km² (Dijkstra, Poelman 2014). The urban core is similarly defined in the methodology of the Organization for Economic Co-operation and Development (OECD) as contiguous or highly interconnected densely inhabited urban areas with a population density of 1000 or 1500 people/km² (OECD 2012). Both methodologies take a cell of 1 km² as a basic unit for identification of urban areas and this unit should, therefore, be understood also as the main element for the modelling purposes of urban freight deliveries.

The topology of the urban road network has been studied by many researchers (Lin, Ban 2017; Strano et al. 2013; Tsiotas, Polyzos 2017). Generally, the following characteristics of urban road networks are noted: heterogeneity (very few road sections interact with each other), planarity (roads tend to be parallel more than forming triangles), validity of power law (the majority of road sections are short in distance, especially in city centres) and density (high density of intersections and many roads in urban areas). One can conclude that an urban road network can be best approximated by a rectangular grid structured city mash. Especially in city centres (where LBs are predominantly needed), the grids are mostly rectangular. Moreover, there are some parts of almost any city (regardless of the type of road pattern), which undergoes to a rectangular structured network (Boeing 2017).

The delivery demand of customers varies considerably in quantity and time. Bigger customers, for example, accommodate a higher number of deliveries and as result, they might call either for a bigger number of LBs or simply bigger capacity of LBs located in their proximity. Data on the actual number of deliveries is hard to find in the literature (are almost non-existing), we, therefore, need to rely on some assumptions and modelling principles (Muñuzuri et al. 2009, 2012). The quantity of demand for modelling deliveries in urban areas mainly relates to the number of inhabitants. Estimations based on existing urban freight studies (Dablanc et al. 2013) are showing on average 0.1 delivery/pick-up per person (inhabitant of an urban area) per day. This may help to assess the number of deliveries if we know the population density of a certain urban area.

1.2. The concept of the model

Considering all the above-mentioned parameters, an idealized urban area model with the following assumptions has been introduced: an urban area is a typical urban grid cell of 1 km² (not the entire city); an urban road network is a grid of equidistant and perpendicular roads (rectangular grid); customers are uniformly but randomly distributed over the perpendicular lattice, using the – Latin Hypercube Sampling (LHS) method, which is described in the sequel. To test the results of optimal number and position of LBs under different conditions the model allows to simulate different parameters: road network density (number of roads and the distance between them); customers’ location and density (number of customers on 1 km²); the number of LBs.

The model is built on a dynamically changing demand, which is in the model presented as a variety of randomly generated positions of customers (instances). The model simulates main features (different sets of parameters) and compares solutions for different instances to provide statistically relevant and robust results yielding the optimal #LBs and positions in any characteristic urban grid cell.

In the model (the overall scheme is presented in Figure 1), one can vary the density of the road network and the density of customers, and provide graphical (scatter plot, contour plot and a 3D plot of probability density surface respectively) and computational results. The results are numerically evaluated with the maximal distance d from the LBs to the customers. Distances are measured with Manhattan (city block) metric, which is probably the best approximation for measuring distances in the real city centres with the rectangular road network.

Clustering procedure serves to provide the optimal location of LBs. The aim is to identify k cluster centres and their positions in the idealized urban area. To find the optimal number of LBs and their positions, we use \( d_{\text{max}} \) as a maximal allowed walking distance from the cluster centre to all the corresponding customers. Maximal allowed walking distance determines the crucial constraint of the problem.

The model is built in the MATLAB software (https://se.mathworks.com). For modelling the distribution and clustering of customers, MATLAB built-in procedures and algorithms are used (see the following sections for details).
1.3. Idealized urban area (city plan) and road network

To construct an idealized urban area model, one has to understand the main characteristics of a typical urban road network, which consist of road sections and intersections. To range those values, we randomly selected samples of several cities and studied their maps. Then we selected a 1 km^2 cell from the city centre, marked all intersections and measured the distances among them (Figure 2).

Note again that 1 km^2 is considered as a typical urban cell and is therefore used as a typical instance in our proposed model where different road network densities are simulated. The road network on 1 km^2 cell is determined by:

- number of roads (city streets) in E–W-direction \( N_x \) and N–S-direction \( N_y \);
- distances between two adjacent roads in E–W-direction \( l_x \) and N–S-direction \( l_y \).

Empirical results reveal that:

- number of intersections \( I_{nr} \) range from 90 to 368;
- maximum distance among the intersections \( D_{\text{max}} \) range from 125 to 240 m;
- minimum distances between the intersections \( D_{\text{min}} \) range from 13 to 80 m;
- average distances among the intersections \( D_{\text{avg}} \) range from 55 to 117 m.

Note that one can barely recognise some common characteristics. Yet, the intersections are more or less evenly distributed in the area, which justifies the approximation of a typical 1 km^2 area with an equidistant rectangular grid network. In reality, we can find highly symmetric and also highly asymmetric rectangular street layouts concerning differences in the actual block spacing, as shown in Figure 3.

For the purpose of this paper, we decided to work with idealized square grid structures and vary the network density to assess its influence on the number and position of LBs. According to data gained through the empirical test, the most suitable choice for idealized rectangular road network are grid structures shown in Figure 4.

The grid structure of:

- 21×21 nodes are relevant for cities with \( D_{\text{avg}} \approx 50 \) m;
- 15×15 nodes are relevant for cities with \( D_{\text{avg}} \approx 72 \) m;
- 11×11 nodes are relevant for cities with \( D_{\text{avg}} \approx 100 \) m;
- 7×7 nodes are relevant for cities with \( D_{\text{avg}} \approx 167 \) m;
- 4×4 nodes are relevant for cities with \( D_{\text{avg}} \approx 333 \) m.

The basic cell of the lattice (grid) has a dimension of: \( \frac{1000}{N_x-1} \times \frac{1000}{N_y-1} \) m. For example, in case when \( N_x = N_y = 11 \), the basic cell is a square of dimension 100×100 m.

1.4. Number and distribution of customers

The number of the customers was assumed as a fraction of the product \( N_x \times N_y \). Example of an idealized city plan with LHS distributed location of customers is presented in Figure 5.

To simulate a realistic situation, customers need to be randomly distributed over the whole (rectangular) area. In our model, we assume independence of both location components for different customers and covering the domain of both components, as good as possible, at the same time. To this end, the range of each variable should be divided into several intervals of equal probability. This can be done best with the LHS method (Janssen 2013).
One of the last improvement of the LHS provided by Husslage et al. (2011) is integrated also into MATLAB and it fits perfectly to our requirements. The MATLAB code uses (by default) the maxi-min criterion, which maximizes the minimum distances between the sample points.

Mathematical-style pseudo-code, for the algorithm providing an idealized city lattice and the LHS distribution of customers, is the following:

**algorithm** generate/print idealized city lattice with the customers

**input:** number of roads in both directions: \( N_x \) and \( N_y \),

distances between two adjacent roads: \( l_x \) and \( l_y \),

number of customers \( \#r \) (related to \( N_x \cdot N_y \)),

**output:** perpendicular road map with \( \#r \) customers on the roads

\( r = 0 \), generate \( X \) as \( 2 \times (\#r) \) a matrix using the LHS method

**while** \( r \leq (\#r) \) **do** \( r \rightarrow r + 1 \)

**if** \( \operatorname{rand}() < 0.4999999999999 \)

locate the customer vertically to lattice:

\[
\text{CustomerLoc} (r, 1) = \text{round} (X(r, 1) \cdot (N_x - 1) \cdot l_x) \quad \text{and} \quad \text{CustomerLoc} (r, 2) = X(r, 2) \cdot (N_y - 1) \cdot l_y
\]

**else**

locate the customer horizontally to lattice:

\[
\text{CustomerLoc} (r, 1) = X(r, 1) \cdot (N_x - 1) \cdot l_x \quad \text{and} \quad \text{CustomerLoc} (r, 2) = \text{round} (X(r, 2) \cdot (N_y - 1) \cdot l_y)
\]

**end**

1.5. Clustering of customers

As already mentioned, the continuous MFLP can be considered as a classical clustering problem, where customers and facility locations correspond to data points and centres. The clustering criterion is to minimize the sum of distances from each data point to the corresponding cluster centre. Potential location of LBs in the model is therefore determined by using Fuzzy \( c \)-Means (FCM) clustering procedure. FCM was chosen to get some flexibility to the model with the possibility to simulate different membership degrees of customers to cluster centres, in our case LBs (Ross 2010).

FCM aims at minimizing the following function (Bezdek 1981):

\[
J_m = \sum_{i=1}^{D} \sum_{j=1}^{N} \mu_{ij}^m \cdot x_i - c_j^2
\]

where: \( D \) – number of customers; \( N \) – number of clusters; \( x_i \) – \( i \)-th data point (each data point corresponds with the location of a particular customer); \( c_j \) – centre of cluster \( j \); \( \mu_{ij} \) – degree of membership of customer \( i \) to cluster \( j \); \( m \) – parameter of fuzziness.

FCM algorithm implemented in MATLAB performs the following steps:

1) randomly initialize the cluster membership values \( \mu_{ij} \);

2) calculate the cluster centres:

\[
\overline{x}_j = \frac{\sum_{i=1}^{D} \mu_{ij}^m \cdot x_i}{\sum_{i=1}^{D} \mu_{ij}^m};
\]

3) update \( \mu_{ij} \) according to the following:
\[ \mu_j = \frac{1}{\sum_{k=1}^{N} \left( \frac{x_j - c_k}{|x_j - c_k|^m} \right)^{m-1}} \]  

(3)

4) calculate the objective function \( J_m \);

5) repeat Steps 2–4 until \( J_m \) improves by less than a specified minimum threshold \( \epsilon \) or until after a specified maximum number of iterations.

In our model we used the FCM routine of MATLAB with the parameter \( m \) specified as 2, the maximum number of iterations being set to 100, with minimum improvement in objective function between two consecutive iterations specified as \( \epsilon = 10^{-5} \). The main output of the algorithm is the position of \( N \) centres \((c_1, c_2, ..., c_N)\) and the partition matrix \( U \).

1.6. Numerical evaluation and graphical presentation of results

The presented model allows to optimize and quantitatively evaluate either the number of LBs or the maximum distances between customers and cluster centres. In addition to that, results of the model can be evaluated also quantitatively based on the graphical presentation of the results.

Mathematical-style pseudo-code, for the algorithm providing clustering of customers and defining \( d_{\text{max}} \) is the following:

**algorithm** form clusters and cluster-centres is

**input:** number of clusters: \( k \),
the lattice and the customers

**output:** \( k \) clusters of customers together with cluster-centres
(such that the sum of the distances from each cluster centre to all customers in the cluster is minimal)

**use** fcm (CustomerLoc, \( k \)),
read CustomerLoc customer locations
use fcm fuzzy clustering algorithm to generate \( k \) clusters
locate each cluster-centre to the nearest road
(vertically or horizontally to lattice)
for \( i = 1 \) to \( k \) do
find the maximal distance \( d_i \) from cluster-centre
to the customers within cluster \( i \)
define \( d_{\text{max}} : = \max(d_i) \)

The values of actual (maximal) walking distance \( d_{\text{max}} \) are in the model presented in terms of the following percentiles: 1, 2, 5, 16, 32, 50, 68, 84, 95, 98 and 99. Such approach allows better evaluation of results in case we decide not to cover all customers. This is a reasonable assumption because setting and maintaining the LB is always related to costs and limited space in urban areas.

Qualitatively, solutions are measured in terms of relevant graphical patterns showing statistically optimal locations of LBs. Three different types of graphical representation are used in the model: scatter plot, contour plot and a 3D plot of probability density surface.

In the scatter plot in Figure 6a, the location of cluster centres for a particular predefined number of clusters and several simulations is presented. In general, some cluster centres might overlap, therefore we decided to provide also the results with a contour plot (Figure 6b). In that case, statistically optimal locations can easily be identified. Finally, the results are presented as a plot of the probability density surface in 3D (for the best position of the cluster centre; see Figure 6c). In this case, some fine details not evident from previous graphical presentations may be recognized.

Note that to present the Probability Density Functions (PDFs) as surfaces in 3D space and as contour plots in 2D space a Conditional Average Estimator Neural Network (CAE NN) approach was used (Peruš et al. 2006), which can successfully incorporate different uncertainties always present in the real-life case scenarios (Peruš et al. 2012). CAE NN uses the Gaussian density function \( a_n \), with a width \( w \), as a weight function. It is centred at each \( n \)-th point in the multi-dimensional space to get the impact of the \( l \)-th model vector at the point of the prediction vector:

\[ a_n = \frac{1}{(2 \cdot \pi)^{D/2} \cdot w^D} \cdot \exp \left( \frac{- \sum_{l=1}^{D} \left( b_{nl} - b_{nl} \right)^2}{2 \cdot w^2} \right), \]  

(4)

where: \( b_{nl} \) is the \( l \)-th input variable of the \( n \)-th model vector in the database \((b_{n1}, b_{n2}, b_{n3}, ..., b_{nD})\); \( b_l \) is the \( l \)-th input variable corresponding to the prediction vector; \( D \) is the number of input variables and defines the dimension of the sample space.

![Figure 6. Scatter plot, contour plot and a 3D plot of probability density surface, an example presentation of simulation results for \( k = 8 \)](image_url)
In the presented case \((D = 2)\), the above equation simplifies to:

\[
a_n = \frac{1}{2 \cdot \pi \cdot w^2} \exp \left\{ -\frac{(x_i - x_{ni})^2}{2 \cdot w^2} - \frac{(y_i - y_{ni})^2}{2 \cdot w^2} \right\}
\]

(5)

where: \(x\) and \(y\) denote the \(x\)- and \(y\)-coordinates (positions) of the points where PDF is estimated by the expression:

\[
\hat{pdf}(x, y) = \frac{1}{N} \sum_{n=1}^{N} a_n,
\]

where: \(N\) is the number of all cluster centres obtained in particular simulation and the sign “\(^\wedge\)” denotes the estimation.

### 2. Simulation and results

Our research aims to evaluate the impact of different parameters on the optimal number and location of LBs. The main idea is to create different instances of orthogonal city structures and to simulate different customer densities in order to identify the existence of statistically relevant patterns, which could be used for identification of optimal locations of LBs. This is the basis of our statistical approach, which generally results in:

- the existence of a statistically relevant concentration of cluster centres (where probability density is high enough), which could serve as potential locations for positioning of LBs;
- theoretical relation between the \(#LBs\) and the value of \(d_{max}\);
- the impact of different density of customers on a number and location of LBs;
- the impact of different network density on a number and location of LBs;
- the reasonable minimal and a maximal number of LBs for particular parameter characteristics.

#### 2.1. Simulation instances

For simulation purposes, different characteristics of an idealized city plan with network densities \(\rho_n (7 \times 7, \ 11 \times 11, \ 15 \times 15 \text{ and } 21 \times 21 \text{ roads})\) has been considered. For each of the city plans, different density of customers \(\rho_c (100, 300, 500, 700, 1000 \text{ customers/km}^2)\) located randomly in terms of the LHS distribution has been simulated. The examples of basic simulation instances are presented in Figure 7.

For each of the instances (different combination of \(\rho_n\) and \(\rho_c\)), 200 simulations have been performed to provide a statistical basis for identification of potential locations of LBs.

#### 2.2. Graphical patterns – optimal number and location of LBs

Graphical patterns have been observed for a different number of LBs \(#LB\) in the range of 1 to 55. For the reason of brevity, only the most interesting and the most important graphical patterns are presented. Consequently, only graphical patterns for \(\rho_n = 7 \times 7 \text{ roads and } \rho_c = 1000 \text{ customers/km}^2\) are shown.

In case of \#LB = 1, the simulation shows a clear central position result. Very interesting results, worth to be mentioned, are in the case of simulating \#LB from 2 to 11 and \#LB = 27.

Results of simulating \#LB = 2 (Figure 8) indicate the pattern tending to 4 equally important locations. It seems like the system naturally tends toward 4 locations but in fact, it represents a solution of the well-known XOR problem. In general, this may be explained as two pairs of diagonally placed (equally important) locations. In practical terms, this means that we can opt for one or another diagonal positions.

Results of simulating \#LB = 3 (Figure 9) indicates a continuous square rounded oval shape of possible locations. In this case, we may understand this as a rotation of three locations, which are almost equidistant. Since the distribution is more or less proportional, this means we can generate an infinite number of solutions of three LBs placed equidistantly around the oval. All possible locations have almost the same values of the PDF.

Results of simulating \#LB = 4 (Figure 10) indicates a clear pattern of 4 locations with exactly the same values of PDF. When simulating \#LB = 5 (Figure 11) a clear pattern of 5 locations could be recognized, but the fifth central placed location is less significant than the four ones placed near the corners.

Results of simulating \#LB = 6 and \#LB = 7 (Figures 12 and 13) indicates a pattern of 8 locations with 4 placed near the corners having considerably higher values of PDF. The other 4 locations placed between them are less significant in terms of the value of PDF.

Results of simulating \#LB = 8 and \#LB = 9 (Figures 14 and 15) both indicate a pattern of 9 locations, with centrally placed location considerably less significant in the case of \#LB = 8.

Results of simulating \#LB = 10 (Figure 16), scatter plot indicates a pattern of 25 locations \((5 \times 5)\) while contour plot and 3D surface both indicate a pattern of 9 locations \((3 \times 3)\). Locations placed near the corners have bigger significance, locations placed between them (on the outer edge) are a bit less significant and the centrally placed locations have the smallest values of PDF. We can also notice the existence of several smaller densities on the 3D surface, which is indicating the beginning of forming the pattern of 25 locations.

Concerning the results of simulating \#LB = 11 (Figure 17), all three types of graphical representation indicate a pattern of 25 locations while locations placed near the corners are still the prevailing ones. This pattern remains relevant until simulating \#LB = 27 (Figure 18), where all three types of graphical representations start to indicate 45 locations, which is the “final number” of LBs for the particular network. This means that from the statistical point of view, increasing of LBs does not change the number and the placement of locations. More precisely the graph...
of the PDF does not change anymore. In this case, we can also recognize that locations near the corners are becoming less significant and centrally placed locations are becoming more statistically relevant.

In-depth analysis of all three kinds of graphical representation of simulation results, for the whole range of up to 55 LBs and for all four network densities, revealed some interesting characteristics, as can be seen from the Table 1.

All three kinds of graphical representation are showing that the optimal number of LBs naturally tends to the perfect squares (4, 9, 16, 25, etc.). This seems to appear as a natural law and should be considered when planning logistics facilities in urban areas. It is also very interesting that in certain cases some perfect squares are skipped, like 36 for example in the case of 7×7 and 11×11 grid road network. In other cases, a #LB locations identified seems to be smaller or bigger than those being simulated. The cognitive reason for both cases is the way we interpret the particular number of locations. Once we identified a certain number of locations for a lower #LB we, as by the law of inertia, kept this number as relevant until the next pattern was identified. It has to be noted that in between the patterns get fuzzy, but once reviewing simulations for a different number of LBs in a sequence we noticed this is the best possible way to present the results.

There are also many cases, especially from #LB sim = 10 onwards, when the #LBs, visually identified, differs between plot types. The scatter plot gives the best representation of LBs dispersion but on the other hand does not provide a good basis for identification of statistically optimal locations. Since one cannot recognize the importance of certain concentrations (it is not evident how many locations overlap), the scatter plot shows a higher number of relevant locations than contour and PDF surface plot.
Some differences in visual interpretation of patterns are recognized also between contour plot and PDF surface plots and in particular when simulating higher number of LBs or the networks with higher network density (15×15 and 21×21). In these cases, the PDF surface plot is indicating higher number of statistically optimal locations than the contour plot. The reason is based in more accurate visual identification of the pattern from the PDF surface plot showing some additional details, which are not evident from other two representations.

Figure 8. Scatter plot, contour plot and a 3D plot of the probability density surface for:
#LB = 2; \( \rho_n = 7 \times 7 \) roads; \( \rho_c = 1000 \) customers/km²

Figure 9. Scatter plot, contour plot and a 3D plot of the probability density surface for:
#LB = 3; \( \rho_n = 7 \times 7 \) roads; \( \rho_c = 1000 \) customers/km²

Figure 10. Scatter plot, contour plot and a 3D plot of the probability density surface for:
#LB = 4; \( \rho_n = 7 \times 7 \) roads; \( \rho_c = 1000 \) customers/km²

Figure 11. Scatter plot, contour plot and a 3D plot of the probability density surface for:
#LB = 5; \( \rho_n = 7 \times 7 \) roads; \( \rho_c = 1000 \) customers/km²
Figure 12. Scatter plot, contour plot and a 3D plot of the probability density surface for:
#LB = 6; \(\rho_n = 7 \times 7\) roads; \(\rho_c = 1000\) customers/km\(^2\)

Figure 13. Scatter plot, contour plot and a 3D plot of the probability density surface for:
#LB = 7; \(\rho_n = 7 \times 7\) roads; \(\rho_c = 1000\) customers/km\(^2\)

Figure 14. Scatter plot, contour plot and a 3D plot of the probability density surface for:
#LB = 8; \(\rho_n = 7 \times 7\) roads; \(\rho_c = 1000\) customers/km\(^2\)

Figure 15. Scatter plot, contour plot and a 3D plot of the probability density surface for:
#LB = 9; \(\rho_n = 7 \times 7\) roads; \(\rho_c = 1000\) customers/km\(^2\)
It is also very important to note that the final number of statistically relevant locations depends on the grid density. In particular, for a $n \times n$ city plan the final number of locations was found to settle at $\#LB = n^2 - 4$. For example, a typical grid of $7 \times 7$ roads tend to 45 relevant locations. Similar holds true for a typical grid of $11 \times 11$ roads, which tends towards 117 statistically relevant locations as seen in Table 1. Presuming these characteristics, a typical grid of $15 \times 15$ would tend to 221 locations and a typical grid of $21 \times 21$ to 437 relevant locations. These two examples are not evident from the Table 1, since the number of LBs simulated was limited to 55. On one hand, this is a consequence of network regularity, but on the other, it indicates also the need to place LBs at each road section of the city grid and specifically in the proximity of interactions. Special cases are locations at the corners of the surveyed area, where LBs are considerably less relevant (they are covering a smaller part of the network than the other ones). In addition to that, it has to be noted that this relation is valid until $d_{\text{max}} \geq l$, where $l$ is the distances between two adjacent roads. In case when $d_{\text{max}}$ get smaller than $l$, different patterns, with more locations would accrue.

Though, our research is limited to uncapacitated MFLP, note that the PDF is not useful only for identification of the optimal placement of the LBs but could be also used to assign different capacities to different LBs. Namely, the locations with higher (lower) values of PDF naturally need a higher (lower) capacity.
2.3. The relation between #LBs and \( d_{\text{max}} \)

The relation between the number of LBs and \( d_{\text{max}} \) has been simulated for different network and customers densities. Furthermore, the relation is presented also for different percentile values, which statistically indicate the rate of estimated \( d_{\text{max}} \) in case of covering a particular percentage of customers (from 1 to 99%). In practical terms, city authority can opt to cover only certain percentage of customers and this kind of presentation allows to determine a reasonable number of LBs assuming that some customers may be served with a reasonably bigger \( d_{\text{max}} \) values.

A typical example of a graph of the functional relation \( d_{\text{max}} = d_{\text{max}}(#\text{LB}) \) for a particular network and customer density and for different percentile values is presented in Figure 19.

The results presented in Figure 19, are showing that the relation between \( d_{\text{max}} \) and #LB is a descending convex curve. In relation to the optimal number of LBs, this means that a smaller number of LBs result in bigger \( d_{\text{max}} \) values and vice versa. If we decide, for example, that \( d_{\text{max}} \) should be lower than 100 m, this would mean (for a particular instance presented above), that we need to set at least 22 LBs to cover 50% of all customers, 34 LBs to cover 68% of all customers and 52 LBs, if we would like to cover 84% of all customers, respectively.

To understand the nature of the relation between \( d_{\text{max}} \) and #LB, linear regression of the corresponding log-log dependence has been performed. In Figure 20, the example of 95 percentile relation for \( \rho_n = 21 \times 21 \) roads and \( \rho_c = 1000 \) customers/km² is presented as a log-log dependence.

The presented log-log dependence results in \( d_{\text{max}} = 18.169(#\text{LB})^{-4.719} \), and reveals how significantly the relation \( d_{\text{max}} = d_{\text{max}}(#\text{LB}) \) fits the power low dependence.
2.4. The dependence of $d_{\text{max}}$ in terms of #LB for different values of $\rho_n$ and $\rho_c$

The relation between $d_{\text{max}}$ and the number of LBs was further analysed to test the dependence on the customers density. One would expect that the $d_{\text{max}}$ varies significantly for different customers densities, but this is not the case, as it can be seen from graphs presented in Figure 21.

The analysis of a fixed customers density $\rho_c$ and several different network densities $\rho_n$ (4×4, 7×7, 11×11, 15×15, 21×21 roads) is presented in Figure 22. It can be as seen from the graphs that also in those cases, the value of $d_{\text{max}}$ does not vary significantly, if the network density is changed.

We see that the dependence $d_{\text{max}} = d_{\text{max}}(#\text{LB})$ is quite robust, when changing either $\rho_c$ or $\rho_n$. This means that the “typology” of a city structure (i.e. different $\rho_c$ or $\rho_n$) does not have a significant impact on the setting of LBs in urban areas with rectangular road networks.

2.5. Minimal and maximal number of LBs

The graph of the descending function $d_{\text{max}} = d_{\text{max}}(#\text{LB})$ for different percentile values can be used also to determine the minimal and maximal number of LBs for a particular value of $d_{\text{max}}$. The minimal number simply depends upon the $d_{\text{max}}$ corresponding to the percentile value, which we consider as “still acceptable”. The maximal LBs has to be recognised as a critical point at which the derivative of $d_{\text{max}} = d_{\text{max}}(#\text{LB})$ becomes “small enough”, which in practice means that any additional LB does not bring any relevant decrease of $d_{\text{max}}$.

To determine this critical point for each particular case a second order approximation of the first derivative of a function $d_{\text{max}} = d_{\text{max}}(#\text{LB})$ was calculated. Our aim was to identify when the derivative becomes small enough. The following second order approximation formula for the first derivative of $F$:

$$F'(x) = \frac{F(x+h) - F(x-h)}{2 \cdot h} + O(h^2).$$  \hspace{1cm} (7)

Results for the interval from 3 to 35 LBs, is presented in the following Figure 23.

As it can be seen in Figure 23, each additional LB results in a considerable decrease of $d_{\text{max}}$ for #LB ≤ 10 and in a very small decrease from #LB ≥ 20. A considerable difference can be observed also between lower and higher percentile values. For lower percentile values (e.g. 5%) any additional LB does not bring a reasonable improvement even for a lower #LBs. On the other hand, one additional LB results in a significant improvement for $d_{\text{max}}$, provided 95% of all customers are to be covered.

Based on these findings, we can finally set the maximum (critical) number of LBs still reasonable to be opened just by defining percentile values and acceptable decrement of $d_{\text{max}}$. If we decide to cover 95% of customers assuming the decrement of $d_{\text{max}}$ is greater than 4 m, this results in a maximum of 20 LBs (Figure 23). This indicates the existence of a threshold in the number of LBs.

3. Interpretation of results

3.1. Practical example

In this sub-section we consider a practical example illustrating the basic concept of the model. Note that defining the threshold of a maximal walking distance from the LB to the customer and consequently the number of LBs in a particular area of the city is always a political decision.
The model presented in this paper is providing the methodology, which may help public authorities and urban planners to understand and use the relation \( d_{\text{max}} = d_{\text{max}}(#\text{LB}) \) to construct the optimal network of LBs, knowing only main parameters of the urban area, without the need for gathering the big amount of data.

Let us consider the area of 1 km\(^2\) with \( \rho_n = 7 \times 7 \) roads and \( \rho_c = 1000 \) customers/km\(^2\). In this case, the city authority should check the relation \( d_{\text{max}} = d_{\text{max}}(#\text{LB}) \) for (before mentioned) urban area characteristics (Figure 24) to make its decision. If, for example, the city authority would decide for \( d_{\text{max}} \leq 200 \) m, this would result with: 55 LBs in case of covering 99\% of all customers, 24 LBs in case of covering 95\% of all customers, 15 LBs in case of covering 84\% of all customers, 10 LBs in case of covering 68\% of all customers and 6 LBs in case of covering 50\% of all customers.

As we can see, the city authority should carefully consider the percentage of customers to be covered, since the bigger percentage considerably increase the required number of LBs.
On the other hand, the city authority might also decide to place a particular number of LBs and check the corresponding $d_{\text{max}}$ values, in order to cover a certain percentage of customers. If, for example, the city authority would decide for $\# \text{LB} = 5$, this would result in $d_{\text{max}} = 470$ m in case of covering 99% of all customers, $d_{\text{max}} = 412$ m in case of covering 95% of all customers, $d_{\text{max}} = 330$ m in case of covering 84% of all customers, $d_{\text{max}} = 272$ m in case of covering 68% of all customers and $d_{\text{max}} = 225$ m in case of covering 50% of all customers.

From the locations point of view, the model provides statistically optimal locations in the form of contour plots. Let us continue with the above example of $\# \text{LB} = 5$. As seen from the Figure 25 (example for the decision of placing 5 LBs), the central contour lines are presenting optimal locations. All other contour lines are representing statistically less but still highly relevant options for setting a particular LB. If the LB is placed on the particular contour line, the quality of a solution does not decrease (i.e. $d_{\text{max}}$ does not increase). In that respect, we have a certain freedom of choosing the most feasible solution.

In the real-life process of final (precise) placement of LBs, various criteria may be considered. In the case when the optimal location is not feasible, we can always find a statistically acceptable solution, which is near the optimum. Beside physical constraints also economic aspect could be considered as a relevant criterion.

### 3.2. Covering of customers with different percentile and $d_{\text{max}}$ values

To show the (on-street) locations of the serviced and un-serviced customers, one has to consider the (convex) graph on the Figure 24 to associate the particular $\# \text{LB}$ with the $d_{\text{max}}$ at a given covering percentile value $c_p$. To explain these relations and to present them graphically we selected $\# \text{LB} = 5$, $\rho_n = 7 \times 7$ roads, $\rho_c = 1000$ customers/km$^2$ and three different covering percentile values: $c_p = 50$ (with $d_{\text{max}} = 225$ m), $c_p = 84$ (with $d_{\text{max}} = 330$ m), and $c_p = 99$ (with $d_{\text{max}} = 470$ m).

As seen on Figure 26, in case $c_p = 50$, the customers on the outer edge and in between LBs (on the distances note reached by the $d_{\text{max}}$) remains un-covered. In the case $c_p = 84$, all the customers in the inner part of the network are covered, except those located in the middle of each side of the outer edge. In the case of $c_p = 99$, the majority of customers are covered and there is (statistically) only 1 percent of customers which are to be considered as uncovered (all located on small segments of the outer edge).

### 3.3. An example with asymmetric city grid transport network

In the real city centres with orthogonal structured transport network, distances between streets are most often not completely symmetric. The model was therefore adopted to simulate (using LHS method) also asymmetric city grids. Two city grids with randomly distributed streets on the area of 1 km$^2$ with $\rho_n = 7 \times 7$ roads, $\rho_c = 1000$ cus-
Figure 26. Covering patterns for customer locations for #LB=5: $\rho_n = 7\times7$ roads, $\rho_c = 1000$ customers/km$^2$ for $cp = 50, 84, 99$; red full squares present LBs, green marked lines covered locations, orange marked lines un-covered locations.

Figure 27. Two examples of randomly generated asymmetric city grids and results of simulating #LB = 5 in the form of scatter plot, contour plot and a 3D plot of probability density surface.

Figure 28. Two examples of randomly generated asymmetric city grids and results of simulating #LB = 9 in the form of scatter plot, contour plot and a 3D plot of probability density surface.
tomers/km² have been produced and simulated. Results of the simulations are presented for #LB = 5 and #LB = 9, which were identified as the most interesting in previous simulations.

Graphical locations of customers are indicating the city grid, which is presented on the far-left side of the Figures 27 and 28. In total 200 simulations have been performed for each city grid. Results are presented in the form of scatter plot, contour plot and a 3D plot of probability density surface (from left to the right).

As shown in Figures 27 and 28, simulations still result with the same number of statistically identified LBs, revealing a pattern similar to the symmetric one, but following the (asymmetric) grid structure. Obviously, the distribution of streets affects the pattern, but its geometric structure remains generally the same as in case of equidistant grid structure. In this sense, the geometric patterns obtained in the Sub-section 2.2 are proven to be robust. Finally note that statistically identified LBs are tending to areas with higher street density. For example, in case of extremely high customer’s density, 49 LBs will be located exactly on the junctions (vertices) of the asymmetric street grid.

Conclusions

A multi-parametric model for determining the optimal #LBs and location of LBs in idealized urban areas is proposed. The model is based on a statistical approach allowing consideration of different urban context and uncertainties in delivery demand. The model provides robust solutions, which are valid and transferable to any other rectangular based city and can be considered as a tool for decision-making process of urban planners worldwide.

One of the crucial input parameters of the model is the maximal allowed walking distance \(d_{\text{max}}\), of the delivery from the LB to the customer. Various simulations reveal the power law dependence between \(d_{\text{max}}\) and #LB. The dependence \(d_{\text{max}} = d_{\text{max}}(#\text{LB})\) is proved to be robust for varying the customers and grid density. In addition to that it was found that for any customers and grid density on the area of 1 km² a threshold of #LB = 20 exist. Each additional LB results in minimal (less than 4 meters) decrease of the maximal walking distance \(d_{\text{max}}\).

In-depth analysis of graphical representation of simulation results (scatter plot, contour plot and the 3D plot of PDF) indicates that for any number of selected LBs their distribution is not random, but always follows specific laws. Moreover, the increasing of the number of LBs reveals the formation of patterns from which it can be concluded that: the optimal number of LBs naturally tends to the perfect squares (4, 9, 16, 25 etc.) and depending on the grid density some perfect squares are skipped and the final numbers are limiting to \(n^2 - 4\). To fully understand all different aspects influencing the number and location of LBs in urban areas, future research should contain: point and linear distribution of customers (simulating higher concentration of demand located at certain points/areas or along a certain road) and different grid shapes of urban areas (rectangular and parallelogram-based shapes of different sizes and dimensions and finally also some irregular shapes and real city network).

It should be mentioned that although simulations presented in this paper have been done for a typical urban grid cell of 1 km² and a rectangular road network, the theoretical framework of the proposed model allows inclusion of the network and the topology of any real city of the world. Last but not least, although the presented research was focused on LBs, the obtained results can be, to some extent, applicable also to optimal number and location of public transport stations (bus, metro), locations of different public services (e.g. waste collection stations, drugstores) and different kinds of collection points (lockers, 3D printer stations).

Author contributions

Tomislav Letnik contributed mostly to the literature review, conceptual design of the model, simulation, analysing the results and writing first draft version of the paper.

Iztok Peruš was responsible for programming in MATLAB and statistical analysis of the results.

Matej Mencinger was responsible for mathematical formulation of the model, design of the algorithms and quality check of the entire paper.

Stane Božičnik contributed mostly to interpretation of results and writing conclusions.

Disclosure statement

This manuscript has not been previously published and is not under consideration in the same or substantially similar form in any other peer-reviewed media. Furthermore, we have no conflicts of interest to disclose.

Acknowledgements

The research of the third and fourth author was supported in part by ARRS (Slovenia), Research program funding No V2-1740 (third author), Research program funding No N1-0063 and Research core funding No P1-0288 (fourth author).

References


https://doi.org/10.1016/j.rtbm.2017.05.002

https://doi.org/10.1016/j.retrec.2017.09.001

https://doi.org/10.1061/(ASCE)UP.1943-5444.0000099

https://doi.org/10.1061/(ASCE)UP.1943-5444.0000011

https://doi.org/10.1787/9789264174108-en

https://doi.org/10.1061/(ASCE)UP.1943-5444.0000406

https://doi.org/10.1002/eqe.584

https://doi.org/10.1016/j.eswa.2011.11.054

https://doi.org/10.1287/trsc.1080.0235

https://doi.org/10.1002/9781119994374

https://doi.org/10.1016/j.sbspro.2012.03.091

https://doi.org/10.1061/(ASCE)UP.1943-5444.0000052

https://doi.org/10.1016/j.rtbm.2014.03.003

https://doi.org/10.1016/j.trpro.2017.05.087