Distinguish Last Mile Transit Problems in Urban Areas

Group 3: Si Chen(sichen12), Hao Sun(haos3), Huajie Shao(hshao5)

University of Illinois at Urbana-Champaign

Github link: https://github.com/shj1987/HCI565_group3.git

Demo link: http://www.leam.illinois.edu/lastmile

Abstract

Last mile problem refers to the disconnection of the end-chain reaching to homes from public transit stops. It has been one of important research topics in urban planning since it directly determines whether the public transit system is accessible for the communities in daily commuting. The prior works mainly adopt buffer or isochrone to judge whether transportation hubs are within homes' walking distance, in order to distinguish the last-mile problem, which failed to consider the commuters'real needs for specific bus lines to reach specific destinations. Different from prior works, this paper develops a novel last mile (LM) system with the help of personal informatics to find the residential cells that have last mile problem in daily transits based the transportation and land use datasets. We use City of Chicago as the example city to evaluate the performance of our designed LM system. Based on the customers' review and rating data from Yelp and public transits system in Chicago, we compute the popular score and coverage score to measure the popular community destinations and detect last mile problem, respectively. The evaluation results show that our proposed LM system can efficiently detect the last mile problem. Our results could work as potential resource for professional planners to make better planning strategies, or for real estate developers and citizens to efficiently find suitable residential areas.

Introduction

Last mile problem refers to the disconnection between homes and public transit [1]. Last mile in transportation determines the end-chain reaching to homes from clostest public transit stops. It is one of the most important design topics in city planning, since it the first step and the final step of people's daily commuting.

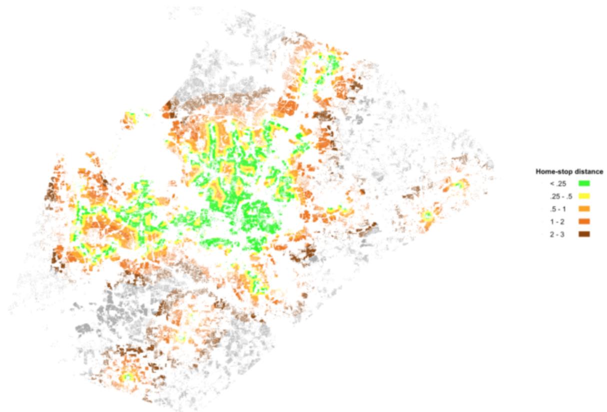


Fig. 1 home-stop distance in last mile problem

And it directly determines whether the public transit system is accessible for the communities. Last mile has a wide range of applications such as collection and delivery, travel and environment protection [1][2]. Solving the last mile problem can bring a lot of benefits to citizens and city planning. For instance, residents can better choose their living areas if they have know the end-chain public transit accessibility from home to workplace. In addition, it is helpful for the real estate companies to select developing areas. Therefore, nowadays a lot of researchers pay more attentions to the last mile problem in recent years [3-5].

In the past few years, several works have been studied on last mile problem [6-10][16]. For example, some researchers obtain commuters' locations in block group unit from Census data (national demography data source) to analyze the last mile problem. However, one problem of this method is that the analyzing unit is too big. Other researchers also adopt buffer or isochrone to analyze whether transportation hubs are within homes' walking distance as shown in Fig. 1. to judge the end-chain connecting level between home and public transit system. However, the problem is that commuters may need specific public transit lines to reach their destinations, rather than the lines with stops within homes' walking distance. In such cases, last mile problem still exists, even there are walking accessible public transit stops from homes.

In this research project, we develop last mile (LM) system to find the residential cells that have last mile problem in daily transits based on city public transit network and residential cells' locations. Specifically, we first delineate the individual residential cells' reachable areas if talking the accessible public transit lines, those have stops within 10-minute walking distance from residential cells. This paper uses 40 minutes from home to eventual destinations as the standard measurement of reachable areas, considering City of Chicago's average commuting time per travel range from 31 to 39 minutes from 2000 to 2017[11]. We further analyze whether sufficient surrounding popular commuting destinations are covered by each residential cell's reachable area using coverage ratio. Here ***coverage ratio*** is defined as the ratio for the number of popular commuting destinations in the reachable areas of residential cells to that in an area with radius of 30-minute driving isochrone. When the coverage ratio is lower than 50%, indicating there is a last mile problem. To better evaluate the performance of our proposed system, this paper chooses City of Chicago as an example considering dataset availability. We use the real-world dataset from Yelp to find popular commuting destinations based on the number of reviews and rating scores. Our experiment results show that some residential cells have last mile problems, especially those in the south areas and north boundary areas in City of Chicago. To deal with this problem, this paper proposes some effective suggestions, such as offering public bicycles and increasing bus stops.

In summary, we make three contributions in this paper:

- We analyze individual residential cells' reachable areas and popular commuting destination based on Yelp dataset to detect the last mile problems.
- We choose City of Chicago as an example to evaluate the performance of our proposed LM system.
- Experiment results show that City of Chicago performs well in most northern areas. However, in south-west areas and north boundary areas, the last mile problems are relatively serious.

The rest of this paper is organized as follows. We introduce the related work about the last problem in Section 2. Section 3 presents the overall architecture of our proposed LM system. In Section 4, we introduce the real-world datasets from Yelp and public transits system. The evaluation results are presented in Section 5. In order to resolve the last mile problem, some suggestions are made in Section 6. Section 7 discusses the limitations and future work of this paper. Finally, we conclude the paper in Section 8.

Related Work

In this section, we are going to review the related work about last mile problem. Last mile problem has attracted a lot of attentions from researchers around the world in recent years. In earlier years, researchers did several similar works about last mile problem. For instance, Bertsimas and Van Ryzin [12] proposed the Dynamic Traveling Repairman Problem for the case of single and multiple vehicles. Baldacci et al. [13] studied the Car Pooling Problem using the similar method in last mile problem.

Recently, researchers study the last mile problem in E-commerce. For example, Boyer et al. [14] attempted to identify the main factors that influence the efficiency of "last mile routes" for package deliveries. In addition, Liying Song [1] addressed the last mile problem about the collection and delivery points. Xuping Wang et al. [15] tried to explore the competitiveness of three "Last mile" delivery modes—attended home delivery (AHD), reception box (RB), and collection-and-delivery points (CDPs). Different from prior works, this paper develops a novel last mile (LM) system to find the residential cells that have last mile problem in daily transits based on city public transit network and residential cells' locations.

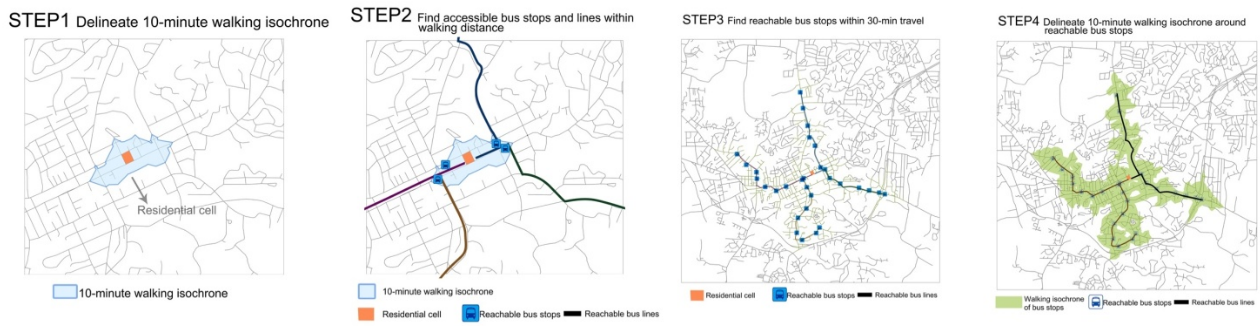


Fig.2, Finding walkable areas

Overview of Last Mile (LM) System

In this section, we are going to introduce the overall architecture of the LM system. Our proposed LM system includes three parts: (i) Finding walkable areas based on public transits; (ii) Defining popular community areas based on Yelp dataset; (iii) last mile detection based on coverage ratio. (iiii) interface design. In the following, we will introduce these modules one by one in detail.

1. Finding walkable areas

In general, a reachable area is defined by 35 minutes' covered areas from home to destinations by walking and by public transits. As shown in Fig.2, this paper adopts four steps to look up the reachable areas. The first step is to delineate 10-minutes walking isochrones. The second step is find the available bus stops and lines within walking distance. The third step is to find reachable stops within 30 minutes' travel. The final step is to delineate 10 minutes' walking isochrones around reachable bus stops. Through the above four steps, we can find the reachable areas of residents in the end.

2. *Defining popular community areas*

In this paper, we use the popular score to define the popular community areas, as shown in Fig. 3.

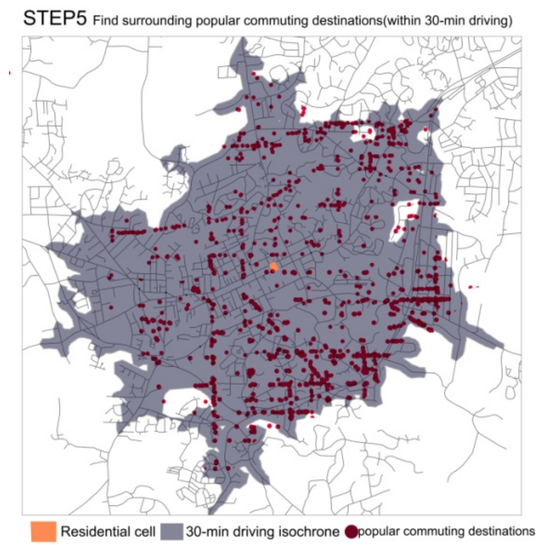


Fig.3, defining popular community areas

The popular score is the sum of weighted counts of reviews and rating for each restaurant or shopping mall on Yelp. Since the scale of the number of counts and rating is different, we first normalize the counts of reviews and rating into a range [0,1]. Then the popular score can be given by

$$popular\ score = w*norm\_counts + (1-w)*norm\_rating \qquad (1)$$

where w is the weight of each term, *norm_counts* denotes the normalized counts of reviews and *norm_rating* is the normalized rating of each facility.

We can rank the popular score to find the most popular community areas in a city. The higher the popular score, the more popular the community areas will be.

3. Last mile detection

After finding the popular areas, the next step is to detect the last mile problem. This paper adopts the coverage ratio as a measurement metric to evaluate the last mile area as shown in Fig.4. Specifically, we compute the ratio for the number of host spots in the walkable areas to that in the reachable area with radius of 15 miles by driving a car. Then we use a threshold to judge which residential cell has last mile problem based on the coverage ratio. If the coverage ratio is less than a threshold, indicating there is a last mile problem.

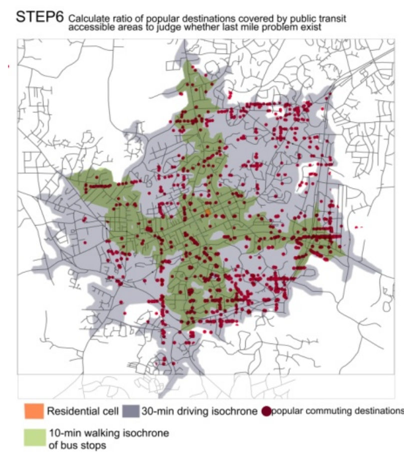


Fig.4, Last mile detection

4. Interface design

We build a web interface with mapping and feedback box, to provide a visualized result for our LM system. Our designed interface could provide useful information for both professional and non-profession users. For professional users, it can help them obtain the problems in public transit system efficiently, rather than manually deal with the tedious transportation data. For common citizens, they could obtain directly visualized last mile problems in the specific areas. In mapping part, considering the issues we address in this project are urban spatial analysis, we adopted Google API as base map to locate the residential cells with graded colors to show different last mile problem level in these areas as the following figure shows. In next step, all USA cities' last mile problem maps would be produced and displayed on the website. So we provide the 3 selection boxes to allow users to narrow down to the areas where they care about last mile problems, from state to county to city.

In feedback part, we currently encourage users to give feedback to our algorithm and interface design through a “comment” box, and report the unobserved last mile problems through a “report” box. These feedbacks would be collected by the system developers and used to improve parameters, threshold and algorithm of our LM model.

Besides these key functions, we did some decorations and further explanations on the website to make it more user friendly and help users to better understand last mile problem maps. Firstly, we designed a logo for LM system to make a sense of identity. Secondly, we showed the legend for the last mile problem maps. Thirdly, we provide our contact information for those potential cooperators or citizens who want to build connection with us.

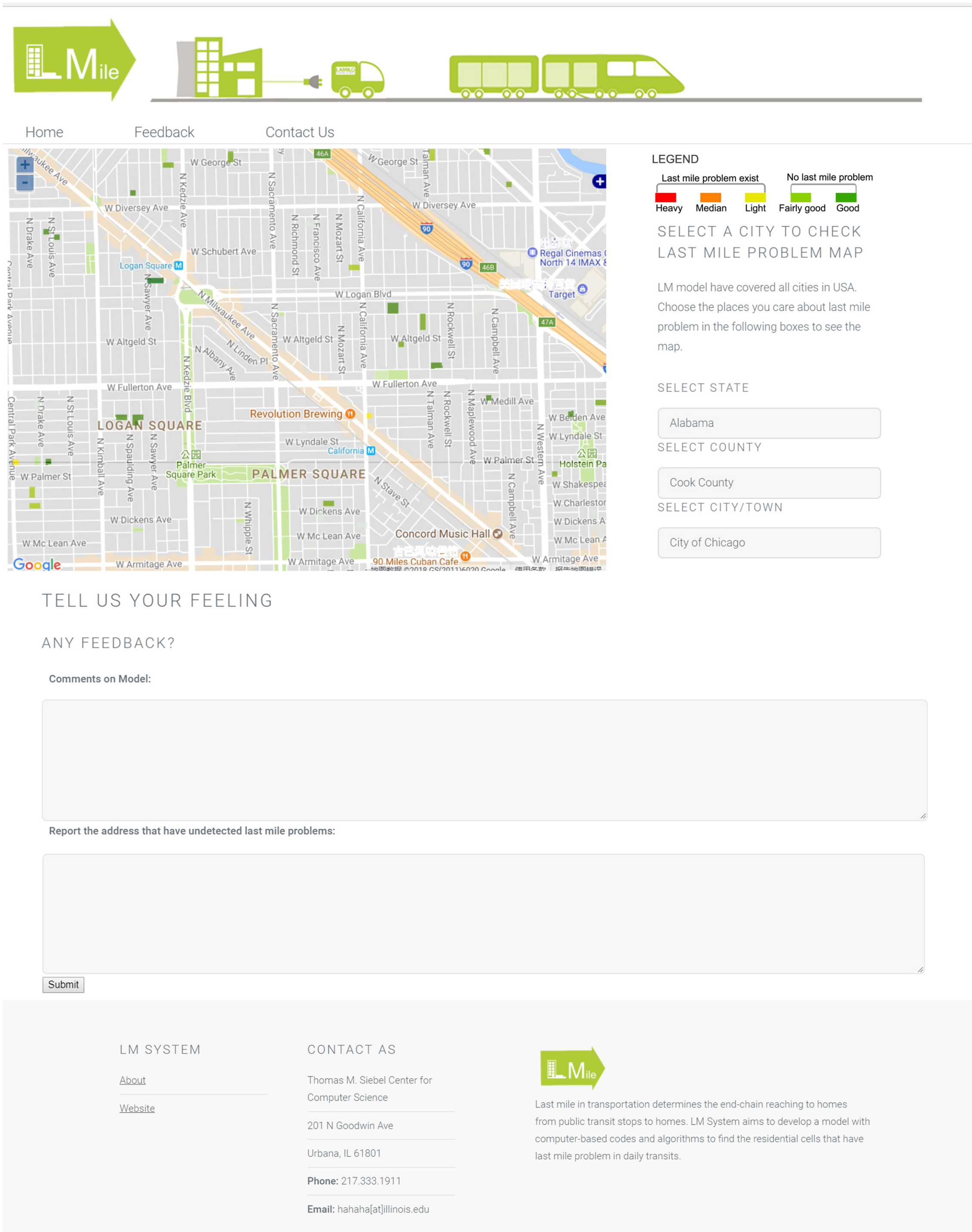


Fig.5, interface of our LM system

Datasets

1. Yelp dataset

In this paper, we use the Yelp dataset about Chicago City to calculate the popular score defined above. The dataset was collected by authors in the Github [17]. In this dataset, there are about 5200 important commuting destinations such as restaurants, hospitals and shopping malls. Each item in the dataset contains locations, category of destination, number of reviews and rating. Based on this dataset, we can compute the popular score using Eq. (1).

2. Land use map

The land map displays the commercial, residential and industrial areas and so on. We use five categories numbered as 1111, 1112, 1130, 1140, 1151 to represent single family, single family, multiple families, mixed single family and multiple families, mixed commercial and residential area, respectively.

3. Public transit system dataset

In this paper, we use public transit system dataset [18] from Chicago city to find reachable destinations. This dataset includes all the lines/routes of public transits, which can help us find the stops across the entire Chicago city.

Evaluation Results

In this section, we conduct the experiment to evaluate the performance of our designed LM system using the real-world dataset from Yelp and Public transit system. The map of residential cell's locations and bus lines/stops in City of Chicago as shown in **Appendix**. In addition, we use the ArcGIS software [19,20] to map, compile geographic data, analyze mapped information, discover geographic information and manage geographic information in a database.

In our experiment, we divide the last mile problem into 5 levels: 0-15%(heavy last mile), 15%-30% (medium), 30%-50%(light), 50%-75%(fairly good), 75%-(good). Fig.6 shows the result of last mile problem in Chicago. We can see that City of Chicago performs well in northern areas. In the most northern part of the city, the problem is relatively severe due to the low density of bus stops. In downtown Chicago, the public transportation condition is very good. Multiple bus lines go through this part of the city. However, in south-west and north boundary areas, the last mile problem is pretty severe, which is highlighted with red color. The density of bus stop is low and the area of blocks is bigger. People need much more time to utilize public transportation. Thus, we make a couple of suggestions to deal with this issue in next section.

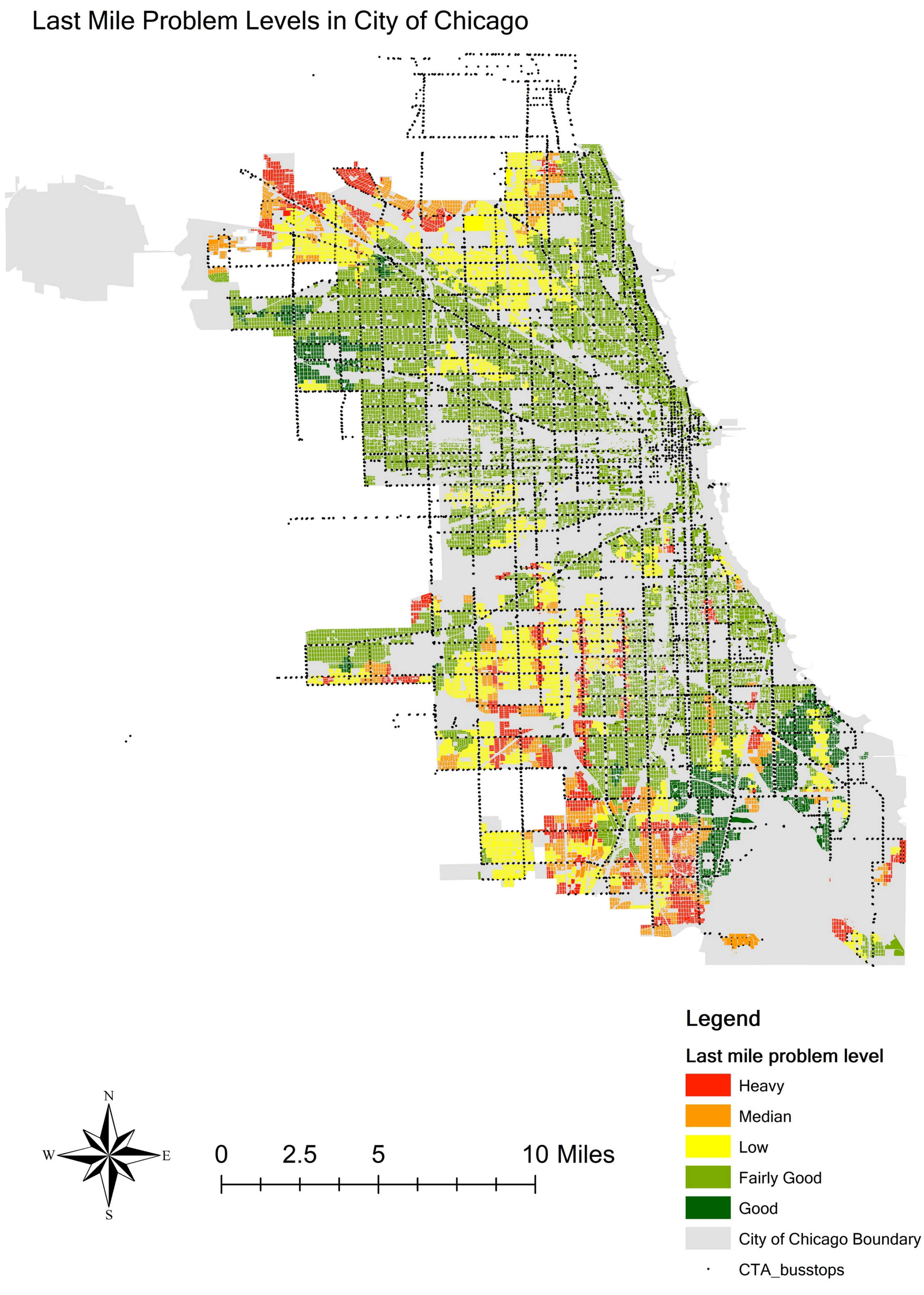


Fig.6, experiment results about last mile in Chicago

Suggestions

From the experiments, we find that there is a last mile problem in some areas in Chicago. Therefore, we develop a couple of ideas to deal with this problem as follows:

- Offering public bicycle services. Since it is hard to change the bus lines in the real-life due to the cost, one possible way is to offer public bicycles in the last miles for residents to reach their destinations. For example, in some big cities such as Beijing and Shanghai in China, companies and government have provided the public bikes to citizens to solve traffic and last mile problem.
- Increasing bus stops. Another way to handle the last mile problem is to increase the number of bus stops in each community area. Using our LM system, we can find the areas that have last mile problem. Then we can suggest the bus service company to increase the bus stops in some areas with serious last mile problem.

Discussion and Future Work

In this paper, we do not consider the public subway lines to find the researchable destinations. Actually, in some big cities such as New York, people prefer to take subway for a long-distance trip. In addition, we also need to consider the transfer from one stop to another in the public transits. In the future, it is necessary for use to take the transfer lines into account for the last mile problem. Due to the complication of our model, we also ignore the transfer between bus lines and subway lines in this paper. To better improve our last mile system, our future work will focus on the transfer problem for public transits.

Conclusions

This paper proposed a new last mile (LM) system to find the residential cells that have last mile problem in daily transits based on city public transit network and residential cells' locations.

Specifically, we first delineate the individual residential cells' reachable areas if talking the accessible public transit lines, those have stops within 10-minute walking distance from residential cells. We defined 40 minutes from home to eventual destinations as the reachable areas. We further analyze whether sufficient surrounding popular commuting destinations are covered by each residential cell's reachable area based on coverage ratio. This paper selected Chicago city as an example to evaluate the performance of the proposed LM system using the real-world datasets from Yelp and public transit system. Our experiment results show that some residential cells have last mile problems, especially those in the south areas and north boundary areas in City of Chicago. Then we make some effective suggestions, such as offering public bicycles and increasing bus stops, to handle this problem.

References:

[1] Song, L., Cherrett, T., McLeod, F., & Guan, W. Addressing the last mile problem: transport impacts of collection and delivery points. Transportation Research Record: Journal of the Transportation Research Board, (2097), 9-18, 2009.

[2] Verdict (2006), Verdict on UK Home Delivery and Fulfillment 2006, Verdict Research, London.

[3] Wang, Xuping and Zhan, Linmin and Ruan, Junhu and Zhang, Jun, How to choose "last mile" delivery modes for E-fulfillment, Mathematical Problems in Engineering, 2014.

[4] Triantafyllou, Maria and Cherrett, Tom and Browne, Michael, Urban Freight Consolidation Centers: Case Study in the UK Retail Sector, Transportation Research Record: Journal of the Transportation Research Board, pp. 34--44, 2014.

[5] Hecht, Brent J and Stephens, Monica, A Tale of Cities: Urban Biases in Volunteered Geographic Information., ICWSM, vol. 14, pp. 197--205, 2014.

[6] ALiu, Zhili, X. Jia, and W. Cheng. "Solving the Last Mile Problem: Ensure the Success of Public Bicycle System in Beijing." Procedia - Social and Behavioral Sciences 43.4(2012):73-78.

[7] Wang, Hai, Ph. D. Massachusetts Institute of Technology. "Design and operation of a last mile transportation system.", 2015.

[8]. Baldacci, R., Maniezzo, V. and Mingozzi, A. An exact method for the car pooling problem based on Lagrangean column generation. Operations Research, 52(3), 422-439, 2014.

[9]. Boyer, K. K., Prud'homme, A. M., & Chung, W. The last mile challenge: evaluating the effects of customer density and delivery window patterns. Journal of Business Logistics, 30(1), 185-201, 2009.

[10] Liu, Z., Jia, X., & Cheng, W.. Solving the last mile problem: Ensure the success of public bicycle system in Beijing. Procedia-Social and Behavioral Sciences, 43, 73-78, 2012.

[11]http://www.dailymail.co.uk/sciencetech/article-4380262/Commute-time-map-reveals-Chicago-longest-commute.html

[12] Bertsimas, D. J. and Van Ryzin, G.. A stochastic and dynamic vehicle routing problem in the Euclidean plane. Operations Research, 39(4), 601-615, 1991.

[13] Baldacci, R., Maniezzo, V. and Mingozzi, A. An exact method for the car pooling problem based on Lagrangean column generation. Operations Research, 52(3), 422-439, 2004.

[14] Boyer, K. K., Prud'homme, A. M., & Chung, W.. The last mile challenge: evaluating the effects of customer density and delivery window patterns. Journal of Business Logistics, 30(1), 185-201, 2009.

[15] MLAWang, Xuping, et al. "How to Choose “Last Mile” Delivery Modes for E-Fulfillment." Mathematical Problems in Engineering, 2014, (2014-6-11) 2014.1(2014).

[16] Balcik, B., Beamon, B. M., & Smilowitz, K. Last mile distribution in humanitarian relief. Journal of Intelligent Transportation Systems, 12(2), 51-63, 2018.

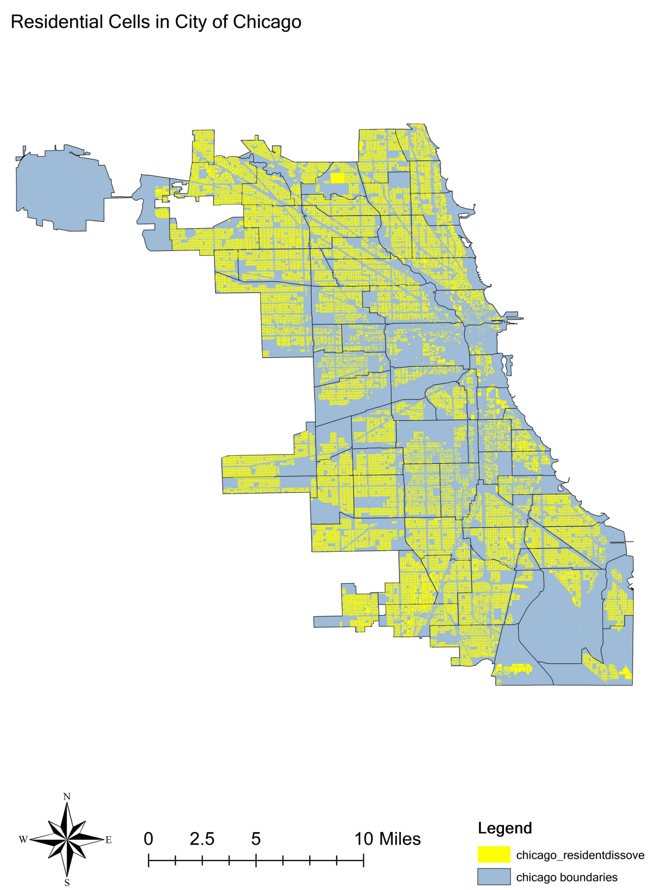
[17] Yelp dataset, https://github.com/jpvelez/restaurant_inspection_analysis/tree/master/data.

[18] Public transit system data, http://www.cmap.illinois.gov/data.

[19] Arc GIS, https://www.arcgis.com/home/index.html.

[20] https://en.wikipedia.org/wiki/ArcGIS

Appendix



Bus stops and lines in City of Chicago