Sign Optimization Model for Rail Transit: A Big Data Approach

ZHUCUI JING AND WEI BAI
School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China
Corresponding author: Zhucui Jing (zcjing@bjtu.edu.cn)

This work was supported by the Fundamental Research Funds for the Central Universities under Grant 2016JBWZ007.

ABSTRACT The subway station has a large passenger flow and strong mobility, and subway-oriented signs have become the necessary core support for the subway system. The paper first takes the complaint data of Beijing Metro as the research object and uses Word2vec methods to effectively prove the phenomenon of subway-oriented identification. Next, it builds a MIP model for the subway-oriented identification system based on optimization theory and analyzes the model. As an example, Dongzhimen Station of the subway has optimized the spatial layout of its guide signs. Dongzhimen station is an important transportation hub, there are some sign problem, such as long transfer distance, narrow transfer space, complicated traffic conditions, and obvious passenger flow aggregation. The amount of induced information in the area is the objective function, and an optimization model for the optimization of signs is established to help scientifically find the optimal placement of guide signs from multiple candidate positions. The model can theoretically reduce the guidance errors caused by the guidance marks, and will be verified by real cases in practical applications, thereby verifying the effectiveness of the model. The research results can provide reference and reference for subway passenger flow guidance.

INDEX TERMS Rail transit, identification, optimization, big data.

I. INTRODUCTION

With the rapid development of urban rail transit, the guidance sign system in the subway has gradually become one of the main ways to meet the needs of passengers. The subway station has a large passenger flow and strong mobility, and subway-oriented signs can guide people to identify the information in the subway safely and quickly, so that people can travel more safely and conveniently in the underground environment. Faced with the huge subway network and the huge daily travel pressure, subway guide signs are becoming more and more important. Therefore, higher requirements should be put on urban subway public signs.

The subway guide sign is a guiding and guiding public facility sign. It bears the important responsibility of allowing passengers to more quickly identify the best ride and transfer routes in a certain underground space environment and to travel safely by car. However, subway guide signs in underground spaces do not meet the needs of all passengers well, and some signs can cause visual misleading to passengers and cause many unnecessary troubles. First of all, in the complex environment of closed underground space, without the comparison of buildings on the ground as a reference, the passenger’s ability to accept information will be greatly weakened, and it is difficult to quickly and accurately distinguish the direction in strange and complex environments. Secondly, the setting of the sign will also have a great impact on the passenger’s acceptance of the information. Excessive information can cause visual confusion, improper and irregular logo settings can also affect passenger boarding, or the contents of the logo cannot be easily understood by passengers, which will confuse passengers. Then, because of factors such as the behavior of passengers, there are also strong uncertainties. Many factors such as age, way of thinking, whether they are locals, whether they are taking the subway for the first time will guide the understanding of the sign information content, which may cause many problems. It can be seen that the guidance signs in subway stations directly affect the smoothness and effectiveness of passenger travel. The reasonable setting of the guide signs has a significant impact on the safe and convenient travel of passengers in the subway.

Sign optimization problems can be summarized as location problems and location-based quantity problems, and MIP model is suitable for the problem analysis. Word2vec converts words from text data to vectors, making it easy to calculate

the identification system can be applied in different places setting, pictograms and icons in the identification system are information Role [1]. Uebele clarified that the color, type-oriented identification systems.

foreign countries have relatively mature rail transit systems, After many years of construction and development, many sections V presents case analysis; the last part is research IV presents optimization model and solution; section II presents literature review; section III presents text mining model; section V presents case analysis; the last part is research conclusion.

II. LITERATURE REVIEW

After many years of construction and development, many foreign countries have relatively mature rail transit systems, and there have been relatively complete studies on rail transit-oriented identification systems.

In theory, Neurath had created a new artificial language in the form of symbols combined with easy-to-understand simple graphics to express related information as early as the 1930s. Since then, logos have begun to spread and express information Role [1]. Uebele clarified that the color, type-setting, pictograms and icons in the identification system are all elements that can be changed and used, and proposed that the identification system can be applied in different places such as airports, libraries, etc [2]. Arthur et al. designed a “pathfinding experiment” to obtain the conclusion that pathfinding is a behavior of a person in order to reach a certain place in space, and to decide how to design the location, content and elements of the pedestrian according to the pathfinding behavior of the pedestrian [3], [4]. Chen designed and optimized the emergency evacuation guide signs in the supermarket through the traditional maximum coverage model (MCLP) [5], and verified the optimization model by means of micro-simulation [6]. Thompson analyzed the guide signs in the airport, and qualitatively found that the setting of the guide signs should integrate the important factors of space environment and passenger characteristics [7].

Duszek analyzed and evaluated the user’s thought process in the Warsaw metro sign system in Poland in detail, and found that the user’s needs were identified by the user-oriented approach to the identification guidance information at different locations [8]. Wakide used Yokohama subway station in Japan as an example to summarize the traffic flow in the hub station, which is conducive to installing such standardized and easy-to-understand traffic guidance signs throughout the station [9]. Zeng designed and established a method for evaluating the performance of subway station identification systems to find the shortcomings and limitations of existing signs at subway stations, and to prepare for the further optimization and improvement of the guidance function of subway signs [10]. Gyu-Yeob Jeon, Won-Hwa Hong studied the changes of human behavior with phosphorescent guidance signs by inviting 103 scientific research subjects to conduct evacuation experiments in an environment with impaired visibility in the subway stations, and concluded that the phosphorescent guidance signs in subway stations were more It is convenient to find the experimental results of the evacuation route [11]. Vilar, because the pedestrian’s pathfinding behavior in other situations is not as good as daily rationality, so he designed a real scene test to verify that the guidance sign has a greater effect on daily pedestrians than in emergency situations [12]. Nassar found that the visual effects of the guidance signs are inversely proportional to the layout distance, and compared the guidance effects of three different guidance sign layout methods to determine the optimal layout of the guidance signs [13]. Wang believed that the attractiveness of research subjects, the distance between pedestrians and research subjects, and the occupancy rate of research subjects in pedestrian fields of vision are important factors that influence the visibility of research subjects [14]. Motamedi et al. used BIM-enabled virtual reality (VR) environments to optimize the placement of signage in public spaces [15]. Tu et al. analyzed the passenger’s abnormal travel path by analyzing the OD data of passengers, and further propose the corresponding subway sign optimization scheme [16]. Yin et al. built a simulation model to capture the passenger’s path selection behavior, considers the impact of the guidance information on passenger behavior, simulates a dynamic passenger flow distribution, and proposes an optimization scheme for passenger flow guidance information [17].

From a practical point of view, the British government proposed to apply the unlined font “railway body” as a standard font to the London Underground Signage System, becoming the first country in the world to use this font [9]. The United States is the world’s earliest country for large-scale logo design. Until 1974, the Graphic Design Institute of the United States designed 34 new traffic signs for transportation hubs [18]. Until 1979, 16 were added to the original basis A total of 50 new traffic signs. Unlike other signs, these new traffic signs are universal, accurate, and internationally recognized. They can be understood by people in various countries mainly by means of graphics, used by transportation hubs,
and facilitated by pedestrians to specific facilities, directions, and transportation. The content is clear at a glance [19]. The US Department of Transportation’s “Complete Collection of National Public Sign Design Principles and Graphics” specifies the principles of public signs in the United States [20]. This standardized graphic sign is used by various countries in the world. In order to facilitate the understanding and use of the signs by foreign passengers in Japan, Japanese, English, Chinese, and Korean characters will be provided on the visualized text. These signs can be used by foreign passengers in many languages to travel smoothly through these standardized languages. In addition, a variety of signs and devices for the visually impaired have been set in the Japanese subway [21]. Signs such as convex images, raised text, and generators with guide signs have greatly facilitated the visually impaired. subway. The Japan Metro has special seats in different colors designed for pregnant women, children, and people with disabilities. The old, young, and disabled signs are also very eye-catching, and there are also “priority seats” on the floor of the special seat car platform.

Identification [22].

Based on the above research, the guidance identification system provides passengers with guidance information services to correctly guide passengers to complete activities such as ticket purchase, ticket checking, waiting, boarding, transfer, exit, etc., to ensure that passengers are safe, orderly, and smoothly travel by car. The rationality, standardization and scientific planning and design of the system directly affect the operating efficiency of the subway. However, current rail transit-oriented signs still have the following deficiencies:

1. The cause of the subway guidance sign problem is unclear. At present, the research focuses on the research of identification optimization, and the reasons for the identification problems are rarely described.

2. The setting of subway signs is not standardized and systematic. The location of subway guide signs is unreasonable, which is likely to cause congestion, and too many signs at the same location lead to excessive information and the accumulation of guide signs is complicated, which affects passengers to obtain accurate information.

Therefore, the article first analyzes the complaint data of Beijing Metro, effectively proves the problem of subway signage, and then builds a MIP model of the subway signage system based on the optimization theory, and analyzes the model; finally, it takes Beijing subway Dongzhimen Station as an example.

III. TEXT MINING IDENTIFICATION TO IDENTIFY PROBLEMS

A. SUBWAY SIGN IDENTIFICATION COMPLAINT

The Beijing Metro Hotline is an important communication channel between Beijing Metro and passengers. It accumulates a large amount of user complaint information every year. Such information often exposes problems in the operation process or discovers potential needs of customers. Adjustment and improvement of services have an important guiding role. On the one hand, the number of complaints is very large, and it takes a lot of manpower to classify the problems, and then distribute them to the relevant departments for resolution and feedback; on the other hand, because the complaints are more complicated, it is sometimes difficult to determine the source of the problem, which reduces the efficiency of problem handling. There are many types of complaints and different ways of expressing passengers. Beijing Metro’s current method of handling complaints one by one lacks comprehensive analysis in the later stages, fails to effectively use a large amount of complaint information, deeply digs passenger needs, and finds Beijing. Lack of subway.

Among the many complaints, there are some complaints about subway signs and labels. These complaints are mainly about the lack of signs and signs in the subway, the contents of the signs, or the signs being damaged or damaged. Misleading, inconvenient information acquisition, feedback to the subway management department through hotline, email, etc. An analysis of these subway sign identification complaint texts will help the subway management department to improve the design and setting of the sign identification, and make it easier for passengers to travel.

This article collects 2040 complaints about subway signs and logos in 2018, and the formed text is shown below:

There are three main steps in the processing of subway sign identification text complaints: the first step is to pre-process the text, including text deduplication, text segmentation, and removal of stop words to form the data format required for text analysis; the second step, Basic statistics of the complaint text. Separate the lines and stations from the complaint text and analyze the lines and stations where the complaint is concentrated. The third step is to analyze the content of the complaint text, including the extraction of topic keywords, syntactic dependency analysis, Word cluster analysis and keyword co-occurrence network analysis. The processing flow is shown in the Figure 2:

B. METRO SIGN COMPLAINT TEXT PREPROCESSING

Because the complaint data is aggregated manually, there is a large amount of duplicate or similar text data in the original text of the complaint text. It is necessary to pre-process the original data and perform duplicate operations on the duplicate data.

FIGURE 1. Subway sign logo complaint collection text.
1) DEDUPLICATION OF RAW DATA
Because the complaint data is aggregated manually, there is a large amount of duplicate or similar text data in the original text of the complaint text. It is necessary to pre-process the original data and perform duplicate operations on the duplicate data.

2) TEXT SEGMENTATION
Words are the basic unit of natural language processing. The processing of text content in text mining is also based on words. This research uses the Jieba word segmentation algorithm. Its main functions include Chinese word segmentation, part-of-speech tagging, recognition of new words, entity name recognition, and so on. The word segmentation accuracy is high, and the recall rate of new word recognition based on role tagging is higher than 90%, and the part-of-speech tagging and word segmentation processing speed reaches 543.5KB/s. The word segmentation stage is to perform Chinese word segmentation, part-of-speech tagging for each text, and output its results in a prescribed format. Because some special vocabulary algorithms cannot be directly identified, in the process of word segmentation, some special words are manually labeled by manual means to establish a word segmentation dictionary.

3) FILTER STOP WORDS
Because in the word collection obtained in the word segmentation stage, many words are meaningless. The impact of these words on the analysis work can be ignored, but if these words are not used as text feature words, Bringing large errors to text classification results, these words are often referred to as stop words in this article. Removed words marked as part of speech: orientation, preposition, quantifier, auxiliary, punctuation, non-morpheme, mood. These words are like, “Yeah, this, if, that, then.” If these words are added to the text analysis, it will increase the cost of our text analysis. Therefore, we introduce a stopword dictionary and remove these stopwords in the process of word segmentation.

According to all available resources, various stopword lists such as “HIT’s Stop Words Thesaurus”, “Sichuan University Machine Learning Intelligent Laboratory Stop Thesaurus”, and Baidu’s Stop Words will be organized to focus on extracting Chinese Words (instead of a large number of English words and Chinese punctuation marks) came out of a more comprehensive vocabulary, with a total of 1598 stop words.

C. ANALYSIS RESULT OF SUBWAY SIGN LOGO COMPLAINT TEXT
1) STATISTICS ON COMPLAINT LINES FOR SUBWAY SIGNS
There is a description of the line in the passenger’s sign identification complaint text, for example: “On December 31, the passenger’s Weibo reflected the stepping off of the line 8 rocket Wanyuan Station stairs.” The text line was obtained by text feature extraction. Relevant keywords and statistics obtained through statistics are shown below:

| Serial number | Site name    | Number of complaints |
|---------------|--------------|----------------------|
| 1             | Nanluoguxiang| 70                   |
| 2             | Sanyuan Bridge| 70                   |
| 3             | Huijialou    | 70                   |
| 4             | Dongzhimen  | 60                   |
| 5             | Shaoyaoju   | 50                   |
| 6             | Dongdan     | 50                   |
| 7             | Bagou       | 50                   |
| 8             | Songjiazhuang| 40                   |
| 9             | Xizhimen    | 40                   |

2) STATISTICS OF COMPLAINT STATIONS IN SUBWAY SIGNS
There is a description of the site in the passenger’s sign identification complaint text, for example: “Weibo reflects that the station logo of the south exit of Huixin West Street of Line 5 has recently changed its font size to be too small to see clearly.” By means of text feature extraction, Get the keywords related to the text site, and get statistics as follows:

3) EXTRACTION OF SUBJECT WORDS FOR COMPLAINT TEXTS
   a: TF-IDF TOPIC FEATURE EXTRACTION ALGORITHM
   TF-IDF (Term Frequency-Invers Document Frequency) is a weighting technique commonly used in information processing and data mining. This technique uses a statistical method to calculate the importance of a word in the entire corpus based on the number of occurrences of the word in the text and the frequency of documents appearing in the entire
corpus. Its advantage is that it can filter out some common but irrelevant words, while retaining important words that affect the entire text. The calculation method is shown in the following formula.

\[
tfidf_{i,j} = tf_{i,j} \times idf_i
\]  

where in the formula, \(tfidf_{i,j}\) represents the product of word frequency \(tf_{i,f}\) and inverted text word frequency. A larger TF-IDF value indicates that the feature word is more important to the text.

TF (Term Frequency) indicates how often a certain keyword appears in the entire article.

IDF (InversDocument Frequency) indicates the frequency of calculating inverted text. Text frequency refers to the number of times a certain keyword appears in all articles in the entire corpus. Inverted document frequency is also called inverse document frequency. It is the inverse of document frequency. It is mainly used to reduce the effect of some common words in all documents that have little effect on the document. The following formula is the calculation formula of TF word frequency.

\[
tf_{i,f} = \frac{n_{i,j}}{\sum_k n_{k,j}}
\]

where \(n_{i,j}\) is the number of times a feature word appears in the text and \(n_{k,j}\) is the number of all feature words in the text. The result of the calculation \(tf_{i,f}\) is the word frequency of a feature word.

The following formula is the calculation formula of IDF.

\[
idf_i = \log \left( \frac{|D|}{|j : t_i \in d_j|} \right)
\]

where \(|D|\) represents the total number of texts in the corpus and the number of feature words contained in the text.

Take this complaint text as an example, “Passengers report that there is a bus information notice board around the station at Exit B of Haidian Wuluju Station on Line 6 with a lot of patches on it, which affects aesthetics. Passengers hope to be replaced with new information in time. Notice board.”

4) ANALYSIS OF SYNTAXIC DEPENDENCIES

In the previous chapter, we obtained the topic keywords in the complaint text. These topic keywords are the object of passenger complaints. How to analyze the problems that occur in these object of complaints, we need to dig out the topic keywords through the syntactic dependency relationship. Related adjectives, verbs, etc.

Dependency is a very widely used grammatical form, which was first proposed by French linguist Terre in his book “Basic Structure Syntax” (1959). Represents an asymmetric...
relationship between two sentence components (generally words). One component is called the dominant component (or head), and the other component is called the dependent component (or modifier). There are many ways to distinguish the dominating and subordinate components. For example: the dominating component determines whether the subordinate component is optional; the dominating component determines the appearance form of the subordinate component; the dominating component specifies the semantic object, and the subordinate component merely supplements. Dependency can be either the grammatical relationship between words in a sentence, or the semantic relationship.

Dependency syntax explains its syntactic structure by analyzing the dependencies between components in a language unit, and claims that the core verb in a sentence is the central component that governs other components. But it is not dominated by any other component, and all dominated components are subordinate to the dominator in some relationship.

Through the visual display of the syntax analysis, the example “On December 31, the passenger’s Weibo reflected the steps of the rocket Wanyuan station on Line 8 and the guidance signs fell off.”

When the subject is the dominator, the syntactic dependency analysis is used to obtain the relational triples of the relevant subject words. From the perspective of obtaining the triads, there are mainly the relationship between the center, the verb and the subject, and the subject-predicate relationship. Among them, the central relationship acquires the nouns related to the subject word, the verb-object relationship acquires the subject-related verbs, and the subject-predicate relationship acquires the adjectives. The relational triples of related adjective extraction are shown in the following table:

### 5) TOPIC KEYWORD CLUSTERING

The most traditional word vector is expressed by One-hot Representation. Each word is numbered, and the length of the word vector is the length of the word list. The word vector of a word is the position of the word number is 1, and the remaining positions are 0. When the word vector formed in this way is very large, the problem of dimensional

---

**TABLE 5. Syntactic dependencies.**

| Relationship type          | Tag    | Description          |
|---------------------------|--------|----------------------|
| Subject-predicate relation| SBV    | subject-verb         |
| Verb-object relationship  | VOB    | verb-objective       |
| Goal relationship         | GOB    | goal-object          |
| Pre-object                | FOB    | fronting-object      |
| Part language             | DBL    | double               |
| Fixed Sino relation       | ATT    | attribute            |
| State-like structure      | ADV    | adverbial            |
| Dynamic complement structure| CMP  | complement          |
| Constellation             | COO    | coordinate           |
| Referee relationship      | POB    | preposition-object   |
| Left attachment           | LAD    | left adjunct         |
| Right additional relation | TAD    | right adjunct        |
| Independent structure     | IS     | independent structure|
| Core relationship         | HED    | head                 |

---

**TABLE 6. The example of syntax analysis.**

| 12 | ATT | month |
|----|-----|-------|
| 31 | ATT | day   |
| day | ADV | reflect |
| passenger | ATT | weibo |
| weibo  | SBV | reflect |
| reflect | HED | ##core## |
| 8   | ATT | number |
| number | ATT | line |
| line | ATT | rocket |
| rocket | ATT | wanyuan |
| wanyuan | SBV | station |
| station | ATT | guide |
| steps | VOB | station |
| ‘s  | RAD | station |
| guide | VOB | reflect |
| sign | SBV | fall of |
| fall of | VOB | guide |
| .   | Punctuation | reflect |

---

**TABLE 7. Relational triples.**

| Fixed Sino relation | Verb-object relation | subject-predicate relationship |
|---------------------|----------------------|-------------------------------|
| Circuit diagram     | location             | identification               |
| information         | Incorrect            | sign                          |
| Circuit diagram     | recogniton           | sign                          |
| sign                | problem              | sign                          |
| arrow               | direction            | Circuit diagram              |
| letter              | capital              | sign                          |
| No.                 | position             | subject-predicate            |
| sign                | direction            | misleading                    |
| sign                | information          | sign                          |
| sign                | missing              | font                          |
| No.                 | font                 | informat                          |
| sign                | font                 | enlarge                       |
| pinyin              | Transliteration      | Circuit diagram              |
| circuit diagram     | paste                | color                         |

---

(VOLUME 8, 2020)
disaster will occur, and the semantic relationship between words cannot be expressed. The word vector obtained by the open source tool Word2Vec uses distributed representation (Distributed Representation). Different from the traditional One-hot method, the basic idea is to train on the experimental corpus, and the words can be expressed as dimensions. The smaller number of word vectors, avoids the dimensional disaster, and makes the distance between the word vectors of words with similar semantics relatively close.

Google in 2013 opened Word2vec’s software tool for training word vectors, which can express words in the form of vectors.

Word2Vec model includes two training models, namely skip-gram model and Continuous Bag-Of-Words (CBOW) model. Among them, the skip-gram model uses the current word to predict the context, so the training input of the skip-gram model is the word vector of a specific word, and the output is the word vector of the specific word. The CBOW model is exactly the opposite of the skip-gram model. The CBOW model predicts the current word by context, so the training input of the CBOW model is the word vector of a specific word before and after, and the output is the word vector of that specific word.

The Skip-Gram model predicts the context based on the current word. Given a word sequence \( w = \{w_1, w_2, \cdots, w_m\} \), the model maximizes the average log probability as:

\[
l(w) = \frac{1}{M} \sum_{m=1}^{M} \sum_{-L \leq i \leq L} \log p(w_{m+i}/w_m)
\]  

(4)

where \( L \) is the size of the context window.

The CBOW model predicts target words by delimiting window words. Given a word sequence, the model maximizes the average log probability as:

\[
l(w) = \frac{1}{M} \sum_{i=L}^{M-L} \log p(w_i/w_{i-L}, \cdots, w_{i+L})
\]  

(5)

where \( L \) is the size of the context window.

Word2vec is used to train the complaint text to convert Chinese words into word vectors. For example, after pretraining, the “sign” is converted into a 256-dimensional word direction, as shown below.

Since high-dimensional word vectors cannot be clustered, the 256-dimensional word vector is reduced to two-dimensional, for example, the word vector of “identification” is reduced to two-dimensional, that is, \([2.11701938e-03, -1.12757953e-02]\)

Cluster the topic keywords, as shown in the Figure 5:

Through keyword cluster analysis, we can divide the complaints of subway sign identification into three categories: unclear identification, incorrect identification, and missing identification:

**a: UNCLEAR IDENTIFICATION**

Unclear signs mainly refer to the improper writing of guide signs in subway stations, unclear fonts, damaged appearance, improper placement of signs, or inconspicuous design of sign colors, affecting passengers’ quick and accurate understanding of signs. Related keywords with unclear identification include: “(not) clear”, “fall off”, “location”, “fading”, etc.

**b: INCORRECT IDENTIFICATION**

The expression of the identification information is unclear, ambiguity or obscurity occurs, and it is easy to make it difficult for passengers to obtain correct information through the identification content, or even only get wrong information. Relevant keywords with incorrect identification include: "fall off", "not clear", "missing", etc.

**c: MISSING IDENTIFICATION**

Missing signs refer to the failure to set up guide signs in subway stations, affecting passengers’ quick and accurate understanding of signs. Related keywords with missing identification include: "missing", "not present", "missing", etc.
include: “Misleading”, “Incorrect”, “(Not) Consistent”, “(Not) Updated”, etc.

c: MISSING LOGO
Lack of necessary guidance signs, passengers have difficulty finding subway entrances and exits in the space outside the station, and lack of environmental guidance information on the ground space of the station. Relevant keywords that identify missing include: “no”, “missing”, “adding” and other words.

According to statistics, there are 530 complaints about unclear signs, 440 complaints about incorrect signs, and 690 complaints about missing signs. Three types of complaints are shown below:

IV. OPTIMIZATION MODEL FOR IDENTIFYING PROBLEMS
The layout design of the guide mark is an important part of the guide mark system. It should be set and optimized reasonably and scientifically according to the starting point of the streamline. The flow pattern and direction of the passengers determines the choice of the deployment points. It is better not to set as many guidance signs as possible. Complex and excessive signs will confuse passengers’ vision and thinking. Passengers will not be able to quickly identify the required information, which will cause stagnation and congestion.

The key to optimizing the logo is to choose the location of the layout of the logo [23], [24]. It is essentially a complex decision-making point optimization problem [25]–[28]. It refers to selecting the most appropriate and optimized point from multiple alternative locations to set the installation logo. Make it more effective to provide sufficient information for passengers. Therefore, this problem is solved by establishing an optimization mathematical model for the placement of the guide marks. The position where passengers need to stay and make a decision is called a decision point. Whether to set a guide mark at a certain decision point is a 0-1 planning problem, and at which decision points to set a guide mark is an optimization model.

A. PROBLEM DESCRIPTION
Figure 7 shows the conceptual model for optimizing the layout of the guide marks.

The layout of the guide signs for urban rail transit stations is based on the passenger flow line, which guides passengers to make decisions at the nodes and autonomously reach the destination point in accordance with the established flow line. The method of optimizing the location of the guide mark placement points of urban rail transit stations is described as follows: Assume that there are m alternative sign placement points (m ∈ guide mark alternative point set), and from m alternative points according to certain constraints The condition calculation screens out p places that meet the

![FIGURE 6. Subway signs identify complaints statistics.](image)

![FIGURE 7. Conceptual model for optimizing the layout location of guide marks.](image)
requirements of the restriction to set the guide mark. The 0-1 function can be used to represent the decision of the location of the guide mark:

\[ f(j) = \begin{cases} 1, & \text{Guide flags at alternative point } j \\ 0, & \text{No guidance sign at alternative point } j \end{cases} \] (6)

where \( j \) is an alternative position, \( j = 1, 2, \ldots, m \).

### B. OPTIMIZATION CONSTRAINTS

1) Assume that guide signs are set at the intersections of passenger flow lines, entrances, stairs, elevators, corners, etc. Therefore, a set of position-oriented mark alternative points such as the intersections of flow lines, entrances, stairs, elevators, and corners;

2) The display information of the guidance signs should not be too far away. Therefore, it is assumed that the guidance signs set at each candidate point only show the guidance information within 3 nodes around the location, and do not display the guidance information beyond the range of 3 nodes;

3) Assume that one level within the station is used as the study area.

### C. BUILDING AN OPTIMIZATION MODEL

1) OBJECTIVE FUNCTION

A guideline identification point optimization model that uses the maximum amount of induced information in the study area as the optimization goal. The relationship of the optimization objective function is:

\[
\text{max } z = \sum_{j=1}^{m} \sum_{i=1}^{n} \frac{f(j)p_{ji}w_{ji}^{p}}{d_{ji}} \quad (7)
\]

In the formula, \( j \) – an alternative position for the guide mark setting, that is, a decision point, \( j = 1, 2, \ldots, m \);

\( i \) -- the destination where the passenger needs to reach, \( i = 1, 2, \ldots, n \);

\( f(j) \) – a function of 0-1, as shown in formula (6), when the guide mark is set at the alternative point \( j \), \( f(j) = 1 \); when the guide mark is not set at the alternative point \( j \), \( f(j) = 0 \);

\( p_{ji} \)– Guidance coefficient (distribution coefficient), assuming that the guidance mark has been set at point \( j \), when the passenger arrives at the target location \( i \), he can receive the guidance sign guidance service at point \( p_{ji} = 1 \), otherwise \( p_{ji} = 0 \);

\( w_{ji} \) – passenger flow, which indicates the passenger flow guided by the guidance sign set at point \( j \) when the passenger arrives at the target location \( i \). Under the same conditions, the larger the passenger flow and the more passengers being guided, the greater the amount of induced information and the higher the effect of the flag setting;

\( \alpha \) – weight coefficient, usually a selected constant, which indicates the balance and weight relationship between the induced passenger flow and the induced distance of the guide mark. When the induced passenger traffic weight is larger, the value of \( \alpha \) will be larger; when the induced distance weight is larger, the value of \( \alpha \) can be slightly smaller, and vice versa; in this study, \( \alpha = 1.5 \).

\( d_{ij} \) – The number of intersections refers to the number of intersections that passengers can pass during the journey from the guidance mark at point \( j \) to the target point \( i \).

From the above optimization model, it can be found that the traffic flow of the guide mark is positively related to the amount of inducement information, that is, under the constraint conditions, the more passenger flows are guided by the guide mark, the more obvious the effect of the guide mark setting, and the more sufficient the level of induction. On the way from the setting point of the guide mark to the target location, if the passenger passes through fewer intersections and is shorter from the target location, the greater the amount of information provided by the guide mark, which indicates that the higher the level of induction of the guide mark.

2) CONSTRAINTS

i. Constraints on the number of guiding signs. Assuming that there are \( p \) suitable points for setting the guiding signs, the constraints are as follows:

\[
\sum_{j=1}^{m} f(j) = p, \quad j = 1, 2, \ldots, m \quad (8)
\]

where \( m \) – the number of candidate positions for the guide mark;

\( p \) – The number of suitable placement points for the guide mark, \( 1 < p < m \) and \( p \) is an integer.

ii. Assuming that all passenger flows have been directed at least once, the constraints are as follows:

\[
\sum_{j=1}^{m} f(j) > 0 \quad (9)
\]

where \( r \) – the actual passenger flow into the subway;

\( I_r \) – The set of alternative points for guiding signs that must be passed after the \( r \) passenger flow enters the subway.

iii. Each target location has been shown at least once by the guide sign, and the constraints are as follows:

\[
\sum_{j=1}^{m} f(j) > 0 \quad (10)
\]

where \( i \)–the actual passenger flow into the subway;

\( I_r \) – The set of alternative points for the guide signs that must be passed after the \( r \) passenger flow enters the subway.

Summarizing the above discussion, the optimal mathematical model is:

\[
\text{max } z = \sum_{j=1}^{m} \sum_{i=1}^{n} \frac{f(j)p_{ji}w_{ji}^{p}}{d_{ji}} \quad (11)
\]
D. MODEL SOLUTION ANALYSIS

The optimization model established in this study was selected to solve using Lingo software. Lingo (Linear Interactive and General Optimizer), which is an “interactive linear and general optimization solver”, is a set of software packages developed by the American LINDO system company specifically for solving optimization problems. Lingo software converts mathematical models into computer programming languages to solve. It is very powerful, convenient, flexible, and fast to execute. It is often used to solve optimization problems.

Compared with other software, Lingo software has a concise programming language, which is basically consistent with its optimization model when entering code programs, and can exchange data with software such as Excel. It can also solve linear programming (LP), nonlinear programming (NLP) and other optimization models. Lingo software has helped solve optimization problems in various fields and has a wide range of applications. It can solve optimization problems such as location selection problems, shortest path problems, assignment problems, allocation problems, minimum spanning tree problems, and drilling layouts.

It can be foreseen that the optimization problem of Lingo software applied to the guidance sign placement point can be accurately and quickly solved, which will inevitably make the choice of the placement point efficient and practical. The steps for establishing an optimization model and using Lingo software are as follows:

1. Identify optimization issues. Through the analysis of the current situation, we find the existing deficiencies and identify the problems that need to be optimized. For example, what needs to be optimized in this research is the layout optimization of subway guide signs.

2. Analysis of influencing factors. According to the problem description, analyze the factors and variables that currently affect the optimization goal.

3. Establish the objective function. The decision problem is transformed into an objective function that can be expressed in a mathematical model, and must begin with “max =” or “min =”.

4. Establish constraints. Propose some kind of “restriction” or “constraint” a condition that the variables in the objective function must satisfy.

5. Determine the parameter value. In the optimization model, some basic parameters are usually used to determine the variables. This step should determine the parameter values according to the actual situation.

6. Use Lingo to solve. The mathematical model is converted into Lingo language input, among which a large number of functions in Lingo software can help solve complex optimization models. The Lingo solver will determine the type of solution based on the input program code. Based on this, different methods are used to solve different types of optimization problems.

V. CASE STUDY OF IDENTIFICATION ISSUES

A. CASE SITUATION

Dongzhimen Metro Station is located in Dongzhimen Bridge, Dongcheng District, Beijing, and it is also a very important traffic conversion center in eastern Beijing. Dongzhimen Station is a first-class comprehensive transportation hub that integrates various modes of transportation, such as subway, bus, taxi, private car, and bicycle.

Select the subway data of passengers who are within 2 minutes and whose passenger frequency is 1 within a day, and perform statistical analysis on the stations that have entered and exited at the same station within 7 days. November 4, 2017 to 2017 On November 10, the station with the most number of problems in the same station was Dongzhimen Station of Metro Line 2. Dongzhimen Station has always been responsible for Beijing’s public transportation hub, and is the largest transportation hub in Asia. The station has a huge passenger flow and extremely complicated traffic conditions. Once the guidance signs are set unreasonably, leading to inaccurate information transmission, it is extremely easy for passengers to stay and give up. Therefore, Dongzhimen Station was selected as the actual case for analysis.

B. OPTIMIZATION MODEL

1. Identify alternative locations, target locations, and foot traffic

   The Figure 9 shows the passenger flow line on the second floor of Dongzhimen Station.

   In order to conveniently describe the behavior of passengers when they get off the bus, such as transfer, exit, etc., passengers are divided into 3 areas after getting off the subway line 2, namely A1, A2, A3. Target locations 1, 2, and 3 respectively indicate that passengers get off the train and transfer to the subway station, airport line, and bus terminal hall. Target location 4 indicates that passengers leave the station and exit the station. Select the intersection of passenger flow lines and decision points in the transfer area of the second basement of Dongzhimen subway station as the alternative location j (j = 1,2, ..., 11) for the placement of the guide marks, and obtain the parameters in the optimization model. The calculation table is shown in the following table:

2. Establishing the objective function

   Establish the relationship of optimization objective function according to formula (7):

   \[
   \text{max } z = \sum_{j=1}^{11} \sum_{i=1}^{4} \frac{f(j)p_{ji}w_{ji}^{1.5}}{d_{ij}}
   \]
Based on the above formula, the optimization objective function can be obtained in (14), as shown at the bottom of this page.

(3) Determine constraints

According to the constraint conditions for the placement of the guide marks according to formula (8), set the number of guide marks to six.

$$\sum_{j=1}^{11} f(j) = 6 \quad (15)$$

According to the constraint condition of formula (9), that is, all passenger flows have been guided at least once. The corresponding passenger flow constraints for the starting points $A_1, A_2,$ and $A_3$ are:

$$A_1 : f(1) + f(2) + f(3) > 0 \quad (16)$$

$$A_2 : f(6) + f(8) + f(9) > 0 \quad (17)$$

$$A_3 : f(8) + f(10) + f(11) > 0 \quad (18)$$

According to the constraint condition of formula (10), each target place is displayed at least once by the guide mark. The

$$\max z = \frac{f(1) \times 135^{1.5}}{1} + \frac{f(2) \times 375^{1.5}}{1} + \frac{f(2) \times 247^{1.5}}{2} + \frac{f(2) \times 157^{1.5}}{1} + \frac{f(3) \times 282^{1.5}}{1} + \frac{f(3) \times 157^{1.5}}{1}$$

$$+ \frac{f(4) \times 273^{1.5}}{2} + \frac{f(5) \times 156^{1.5}}{1} + \frac{f(6) \times 238^{1.5}}{2} + \frac{f(6) \times 192^{1.5}}{1} + \frac{f(6) \times 157^{1.5}}{1}$$

$$+ \frac{f(7) \times 35^{1.5}}{1} + \frac{f(7) \times 157^{1.5}}{1} + \frac{f(8) \times 242^{1.5}}{2} + \frac{f(8) \times 105^{1.5}}{3} + \frac{f(8) \times 51^{1.5}}{2} + \frac{f(8) \times 8^{1.5}}{1}$$

$$+ \frac{f(9) \times 116^{1.5}}{1} + \frac{f(9) \times 8^{1.5}}{1} + \frac{f(10) \times 261^{1.5}}{3} + \frac{f(10) \times 188^{1.5}}{4}$$

$$+ \frac{f(10) \times 65^{1.5}}{3} + \frac{f(10) \times 73^{1.5}}{2} + \frac{f(11) \times 65^{1.5}}{2} + \frac{f(11) \times 8^{1.5}}{1}$$

(14)
constraints of passenger flow at target locations 1, 2, 3, and 4 are:

Target location 1:
\[ f(1) + f(2) + f(8) + f(9) + f(10) > 0 \] (19)

Target location 2:
\[ f(2) + f(3) + f(6) + f(7) + f(8) + f(9) + f(10) > 0 \] (20)

Target location 3:
\[ f(2) + f(3) + f(4) + f(5) + f(6) + f(7) + f(8) + f(9) + f(10) + f(11) > 0 \] (21)

Target location 4:
\[ f(8) + f(10) + f(11) > 0 \] (22)

C. MODEL SOLUTION ANALYSIS

The optimization model in this paper is selected to use Lingo software to solve. The objective function relationship and constraints in the optimization model are written into Lingo software programming language. The location of the subway station’s key guidance signs is 2, 3, 4, 6, 8, 10, and the maximum amount of induced information in the area is 33980.67.

According to the calculation results and combined with the passenger flow line, it can be observed that the guide mark setting points 2, 3, 4, 6, 8, and 10 are all locations where the passenger flow is concentrated and the flow lines cross, of which the settings are at 2, 3 4:00 and 4:00 are because these three points are close to the target location and the passenger flow is large. Choose 2, 6, 8, and 10 because these points are at the intersection of the flow lines or corners, and passengers need to make a distinction in the direction of travel at this location.

According to the description, we can understand that the guidance signs for the interchange between Dongzhimen Metro Line 2 and other public transportation are not perfect. Therefore, the optimization model calculation results show that the alternative positions 3 and 4 of the signs are Metro No. 2 respectively. The important nodes at the intersection of the passenger express line during the transfer between the airport express rail and the bus can be provided with wall-attached, ground-based, and guidance-oriented signs, as shown in Figure 10 below, which is beneficial to passengers during the transfer process. The guide signs are set and optimized here, which can meet the needs of passengers’ correct and smooth transfer, and strengthen the role of subway management departments in channeling people.

In addition, based on the importance of passenger flow, enhancing the guidance of the guidance sign to important core passenger flow lines is another important goal of optimizing the guidance sign. The passenger flow during the transfer between Metro Line 2 and Metro Line 13 is greater than the transfer between other lines. Once the guide signs cannot guide passengers to transfer in time, this will increase the waiting time of passengers and cause congestion problems. According to the results of the optimization model, the marker set points 2, 6, 8, and 10 are the inevitable places for passengers to take Line 2 to Dongzhimen Station and get off at Line 13. It is also the intersection of dense passenger flow lines. In such a high-traffic environment, aerial suspension, attachment and guidance signs should be set, as shown in the Figure 11 and Figure 12 below. Optimizing the setting of the guide signs at these four points can avoid herdness and blindness in the transfer process and enhance passenger mobility.

The 6 positions of the guidance signs obtained are in line with the actual situation of Dongzhimen Station, which not only satisfies passengers to find guidance signs within a certain range, but also reflects the continuity of the signs, and a certain number of signs can accurately provide guidance information for passengers. To maximize the amount of guidance information. The model has been effectively verified by real cases, avoiding excessive and complicated signs that make passengers’ thinking confused and affect travel. Reasonably and scientifically setting the guidance signs in this complicated situation is the necessary guarantee to guide passengers to travel smoothly. Therefore, the optimization
model also has real validity, which greatly plays the important role of the guide sign, improving passenger transfer efficiency and subway operation efficiency.

VI. CONCLUSION
With the acceleration of the modernization process, in order to ease the traffic pressure in cities, the role of rail transit has become more and more powerful. As the most critical part of public transportation, the guidance sign cannot be ignored. Guidance signs provide passengers with clear and unambiguous guidance and direction information, help passengers complete passenger, transfer and other behaviors, and ensure safe and efficient travel efficiency. The imperfection of the current guidance signs has caused many guidance problems for passengers in their daily use, causing passengers to face many difficulties, which directly affects the operational efficiency of rail transit stations. This study provides an optimized method for the layout of the guidance sign system in urban rail transit stations, making the operation and management of the subway sign guidance system more scientific and accurate.

Of course, this paper still has some shortcomings needs to be improved in the future research: the values in the optimization model have a certain randomness, and the value of the data should also be studied in depth; the mathematical model is optimized according to the internal passenger flow of the subway, and is nearly suitable for the signs in the internal subway station, which is not suitable for external subway spaces. Therefore, the subway optimization model should be thoroughly researched and improved, so as to obtain more practical guidance sign placement.

REFERENCES
[1] O. Neurath, *International Picture Language: The First Rules of Isotype*, K. Paul Ed. Whitecaklington, U.K.: Trench, Trubner. 1936.
[2] A. Uebele, “Signage systems & information graphics: A professional sourcebook,” *Step Inside Des.*, vol. 8, no. 6, p. 109, 2007.
[3] P. Arthur and R. Passini, *Wayfinding: People, Signs and Architecture*. New York, NY, USA: McGraw-Hill, 1992.
[4] R. Passini, “Wayfinding design: Logic, application and some thoughts on universality,” *Des. Stud.*, vol. 17, no. 3, pp. 319–331, Jul. 1996.
[5] M. Tang, D. Gong, S. Liu, and X. Lu, “Finding key factors affecting the locations of electric vehicle charging stations: A simulation and ANOVA approach,” *Int. J. Simul. Model.*, vol. 16, no. 3, pp. 541–554, Sep. 2017.
[6] R. Church and C. R. Velle, “The maximal covering location problem,” *Papers Regional Sci.*, vol. 32, no. 1, pp. 101–118, 1974.
[7] N. Thompson, “Way finding and airport terminal design,” *J. Navigat.*, vol. 54, no. 2, pp. 177–184, 2009.
[8] R. Duszek, “The warsaw subway signage system,” *Inf. Des. J.*, vol. 12, no. 3, pp. 181–186, Jan. 2004.
[9] I. Wakide, “Standardized guide signs at Yokohama station,” *Jpn. Railway Eng.*, vol. 154, no. 1, pp. 16–18, 2005.
[10] Q. Zeng, “Performance evaluation of Signage system in subway stations,” *Civil Eng.*, vol. 12, no. 3, pp. 34–38, 2011.
[11] G.-Y. Jeon and W.-H. Hong, “An experimental study on how phosphorescent guidance equipment influences on evacuation in impaired visibility,” *J. Loss Prevention Process Industries*, vol. 22, no. 6, pp. 934–942, Nov. 2009.
[12] E. Vilar, F. Rebelo, P. Noriega, E. Duarte, and C. B. Mayhorn, “Effects of competing environmental variables and signage on route-choices in simulated everyday and emergency wayfinding situations,” *Ergonomics*, vol. 57, no. 4, pp. 511–524, Apr. 2014.
[13] K. Nassar, “Sign visibility for pedestrians assessed with agent-based simulation,” *Transp. Res. Rec.*, J. Transp. Res. Board, vol. 2264, no. 1, pp. 18–26, Jan. 2011.
[14] W. L. Wang, S. M. Lo, S. B. Liu, and H. Kuang, “Microscopic modeling of pedestrian movement behavior: Interacting with visual attractors in the environment,” *Transp. Res. C, Emerg. Technol.*, vol. 44, pp. 21–33, Jul. 2014.
[15] A. Motamedi, Z. Wang, N. Yabuki, T. Fukuda, and T. Michikawa, “Signage visibility analysis and optimization system using BIM-enabled virtual reality (VR) environments,” *Adv. Eng. Informat.*, vol. 32, pp. 248–262, Apr. 2017.
[16] Q. Tu, Z. Zhang, and Q. Zhang, “Directional signage location optimization of subway station based on big data,” *IEEE Access*, vol. 8, pp. 12354–12363, 2020.
[17] H. Yin, J. Wu, Z. Liu, X. Yang, Y. Qu, and H. Sun, “Optimizing the release of passenger flow guidance information in urban rail transit network via agent-based simulation,” *Appl. Math. Model.*, vol. 72, pp. 337–355, Aug. 2019.
[18] L. Liu, *Languages in the Track: An Analysis of the Guidance System of Beijing Metro*. Beijing, China: Beijing Jiaotong Univ., 2009.
[19] T. Su, *Preliminary Study on Design of Space Guidance System in Beijing Metro Station*. Beijing, China: Beijing Jiaotong Univ., 2010.
[20] United States Federal Department of Transportation, *National Public Sign Design Principles and Graphic Collection*. Beijing, China: Beijing Press. 2003.
[21] H. Kang, “Design concept and application of metro signs in developed countries,” *World Rail Transit*, vol. 2, no. 4, pp. 53–55, 2007.
[22] R. Wang and T. Wei, “Enlightenment of the humanized construction of Japanese subway on Xi’an subway,” *Beauty Times (Urban Ed.)*, vol. 5, no. 2, pp. 50–51, 2016.
[23] D. Gong, M. Tang, S. Liu, and Q. Li, “A reconsidering production coordination: A principal-agent theory-based analysis,” *Adv. Prod. Eng. Manage.*, vol. 12, no. 1, pp. 51–61, Mar. 2017.
[24] D. Gong, M. Tang, S. Liu, G. Xue, and L. Wang, “Achieving sustainable transport through resource scheduling: A case study for electric vehicle charging stations,” *Adv. Prod. Eng. Manage.*, vol. 14, no. 1, pp. 65–79, Mar. 2019.
[25] P. X. Zhao, W. H. Luo, and X. Han, “Time-dependent and bi-objective vehicle routing problem with time windows,” *Adv. Prod. Eng. Manage.*, vol. 14, no. 2, pp. 201–212, Jun. 2019.
[26] D. Gong, S. Liu, J. Liu, and L. Ren, “Who benefits from online financing? A sharing economy E-tailing platform perspective,” *Int. J. Prod. Econ.*, vol. 222, Apr. 2020, Art. no. 107490.
[27] P. X. Zhao, Q. Gao, X. Han, and W. H. Luo, “Bi-objective collaborative scheduling optimization of airport ferry vehicle and tractor,” *Int. J. Simul. Model.*, vol. 18, no. 2, pp. 355–365, Jun. 2019.
[28] A. Burinskiene, “Pharma supply chain: Efficiency modelling approach,” *J. Syst. Manage. Sci.*, vol. 8, no. 2, pp. 65–73, 2018.
ZHUCUI JING received the Ph.D. degree in computational mathematics from the Chinese Academy of Sciences, China. She is currently an Associate Professor with the School of Economics and Management, Beijing Jiaotong University. Her current research interests involve in innovation, leadership, and health industry.

WEI BAI is currently pursuing the degree with the School of Economics and Management, Beijing Jiaotong University. His current research interests involve in innovation, leadership, and numerical methods.

* * *