A GEOSPATIAL ANALYSIS ON THE ACCESSIBILITY CHARACTERISTICS OF TRANSIT TRANSFERS

by

Kan Tao

A thesis submitted in partial fulfillment of

the requirements for the degree of

Master of Science

(Civil and Environmental Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2013

Table of Contents

List of Figures

List of Tables

Abstract

This Master's thesis presents an empirical analysis for the accessibility characteristics of the transit transfer based on the geospatial and topological features of the transit system.

Transit transfer activities have been widely investigated around the world in recent decades. Considering transfers actually duplicate the number of passengers within the complete travelling pattern, we are interested in finding what kind of characteristics will have a significant impact on the accessibility of transit transfers over variously located bus stops. In this way, based on the transit faring records and the layout of transit networking infrastructures, an empirical study has been implemented in a macroscopic scope. The findings reveal the inner relationship between the distributions of transit transfer and the geospatial and topological features of the transit system, which indicates some potential improvements of the transit level of service in terms of the transfer activities.

The major contributions of this empirical study are 1) proposed and defined a list of influencing accessibility characteristics towards the transit transfer in both geospatial and topological measures; 2) summarized and described how the proposed characteristics will influence the actual transit transfer accessibility; 3) implemented a linear regression based analysis framework towards the proposed characteristics while using a stepwise AIC to eliminate the irrelevant ones; 4) proposed an approximation algorithm for the detection of transfer activities while no swiping card record is available when the passengers alight the buses.

1 Introduction

1.1 Research Background

Transfers are the transit activities in which riders switch from one transit mode to another, or switch from one route to another. Considering the current level of service (LOS) of the transit system, transfers are almost inevitable within the majority of the urban areas. In most large transit systems in North America, at least 10% of riders make one or more transfers to reach their final destination. For example, in Boston, Massachusetts, 24% of subway trips involve at least one transfer, whereas in Chicago, Illinois, more than 50% of Chicago Transit Authority (CTA) passengers transfer during their typical trips [1].

In spite of the high percentage in ridership, transfers also play a significant role in daily transit operations in relation to the cost-effectiveness and customer satisfaction [2]. On one side, they connect different transit modes and routes, which improve both of the accessibility and usability of the transit systems significantly with no extra cost. Considering the combinations of transit modes and routes using transfers as a viable path for the riders, the number of available choices will grow exponentially if the maximum number of transfers allowed is increased. On the other side, riders usually have a negative perception of transfers because of the inconvenience caused by the transfer waiting time and walking time [2]. This makes transfer an undesirable connection activity which always results in a disutility for the corresponding transit mode.

Furthermore, transfer activities also duplicate part of the ridership, which contribute to a higher work load for the transit system. Different from the stable direct transit demand which can be effectively estimated using OD survey, transfer activities are hard to formulate and are less resistant to the modification of the transit infrastructures. The clustered transfer activities can swiftly shift from one bus stop to another if some of bus routes are changed or a new transfer point is developed.

In that sense, understanding the distributions of the transfer activities can have significant implications for the transit department, helping to identify which accessibility characteristics are most significant to the transfers. Then, transit planners can construct more accurate models to estimate the transfer ridership for each bus stop while they are designing the transit network. The transit operators will also be able to precisely evaluate the LOS of the transit network when they are proposing different bus scheduling plans.

Even though transit transfer has been widely discussed within myriads of researches which are focused on 1) transfer optimization within urban transit network planning and design, 2) transit closest path choice considering viable routes with multiple transfers 3) transit assignment considering transfer activities within multi-modal network, few researches have been focused on the geospatial distribution of the transfer activities.

Transit accessibility is a broadly recognized concept in urban transportation planning. In general terms, accessibility is a measurement of the spatial distribution of activities about a point, adjusted for the ability and the desire of people or firms to overcome spatial separation [3]. In this Master's research, we proposed and defined several accessibility characteristics to formulate the distribution of the transfers in both of the geospatial measure and topological measure within a macroscopic scope. The significance of these proposed accessibility characteristics are further evaluated through an empirical study based on the actual transit operation condition in the City of Madison.

1.2 Problem Statement

The problem researched in this empirical study is to evaluate the impacts from several important geospatial and topological measures on the distributions of transit transfer. As the transfer activities are clustered at the bus stop, the characteristics are prepared by taking bus stops as the basic evaluation units and enumerated in the following discussion.

Geospatial Measure:

- Transfer point or not.

Transfer points are the transport harbors specially planned and designed for transfer activities. Considering both of the transit infrastructural layouts and operational scheduling, timed transfers are widely implemented, which make transfers much easier around these areas. Therefore, transfer point bus stops are assumed to possess a higher transfer accessibility value comparing the ordinary ones.

- Distance from the current bus stop to the nearest transfer point.

As transfer points can be considered as major transfer attraction regions, the bus stops which are close to these areas may also be affected. Therefore, the distance from the current bus stop to the nearest transfer point is of great interests to be evaluated as another geospatial measure.

- Land use category of the adjacent neighborhood.

As a special geospatial measure, the land use category always describes the current functioning and development of the neighborhood. In this way, the land use can be regarded as a social economic descriptor to the environment of the community. City planners have continually emphasized the far reaching interactions between the transit accessibility and the development of land [3]. Therefore, the land use variable is also deemed to be relevant to the distribution of transfers.

Topological Measure:

- Number of bus routes passing through the current bus stop.

In graph theory, the degree of a vertex is the number of edges incident to the vertex [4], which can be regarded as a measure of the connectivity of the local network. Similarly, by formulating the transit network as a directed graph, the number of bus routes passing through the current bus stop can be seen as the degree of the bus stop, which is also a measure to the connectivity of the local transit network. Considering transfer riders may prefer more route choices at the bus stop with high connectivity, this variable is incorporated as one of the topological measures.

- Number of bus stops directly reachable from the current bus stop.

On some occasions, the number of bus routes passing through the current bus stop is not a very accurate measure to the connectivity. For some of bus stops which are located at the very end of the route, even if the ‘degree’ of the bus stop is comparatively high, the actual connectivity is restricted. Therefore, the number of bus stops directly reachable from the current bus stop can be regarded as an adjustment to the previous characteristic.

- Number of bus stops reachable from the current bus stop while one transfer is required.

Multiple transfers are sometimes required within complicated urban transit network. Therefore, as a step further from the variable denoted above, the number of bus stops reachable from the current bus stop while one transfer is required is proposed as a higher order term to describe the transit connectivity.

Another problem evolved in this Master's thesis is about the data processing. As no alighting transit record is available within the current faring system of Madison Metro, the transfer activity identification gives rise to another major problem. However, based on some spatial and temporal features, an approximation algorithm is proposed to tackle this problem.

Therefore, the research work in this Master's thesis will focus on the following problems. First, design an approximation algorithm is to identify the transfer activities from the faring records. Then the number of transfers and accessibility characteristic are prepared for each bus stop. In the end, by deploying a multiple linear regression model, the significances of the variables are evaluated.

1.3 Thesis Organization

The structure of the Master thesis can be generally divided into the following eight parts: 1) brief introduction of the background and the problem statement of the Master's research; 2) overview of the existing work which is relevant to the study of transit transfers; 3) introduction of the major methodologies implemented during this research; 4) description of the variable specification; 5) description of the data source, data processing and data characteristic; 6) detailed analysis of the model calibration and variable interpretation; 7) summarization of the conclusions and future work; 8) reference and appendix.

2 Literature Review

In this Chapter, some important researches are introduced and summarized to provide a general introduction of the existing studies which are relevant to transit transfer. Three major researches have been typically addressed within the following areas: 1) transfer optimization within urban transit network planning and design, 2) transit closest path choice considering viable routes with multiple transfers 3) transit assignment considering transfer activities within multi-modal network

2.1 Transit Network Design

Transit transfers generate additional waiting time and even walking time to the passengers who have to shift between different transit routes. This will in turn decrease the efficiency and deteriorate the level of service of the whole transit system. Therefore, as part of the objective function, transit planners and operators attempt to minimize this overall inconvenience when they are dealing with these urban transit network design problem (UTNDP) [5].

The UTNDP could be subdivided into two major components, namely the urban transit routing problem (UTRP) and urban transit scheduling problem (UTSP) [6]. As for UTRP, within large scale transit systems, systematic optimization solvers or searching algorithms are widely used while setting transit transfers as penalties in various forms. [7]. Baaj and Mahmassani discussed the sources of complexity of the TNDP and existing solution approaches to this problem and their limitations. They also presented the framework of an Artificial Intelligence (AI)/ Operations Research (OR) hybrid solution approach that combined AI search concepts with familiar constructs from OR vehicle routing heuristics and transit systems analysis methods [8]. A route generation design algorithm (RGA) has been implemented by using bus routes and frequencies as design variables to generate sets of routes corresponding to different trade-offs [9]. Zhao and Gan provided an effective computational tool for the optimization of a large scale transit route network while attempting to find the optimal transfer, route directness, and ridership coverage [10]. Influenced by the rapid growth in computing power, metaheuristic techniques have become more popular for solving combinatorial problems. Methods such as genetic algorithms (GA), tabu search (TS) and simulated annealing (SA) have all played important roles in recent research on the UTNDP [5]. Pattnaik et al. [11] Chakroborty [6] and Chien et al. [12] highlighted the effectiveness of genetic algorithm (GA) in solving TRP which required substantially low computation efforts through a more efficient representation of the discrete variables in the routing. For the other metaheuristic methods, Fan and Machemehl[13][14], Zhao and Zeng [15] utilized their solution methodology with tabu search and simulated annealing to solve specialized UTNDP problems by setting the average number of boarding as the objective function. In recent studies, Yun et al. proposed ant colony optimization (ACO) algorithm which extended transit network design included both of the demand density of the direct demands and transfers. It was indicated that the solution quality can be improved if the transfer coefficient is reasonably set [16].

In UTSP, there are basically two ways to reduce waiting time for transfers in a transit network. The first one, decreasing the headway on certain routes, involves an increase in the numbers of buses and drivers required to operate these routes. The second alternative, transfer coordination, includes two possibilities: timed transfer and transfer optimization [17].

Ceder et al. attempted to decrease the passengers' total waiting time while creating bus timetables to obtain a maximal synchronization at the transfer points. Considering the complexity of the problem, a heuristic algorithm has been developed to solve the mixed inter linear programming problem in polynomial time [18]. However, Bookbinder and Desilets [17] suggested that timed transfers were inappropriate for the large transit network with decentralized transfers as additional layovers could be introduced. Instead, they showed the importance of taking randomness of travel times into account when optimizing transfers. Zhao and Zeng proposed another metaheuristic methodology which was based on an iteratively defined local solution search space combined with an integrated simulated annealing, tabu, and greedy search algorithm (ISTG). The objective function aimed to be minimized was divided into three components: walking time (from a stop on the first route to one on the second route), transfer penalty time, and waiting time to board a vehicle on the second route [19]. Lee and Vuchic exhibited the iterative approach could flexibly handle the dynamic characteristics of the relationship between variable transit trip demand the optimal transit network design while regarding number of transfers as part of the penalty.

2.2 Transit Path Choice and Assignment

Transfers and waiting times are significant factors in public transport assignment. Some passengers may choose routes in order to minimize the number of transfers, while other minimizes travel times [20]. Therefore, a great number of researches have been focused on investigating algorithms finding shortest viable path considering transit transfers. Lozano and Storchi formulated a way to find a set of the shortest viable paths in a multi-modal environment while setting the maximum number of transfers as a constraint [21]. The algorithm was a modified version of Chrono-SPT, a chronological algorithm proposed by Pallotino and Scutella [22]. Following a divide and conquer strategy, Abdelghany devised a multi-labeling approach to generate the route choice set considering multi-modal transferring, in which non-dominated sub-paths were combined to form the paths from origins to destinations [23] [24]. Based on an efficient time dependent least time path algorithm [25], Ziliaskopoulos and Wardell [26], Chang and Ziliaskopoulos [27] proposed several time dependent intermodal least time path algorithm considering dynamic arc travel times and mode transferring delays.

Based on these path choice algorithms, various transit assignment strategies and frameworks have been proposed and evaluated. Nielsen presented a framework for public traffic assignment which builds on the probit-based model, in which transfers were modeled by setting coefficients on waiting time and associated error components to represent passengers' preferences against transfers [20]. Lo et al. developed a state-augmented multi-modal (SAM) network together with nested logit (NL), to explicitly consider the number and kinds of transfers and non-linear fare and utility structures within the multi-modal network [28][29]. The network structure was developed to limit the modal transfers to a set of probable ones by embedding the possible transfer state transitions. By assuming the link utility is additive, the non-linear fare structure is captured by summing over all the link utility over each route. In the work of Nuzzulo et al., a schedule-based path choice model was proposed for high-frequency transit networks, which incorporated a random utility model to formulate the transfer time and waiting time [30]. In the recent work developed by Zhou et al., a dynamic micro-assignment modeling approach was proposed for the multimodal urban corridor. It implemented the efficient time-dependent least cost path algorithm to generate an intermodal route choice set which recognized time-dependent mode transfer costs and feasible mode transfer sequences [31].

2.3 Transit Accessibility Measurement

Transit accessibility is a broadly recognized concept in urban transportation planning. In general terms, accessibility is a measurement of the spatial distribution of activities about a point, adjusted for the ability and the desire of people or firms to overcome spatial separation [3].

Some well-known definitions included 'the potential of opportunities for interaction' [3], 'the ease with which any land-use activity can be reached from a location using a particular transport system' [32]. From various aspects of location accessibility, individual accessibility, economic benefits of accessibility, several researchers have worked on review articles focusing on accessibility measure. In recent studies, Geurs and Wee reviewed and summarized the usability of accessibility measures in evaluations of both land-use and transport changes, and related social and economic impacts [33].

3 Methodology

3.1 Modeling Methodology

3.1.1 Multiple Linear Regression Model

As this research is focused on investigating the relationship between the transfer accessibility index and several corresponding topological and geographical features, we adopt a multiple linear regression model to estimate the influence from each variable, by assuming a linear relationship between the dependent variable y and explanatory variables X. The multiple linear regression model takes the form as follows:

$$y_i = X_i^T \beta + \varepsilon_i \tag{3.1}$$

Where:

y_i : Number of transfer at bus stop i,

X_i: Topological features and geo-spatial features of bus stop i,

ε_i : Error term.

3.1.2 Model Reduction Based on Information Entropy

At this stage, given the initial analysis results, some evaluation criteria should be adopted to prevent the model from over-fitting by including too many variables which are not significantly relevant to the transit transfer accessibility we've proposed. We adopt the strategy, by starting the model calibration from the full model, the variable will be added or eliminated until all the variables left are significantly relevant enough. This brings up the issue of model comparisons, which compares the full model with the restricted version by setting one of its coefficients to zero.

Within the myriad of model comparison criteria, some of them are widely recognized and applied, such as error measurement based evaluation (mean squared error, mean absolute percentage error, adjusted R square), hypothesis test based evaluation (t test, F test) and information theory based information criteria evaluation (Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC)). In this Master's research, we first choose the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which compare the fitness of the models when arbitrary variables are included and eliminated. The AIC and BIC are grounded in the concept of information entropy, in effect offering a relative measure of the information lost when a given model is used to describe reality. It can be said to describe the tradeoff between bias and variance in model construction, or loosely speaking between accuracy and complexity of the model [34].

Given a set of candidate models for the data, the preferred model is the one with the minimum AIC/BIC value. Hence AIC not only rewards goodness of fit, but also includes a penalty that is an increasing function of the number of estimated parameters. This penalty discourages over-fitting (increasing the number of free parameters in the model improves the goodness of the fit, regardless of the number of free parameters in the data-generating process).In spite of the fitness of the model measured by AIC and BIC, the two information criteria are asymptotically equivalent to the K fold cross-validation [35], which also guarantees the robustness of the model selected using this criterion.

The expressions of AIC and BIC are listed as follows [36]:

$$AIC_m = -2 \times \ln(L) + 2 \times k \quad (3.2)$$

$$BIC_m = -2 \times \ln(L) + \ln(n) \times k \quad (3.3)$$

Where:

L: Maximized value of the likelihood function for the estimation model

n: Sample size.

k: Number of parameters

With respect to the two expressions, as we've just mentioned, they both include a penalty on the number of voice measures in the selected subset, offering a compromise between sample error and model complexity. Therefore, in detailed implementation, Model 1 is selected over Model 2 if $AIC_1 < AIC_2$, otherwise, Model 2 will be selected. Similar selection rule could be also implemented for the BIC. However, different from AIC, BIC plays a larger penalty on the number of estimated parameters and is more parsimonious when $\ln(n) > 2$. Considering the sample size of our model is quite large, and we do not want to impose a too strict penalty on the number of variables, AIC is preferred in our on-going analysis.

3.2 Measurement of Effectiveness

3.2.1 Root Mean Square Error (RMSE)

The root mean square error (RMSE) or root mean square deviation (RMSD) is a frequently used measure of the differences between values predicted by a model or an estimator and the values actually observed. These individual differences are called residuals when the calculations are performed over the data sample that was used for estimation, and are called prediction errors when computed out-of-sample. The RMSE serves to aggregate the magnitudes of the errors in predictions for various times into a single measure of predictive power. It is a good measure of accuracy, but only to compare forecasting errors of different models for a particular variable and not between variables, as it is scale-dependent [37].

The RMSE of predicted values f_i for times i of a regression's dependent variable y is computed for n different predictions as the square root of the mean of the squares of the deviations:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - f_i)^2}{n}} \tag{3.4}$$

3.2.2 R Square and Adjusted R Square

Considering the RMSE is a scale dependent measurement of effectiveness, it doesn't provide a straightforward measurement of the goodness of fit of the model. In this section, we use R

Square (R^2) and Adjusted R Square ($\bar{R}^2$), which denote the statistical measure of how well the regression line approximates the real data points.

In a general form, R^2 compares the unexplained variance (variance of the model's errors) with the total variance (of the data) [38].

$$R^2 = 1 - \frac{SS_{err}}{SS_{tot}} \tag{3.5}$$

$$SS_{err} = \sum_i (y_i - f_i)^2 \tag{3.6}$$

$$SS_{tot} = \sum_i (y_i - \bar{y})^2 \tag{3.7}$$

Where:

SS_{err}: The sum of squares of residuals,

SS_{tot}: The total sum of squares,

f_i: The associated modeled value of each y_i.

However, in multiple linear regression model, R^2 is weakly increasing in the number of variables in the model. Therefore, R^2 alone cannot be used as a comprehensive comparison of models with very different number of variables.

Adjusted R square ($\bar{R}^2$) is an attempt to take account of the phenomenon of the R^2 automatically and spuriously increasing when extra explanatory variables are added to the model. It is a modification that adjusts for the number of explanatory terms in a model relative to the number of data points. [38]. The expression of $\bar{R}^2$ is listed as follows:

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n - 1}{n - p - 1} \tag{3.8}$$

Where:

n: The number of observations,

p: The number of variables.

4 Variable Specification

An accessibility measure should firstly be sensitive to changes in the transport system, i.e. the ease or disutility for an individual to cover the distance between an origin and a destination with a specific transport mode, including the amount of time, costs and effort. Secondly, an accessibility measure should be sensitive to changes in the land-use system, i.e. the amount, quality and spatial distribution of supplied opportunities, and the spatial distribution of the demand for those opportunities, and the confrontation between demand and supply [33]. Followed by this instruction, the proposed geospatial measures and topological measures are proposed and defined in this chapter. As the transfer activities are clustered at the bus stop, the characteristics are prepared by taking bus stops as the basic evaluation units in the following discussion.

4.1 Geospatial Measures

Geospatial features are important characteristics which measure the transit transfer accessibility from the aspects of transit infrastructure and land use environment. Among the transit facilities, some of the bus stops are designed for transfer activities, in which the layout and the infrastructure are specially arranged. Other than these transfer oriented areas, some neighborhoods seem to be quite transferring friendly which also generate a high transfer rate than the other ordinary stops. Therefore, we propose a study over the relationship between the transit transfer accessibility and these geospatial features within this section.

4.1.1 Transfer Point or Not

Considered as a transportation infrastructure measurement, the first factor we would like to propose is the fact if the bus stop is a transfer point or not. In actual operations and organizations, transfer points are serving as the transitional stations, where some of the buses will be switching from one route to another. As a transit scheduling strategy, timed transfers are widely implemented. Buses coming from different directions are largely scheduled to arrive simultaneously at the transfer points. This actually optimizes the transit transfers around the transfer points, because the intersection of the connecting time windows of the buses has been maximized.

Moreover, the directions of the routes stemming from transfer points seem to be designed to spread more widely across the whole network, which greatly extend the service areas of the transit system. In addition, most of the transfer points are equipped with shelters and couches, which bring more comfort to the transfer activities. Therefore, even if the numbers of routes, number of bus stops reachable directly or indirectly is within comparatively the same level of the other ordinary bus stops, the transit transfer accessibilities at these locations are significantly augmented.

Hence, a dummy variable indicating whether the current bus stop is a transfer point will be incorporated within the model.

4.1.2 Distance from the Bus Stop to the Nearest Transfer Point

As demonstrated above, the transfer points categorized as major transferring infrastructures act like attraction areas within the whole transit network. Then, how will these transferring oriented bus stops affect the other ordinary bus stops would be another question worthy of investigating. Considering the distance between ordinary bus stops and transfer points may bring up the major influence, we hereby adopt a location-based accessibility measure as our analyzing approach.

Within the range of location-based accessibility measurement, potential accessibility measures (also interpreted as gravity based measures) have been widely used in urban and geographical studies The potential accessibility measure estimates the accessibility of opportunities in zone i to all other zones (n) in which smaller and/or more distant opportunities provide diminishing influences [33]. Following the similar logic, we expect farther the bus stop stays away from the transfer point, less significant influence it will gain from the transfer attraction points. Therefore, a negative exponential term is introduced for the distance between the ordinary bus stop and the transfer points. Regarding the bus stops will be mainly influenced by only one of the transfer points, for simplicity, the distance variable is roughly parameterized as the distance between the bus stop and its nearest transfer point.

4.1.3 Land Use Category of Adjacent Neighborhood

City planners have continually emphasized the far reaching interactions between the transit accessibility and the development of land [3]. Land use measurement reflects the land use system, consisting of a) the amount, quality and spatial distribution opportunities supplied at each destination (jobs, shops, health, social and recreational facilities, etc.), and b) the demand for these opportunities at origin locations (e.g. where inhabitants live), c) the confrontation of supply and demand for opportunities, which may result in competition for activities with restricted capacity such as job and school vacancies and hospital beds [33].

As a special geospatial measure, the land use category always describes the current functioning and development of the neighborhood. In this way, the land use can be regarded as a social economic descriptor to the environment of the community. Therefore, the land use variable is also introduced to test if they are relevant to the distribution of transfers.

4.2 Topological Measures

The topological measures here refer to the attributes induced from the bus stops and routes while considering the transit network as a pure connected graph. In detailed implementation, bus stops are simulated as nodes and the bus routes are formulated as directed links. The whole transit network could then be expressed as a directed graph, in which every pair of connected bus stops are linked by one or multiple links. By discovering the relationship between the transfer distribution and the topological features listed below, the focus could be narrowed down to investigate how the combination of the nodes and links will dense or dilute the transfer density in its neighborhood. In the following illustration, the terms of bus stop and node, bus route and link are interchangeably used.

4.2.1 Number of Bus Routes Passing through the Current Bus Stop

In the equivalent directed graph model, each section of the bus routes is represented by the link going from or to the nodes. Gaining from experience, the more links are pointed to the node, a higher probability that the passengers will arrive. On the other hand, given a larger number of outgoing links, the passengers could seem to obtain a more efficient transfer routing among more alternatives. This is consistent with the fact that, in graph theory, the degree of a vertex is the number of edges incident to the vertex [4], which can be regarded as a measurement of the connectivity of the local network.

Therefore, the number of ingoing and outgoing links of each node should be considered as one of the important characteristics measuring the network connectivity, which directly influences the number of transfer passengers it will attract. As the starting and ending bus stops are the only sources and sinks in the graph, the numbers of ingoing and outgoing links are equal to most of the bus stops. Therefore, the two features are succinctly modeled as one parameter by using the number of bus routes passing through the current bus stop.

4.2.2 Number of Bus Stops Directly Reachable from the Current Bus Stop (R_0)

Even if the passing through bus routes have been considered, the bus stop could still be misinterpreted if we ignore its positioning information in the network. Let's suppose there is a bus stop which locates at certain position that almost all the passing through bus routes are approaching to their terminuses, there will be little possibility that passengers would like to make a transfer at that point even if quite a number of bus routes are running over that stop. Therefore, we introduce another topological feature, number of bus stops reachable from current stop, to describe the transit connectivity as an adjustment.

Let's suppose the total number of bus stops is N, then an $N \times N$ connectivity matrix M_0 will be constructed, in which $M_0[i][j]$ equals to 1 if bus stop j is reachable from i, and 0 otherwise. The following part demonstrates the pseudo code of the connectivity matrix generating algorithm:

```
for each bus route r ∈ R{
        for (node i = r(1); i < r(size(r)); i++){
                for (node j = r(i + 1); j ≤ r(n); j++){
                        if (M_0[i][j] == 0)
                                M_0[i][j] = 1;
                }
        }
}
```

Then the total number of bus stops reachable from node i could be calculated as:

$$R_0[i] = \sum_{j=1}^{N} M_0[i][j]. \tag{4.1}$$

Note:

- $r(i)$ refers to the i_{th} bus stop on route r, in which the sequence of the bus stops are their partial order in the graph model.
- $size(r)$ refers to the number of bus stops on route r.

4.2.3 Number of Bus Stops Reachable while One Transfer is Required (R_1)

Multiple transfers are sometimes required within complicated urban transit network. Therefore, the number of bus stops reachable from current bus stop while one transfer (R_1) is required is also introduced in the analysis framework to better represent the connectivity of the transit network.

Let's rephrase the passenger's routing decision as an optimization problem. Under certain circumstances, if the passengers' objective is focused on minimizing the total distance travelled, the optimal value considering multiple transfers will definitely be a valid competitor to the one while only one transfer is counted. So, the number of bus stops reachable while one transfer is required can serve as another important factor which influences the transit transfer accessibility of the bus stop. With this observation, R_1 is hereby incorporated to characterize this indirect connectivity, which will in turn affect the number of transfer activities generated by the multiple-transfer routing strategy.

Stemming from the predefined transit connectivity matrix M_0 , the indirect connectivity matrix M_1 could be calculated based on the following criteria:

$$If: \quad \exists\, k, 1 \leq k \leq N, M_0[i][k] = 1 \; and \; M_0[k][j] = 1 \; and \; M_0[i][j] = 0,$$

$$M_1[i][j] = 1$$

$$Else : M_1[i][j] = 0$$

Observing that all the entries in M_0 are non-negative, the encoding protocol can be more gracefully expressed in a two stepped algorithm listed below:

Step 1: Generate an intermediate matrix $M_1' = M_0 \times M_0$;

Step 2: Go through all the entries in M_1', M_1 and M_0, mark $M_1[i][j] = 1$, if $M_1'[i][j] > 0$ and $M_0[i][j] = 0$.

Then the number of bus stops reachable from current bus stop while one transfer is required can be calculated as:

$$R_1[i] = \sum_{j=1}^{N} M_1[i][j]. \tag{4.2}$$

5 Data Processing and Data Characteristics

5.1 Data Sources

5.1.1 IC Card Data

In this study, empirical transit data is provided by Metro Transit of Madison, Wisconsin. The IC card data used in this paper was collected by the Automatic Fare Collection (AFC) system. AFC is an electronic payment method which has been widely recognized within the intelligent public transport systems, such as buses, railways, metros, taxis, parking and so on so forth. When passengers are using the prepaid IC cards or tickets to pass through the terminals installed on the buses or railway stations, systems will automatically collect the information from the card while allowing commuters the access to transport facilities. Therefore, the AFC system operated by Madison Metro is able to provide individuals' boarding information including fare card ID, fare card type, together with the card swiping time, transit route number, vehicle number and direction. Cooperated with the Automatic Vehicle Location (AVL) system, which is functioning based on the Global Positioning System (GPS) technology, the faring system will also be able to provide the geographic information of the current bus location at the specific time point while the passengers are boarding the buses.

As a summary, the IC card dataset covers all the boarding information from April 5, 2010 to May 5, 2010. A total of over one million transit records within the 31 days (23 weekdays) are included within this research with no special events or holidays.

5.1.2 Transit Network and Infrastructure Data

In spite of the IC card data, the transit network and infrastructure data are another major data sources, which describe the framework of the transit system. As a whole, there are 52 bus routes operated over weekdays and weekends, which go across 1811 active bus stops. Among all of the active bus stops, there are four transfer points (North Transfer Point, South Transfer Point, West Transfer Point and East Transfer Pint) which are located on the four corners of the city of Madison [39].

The transit network data generally refers to the topological layout of the transit system. It addresses the relationship among the three main elements, bus routes, bus stops and directional trip pattern. As there may exist different travelling patterns for the ordinary and 'extra' buses for the same route in either direction, the Directional Trip Travelling Patterns (DTTP) are hereby introduced to mark these differences across the various patterns. In details, the DTTPs are modeled as links within the actual data tables. Over each link, the bus stops are lined up in the order of the travelled distance from the trip origin up to current bus stop. Therefore, by going through all the directional trip patterns for each bus routes, all the topological information of the transit network could be obtained.

The infrastructure data mainly refers to the categorical and geographical information about each bus stop. As there are two types of bus stops (transfer point bus stops and ordinary bus stops) within the transit system, the categorical information is required to distinguish between the two subtypes. Moreover, as all the transit records are not automatically matched to the bus stops where the passengers get on the buses, these missing connections needed to be screened out using a spatial join according to the geographical coordinates. Therefore, the geographical coordinates of the bus stops are also requested as a major data component of the geo-spatial layout.

The transit network data and infrastructure data are also provided by Madison Metro. In addition, some of the missing relations about the network layout are obtained from the Web Watch on the websites of Metro Transit Tracker [40].

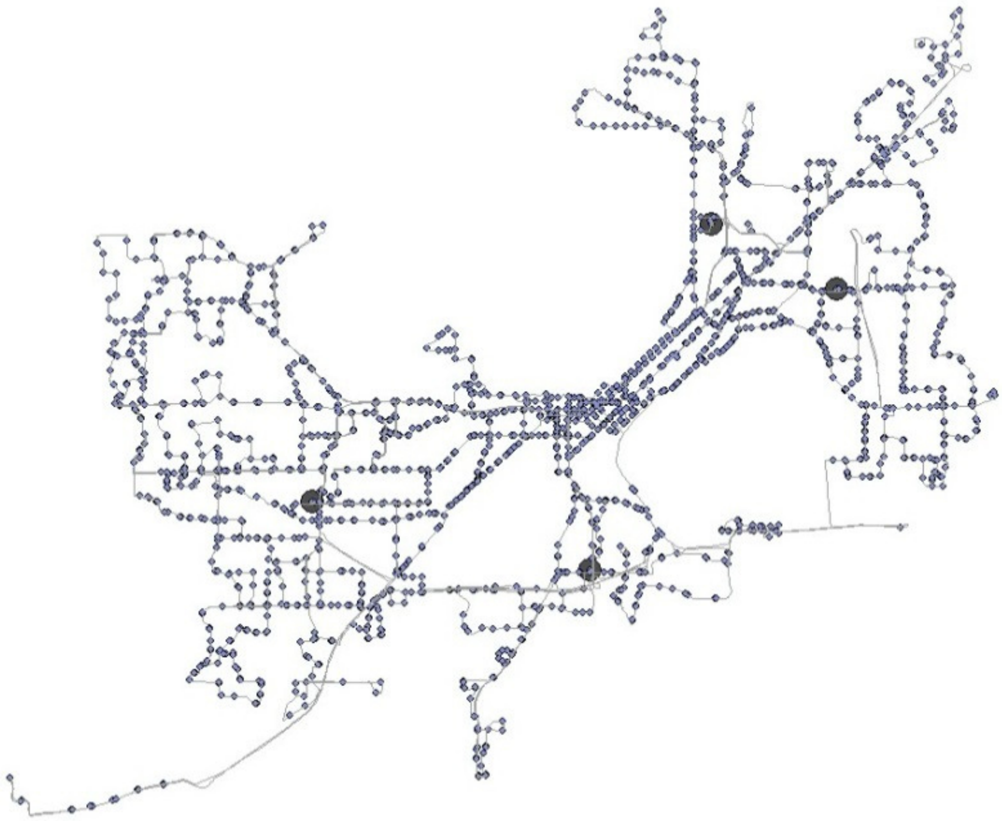


Figure 5-1 Locations of Bus Stops in the City of Madison

5.1.3 Land Use Data

As we've discussed in the previous chapter, the land use measure is another potential influencing factor which may impact the passengers while they are making decisions about transit transfer activities. Land use category is hereby introduced to act as a categorical descriptor of the adjacent neighborhood of each bus stop.

In this Master research, the land use geographical data we have used is the 2005 Land Use Map of Dane County, in which our analysis target, City of Madison, is the county seat. Dane County is also the second most populous county in Wisconsin, which has a population of 488,073 according to the 2010 census. The land use map covers approximately 1,238 square miles (3,206 square kilometers). The map generally divides all the land use into ten major land use categories (residential, industrial, transportation, communication/utilities, commercial retail sales, commercial retail services, institutional/government, outdoor recreation, agriculture/silviculture and other) and 83 minor categories.

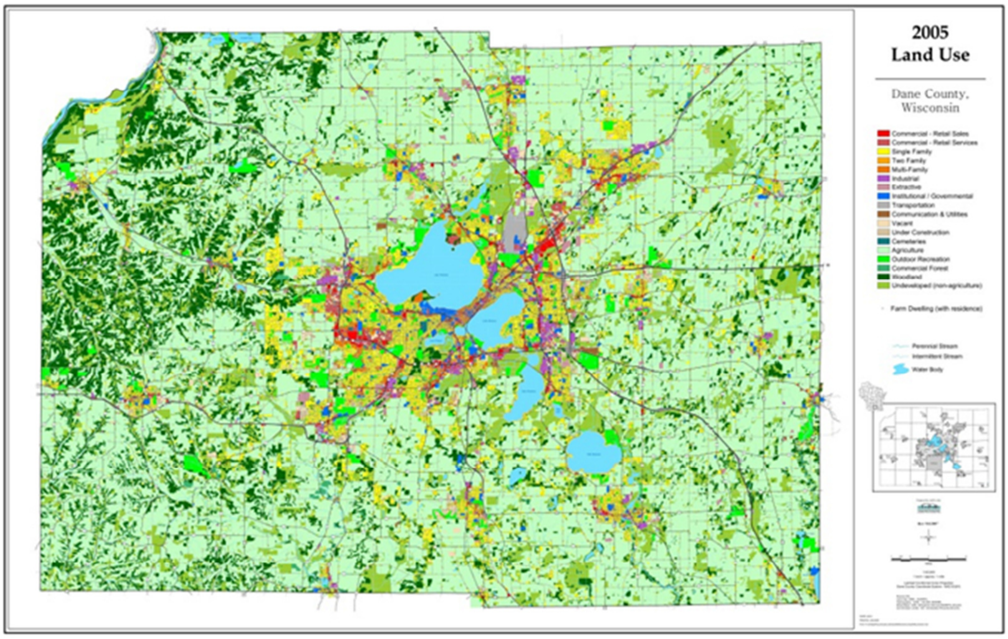


Figure 5-2 2005 Land Use Map of Dane County

5.2 Data Processing

Based on the three major data sources, the data processing can be generally divided into three components, transfer activity identification, transit topological feature generation and transit geospatial feature generation. As no alighting records are available from the current faring system of Madison Metro, the transfer activities could not be identified by using the temporal and spatial gaps between the two consecutive alighting and boarding records, which result in a critical challenge within transfer activity identification. In order to overcome this difficulty, we've designed an approximation algorithm, which uses some other passengers' boarding time as the approximation of the target passenger's alighting time.

To improve the accuracy of the approximated identification algorithm, we further narrow down the analysis scope within transit commuting trips. Comparing with the other transit users, the commuters' trip purposes, trip origins and destinations are always fixed. Therefore, learning from a more stable travelling pattern, we are able to prevent from mistaking two separated transit trips for one transfer activity with a high probability, even if the two consecutive activities are qualified for the temporal and spatial criteria.

Therefore, as a preprocessing step, we first impose a series of constraints on the raw data to filter out the commuting trips by setting up thresholds on the passengers' travelling frequencies and spatial ranges of the origins and destinations. Then, based on the identification algorithm we have introduced earlier, the transfer activity will be finally screened out.

As bus pass ID is the unique identifier in tracking each specific passenger's transit records, passengers who use bus passes are only considered in this research. Details of the transfer identification algorithms are explained in the following sections.

5.2.1 Starting Trip Identification

Considering the general traffic conditions of the City of Madison and the 8 hour working regulations, trips leaving for work or school are usually generated before 9:30AM, and trips leaving for home seldom occur prior to 3:00PM. Therefore, the morning and afternoon start-up trips are identified as the earliest record before 9:30 AM or after 3:00 PM. In addition, the boarding locations of these trips should be different from the following ones.

5.2.2 Commuting Trip Pair Identification

Considering the ordinary trip schedule for a specific commuter, the starting and ending trips will always be the commuting trips. Therefore, the commuting trips will generally appear in pairs. To consolidate this cyclic feature, for a qualified commuting trip pair, both of the starting trips in the morning and the ones in the afternoon should be observed from the passenger's daily trip pattern. In that context, if either of the two starting trips is not observed, the trip set will not be considered as a valid candidate for the commuting trip collection. Then the set will be removed from further analysis.

5.2.3 Origin and Destination Identification

According to the cyclic characteristic of the commuting trips discussed above, the origins and destinations will get switched for the trips in the morning and afternoon. Using the symmetries of the OD pair, the origins and destinations could be directly obtained by locating the two starting trip boarding positions within the same commuting trip pair.

5.2.4 Qualified Commuter Identification

As commuting trips are related with trip purposes of going to work or school, these everyday activities should be observed with comparatively high frequencies within the entire trip collections. Therefore, for each qualified commuter, riders' origins and destinations of the commuting trips should be frequently duplicated within the sample data. In this way, to consolidate the validation of this repetitive characteristics, a travelling frequency of 12 Commuting Trip Days (CTD) for the same origin and destination regions (approximately 3 CTD per week) is suggested as an threshold for this research, considering people can switch to the other transportation modes occasionally. Assuming the commuters may have alternative route choices across the commuting activities, the corresponding boarding bus stop may actually vary a little bit in its geographic locations even if the same OD pair still apply. Therefore, a $500\text{m} \times 500\text{m}$ squared region is predefined as a geographic boundary in identifying the origins and destinations.

5.2.5 Transfer Activity Identification

Within transfer activities, a time gap between the two consecutive transit trips is always unavoidable even if the activity is generated at a highly transit oriented district. The time gap can be generally divided into two components, walking time between the bus stops and waiting time for the upcoming buses. However, as a disutility, the time gap will decrease the attractiveness of a transfer when it becomes larger and larger. Therefore, only when the time gap is within a reasonable scale, the following trip record (also referred as potential transfer trip record) could be regarded as a valid transfer trip. However, in the actual data process, as boarding records are only available from the faring system, an approximation method is designed for estimating the same passenger's alighting time point.

In reality, passengers will try to optimize the transfer activity by minimizing the walking distance. Let's suppose passenger P_i gets off at the nearest bus stop S_m at t_m^i and makes a transfer on S_n at t_n^i . Then t_m^i could be approximated using some other passenger j's boarding time t_m^j at the same bus stop S_m on the same transit trip T operated by the same bus. However, as the passengers' boarding records are not uniformly distributed with respect to the bus stops along transit trip T, it is very likely that no boarding records could be found at S_m. To solve this problem, instead of using the nearest bus stop as the unique option, we add an additional spatial constraint D to find a collection of P_i's potential alighting bus stops. By using the qualified boarding record on the nearest bus stop available, we greatly improve the probability to find a valid approximation to t_m^i. Hence, by taking P_i's following boarding location as S_n, another amended algorithm is applied to first find all the qualified boarding records $\{t_k, S_k\}$ which are on the same bus trip T and within a distance of D to S_n. Then the nearest boarding record will be selected as the approximation to the alighting time t_m^i'. Considering the transit network density and operating condition in the City of Madison, a 2500 meter distance (5 min traveling distance at the speed of 20 mph) is proposed as the spatial constraint in searching for the potential alighting bus stops.

At this stage, the transfer trips can now be evaluated by measuring the time gap between the approximated alighting time t_m^i' and the following up boarding time t_n^i. As few studies has been focused on the analysis of transfer patterns in the Madison area, little historical data could be obtained as a benchmark of the transfer walking time and waiting time. Therefore, a 30 minutes' time interval is recommended as the maximal transfer gap while considering few activities could be finished within a 30 minute interval. Therefore, the temporal constraint of a qualified transfer trip should satisfy the following equation:

$$t_m^i{}' - t_n^i \in [-30,0] \tag{5.1}$$

However, as the alighting time is estimated based on the nearest boarding time within a distance of 2500 meter, it is very likely that the approximated alighting time will be posterior to the transfer boarding time. Therefore, considering the 2500 meters could be covered within 5 minutes, a five minute offset is set to the transfer identification time window. The following equation shows how the equation is changed.

$$t_m^i{}' - t_n^i \in [-25,5] \tag{5.2}$$

So, if there exists such a pair of starting trip record and following up record approved by the algorithm described above, the connection of the two trips should be regarded as a transfer activity.

As timing constraint is not too restrict for the trip back home, people tend to have a more flexible time schedule in the afternoon. In addition, considering the commuting trips will be intermingled with many other activities in the afternoon, the deteriorative traffic condition and transit LOS will probably result in a longer waiting time. According to the bus transit data, the average time dispersion between starting record and its following up record is about 52 minutes in the afternoon, which is about 1.3 times of the 41 minute time dispersion in the morning. To better capture the transfer activities in the afternoon, the transfer identification time window is enlarged by multiplying a scalar of 1.3. The following equation shows how the timing criteria are changed for the transfer in the afternoon.

$$t_m^i{}' - t_n^i \in [-32.5, 6.5] \tag{5.3}$$

The flow chart of this transfer activity identification algorithm is provided in Figure 5-3.

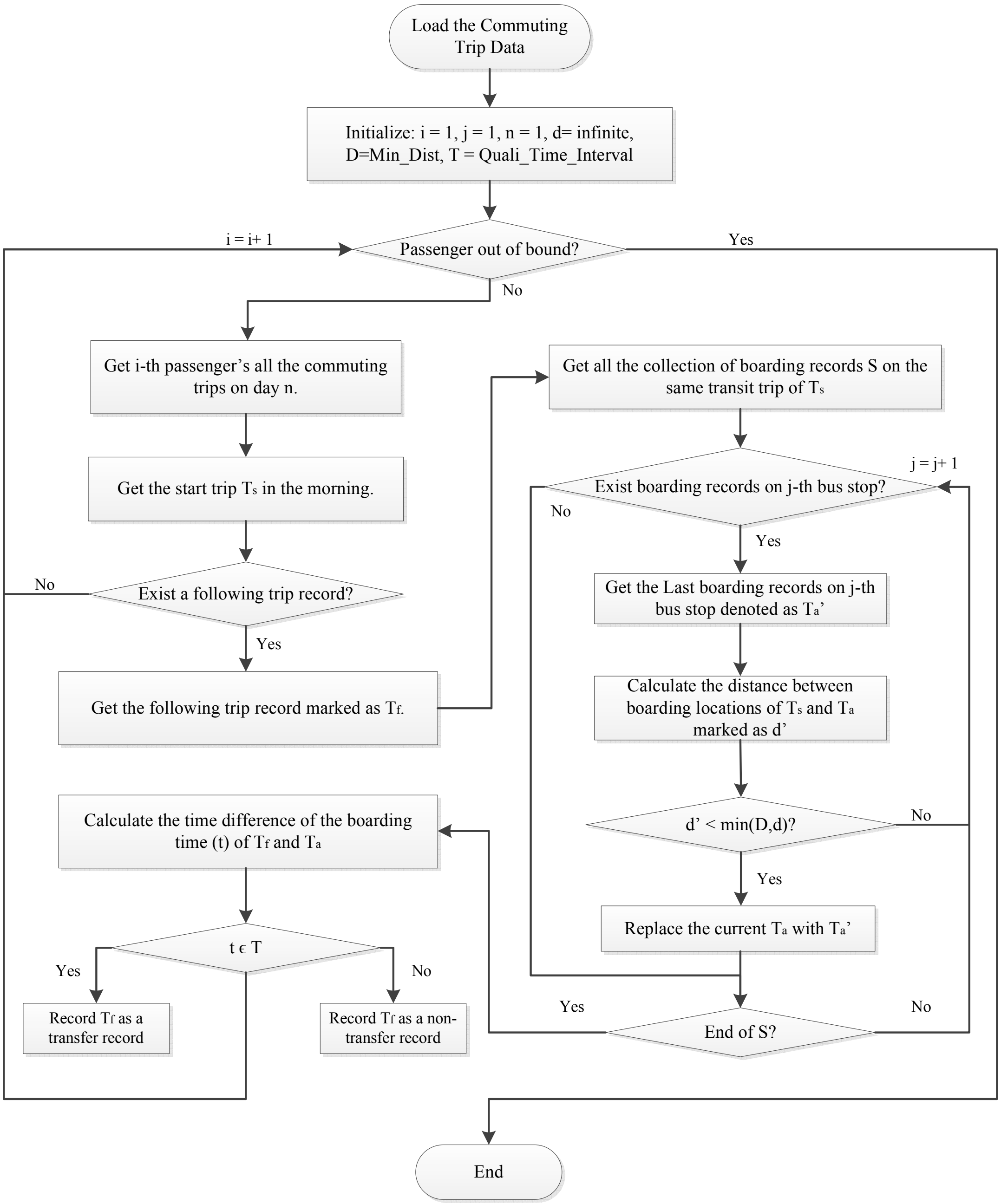


Figure 5-3 Flow Chart of Transfer Activity Identification Algorithm

5.3 Data Characteristics

5.3.1 Number of Transfers at the Bus Stop

Followed by the identification algorithm, the transfer activities are screened out from the dataset of the commuting trip records. As a result, 1707 and 1375 transfer records are filtered out from the 34692 commuting trips, comprising both of the morning and afternoon transit records. In the following paragraphs, some general data validations have been implemented by checking if the sample data are consistent with the real situations.

By looking at the type of the bus pass, some differences could be discovered from the categorized passengers' transit patterns in forms of commuting trip frequency and transfer frequency. The majority of the commuters within the sample data are from the pass holders of Summer Youth Pass, UW Employee Pass and UW ASM Pass. Considering the students and employees from UW-Madison comprise the majority of the transit users, the distribution of the commuters are consistent with the facts.

Table 5-1 Descriptive Statistical Analysis for the Commuting Data Set

Bus Pass Code	Pass Type	# of People	# of Trips (A/PM)	# of AM Transfers	AM Transfer Rate (%)	# of PM Transfers	PM Transfer Rate (%)
21	Summer Youth Pass	551	7987	1001	12.53	787	9.85
16	UW Employee Pass	395	5419	309	5.70	302	5.57
22	UW ASM Pass	153	2076	101	4.87	93	4.48
37	31 Day Rolling	76	986	206	20.89	108	10.95
23	MAD Employee Pass	27	367	16	4.36	26	7.08
18	MATC SUM Pass	15	201	49	24.38	26	12.94
45	Senior/Disable 31 Day	11	144	15	10.42	26	18.06
19	Commuter Pass	7	95	10	10.53	7	7.37
17	Meriter Employee Pass	3	41	0	0.00	0	0.00
24	Edgewood Pass	1	16	0	0.00	0	0.00
20	ST Mary Ann Pass	1	14	0	0.00	0	0.00
/	Total	1240	17346	1707	9.84	1375	7.93

In the temporal dimension, the commuting trips and the corresponding transfers are roughly categorized based on the trip starting time. The temporal distributions of the transfer activities are presented below, together with a calculation of the transfer rate. The temporal distributions are generally consistent with the fact that most of the commuting trips are generated from 7:00 a.m. to 8:00 a.m. in the morning and from 15:00 p.m. to 16:00 p.m. in the afternoon.

Table 5-2 Distribution of Transfer Records w.r.t. Starting Time

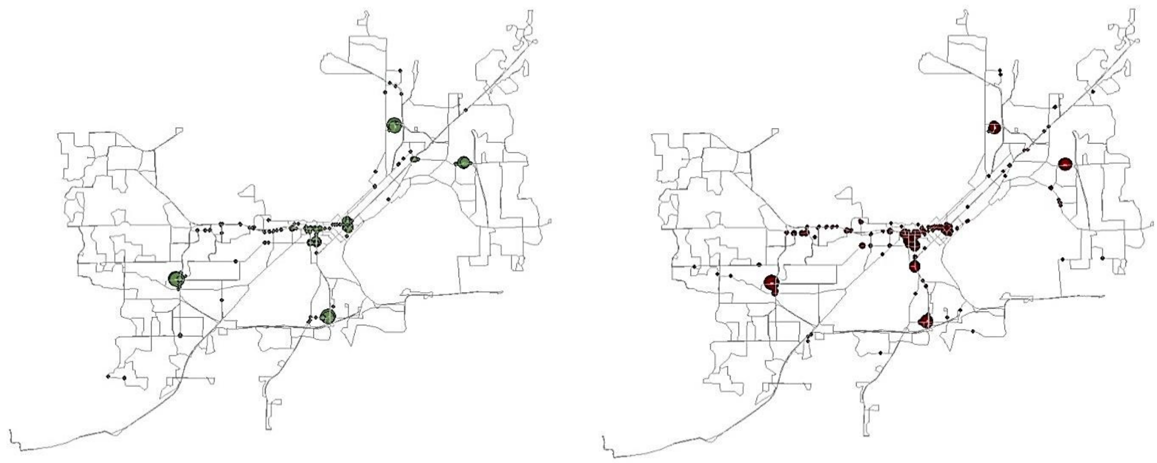
(a) Morning Transfer Record in Different Trip Starting Time

Start Time	# of Commuting Trips	# of Transfer	Transfer Rate (%)
5:00-6:00	404	99	24.50
6:00-7:00	3026	429	14.18
7:00-8:00	11130	959	8.62
8:00-9:00	2328	200	8.59
9:00-9:30	458	20	4.37
Total	17346	1707	9.84

(b) Afternoon Transfer Record in Different Trip Starting Time

Start Time	# of Commuting Trips	# of Transfer	Transfer Rate (%)
15:00-16:00	9466	680	7.18
16:00-17:00	4264	352	8.26
17:00-18:00	2481	246	9.92
18:00-19:00	547	59	10.79
19:00-20:00	195	6	3.08
20:00-21:00	172	18	10.47
21:00-22:00	100	13	13.00
22:00-23:00	81	1	1.23
23:00-24:00	40	0	0.00
Total	17346	1375	7.93

Based on the locations where the transfer activities are generated, transfer frequency maps are plotted using the coordinates of the corresponding bus stops. Considering the patterns of transit activities can be different across the time of day, two separated frequency maps are constituted for both of the morning and afternoon transfer activities. Directly seen from the figures, the transfer activities are largely generated around the transfer point areas. However, some of the transfer activities are also emerged around the campus area and capitol area which are the centers of the City of Madison. As a consequence, little differences could be discovered between the distributions of the transfer activities within the two frequency maps. Therefore, in the following modeling stage, a universal model structure, which takes the number of transfers within one day as a whole, is recommended instead of modeling the morning and afternoon transfer activities using two separated ones.



(a) Morning Transfer Frequency Map **(b) Afternoon Transfer Frequency Map**

Figure 5-4 Frequency Map for Morning and Afternoon Transfer Activities

5.3.2 Transfer Point or Not

As a dummy variable, transfer point or not will indicate if the current bus stop is located within the transfer points. In the City of Madison, there are four transfer points, North Transfer Point, South Transfer Point, West Transfer Point and East Transfer Point. For each one of the transfer point, there are two bus stops located on each direction. The detailed geographical information is listed in the table below.

Table 5-3 Geographical Information of the Transfer Point Bus Stops

GeoNode_ID	GeoNode_Name	Latitude	Longitude	Altitude
1594	SOUTH TFR PT @ BADGER [4100]	43.0390	-89.3947	356
1595	ARIV SOUTH TP @ BADGER [4101]	43.0386	-89.3949	356
1722	NORTH TFR PT @ HUXLEY [5100]	43.1127	-89.3586	85
1723	ARIV NORTH TP @ HUXLEY [5101]	43.1133	-89.3589	85
1842	WEST TFR PT @ TOKAY [6100]	43.0539	-89.4754	268
1843	ARIV WEST TP @ TOKAY [6101]	43.0537	-89.4748	268
2168	EAST TFR PT @ CORPORATE [7100]	43.0986	-89.3223	271
2169	ARIV EAST TP @ CORPORATE [7101]	43.0986	-89.3217	271

As a general overview, the layout of the transfer points and the distribution of transfer activities (both morning and afternoon) are provided in the figures below:

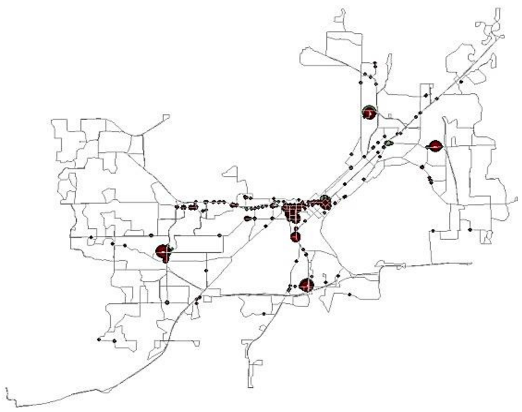


Figure 5-5 Transfer Frequency Map

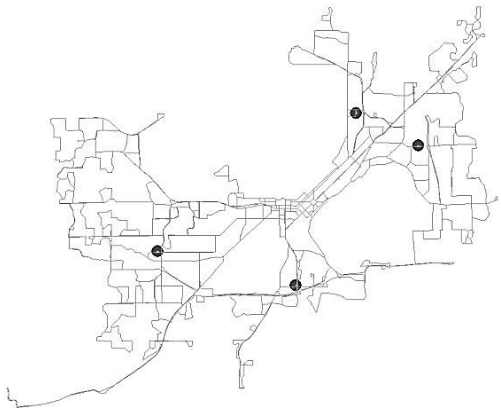


Figure 5-6 Transfer Point Layout

As the transfer points are planned and designed for transfer activities, the numbers of bus routes are assumed to be comparatively high. This has been confirmed through the observation over the two figures above. In the following sections, these topological features which describe the network characteristics for the transfer point bus stops will be addressed.

5.3.3 Distance from Bus Stop to the Nearest Transfer Point

Considering the roadways are planned along the roadways in the city blocks, the minimum distance from current bus stop to any transfer point is calculated using Manhattan distance, which is also named as city block distance.

The Manhattan distance, d_1, between two vectors $\boldsymbol{p}$, $\boldsymbol{q}$ in an *n*-dimensional real vector space with fixed Cartesian coordinate system, is the sum of the lengths of the projections of the line segment between the points onto the coordinate axes. More formally expressed as: [41],

$$d_1(\boldsymbol{p}, \boldsymbol{q}) = \|\boldsymbol{p} - \boldsymbol{q}\|_1 = \sum_{i=1}^{n} |p_i - q_i| \tag{5.4}$$

Where:

$$\boldsymbol{p} = (p_1, p_2, \dots, p_n) \tag{5.5}$$

$$\boldsymbol{q} = (q_1, q_2, \dots, q_n) \tag{5.6}$$

As $\boldsymbol{p}$ and $\boldsymbol{q}$ refer to the geo-coordinates of the bus stops, n takes the value of 2 while we ignoring the distance between the altitudes of the stops. As a result, some simple statistics about the distance from current bus stop to the nearest transfer point is listed as follows:

Table 5-4 Stats of the Distance from Current Bus Stop to the Nearest Transfer Point

Statistics	Distance (Meter)
Average Value	3959.39
Standard Deviation	2105.76
Min	0
Max	14027.71

To better visualize the distribution of bus stops with respect to the distance to the nearest transfer point, a bar chart of the general distribution and a cumulative distribution function (CDF) are provided in the plot below. The stops are first divided into groups based on the distance, and then the number of bus stops is calculated.

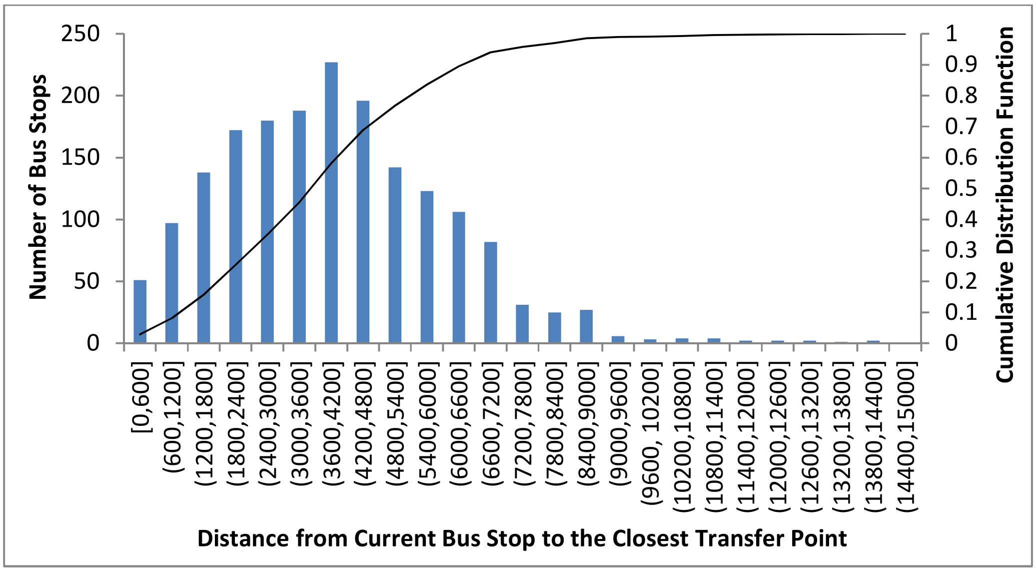


Figure 5-7 Number of Bus Stops w.r.t. Distance to the Nearest Transfer Point

The distribution is more like a Gumbel distribution, which is centered on the first half with a long tail. It could be observed that about 85% of the bus stops are centered on the area which is within the distance not more than 6000 meters away from the transfer points.

By linking with the transfer activities, the bus stops can be divided into groups based on the density of the transfer activities. However, not a very clear relationship can be observed from the table. However, it is still very interesting to see that the most popular transfer area is much closer to the transfer points other than the other districts.

Table 5-5 Average Distance to the Nearest Transfer Point w.r.t. Transfer Frequency

Transfer Frequency	Average Distance (Meter)
{0}	4035.972
(0,5]	3429.37
(5,10]	3975.15
(10,15]	4454.258
(15,20]	4095.802
(20,+∞)	2843.918

5.3.4 Land Use Category of Adjacent Neighborhood

As most of the bus stops are built at roadsides, the land use of transportation its located block is not meaningful to our research. In our study, we use the most frequent land use category of the neighborhood (100 meters as the diameter) as a replacement.

As a consequence, 39 land use categories have been found for the neighborhood of the bus stops within our research. The detailed information for the number of bus stops, number of transfer points, number of transfer records, number of transfers on non-transfer point bus stops and transfer ratio have been calculated for each land use category and listed in Table 5-6. The references of the major and minor land use codes are attached in Appendix A.

Table 5-6 Bus Stops and Transfer Records w.r.t. Major and Minor Land Use

(a) Land Use Code from 1 to 3

Major	Minor	# Bus Stops	# TP	# Transfers	# Transfers (not TP)	Transfer Ratio
1	111	885	1	552	350	0.40
1	113	93	0	29	29	0.31
1	115	258	0	301	301	1.17
1	129	10	0	24	24	2.40
1	142	2	0	0	0	0.00
2	36	1	0	0	0	0.00
2	38	18	0	1	1	0.06
2	39	1	0	0	0	0.00
3	41	4	0	0	0	0.00
3	45	151	2	572	6	0.04
3	46	16	2	501	25	1.56

(b) Land Use Code from 5 to 10

Major	Minor	# Bus Stops	# TP	# Transfers	# Transfers (not TP)	Transfer Ratio
5	9	1	0	0	0	0.00
5	53	2	0	2	2	1.00
5	54	2	0	0	0	0.00
5	57	4	0	1	1	0.25
5	58	31	0	156	156	5.03
5	59	37	0	33	33	0.89
5	531	4	0	2	2	0.50
5	551	12	0	0	0	0.00
5	552	2	0	3	3	1.50
5	553	12	0	12	12	1.00
5	559	2	0	0	0	0.00
6	61	14	0	2	2	0.14
6	62	1	0	0	0	0.00
6	63	74	0	140	140	1.89
6	65	12	0	25	25	2.08
6	67	4	0	0	0	0.00
6	68	1	0	0	0	0.00
7	71	44	0	137	137	3.11
7	72	4	0	2	2	0.50
7	73	8	0	2	2	0.25
7	74	1	0	1	1	1.00
7	75	5	0	31	31	6.20
7	76	2	0	0	0	0.00
7	79	1	0	0	0	0.00
8	83	17	0	1	1	0.06
10	97	22	1	150	0	0.00
10	98	48	2	824	39	0.81
10	99	3	0	0	0	0.00

Observed from the table above, for some of the land use type, transfer generation rate is comparatively high even if the number of bus stops within that category is not very high. To better visualize this distribution, the figure below vividly demonstrates how the two dimensioned statistics are changed with respect to the land use category.

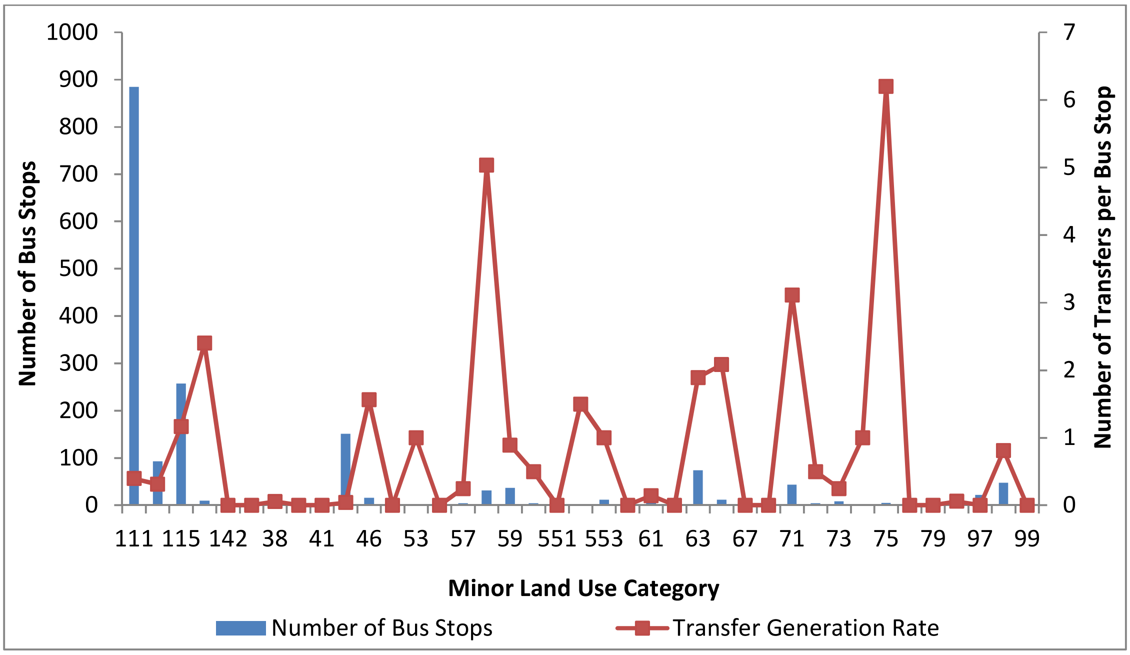


Figure 5-8 Bus Stops and Corresponding Number of Transfers w.r.t. Minor Land Use

By comparing the two trend lines, the levels of transfer generation rate are not quite the same across all the land use types. For some of the minor land use categories, such as eating and drinking, education and business service, the transfer rate are comparatively high comparing with the others.

5.3.5 Number of Bus Routes Passing through the Current Bus Stop

Based on the variable formulation, number of bus routes passing through the current bus stop is calculated for each bus stop. As a result, the average number and standard deviation of the number of bus routes are listed in the table below.

Table 5-7 General Stats about the Number of Bus Routes Passing through the Bus Stop

	Non-Transfer Point Bus Stop	Transfer Point Bus Stop
Min	1	4
Max	25	15
Average	2.16	10.25
Standard Deviation	2.53	3.88

We see that the average number of bus routes passing through the transfer point is much higher than for the non-transfer point bus stops. This is consistent with the fact that transit planners tend to include more bus routes within the transfer point areas. However, the maximum number of 25 bus routes is obtained in some non-transfer point area. This result is mainly caused by the fact that many bus routes overlap on some major corridors. Then, the maximum number of bus route will be reached at some intersection of these corridors.

To better visualize the distribution of the number of bus routes across all the bus stops, a bar chart of the general distribution of the number of bus routes and a cumulative distribution function (CDF) are provided in the plot below. As could be seen from the figure, the distribution of the number of bus routes at each bus stop is fairy heterogeneous. About half of the bus stops only have 1 bus route passing through. Less than 10% of the bus stops possess more than 5 bus routes.

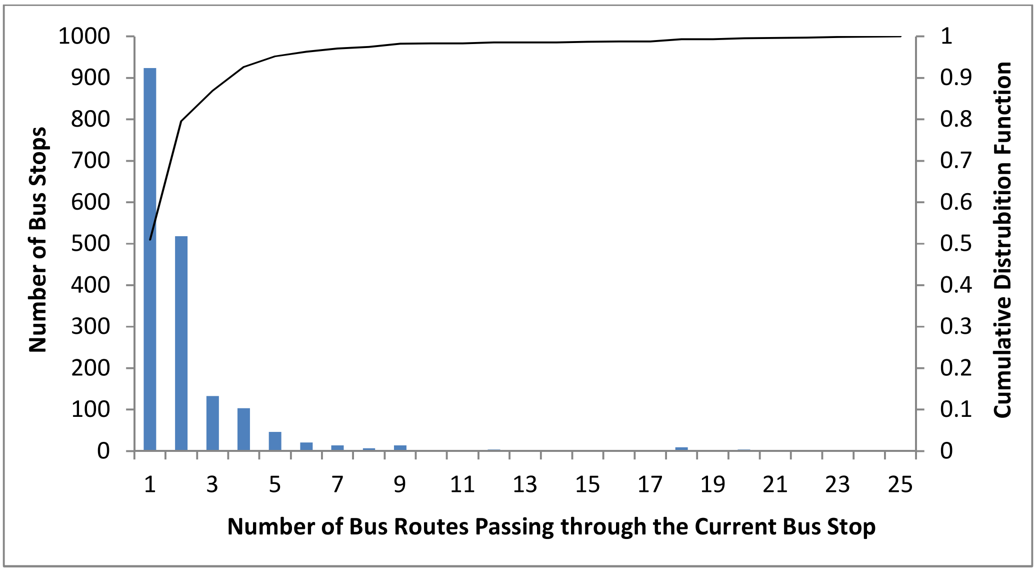


Figure 5-9 Number of Bus Stops w.r.t. Number of Bus Routes

Together with the number of transfers at each bus stop, the relationship between the number of transfers and this explanatory variable has been roughly tested by the plot provided below. In the detailed manipulations, bus stops are first clustered based on the number of routes at each bus stop. Then the average number of transit transfers is calculated within each cluster. As the bus stops in transfer points possess exceptionally high values of number of bus routes, they are simply excluded from the plot to make the graph fit in scale.

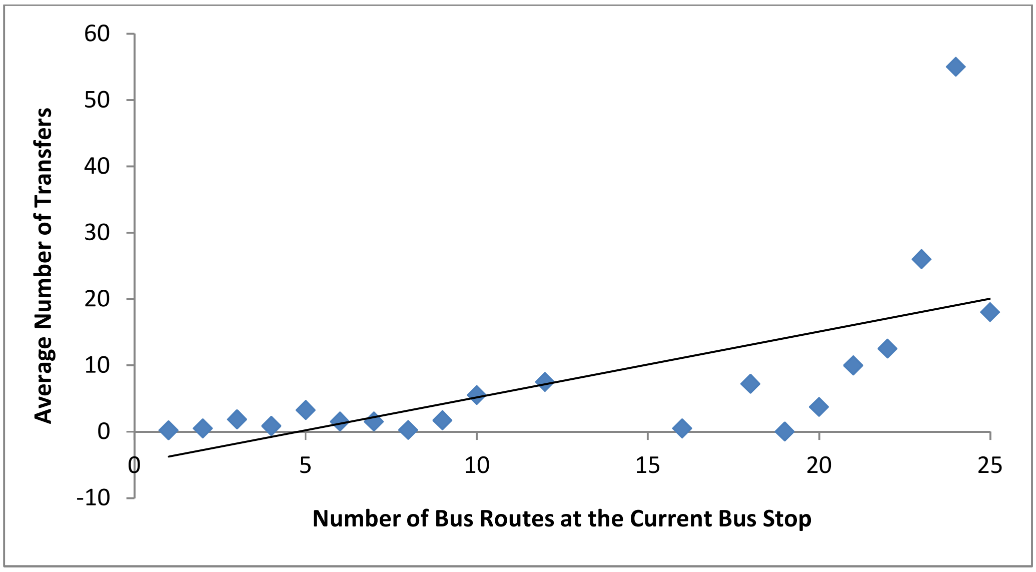


Figure 5-10 Average Number of Transfers vs. Number of Bus Routes

In general, a linear relationship could be observed between the dependent variable and this explanatory variable. According to the fitted line using a simple linear regression, two variables are positively related with each other in general. This indicates that the increase in the number of bus stops will be likely to contribute more transfer activities. This is consistent with the fact that passengers tend to make a transfer at a more transit oriented area.

5.3.6 Number of Bus Stops Directly Reachable from the Current Bus Stop (R_0)

Following the similar steps, an empirical analysis has been made upon the second explanatory variable. Some general statistical information including the minimum, maximum, average value and the standard deviation for both non-transfer point bus stop and transfer point bus stop has been listed below.

Table 5-8 General Stats of the Number of Bus Stops Directly Reachable

	Non-Transfer Point Bus Stop	Transfer Point Bus Stop
Min	0	66
Max	565	500
Average	51.09	237.38
Standard Deviation	55.02	160.82

Similar with the previous variable, the maximum value is still obtained within the non-transfer point bus stops. In average, the level of number of bus stops directly reachable within transfer point bus stops is comparatively much higher than that of the non-transfer bus stops.

A bar chart of the general distribution of the number of bus stops directly reachable is also provided in the figure below. As could be observed, the distribution of the number of bus stops directly reachable is more evenly comparing with the variable of number of bus routes. About half of the bus stops can reach at most 40 other bus stops directly. Only about 10% of the bus stops can directly reach more than 100 bus stops.

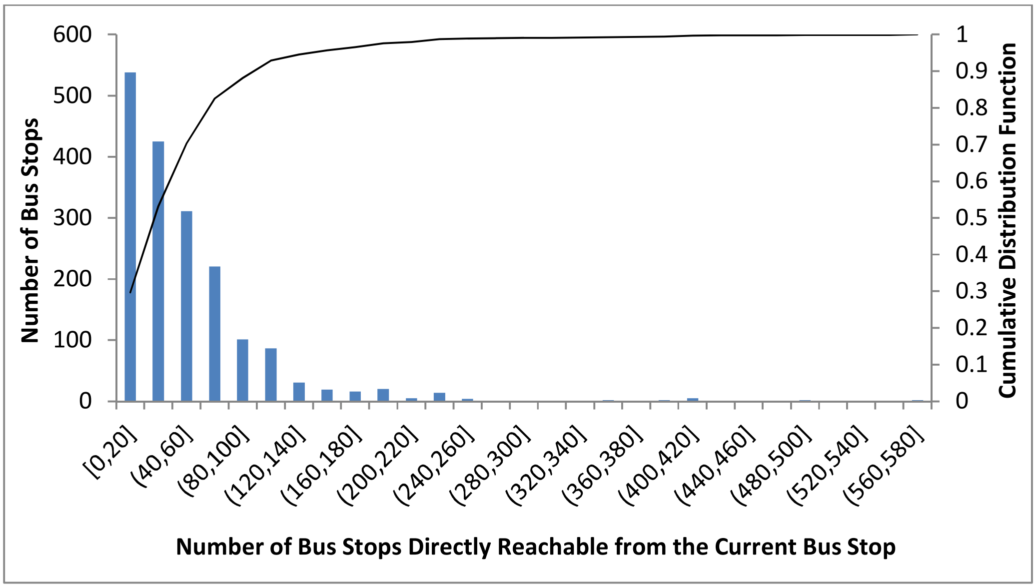


Figure 5-11 Number of Bus Stops w.r.t. Number of Bus Stops Directly Reachable

Similarly, the relationship between the dependent variable of number of transfers and this explanatory variable has also been roughly tested through a simple plot. The bus stops are first clustered using the similar methodology. Then, the average number of transfers is calculated for each group. As the transfer points are very special bus stops, they are not included within this plot.

Even though the plot diverges at the high end, the trend line indicates that a linear relationship could still be fit very well between the two positively related variables. The slope is smaller than the previous variable as the number of bus stops directly reachable has a much larger scale in its number.

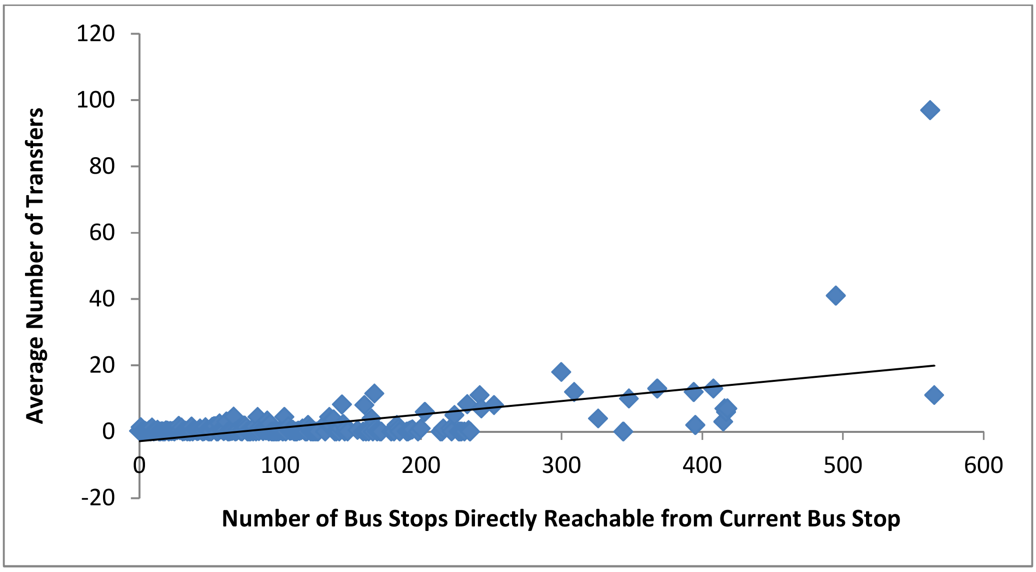


Figure 5-12 Average Number of Transfers vs. Number of Bus Stops Directly Reachable

This positive linear relationship also suggests that a higher value in R_0 will generally result in more transfer activities, which is similar to the previous variable.

5.3.7 Number of Bus Stops Reachable while One Transfer is Required (R_1)

In the similar way, the statistics are prepared for R_1. As an example, the bus stops directly reachable (dark green) and the bus stops indirectly reachable (light green) are calculated and plotted for the bus stop, University Avenue at Park Street, in Figure 5-13.

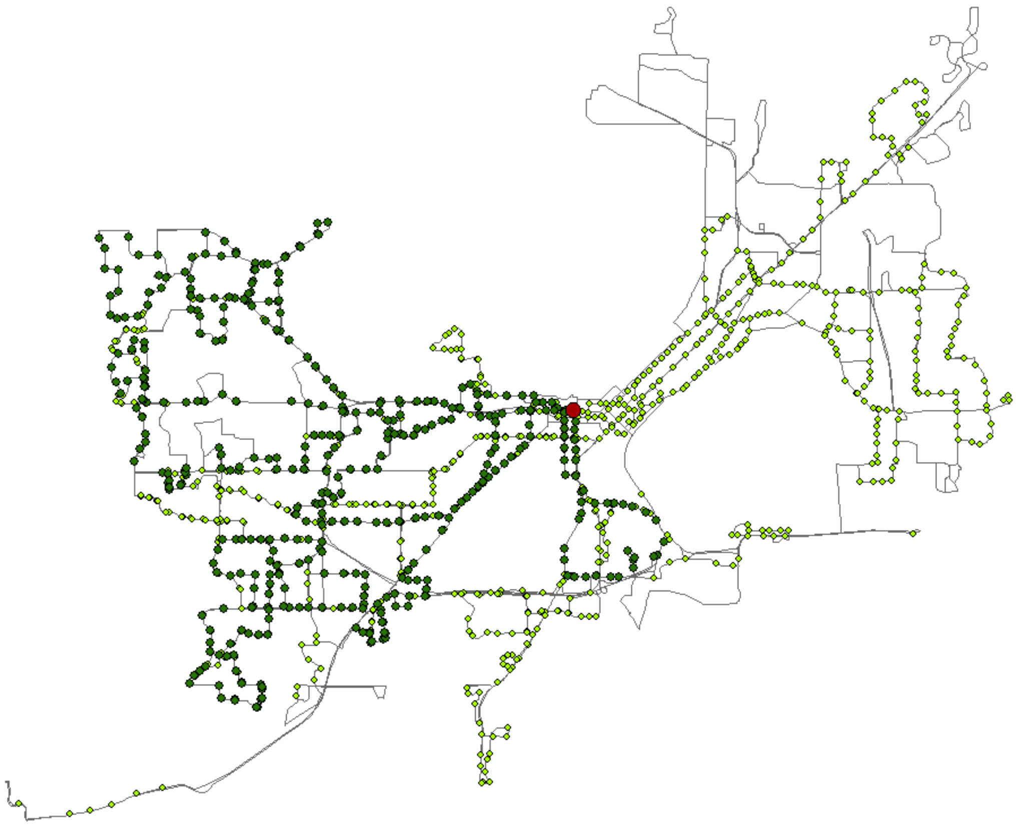


Figure 5-13 Visualization of R_0 and R_1 for University Avenue at Park Street

(Red: Current Bus Stop; Dark Green: bus stops directly reachable;

Light Green: bus stops indirectly reachable if 1 transfer is required)

From this figure, we can see that R_1 covers a much broader range (808 bus stops) comparing with R_0 (408 bus stops). For the general statistics, the maximum value is still obtained within the non-transfer point bus stops while the transfer point category shares a higher average value.

Table 5-9 General Stats of Number of Bus Stops Indirectly Reachable

	Non-Transfer Point Bus Stop	Transfer Point Bus Stop
Min	0	218
Max	1112	1041
Average	436.67	743.38
Standard Deviation	333.41	287.56

According to the distribution from both of the bar chart and the CDF, R_1 is more uniformly distributed than the two previous variables. This is because, through one transfer, bus stop within the low transit density area could be easily linked with a highly transit oriented bus stop. Therefore, by making a transfer, the area indirectly reachable can be dramatically extended. On the other hand, this actually reflects the mitigation from the transfer activities, which redistributes the existing transit resources.

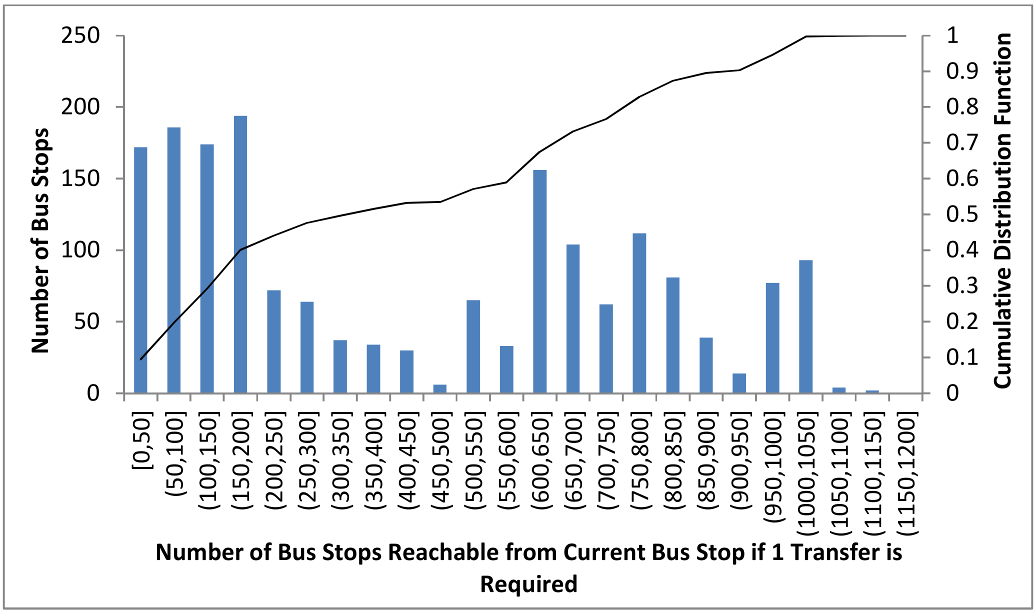


Figure 5-14 Number of Bus Stops w.r.t. Number of Bus Stops Indirectly Reachable

From the plot below, a positive linear relationship can still be observed between R_1 and the average number of transfers for each bus stop group, even though a smaller coefficient is indicated by the trend line. Therefore, we still expect that, the higher the number of R_1 is, the more attractions of transfer activities it will bring about.

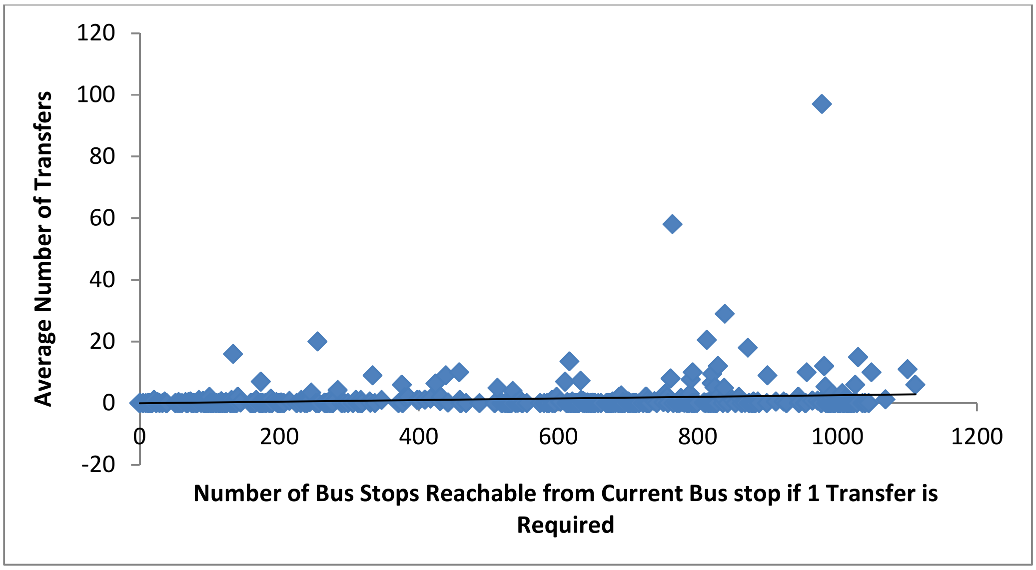


Figure 5-15 Average Number of Transfers vs. Number of Bus Stops Indirectly Reachable

6 Behavior Modeling and Results Analysis

6.1 Model Calibration

As discussed in chapter 2, a multiple linear regression model is used as our main approach to analyze the effects of the geo-spatial and topological characteristics towards the transit transfer accessibility. By assuming a linear relationship between the dependent variable and the explanatory variables, the coefficient for each variable will be calculated to fit the data, which is based on the criteria of minimizing the sum of the residuals S.

$$S = \sum_{i=1}^{n} r_i^2 \tag{6.1}$$

$$r_i = y_i - f(x_i, \beta) \tag{6.2}$$

In this chapter, an independence test will be first carried out to analyze the co-linearity among the explanatory variables. Then a full model will be evaluated by incorporating all the hand on features. In the third step, an information entropy based test will be conducted will simplify the model into a more robust form.

6.1.1 Independence Test

In this section, the correlation coefficients of all the variables will be first calculated to validate the independence among all the variables. After we have included all the minor land use categories as dummy variables in the full model, a total number of 44 variables are incorporated at the initial stage. All the variables included are listed in the table below:

Table 6-1 Description of Explanatory Variables

(a) Variables from X1 to X35

Variable	Description
X1	Transfer Point or Not
X2	Negative Exponent of Dist. from Current Bus Stop to the Nearest Transfer Point
X3	Number of Bus Routes Passing through Current Bus Stop
X4	Number of Bus Stops Directly Reachable from Current Bus Stop
X5	Number of Bus Stops Reachable from Current Bus Stop while 1 Transfer is Required
X6	Single Family or Not
X7	Two Family or Not or Not
X8	Multiple Family (3 or more dwelling units per building) or Not
X9	Group Quarters (dormitories retirement facilities) or Not
X10	Mobile Home Park or Not
X11	Chemicals and Allied Products Manufacture Miscellaneous Industrial or Not
X12	Wholesale Trade Rubber and Miscellaneous Plastic Products Manufacturing or Not
X13	Stone Clay and Glass Products Manufacturing Extractive or Not
X14	Railroad or Not
X15	Street Road Highway Right of Way or Not
X16	Automobile Parking (parking lots or ramps separate from other buildings) or Not
X17	Shopping Center or Not
X18	General Merchandise or Not
X19	Food (bakeries, grocery stores, liquor stores, butcher shops..) or Not
X20	Furniture, Home Furnishings and Equipment or Not
X21	Eating and Drinking (including restaurants) or Not
X22	Commercial Retail Sales Other or Not
X23	General Repair and Maintenance or Not
X24	Motor Vehicles or Not
X25	Tires, Batteries and Accessories or Not
X26	Gasoline Service Stations or Not
X27	Other Retail Trade - Automotive, Marine Craft, Aircraft and Access or Not
X28	Finance, Insurance and Real Estate or Not
X29	Personal Services or Not
X30	Business Services or Not
X31	Professional Services or Not
X32	Transient Lodging or Not
X33	Amusement or Not
X34	Education (schools, universities) or Not
X35	Health or Not

(b) Variables from X36 to X44

Variable	Description
X36	Administrative or Not
X37	Safety or Not
X38	Assembly or Not
X39	Cemeteries or Not
X40	Institutional/Government, N.E.C. or Not
X41	Other Land-Based Outdoor Recreation or Not
X42	Other Open or Vacant Land or Not
X43	Vacant, Unused Land or Not
X44	Woodlands or Not

Therefore, a 44×44 correlation coefficient matrix is calculated and presented in Appendix B. Among the matrix, the most controversial part is in the upper left rectangle, which is the correlation coefficient matrix of the first five variables:

Table 6-2 Correlation Coefficient Matrix

	X1	X2	X3	X4	X5
X1	/	0.71	0.21	0.22	0.06
X2	0.71	/	0.14	0.23	0.06
X3	0.21	0.14	/	0.81	0.31
X4	0.22	0.23	0.81	/	0.61
X5	0.06	0.06	0.31	0.61	/

High co-linearity could be observed among the pairs of variables, $X1$ and $X2$, $X3$ and $X4$. Fortunately, the maximum correlation coefficient does not exceed the critical value of 0.9, which is still admissible for the multiple linear regression.

6.1.2 Full Model Calibration

As a preprocessing step, all the variables are normalized so that the coefficients become easier to interpret. Then, based on the principles of the multiple linear regression, the initial estimation results for the full model are listed in Table 6-5.

Table 6-3 Estimation Results of the Complete Regression Model

(a) Variables from X1 to X35

Variable	Coefficient	Std Dev	p value	t value	Significance
Intercept	0.689	4.800	0.886	0.144	
X1	219.227	3.509	0.000	62.472	***
X2	4.381	0.233	0.000	18.838	***
X3	1.336	0.310	0.000	4.311	***
X4	0.717	0.370	0.053	1.940	*
X5	-0.237	0.224	0.291	-1.055	
X6	0.026	4.805	0.996	0.005	
X7	-0.093	4.852	0.985	-0.019	
X8	0.592	4.819	0.902	0.123	
X9	-0.147	5.268	0.978	-0.028	
X10	0.373	6.788	0.956	0.055	
X11	0.309	8.313	0.970	0.037	
X12	0.126	5.059	0.980	0.025	
X13	0.470	8.315	0.955	0.056	
X14	-0.023	5.879	0.997	-0.004	
X15	0.012	4.832	0.998	0.003	
X16	-4.907	5.105	0.337	-0.961	
X17	0.136	8.315	0.987	0.016	
X18	0.369	6.794	0.957	0.054	
X19	-0.171	6.789	0.980	-0.025	
X20	-0.392	5.880	0.947	-0.067	
X21	2.173	4.959	0.661	0.438	
X22	0.160	4.928	0.974	0.033	
X23	-1.240	5.883	0.833	-0.211	
X24	0.269	5.185	0.959	0.052	
X25	1.955	6.789	0.773	0.288	
X26	0.594	5.184	0.909	0.114	
X27	0.372	6.789	0.956	0.055	
X28	0.279	5.131	0.957	0.054	
X29	0.083	8.314	0.992	0.010	
X30	1.294	4.865	0.790	0.266	
X31	0.780	5.186	0.880	0.150	
X32	-0.075	5.881	0.990	-0.013	
X33	-0.294	8.326	0.972	-0.035	
X34	0.274	4.916	0.955	0.056	
X35	-0.663	5.882	0.910	-0.113	

(b) Variables from X36 to X44

Variable	Coefficient	Std Dev	p value	t value	Significance
X36	-0.360	5.367	0.947	-0.067	
X37	1.067	8.315	0.898	0.128	
X38	3.350	5.688	0.556	0.589	
X39	0.232	6.788	0.973	0.034	
X40	-1.034	8.315	0.901	-0.124	
X41	0.020	5.074	0.997	0.004	
X42	-3.560	5.016	0.478	-0.710	
X43	5.669	4.900	0.247	1.157	
X44	0.231	6.197	0.970	0.037	

Note: ***: p value < 0.001; **: p value < 0.05; *: p value < 0.1

Obviously, quite a number of the variables are not statistically significant indicated by the t value and p value. This further suggests that these variables are likely to be redundant which will result in an over-fitting issue. Therefore, in the following stage, the model reduction will be implemented to get rid of these redundancies and make the model more robust.

6.1.3 Model Simplification

In this section, an information entropy based Akaike information criterion (AIC) is implemented to eliminate the less significant variables from the full model. The reduction is conducted using a bidirectional stepwise algorithm, which tests at each step for variables to be included or excluded. At each stage in the process, after a new variable is added, a test is made to check if some variables can be deleted without appreciably increasing the residual sum of squares (RSS). The procedure terminates when the measure is maximized, or when the available improvement falls below some critical value (43). This provides a fairy good approximation for the globally optimal value which is obtained by traversing all the 2^m combinations of all the variables, where m refers to the number of variables in the full model. After a number of 35 stepwise iterations, the final model recommended by AIC is listed below:

Table 6-4 Estimation Results for the Regression Model Proposed by AIC

Variable	Coefficient	Std Dev	p value	t value	Significance
Intercept	0.825	0.168	0.000	4.923	***
x1	218.950	3.469	0.000	63.111	***
x2	4.416	0.228	0.000	19.344	***
x3	1.503	0.272	0.000	5.517	***
x4	0.447	0.274	0.103	1.630	
x16	-5.079	1.725	0.003	-2.944	**
x21	2.116	1.238	0.088	1.709	*
x30	1.182	0.801	0.140	1.476	
x42	-3.645	1.453	0.012	-2.509	**
x43	5.543	0.991	0.000	5.593	***

Note: ***: p value < 0.001; **: p value < 0.05; *: p value < 0.1.

As could be seen from the results, quite a number of redundant variables are excluded from the full model. The final model recommended by AIC is more robust, which only takes 10 variables including the intercept.

6.2 Results Analysis

6.2.1 Measurement of Effectiveness (MOE)

Based on the model calibration results, the RMSE, R^2 and $\bar{R}^2$ are calculated for both of the full model and the simplified model recommended by AIC.

Table 6-5 MOE of the Regression Models

Statistics	Full Model	AIC
Number of observations	1811	1811
Number of variables	45	10
RMSE	6.788	6.730
R^2	0.883	0.883
$\bar{R}^2$	0.880	0.882

Both of the full model and the model proposed by AIC achieve the same value in R^2. This indicates they both gracefully explain about 88.3% of the total variance of the sample data. However, the model recommended by AIC performs even better with respect to the RMSE and $\bar{R}^2$, as fewer explanatory variables are incorporated. Therefore, considering both of the MOE and the over-fitting issue which might be brought by including too many variables, the model proposed by a stepwise AIC is preferred as the ultimate choice.

6.2.2 Empirical Interpretation

6.2.2.1 Transfer Point or Not

The dummy variable indicates if the bus stop is a transfer point or not is one of the most important features with regard to the calibration results. The sign of the coefficient is positive and the magnitude is fairly large comparing with the others. This further confirms the cognition that transfer points are planned and designed for transfer activities in weighing both of the transit operational scheduling and infrastructural design. Therefore, transfer point bus stops will possess a higher transfer accessibility value than the ordinary ones.

6.2.2.2 Distance from the Bus Stop to the Nearest Transfer Point

The variable of the distance from current bus stop to the nearest transfer point is the second most significant feature. Together with the positive sign of the coefficient, our presumed negative exponential relationship has been verified. Therefore, the model indicates that transfer points also improve the transit transfer accessibility of the surrounding bus stops. However, this local influence is restricted only within a small area. Suggested by the negative exponential form of this variable, the influence will decrease dramatically when the distance increases.

6.2.2.3 Land Use Category of Adjacent Neighborhood

After the model simplification through AIC, some of the land use categorical variables are still preserved within the ultimate model while the others have been eliminated. Detailed information about the coefficients and general statistics of the features is listed in Table 6-6.

Table 6-6 General Stats for Land Use Variables

Variable	Description	Coeff.	# Bus Stops	Transfer Rate	Distance	# Routes	# Stops Reachable
x16	Automobile Parking	-5.079	16	1.56	3465.093	5.625	120.750
x21	Eating and Drinking	2.116	31	5.03	4305.969	5.290	116.290
x30	Business Services	1.182	74	1.89	4309.428	2.257	59.392
x42	Open or Vacant Land	-3.645	22	0	5257.749	2.091	46.273
x43	Vacant, Unused Land	5.543	48	0.81	6393.646	2.146	46.125
	Average Value	/	/	1.93	3959.386	2.198	51.918

Interestingly, the signs and magnitudes of the coefficients are quite different. Some of them have a positive sign while the others have a negative one. This further convinces us that different environments of the neighborhood will have different impacts towards the transfer accessibility of the bus stops.

For the land use categories of automobile parking and open vacant land, the negative coefficient indicates that individuals tend not to make a transfer at these areas even if the values of the other influential factors maintain the same level. In a more comprehensive analysis, by comparing the local average values of the distances, number of routes and number of bus stops directly reachable, the automobile parking areas possess a much higher standard than the average values. Hereby, the coefficient of this land use category amends this superiority by introducing a very large negative coefficient. In real situations, the automobile parking facilities are mainly designed for park and ride, which initiate more starting trip records. Therefore, it is not surprised to see that the transfer rate around these areas is below the average value. Similar explanations could be applied to the land use category of open or vacant land.

Among the other three land use categories, a positive coefficient is assigned. For the land use category of eating and drinking, a quite large coefficient is observed even if the average values of the other three descriptive variables are already above the average standard. As the transfer waiting time has been limited within an interval of 30 minutes, most of the dining activities could not be finished during that time. Therefore, we consider the high transfer accessibility is not attributed by the additional dining activities. By looking at the locations of these eating and drinking districts, we think people tend to make transfers around these neighborhoods are mainly attracted by the local social environment. As in most cases, the restaurants are located in some prosperous neighborhoods with an open and friendly atmosphere. Different from the other communities such as the residential areas, these districts seem to be more public activity oriented. Lively crowds and entertaining elements will distract passengers' anxieties when they are waiting for the upcoming buses. People can even enjoy a cup of coffee while seating under the shelters in case of the unpleasant sunshine and storms. In addition, the safety conditions within these areas are also deemed to be better considering a higher population density. Therefore, making a transfer at these neighborhoods seems to be much more comfortable comparing with the other areas. Then, followed by the similar logic, it is not surprised to see that the coefficient pertained to the land use category of business service also has a positive sign.

As a result, based on the analysis above, the land use categories can be regarded as the social economic measure towards the transit transfer accessibility of the bus stops. It captures the discrepancies of the accessibility indices among different communities, which cannot be perfectly explained by simply using the transit connectivity as the general descriptive features.

6.2.2.4 Number of Bus Routes Passing through the Current Bus Stop

The number of bus routes passing through the current bus stops is the most significant transit topological feature. The positive sign of the coefficient reveals the fact that the larger number of bus routes will contribute to a higher value of transit transfer accessibility. This is consistent with the cognition that people tend to make transfers at a highly populated transit area, which indicate more interactions among the bus routes. Therefore, these bus stops can offer a higher level of service to the transferring passengers while providing a comparatively shorter waiting time and smaller walking distance.

6.2.2.5 Number of Bus Stops Directly Reachable from the Current Bus Stop (R_0)

Number of bus stops directly reachable from the current bus stop is another topological feature recommended by the stepwise AIC. The positive sign suggests that a higher value of R_0 can still contribute to a greater value of the transit transfer accessibility. Even if this variable is not statistically significant enough comparing with the other variables, it is still preserved as an adjustment to the previous.

One explanation could be directly derived from the observation that R_0 provides complementary description about the network connectivity focused more on the number of actual bus stops directly reachable other than simply using the number of routes as the measure. As an example, suppose one bus stop is located near the end along its direction, it could only be reached by few bus stops even if many bus routes pass through that bus stop. As another example, a bus stop can still be directly connected with a large number of bus stops provided it is passed by any one of the highly connected bus routes. Hence, the transit network connectivity cannot be well represented only by the number of bus routes over each single bus itself. Therefore, R_0 is incorporated into the model to help describe the relationship between the transit transfer accessibility and the transit networking topological features.

6.2.2.6 Number of Bus Stops Reachable while One Transfer is Required (R_1)

R_1 has been eliminated from the ultimate model proposed by AIC. This suggests that, as a higher order term which describes the connectivity of the transit network, R_1 is not strongly related to the transfer accessibility. Two reasons are able to explain this result. First, considering the scale of the City of Madison and the LOS of Madison Metro, few trips are carried out by more than one transfer activity. Therefore, the number of bus stops reachable while one transfer is required does not provide a useful reference to the passengers whey they choose the transferring bus stop. Second, as being demonstrated earlier in the preliminary analysis, R_1 has a more uniform distribution comparing with the other two topological measures. In other words, all the bus stops are within comparatively the same scale in terms of the value of R_1. Therefore, this variable can only provide an indiscriminate description towards all the bus stops, which contribute purely as an intercept to the multiple linear regression model.

7 Conclusions and Future Work

7.1 Conclusions

In this Master's research, an empirical study has been conducted towards the accessibility characteristics of transit transfers. Based on the layout of the transit network system, several geospatial and topological features have been defined and proposed as the potential measures towards the transit transfer accessibility. A multiple linear regression model and an information entropy based model simplification method have been deployed as the methodology to analyze the effects of these predefined measures.

As a result, among the geospatial measurements, the transfer point bus stops generally possess a higher standard of transfer accessibility comparing with the other ordinary bus stops. This is consistent with the fact that, within transfer points, the infrastructure of the bus stops are specially designed and the scheduling of the bus routes are typically accommodated for transfer activities. In addition, the transfer points also improve the transfer accessibility of the surrounding bus stops as they attract more transfer activities to the adjacent communities. However, this effect drops in a negative exponential order when the distance between the bus stop and the nearest transfer point increases.

Among the topological measures, the number of bus routes passing through the current bus stop and the number of bus stops directly reachable from the current bus stop are the two most significant characteristics as they generally represent the level of connectivity of the local transit network. The modeling result further indicates that, the higher value being assigned for the two terms will contribute to a higher transfer accessibility ratio for the corresponding bus stops.

Except for the two descriptive characteristics, the land uses of the bus stops serve as categorical variables which measure the discrepancies of the transfer accessibilities among different functioned communities. The estimation result suggests that for some of the land uses, such as eating, drinking, and business service, the overall level of transfer accessibility is generally higher even if the geo-spatial and topological measures are within the same level of values. However, for the other land use categories such as automobile parking and vacant land, the level of transfer accessibility is comparatively lower. This finding further concludes that, the environment of the bus stops will also affect the passengers when they are choosing the bus stop to make the transfer. This will in turn affect the transit transfer accessibility of the bus stop as a consequence.

7.2 Future Work

As no swiping card record is available for the alighting activities from the faring system of Madison Metro, an approximation method is implemented to determine if the two consecutive transit records are connected by a transfer activity. However, in order to secure the accuracy of the algorithm, the research target is restricted within the commuting trips to obtain a comparatively stable trip pattern. Therefore, the findings towards the transit transfer accessibility are somehow limited to commuting trip activities. In the future work, the research scope should be extended to embrace more transit activities. In that sense, the conclusions can be generalized to apply for a broader usage.

Another emphasis of the future work is the granularity of the measurement features. At current stage, the transit network connectivity is modeled by the number of bus routes over each bus stop and the number of bus stops directly reachable from the current bus stop. However, this is only a measurement towards the network connectivity within the spatial dimension. Considering the frequency of the bus route is also an influencing factor, more investigations can be focused on the temporal dimension of the transit scheduling. Moreover, even though the land use categories are proved to be significant to the transfer accessibility, the underlying influential factors still remain unknown. In the future work, more substantial measures based on the social economic features of the land use are recommended to get included while evaluating the discrepancies of transfer accessibility across different land used neighborhoods.

8 References

1. Crockett, C. Process for Improving Transit Service Connectivity. M.S. thesis. Massachusetts Institute of Technology, Cambridge, 2002.
2. Guo, Zhan, and Nigel HM Wilson. "Assessment of the Transfer Penalty for Transit Trips Geographic Information System-Based Disaggregate Modeling Approach." Transportation Research Record: Journal of the Transportation Research Board 1872.1 (2004): 10-18.
3. Hansen, Walter G. "How accessibility shapes land use." Journal of the American Institute of Planners 25.2 (1959): 73-76.
4. Degree (graph theory), www.en.wikipedia.org/wiki/Degree_(graph_theory)
5. Fan, Lang, and Christine L. Mumford. "A metaheuristic approach to the urban transit routing problem." Journal of Heuristics 16.3 (2010): 353-372.
6. Chakroborty, Partha. "Genetic algorithms for optimal urban transit network design." Computer - Aided Civil and Infrastructure Engineering 18.3 (2003): 184-200.
7. Zhao, Fang, and I. Ubaka. "Transit network optimization-minimizing transfers and optimizing route directness." Journal of Public Transportation 7.1 (2004): 63-82.
8. Baaj, M. Hadi, and Hani S. Mahmassani. "An AI - based approach for transit route system planning and design." Journal of Advanced Transportation 25.2 (1991): 187-209.
9. Baaj, M. Hadi, and Hani S. Mahmassani. "Hybrid route generation heuristic algorithm for the design of transit networks." Transportation Research Part C: Emerging Technologies 3.1 (1995): 31-50.

10. Zhao, Fang, and Albert Gan. Optimization of transit network to minimize transfers. No. Final Report. 2003.

11. Pattnaik, S. B., S. Mohan, and V. M. Tom. "Urban bus transit route network design using genetic algorithm." Journal of Transportation Engineering 124.4 (1998): 368-375.

12. Chien, Steven, Zhaowei Yang, and Edwin Hou. "Genetic algorithm approach for transit route planning and design." Journal of transportation engineering 127.3 (2001): 200-207.

13. Fan, W., Machemehl, R.B.: A tabu search based heuristic method for the transit route network design problem. In: The 9th International Conference on Computer-Aided Scheduling of Public Transport, San Diego, California (2004)

14. Fan,W.,Machemehl, R.B.: Using a simulated annealing algorithm to solve the transit route network design problem. J. Transp. Eng. 122–132 (2006).

15. Zhao, Fang, and Xiaogang Zeng. "Simulated annealing–genetic algorithm for transit network optimization." Journal of Computing in Civil Engineering 20.1 (2006): 57-68.

16. YANG, Zhong-Zhen, et al. "Transit route network design-maximizing direct and transfer demand density." Transportation research. Part C, Emerging technologies 22 (2012): 58-75.

17. Bookbinder, James H., and Alain Desilets. "Transfer optimization in a transit network." Transportation science 26.2 (1992): 106-118.

18. Ceder, Avishai, B. Golany, and O. Tal. "Creating bus timetables with maximal synchronization." Transportation Research Part A: Policy and Practice 35.10 (2001): 913-928.

19. Zhao, Fang, and Xiaogang Zeng. "Optimization of transit route network, vehicle headways and timetables for large-scale transit networks." European Journal of Operational Research 186.2 (2008): 841-855.

20. Nielsen, Otto Anker. "A stochastic transit assignment model considering differences in passengers utility functions." Transportation Research Part B: Methodological 34.5 (2000): 377-402.

21. Lozano, Angelica, and Giovanni Storchi. "Shortest viable path algorithm in multimodal networks." Transportation Research Part A: Policy and Practice 35.3 (2001): 225-241.

22. Pallottino, Stefano, and Maria Grazia Scutella. "Shortest path algorithms in transportation models: classical and innovative aspects." Equilibrium and advanced transportation modelling 245 (1998): 281.

23. Abdelghany, Khaled F., and Hani S. Mahmassani. "Dynamic trip assignment-simulation model for intermodal transportation networks." Transportation Research Record: Journal of the Transportation Research Board 1771.1 (2001): 52-60.

24. Abdelghany, Ahmed F., and Hani S. Mahmassani. "Temporal-spatial microassignment and sequencing of travel demand with activity-trip chains." Transportation Research Record: Journal of the Transportation Research Board 1831.1 (2003): 89-97.

25. Ziliaskopoulos, Athanasios K., and Hani S. Mahmassani. Time-dependent, shortest-path algorithm for real-time intelligent vehicle highway system applications. No. 1408. 1993.

26. Ziliaskopoulos, Athanasios, and Whitney Wardell. "An intermodal optimum path algorithm for multimodal networks with dynamic arc travel times and switching delays." European Journal of Operational Research 125.3 (2000): 486-502.

27. Chang, Elaine, Evangelos Floros, and Athanasios Ziliaskopoulos. "An Intermodal Time-Dependent Minimum Cost Path Algorithm." Dynamic Fleet Management. Springer US, 2007. 113-132.

28. Lo, Hong K., C. W. Yip, and K. H. Wan. "Modeling transfer and non-linear fare structure in multi-modal network." Transportation Research Part B: Methodological 37.2 (2003): 149-170.

29. Lo, Hong K., Chun-Wing Yip, and Quentin K. Wan. "Modeling competitive multi-modal transit services: a nested logit approach." Transportation Research Part C: Emerging Technologies 12.3 (2004): 251-272.

30. Nuzzolo, Agostino, Francesco Russo, and Umberto Crisalli. "A doubly dynamic schedule-based assignment model for transit networks." Transportation Science 35.3 (2001): 268-285.

31. Zhou, Xuesong, Hani S. Mahmassani, and Kuilin Zhang. "Dynamic micro-assignment modeling approach for integrated multimodal urban corridor management." Transportation Research Part C: Emerging Technologies 16.2 (2008): 167-186.

32. Dalvi, M. Quasim, and K. M. Martin. "The measurement of accessibility: some preliminary results." Transportation 5.1 (1976): 17-42.

33. Geurs, Karst T., and Bert van Wee. "Accessibility evaluation of land-use and transport strategies: review and research directions." Journal of Transport geography 12.2 (2004): 127-140.

34. Akaike information criterion, www.en.wikipedia.org/wiki/Akaike_information_criterion.

35. Fang, Yixin. "Asymptotic equivalence between cross-validations and Akaike Information Criteria in mixed-effects models." J. Data Sci 9 (2011): 15-21.

36. Akaike, Hirotugu. "Fitting autoregressive models for prediction." Annals of the institute of Statistical Mathematics 21.1 (1969): 243-247.

37. Root-mean-square deviation, www.en.wikipedia.org/wiki/Root-mean-square_deviation.

38. Coefficient of determination, www.en.wikipedia.org/wiki/Coefficient_of_determination.

39. Lai, J. The Study of Experimental Design for Transfer Behavior in Transit Commuting Trips. M.S. thesis. University of Wisconsin-Madison, 2011.

40. Metro Transit Tracker, www.webwatch.cityofmadison.com/webwatch/default.aspx.

41. Taxicab geometry, www. en.wikipedia.org/wiki/Taxicab_geometry.

42. Stepwise regression, www.en.wikipedia.org/wiki/Stepwise_regression.

Appendix A

Reference of Land Use Code

Table A-1 Reference of Major Land Use Code

Major Land Use Code	Land Use Type
1	RESIDENTIAL
2	INDUSTRIAL
3	TRANSPORTATION
5	COMMERCIAL RETAIL SALES
6	COMMERCIAL RETAIL SERVICES
7	INSTITUTIONAL/GOVERNMENT
8	OUTDOOR RECREATION
10	OTHER

Table A-2 Reference of Minor Land Use Code

(a) Major Land Use Code from 1 to 5

Major	Minor	Land Use Type
1	111	Single Family
1	113	Two Family
1	115	Multiple Family (3 or more dwelling units per building)
1	129	Group Quarters (dormitories retirement facilities)
1	142	Mobile Home Park
2	36	Chemicals and Allied Products Manufacturing Miscellaneous Industrial
2	38	Wholesale Trade Rubber, Miscellaneous Plastic Products Manufacturing
2	39	Stone Clay and Glass Products Manufacturing Extractive
3	41	Railroad
3	45	Street Road Highway Right of Way
3	46	Automobile Parking (parking lots or ramps separate from other buildings)
5	9	Shopping Center
5	53	General Merchandise
5	54	Food (bakeries, grocery stores, liquor stores, butcher shops..)
5	57	Furniture, Home Furnishings and Equipment
5	58	Eating and Drinking (including restaurants)
5	59	Other
5	531	General Repair and Maintenance
5	551	Motor Vehicles
5	552	Tires, Batteries and Accessories
5	553	Gasoline Service Stations
5	559	Other Retail Trade - Automotive, Marine Craft, Aircraft and Access

(b) Major Land Use Code from 6 to 10

Major	Minor	Land Use Type
6	61	Finance, Insurance and Real Estate
6	62	Personal Services
6	63	Business Services
6	65	Professional Services
6	67	Transient Lodging
6	68	Amusement
7	71	Education (schools, universities)
7	72	Health
7	73	Administrative
7	74	Safety
7	75	Assembly
7	76	Cemeteries
7	79	Institutional/Government, N.E.C.
8	83	Other Land-Based Outdoor Recreation
10	97	Other Open or Vacant Land
10	98	Vacant, Unused Land
10	99	Woodlands

Appendix B

Correlation Coefficient Matrix of the Full Model

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	/	0.71	0.21	0.22	0.06	-0.05	-0.02	-0.03	0.00	0.00	0.00	-0.01
X2	0.71	/	0.14	0.23	0.06	-0.02	-0.01	-0.02	0.00	0.00	0.00	0.00
X3	0.21	0.14	/	0.81	0.31	-0.14	-0.01	0.03	0.10	-0.02	-0.01	-0.03
X4	0.22	0.23	0.81	/	0.61	-0.07	0.00	0.02	0.03	-0.02	-0.01	-0.04
X5	0.06	0.06	0.31	0.61	/	0.00	0.04	0.05	-0.01	-0.02	-0.01	-0.06
X6	-0.05	-0.02	-0.14	-0.07	0.00	/	-0.23	-0.40	-0.07	-0.03	-0.02	-0.10
X7	-0.02	-0.01	-0.01	0.00	0.04	-0.23	/	-0.09	-0.02	-0.01	-0.01	-0.02
X8	-0.03	-0.02	0.03	0.02	0.05	-0.40	-0.09	/	-0.03	-0.01	-0.01	-0.04
X9	0.00	0.00	0.10	0.03	-0.01	-0.07	-0.02	-0.03	/	0.00	0.00	-0.01
X10	0.00	0.00	-0.02	-0.02	-0.02	-0.03	-0.01	-0.01	0.00	/	0.00	0.00
X11	0.00	0.00	-0.01	-0.01	-0.01	-0.02	-0.01	-0.01	0.00	0.00	/	0.00
X12	-0.01	0.00	-0.03	-0.04	-0.06	-0.10	-0.02	-0.04	-0.01	0.00	0.00	/
X13	0.00	0.00	-0.01	-0.02	-0.03	-0.02	-0.01	-0.01	0.00	0.00	0.00	0.00
X14	0.00	0.00	-0.01	-0.01	0.00	-0.05	-0.01	-0.02	0.00	0.00	0.00	0.00
X15	0.04	0.03	-0.03	-0.06	-0.04	-0.29	-0.07	-0.12	-0.02	-0.01	-0.01	-0.03
X16	0.17	0.12	0.13	0.11	0.05	-0.09	-0.02	-0.04	-0.01	0.00	0.00	-0.01
X17	0.00	0.00	-0.01	-0.01	-0.02	-0.02	-0.01	-0.01	0.00	0.00	0.00	0.00
X18	0.00	0.00	0.01	-0.03	-0.04	-0.03	-0.01	-0.01	0.00	0.00	0.00	0.00
X19	0.00	0.00	0.00	-0.02	-0.03	-0.03	-0.01	-0.01	0.00	0.00	0.00	0.00
X20	0.00	0.00	0.00	0.00	-0.04	-0.05	-0.01	-0.02	0.00	0.00	0.00	0.00
X21	-0.01	-0.01	0.16	0.15	0.02	-0.13	-0.03	-0.05	-0.01	0.00	0.00	-0.01
X22	-0.01	-0.01	0.02	0.01	-0.02	-0.14	-0.03	-0.06	-0.01	0.00	0.00	-0.01
X23	0.00	0.00	0.02	0.03	-0.02	-0.05	-0.01	-0.02	0.00	0.00	0.00	0.00
X24	-0.01	0.00	-0.03	-0.04	-0.06	-0.08	-0.02	-0.03	-0.01	0.00	0.00	-0.01
X25	0.00	0.00	-0.02	-0.01	0.01	-0.03	-0.01	-0.01	0.00	0.00	0.00	0.00
X26	-0.01	0.00	-0.01	0.00	-0.01	-0.08	-0.02	-0.03	-0.01	0.00	0.00	-0.01
X27	0.00	0.00	-0.02	-0.02	-0.04	-0.03	-0.01	-0.01	0.00	0.00	0.00	0.00
X28	-0.01	0.00	-0.03	-0.02	-0.02	-0.09	-0.02	-0.04	-0.01	0.00	0.00	-0.01
X29	0.00	0.00	-0.01	0.00	-0.01	-0.02	-0.01	-0.01	0.00	0.00	0.00	0.00
X30	-0.01	-0.01	0.00	0.03	0.01	-0.20	-0.05	-0.08	-0.02	-0.01	0.00	-0.02
X31	-0.01	0.00	0.04	0.03	0.02	-0.08	-0.02	-0.03	-0.01	0.00	0.00	-0.01
X32	0.00	0.00	-0.02	0.01	0.02	-0.05	-0.01	-0.02	0.00	0.00	0.00	0.00
X33	0.00	0.00	-0.01	0.03	0.04	-0.02	-0.01	-0.01	0.00	0.00	0.00	0.00
X34	-0.01	-0.01	0.21	0.14	0.05	-0.15	-0.04	-0.06	-0.01	-0.01	0.00	-0.02
X35	0.00	0.00	0.02	0.01	0.02	-0.05	-0.01	-0.02	0.00	0.00	0.00	0.00
X36	0.00	0.00	0.00	0.01	0.01	-0.07	-0.02	-0.03	0.00	0.00	0.00	-0.01
X37	0.00	0.00	-0.01	0.00	-0.01	-0.02	-0.01	-0.01	0.00	0.00	0.00	0.00
X38	0.00	0.00	0.07	0.03	-0.01	-0.05	-0.01	-0.02	0.00	0.00	0.00	-0.01

X39	0.00	0.00	-0.02	0.00	0.00	-0.03	-0.01	-0.01	0.00	0.00	0.00	0.00
X40	0.00	0.00	0.01	0.00	-0.02	-0.02	-0.01	-0.01	0.00	0.00	0.00	0.00
X41	-0.01	0.00	-0.02	-0.02	-0.01	-0.10	-0.02	-0.04	-0.01	0.00	0.00	-0.01
X42	0.07	-0.01	0.00	-0.01	-0.03	-0.11	-0.03	-0.05	-0.01	0.00	0.00	-0.01
X43	0.09	0.07	0.00	-0.02	-0.03	-0.16	-0.04	-0.07	-0.01	-0.01	0.00	-0.02
X44	0.00	0.00	-0.01	-0.03	-0.04	-0.04	-0.01	-0.02	0.00	0.00	0.00	0.00
	X13	X14	X15	X16	X17	X18	X19	X20	X21	X22	X23	X24
X1	0.00	0.00	0.04	0.17	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01
X2	0.00	0.00	0.03	0.12	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	0.00
X3	-0.01	-0.01	-0.03	0.13	-0.01	0.01	0.00	0.00	0.16	0.02	0.02	-0.03
X4	-0.02	-0.01	-0.06	0.11	-0.01	-0.03	-0.02	0.00	0.15	0.01	0.03	-0.04
X5	-0.03	0.00	-0.04	0.05	-0.02	-0.04	-0.03	-0.04	0.02	-0.02	-0.02	-0.06
X6	-0.02	-0.05	-0.29	-0.09	-0.02	-0.03	-0.03	-0.05	-0.13	-0.14	-0.05	-0.08
X7	-0.01	-0.01	-0.07	-0.02	-0.01	-0.01	-0.01	-0.01	-0.03	-0.03	-0.01	-0.02
X8	-0.01	-0.02	-0.12	-0.04	-0.01	-0.01	-0.01	-0.02	-0.05	-0.06	-0.02	-0.03
X9	0.00	0.00	-0.02	-0.01	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01
X10	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X11	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X12	0.00	0.00	-0.03	-0.01	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01
X13	/	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X14	0.00	/	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	0.00
X15	-0.01	-0.01	/	-0.03	-0.01	-0.01	-0.01	-0.01	-0.04	-0.04	-0.01	-0.02
X16	0.00	0.00	-0.03	/	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01
X17	0.00	0.00	-0.01	0.00	/	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X18	0.00	0.00	-0.01	0.00	0.00	/	0.00	0.00	0.00	0.00	0.00	0.00
X19	0.00	0.00	-0.01	0.00	0.00	0.00	/	0.00	0.00	0.00	0.00	0.00
X20	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	/	-0.01	-0.01	0.00	0.00
X21	0.00	-0.01	-0.04	-0.01	0.00	0.00	0.00	-0.01	/	-0.02	-0.01	-0.01
X22	0.00	-0.01	-0.04	-0.01	0.00	0.00	0.00	-0.01	-0.02	/	-0.01	-0.01
X23	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	/	0.00
X24	0.00	0.00	-0.02	-0.01	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	/
X25	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X26	0.00	0.00	-0.02	-0.01	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01
X27	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X28	0.00	0.00	-0.03	-0.01	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01
X29	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X30	0.00	-0.01	-0.06	-0.02	0.00	-0.01	-0.01	-0.01	-0.03	-0.03	-0.01	-0.02
X31	0.00	0.00	-0.02	-0.01	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01

X32	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	0.00
X33	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X34	0.00	-0.01	-0.05	-0.01	0.00	-0.01	-0.01	-0.01	-0.02	-0.02	-0.01	-0.01
X35	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	0.00
X36	0.00	0.00	-0.02	-0.01	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01
X37	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X38	0.00	0.00	-0.02	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	0.00
X39	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X40	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X41	0.00	0.00	-0.03	-0.01	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	-0.01
X42	0.00	-0.01	-0.03	-0.01	0.00	0.00	0.00	-0.01	-0.01	-0.02	-0.01	-0.01
X43	0.00	-0.01	-0.05	-0.02	0.00	-0.01	-0.01	-0.01	-0.02	-0.02	-0.01	-0.01
X44	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00	0.00
	X25	X26	X27	X28	X29	X30	X31	X32	X33	X34	X35	X36
X1	0.00	-0.01	0.00	-0.01	0.00	-0.01	-0.01	0.00	0.00	-0.01	0.00	0.00
X2	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X3	-0.02	-0.01	-0.02	-0.03	-0.01	0.00	0.04	-0.02	-0.01	0.21	0.02	0.00
X4	-0.01	0.00	-0.02	-0.02	0.00	0.03	0.03	0.01	0.03	0.14	0.01	0.01
X5	0.01	-0.01	-0.04	-0.02	-0.01	0.01	0.02	0.02	0.04	0.05	0.02	0.01
X6	-0.03	-0.08	-0.03	-0.09	-0.02	-0.20	-0.08	-0.05	-0.02	-0.15	-0.05	-0.07
X7	-0.01	-0.02	-0.01	-0.02	-0.01	-0.05	-0.02	-0.01	-0.01	-0.04	-0.01	-0.02
X8	-0.01	-0.03	-0.01	-0.04	-0.01	-0.08	-0.03	-0.02	-0.01	-0.06	-0.02	-0.03
X9	0.00	-0.01	0.00	-0.01	0.00	-0.02	-0.01	0.00	0.00	-0.01	0.00	0.00
X10	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X12	0.00	-0.01	0.00	-0.01	0.00	-0.02	-0.01	0.00	0.00	-0.02	0.00	-0.01
X13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X14	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X15	-0.01	-0.02	-0.01	-0.03	-0.01	-0.06	-0.02	-0.01	-0.01	-0.05	-0.01	-0.02
X16	0.00	-0.01	0.00	-0.01	0.00	-0.02	-0.01	0.00	0.00	-0.01	0.00	-0.01
X17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X18	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X19	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X20	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X21	0.00	-0.01	0.00	-0.01	0.00	-0.03	-0.01	-0.01	0.00	-0.02	-0.01	-0.01
X22	0.00	-0.01	0.00	-0.01	0.00	-0.03	-0.01	-0.01	0.00	-0.02	-0.01	-0.01
X23	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X24	0.00	-0.01	0.00	-0.01	0.00	-0.02	-0.01	0.00	0.00	-0.01	0.00	-0.01

X25	/	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X26	0.00	/	0.00	-0.01	0.00	-0.02	-0.01	0.00	0.00	-0.01	0.00	-0.01
X27	0.00	0.00	/	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X28	0.00	-0.01	0.00	/	0.00	-0.02	-0.01	0.00	0.00	-0.01	0.00	-0.01
X29	0.00	0.00	0.00	0.00	/	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X30	-0.01	-0.02	-0.01	-0.02	0.00	/	-0.02	-0.01	0.00	-0.03	-0.01	-0.01
X31	0.00	-0.01	0.00	-0.01	0.00	-0.02	/	0.00	0.00	-0.01	0.00	-0.01
X32	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	/	0.00	-0.01	0.00	0.00
X33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	/	0.00	0.00	0.00
X34	-0.01	-0.01	-0.01	-0.01	0.00	-0.03	-0.01	-0.01	0.00	/	-0.01	-0.01
X35	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	/	0.00
X36	0.00	-0.01	0.00	-0.01	0.00	-0.01	-0.01	0.00	0.00	-0.01	0.00	/
X37	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X38	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X39	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
X40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
X41	0.00	-0.01	0.00	-0.01	0.00	-0.02	-0.01	0.00	0.00	-0.02	0.00	-0.01
X42	0.00	-0.01	0.00	-0.01	0.00	-0.02	-0.01	-0.01	0.00	-0.02	-0.01	-0.01
X43	-0.01	-0.01	-0.01	-0.01	0.00	-0.03	-0.01	-0.01	0.00	-0.03	-0.01	-0.01
X44	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00	0.00
	X37	X38	X39	X40	X41	X42	X43	X44				
X1	0.00	0.00	0.00	0.00	-0.01	0.07	0.09	0.00				
X2	0.00	0.00	0.00	0.00	0.00	-0.01	0.07	0.00				
X3	-0.01	0.07	-0.02	0.01	-0.02	0.00	0.00	-0.01				
X4	0.00	0.03	0.00	0.00	-0.02	-0.01	-0.02	-0.03				
X5	-0.01	-0.01	0.00	-0.02	-0.01	-0.03	-0.03	-0.04				
X6	-0.02	-0.05	-0.03	-0.02	-0.10	-0.11	-0.16	-0.04				
X7	-0.01	-0.01	-0.01	-0.01	-0.02	-0.03	-0.04	-0.01				
X8	-0.01	-0.02	-0.01	-0.01	-0.04	-0.05	-0.07	-0.02				
X9	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.00				
X10	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00				
X11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
X12	0.00	-0.01	0.00	0.00	-0.01	-0.01	-0.02	0.00				
X13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
X14	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00				
X15	-0.01	-0.02	-0.01	-0.01	-0.03	-0.03	-0.05	-0.01				
X16	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.02	0.00				
X17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				

X18	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00				
X19	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00				
X20	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00				
X21	0.00	-0.01	0.00	0.00	-0.01	-0.01	-0.02	-0.01				
X22	0.00	-0.01	0.00	0.00	-0.01	-0.02	-0.02	-0.01				
X23	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00				
X24	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.00				
X25	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00				
X26	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.00				
X27	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00				
X28	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.00				
X29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
X30	0.00	-0.01	-0.01	0.00	-0.02	-0.02	-0.03	-0.01				
X31	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.00				
X32	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00				
X33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
X34	0.00	-0.01	-0.01	0.00	-0.02	-0.02	-0.03	-0.01				
X35	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	0.00				
X36	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01	0.00				
X37	/	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
X38	0.00	/	0.00	0.00	-0.01	-0.01	-0.01	0.00				
X39	0.00	0.00	/	0.00	0.00	0.00	-0.01	0.00				
X40	0.00	0.00	0.00	/	0.00	0.00	0.00	0.00				
X41	0.00	-0.01	0.00	0.00	/	-0.01	-0.02	0.00				
X42	0.00	-0.01	0.00	0.00	-0.01	/	-0.02	0.00				
X43	0.00	-0.01	-0.01	0.00	-0.02	-0.02	/	-0.01				
X44	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	/				