Optimal Planning of Real-Time Bus Information System for User-Switching Behavior

Zoonky Lee 1, Sewoong Hwang 1 and Jonghyuk Kim 2,*

1 Graduate School of Information, Yonsei University, 50, Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea; zlee@yonsei.ac.kr (Z.L.); indimoa@gmail.com (S.H.)
2 Division of Computer Science and Engineering, Sunmoon University, 70, Sunmoon-ro221beon-gil, Tangjeong-myeon, Asan-si, Chungcheongnam-do 31460, Korea
* Correspondence: jonghyuk@sunmoon.ac.kr; Tel.: +82-41-530-2266

Received: 28 September 2020; Accepted: 11 November 2020; Published: 13 November 2020

Abstract: Seoul Metropolitan City’s buses cater to more than 50% of the average daily public transportation use, and they are the most important transportation mode in Korea, together with the subway. Since 2004, all public transportation records of passengers have been stored in Seoul, using smart transportation cards. This study explores the environmental and psychological factors in implementing a smart transportation system. We analyze the switching behavior of traffic users according to traffic congestion time and number of transfers based on public transportation data and show that bus-use behavior differs according to the traffic information of users and the degree of traffic congestion. Information-based switching behavior of people living near bus stops induces people to change routes during traffic congestion. However, in non-congested situations, the original routes are used. These results can guide the formulation of policy measures on bus routes. We made it possible to continuously change the routes for certain buses, which were temporarily implemented due to traffic congestion. Moreover, we added a service that posts the estimated arrival time to major stops while reflecting real-time traffic conditions in addition to the bus location and arrival time information through the global positioning system.

Keywords: bus information system (BIS); real-time; user-switching behavior; integrated transportation card; hybrid model; smart city

1. Introduction

Many local governments have started building smart transportation systems as part of smart city projects. In Seoul Metropolitan City, Korea, the research and development budget for land transportation was over $426 million in 2020. Moreover, corporates began to bid for transportation fund public offerings, and institutes investing in startups in the field of land, infrastructure, and transport emerged [1]. Smart transportation systems can be considered a part of the logistics and public transportation industry. In the case of logistics, the World Economic Forum predicts that, by 2030, the demand for logistics delivery in the world’s top 100 cities will increase by 80%, and the number of delivery vehicles will increase by more than 35% [2]. An increase in vehicles results in numerous traffic problems. Developed countries have made several efforts toward solving traffic problems in urban areas, such as establishment of intelligent transport systems, management of traffic demand, introduction of eco-friendly vehicles, and development of sustainable transportation policies [3]. Smart transportation is being established worldwide, to proactively respond to the Fourth Industrial Revolution and create new economic growth engines, along with efficient solutions to urban problems. Various new business concepts, such as self-driving, sharing economy, and smart energy, are being
implemented in smart transportation systems. Such systems create new businesses and provide a stepping stone for future growth.

Unlike physical logistics involving transportation of goods, public transportation is the most representative industry for human logistics. With more diverse and user-friendly new businesses emerging as part of smart cities, existing systems such as the transfer and bus information service systems are trying to upgrade themselves by providing real-time services using Internet of Things (IoT). Seoul’s public transportation system, consisting of buses and subways, covers more than 60% of the city’s total traffic. This accounts for a daily average of more than 10 million people. This reveals the importance of public transportation. The Seoul Metropolitan Government’s bus system caters to more than 50% of the average daily passengers using public transportation, making it the most important mode of transportation in Korea, along with the subway. More than 600 bus routes and about 36,000 bus stops in Seoul and its metropolitan area constitute an integral part of the citizens’ daily lives [4]. We cannot convert all of this into a smart transportation system only by considering passenger flow. To establish a smart transportation system, we need a comprehensive understanding of policies related to land use, structure of urban transportation networks, and exchanges with people. This study explores the environmental and psychological factors in implementing such a system.

With developments in information and communication technology (ICT), various smart technologies are now being incorporated into urban spaces and citizens’ daily lives. Various apps provide time and space information, and the collected information is run in big data systems, such as Cloud and Hadoop. These data can be used to obtain information that was difficult to ascertain in the past, such as route and transit time of passengers. Therefore, there is now a lot of research interest in the field of traffic information using big data. The Seoul Metropolitan Government has been storing traffic records through smart cards since 2004 [5] and provides real-time bus arrival information and other services, such as real-time route information, using the collected traffic data. The data contain both time and space information on the movement of people using public transportation, and they are used in research studies, software, and algorithms for traffic congestion, transfer, accessibility, and time-accuracy management [6]. Despite progress in many of these areas, the inherent complexity of the transportation environment makes the prediction of arrival times and accuracy of time calculations challenging. Moreover, apart from the long history and accumulated data of the smart transportation card, we proceed in this study with only recent traffic data. As shown in Figure 1, the use of smart cards by bus passengers during this period is considerably higher than the number of bus passengers. This is because, when passengers use smart cards, they do not use the usual plastic cards; they mostly use chips embedded in commercial credit cards, as shown in Figure 2. This eliminates the inconvenience of owning a separate plastic card; thus, we can assume that almost all Seoul Metropolitan citizens use smart cards.

![Figure 1. Number of bus passengers and use ratio of smart cards in Seoul Metropolitan City.](image-url)
This study explores the various environmental and psychological factors in implementing smart transportation. Using the real-time public application programming interface (API) and smart transportation card data provided by the Seoul Metropolitan Government, we study the factors affecting bus users’ behavior. This study first summarizes previous studies, including the development of a bus information system (BIS) for governing the Seoul Metropolitan Government’s public transportation system, various theories on bus information and signal tracking systems, signal system theory for real-time implementation, and theories on consumer psychology. Moreover, we organize the actual experimental environment through public API data and customer data of companies executing different parts of the project, such as the smart card software and equipment provider in Seoul Metropolitan City. We expect the conclusions of our experiment to help guide government policies, corporate decisions, and civil societies’ actions. We also expect to make practical contributions to the research community based on a strong theoretical foundation. Moreover, we review the existing literature, to compare the differences and commonalities of public transportation systems by continent to find policy alternatives that can be applied to transportation systems worldwide. We find that traffic congestion occurs during rush hour across Asia, the Americas, Africa, and Europe, and all countries are carrying out various policies, implementations, and studies to solve this traffic problem. Ultimately, we develop a real-time optimization algorithm by bus time that can be applied globally based on the commonalities of national traffic problems.

The remainder of this paper consists of four sections. Section 2 covers the research on traffic-related technical elements and time-space forecasting algorithms, bus policy governance, new business elements related to transportation for smart cities, and factors of the psychological transformation of traffic users. The number of real-time traffic search services for users through websites or mobile apps has been increasing of late. We examine the existing theories on the impact of this information on traffic consumers, change in user behavior, and individual conversion. Section 3 presents the hypotheses for this study and outlines the quantitative analysis, using actual user traffic data. In Section 4, we evaluate the academic contribution of this study, along with a summary. Finally, Section 5 concludes by revealing the implications, limitations, and directions for future research.

2. Background

2.1. Study on the BIS and Integrated Transportation Card System

The integrated public transportation system of the city of Seoul can identify which transportation modes are used by customers. It stores all personal information in the city’s transportation database. In other words, when citizens travel on a certain route, using the subway or a bus, their boarding and disembarking records are stored in the Seoul city database through their transportation card. This database includes integrated, spatiotemporal, and comprehensive information on citizens’ movements. As shown in Figure 3, the tagging algorithm of the transportation card and database conversion of the user signal are processed continuously until the starting, transfer, and destination points are reached, and the process is saved [7].

Figure 2. The form of transportation smart chips embedded in commercial credit cards.
This simple algorithm is important because it allows the government and public institutions, such as subway works, smart card service companies, private bus transportation companies, and telecommunication companies, to distribute transportation revenues. Moreover, this information guides important policy decisions, such as setting of transportation charges by each agency. Seoul’s public transportation system has an integrated fare system. Table 1 shows the system of bus combinations (wide area, circulation, branch line, trunk line, town, and late-night bus) divided into the five parts of Seoul.

Table 1. Seoul Metropolitan City’s bus operation system (2019).

| Bus Types                  | Purpose                                           | # of Routes | # of Operating | # of Companies |
|----------------------------|---------------------------------------------------|-------------|----------------|----------------|
| Branch-line (blue) bus     | Linkage between regions such as the suburbs, downtown and sub-centers | 148         | 3842           | 58             |
| Trunk-line (green) bus     | Linked transfer with blue bus or subway           | 221         | 3284           | 62             |
| Circulation (yellow) bus   | Branch lines in downtown and sub-centers          | 8           | 23             | 6              |
| Wide area (red) bus        | Linking the metropolitan area and the city center | 18          | 257            | 9              |
| Town (green) bus           | Auxiliary role of green bus                       | 276         | 1794           | 152            |
| Late-night bus             | Late-night traffic demand                         | 12          | 59             | 18             |

Seoul city buses operate in proportion to the integrated distance, and within 10 km, free transfers are possible at a basic fare; for distances beyond that, a fare of 100 won is charged every 5 km. When a passenger transfers from one vehicle to another via a transportation card, the card must be tagged every time the passenger gets on or off. Users can board a total of five modes of transportation for free.

Table 2 shows the structure of the transportation card data stored in the card database. The law on data information protection also applies to transportation card data; hence, the card’s serial number and customer information are encrypted to protect user information. The table shows that the card contains information on each bus stop, as well as the entire bus route of the Seoul city bus network. It also has several attribute values that provide important information for calculating variables, such as the transfer relationship between the bus and subway, time, distance, and time to stop. Through the 19 data attributes shown in Table 2, we can determine the traffic of more than 16 million users, as daily average, for the weekdays.

Table 2. Structure of transportation card data.

| ID Attribute               | Data Form | Digit | Description                                      |
|----------------------------|-----------|-------|-------------------------------------------------|
| PCARD_NO                   | Integer   | 10    | Personal card information                        |
| RIDE_DTIME                 | String    | 14    | First boarding time                             |
| RIDE_STA_ID                | String    | 7     | First boarding stop                             |
| TRANS_ID                   | String    | 3     | Transfer stop                                   |
| TRAF_FREQ                  | Integer   | 3     | Number of transfers                             |
| TRANSP_METHOD_CD           | String    | 3     | Transfer method (transportation type)            |
| BUS_ROUTE_ID               | String    | 8     | Classification of bus circulation section        |
Table 2. Cont.

| ID Attribute          | Data Form | Digit | Description               |
|-----------------------|-----------|-------|---------------------------|
| VEHC_ID               | String    | 9     | Bus number                |
| RUN_DEPART_DTIME      | String    | 14    | Total running time        |
| PCARD_USER_CLASS_CD   | String    | 2     | Personal classification   |
| RIDE_TRANSB_BIZR_ID   | String    | 9     | Bus type                  |
| ALIGHT_STA_ID         | String    | 7     | Drop-off station          |
| ALIGHT_DTIME          | String    | 15    | Drop-off time             |
| PASGR_NUM             | Integer   | 3     | Total number of users on board |
| RIDE_AMT              | Integer   | 10    | Transfer amount           |
| RIDE_PENAL_AMT        | Integer   | 10    | Untagged amount on board  |
| ALIGHT_AMT            | Integer   | 10    | Drop-off amount           |
| ALIGHT_PENAL_AMT      | Integer   | 10    | Untagged drop-off amount  |
| STAND_DT              | String    | 8     | Dwell time                |

One commonly used method for tagging transportation cards is radio frequency identification (RFID). Now, the near field communication (NFC) method through smartphone tagging is also being used; however, it can be broadly considered a part of the RFID method. RFID-sensing technology is a method of selecting data by identifying an object while exchanging wireless signals. The RFID sensor detects the type of object, using a tag planted on a small integrated chip on a plastic object, such as a card or other intermediate objects. The two main devices of RFID are RFID tags and RFID readers that utilize radio frequencies; both are very useful for transportation cards using wireless data exchange between the two devices [8]. RFID is divided into three types, according to frequency—low, high, and ultra-high—with differences in terms of recognition distance, influence on environment, and tag price. Therefore, as shown in Table 3, it is very important to determine an appropriate RFID frequency according to the purpose of use in the selected environment.

Table 3. Types of radio frequency identification (RFID) for transportation.

| Frequency        | Low Frequency (LF) | High Frequency (HF) | Ultra HF (UHF) | Microwave |
|------------------|--------------------|---------------------|---------------|-----------|
| (Hz)             | 125–134 kHz        | 13.56 MHz           | 433–960 MHz   | 2.45 GHz  |
| Recognition      | 60 cm or less      | 60 cm–1 m           | 50 cm–10 m    | 10 m or more |
| Operation method | Passive            | Passive             | Active/passive| Active/passive |

The NFC method was first developed to enable electronic payment services and highway toll payments that would replace the mobile and transportation cards supported by existing radio frequency devices [9]. File transfer is also possible through means such as Bluetooth; hence, there is a functionality for automatically performing a specific operation when one gets close to the NFC tag. In recent years, the NFC has been widely used in devices operated with door lock functions, alarm functions, and electronic cards for home IoT or office IoT technology. Google has launched a payment service, using the NFC function of Android smartphones, called Google Wallet. Moreover, through the Smart Lock feature on Android 5.0 Lollipop, users can move to the home screen by simply touching the registered NFC tag. A smartphone with a built-in NFC function acts as a tag when using the card mode; on touching the NFC tag or checking the balance of other transportation cards, it functions as a reader [10].

2.2. Studies on the Bus-Arrival-Time Estimation Algorithm

There are many quantitative studies on the implementation of bus routes and algorithms related to these changes in various industrial domains (psychology, business administration, sociology, and transportation engineering). Studies that presented a predictive model related to bus departures and arrivals mainly used time data as the dependent (result) variable. In other words, most studies have
suggested an important model for accurate prediction of bus travel times along with various causal variables: the local floating population, transfer modes, commute time, and traffic jams. Studies on bus routes that consider the section in which the user moves constitute a very important field in traffic research [11]. In recent years, many practical studies have predicted bus departure and arrival times to devise actual systems, using the weighted moving average method and Kalman filter model [12]. Research methods using the weighted moving average method view bus routes as moving in the same direction as one research unit, or combine the bus routes passing through a specific area [13].

Many studies have used the Kalman filter model for the duration time calculation by calculating the difference between the departure time and arrival time between stops [14]. A fusion algorithm has also been developed that compares the accuracy of each methodology by combining several models. A traffic study has estimated the arrival times of buses by considering the signal intersection as a variable, calculating the link travel time through the global positioning system (GPS) data and summing the signal waiting time [15]. Other research methodologies include weighted moving averages, genetic algorithms, artificial neural networks, linear regression models, and K-nearest neighbor [16].

For example, the travel time between bus stops is divided by day, hour, and section length, and bus arrival time is predicted by exponential smoothing. Moreover, factors affecting travel time are classified according to weather and time; bus arrival time is predicted by using genetic algorithms and the support vector machine [17]. The data fusion approach, which integrates multiple data sources, has been applied to various fields, such as military, marketing, and intelligent transportation systems. For example, the US National Household Travel Survey developed a trip purpose imputation method, using GPS data. This survey estimated an individual’s travel purpose, which was not directly observed, through objective personal GPS data [18]. Here, the concept of data fusion involves estimating the behavioral attributes of individuals through loading data, such as the passenger’s departure place, destination, and corresponding time [19]. Lastly, some studies have attempted to find a methodology in which the user’s traffic pattern, such as a bus transfer, can be predicted accurately to improve the error rate.

Many studies have used delay time at the bus stop as part of the algorithm for arrival time. In other words, the error in estimating bus arrival time is significantly reduced by using boarding and disembarking times, signal period, and the existence of dedicated bus lanes as variables [20]. This method separately calculates the predicted time for each stop to reflect the various environmental factors that cause errors. Many studies have also used a Markov chain process, which is a probabilistic approach, in the model to stop delays. Moreover, there is a quantitative analysis model, called the accessibility calculation model, that considers land location, using the aforementioned data attributes [21]. Previous studies used an index indicating the degree of accessibility of junctions connected to a given transportation network. Moreover, to accurately calculate the approach used in these studies, the relative utilization value of land at a specific point, appropriate land use, and future location plans are also quantified and analyzed [22]. With the recent developments in information technology, the accessibility calculation model has evolved into a very useful model for planning and analyzing traffic, location, land use, and spatial structure of a region [23]. To set the environment, it is necessary to apply the accessibility calculation model to a stable area, not to an area where the transportation network changes within a short time or where the surrounding environment of the domestic bus system frequently changes.

2.3. Studies on Introduction of Advanced Transportation Systems for Smart Cities

Smart cities can be defined as high-tech cities that provide services for users through various devices anytime, anywhere using ICT. Mark Weiser conceived of smart cities in 1988 by introducing the concept of ubiquitous computing into the environment and city [24]. This concept can be considered as a third space created by the fusion of electronic space and physical space that is centered on ICT, such as through sensing and networking. In other words, smart cities create new services by uniting various traditional sectors, such as medical, distribution, construction, and manufacturing, and they
are centered on ICT. Different scholars have defined the concept of smart cities in different ways, but it can be classified into three major categories: improving the quality of life, economic growth and job creation, and creation of a new industrial ecosystem. There is consensus that governance should be carried out in a manner in which smart cities place more emphasis on the role of social capital and relations in the urban development process. Andrea Caragliu, who established the European smart city strategy, argued that smart cities play an important role in achieving sustainable economic growth and high quality of life through investments in human and social capital and traditional and modern infrastructure. He also stated that wise management of natural resources and participation-oriented government are emphasized in such a process [25]. The British Standards Institution explained that smart cities create an effective environment for citizens to ensure a sustainable and inclusive future. In other words, the smart city concept is becoming more comprehensive and complex, and it goes beyond simply spreading urban habitability and activity, using ICT. It is not simply a pursuit of functional achievement, but a simultaneous pursuit of cooperation and system integration between each silo that constitutes a city. Moreover, it improves urban management efficiency and quality of life for citizens, as well as achieves sustainable growth, thus leading to overall progress [26].

The research on transportation services for smart cities can be divided into four categories. The first is related to autonomous driving. The concept of an autonomous shuttle bus refers to a small bus that combines the vehicle body design and system configuration for public transportation without a driver—a completely unmanned autonomous vehicle. The circular road bus, Navly, in Lyon, France, has been in service since 2016. It has a running speed of 20 km/h and operates from 7:00 a.m. to 7:00 p.m. Self-driving buses started operating in Las Vegas, USA, in 2017 and are still in use. The park shuttle in Rotterdam, the Netherlands, has been operating since 2008 and is the fastest self-driving bus running at a maximum speed of 32 km/h. Moreover, on the Wagening University campus in the Netherlands, a shuttle called WEpods operates within the university at a speed of 25 km/h [27]. The second category is related to smart parking. Smart parking services allow drivers to check and use empty parking spaces in real time through mobile and web interfaces, and they contribute toward reducing traffic congestion in the city. These interfaces provide convenient and efficient parking information with a parking search function (address, parking fee, and usage time) and real-time parking space information, contributing to reducing traffic inconvenience and saving energy. Smart parking services are being developed and introduced in many countries, including the USA, the UK, and Singapore. The market size of the parking management sector has grown at an average annual rate of 13.2% since 2012, reaching 195 million dollars in 2020 [28]. The third category is related to personal mobility services, which can also be considered as future transportation infrastructure for smart cities. Personal mobility refers to a small personal transportation mode that uses eco-friendly fuel and can accommodate one or two people; it includes single-person electric vehicles, electric bicycles, and mid/low-speed electric vehicles. Personal mobility devices can usually reach speeds of 10–20 km/h and are suitable for traveling distances too long to walk or too close to drive. In particular, these devices are attracting attention as a next-generation mode of transportation for the elderly. However, in many countries, it is not possible to use such devices on roadways, sidewalks, and parks, owing to road laws [29]. The laws relate to the smart crosswalk service, which is a pedestrian detection and car stop system to prevent traffic accidents near crosswalks. This service reduces social and economic losses by lowering the number of traffic accidents and death rates. Smart crosswalks provide warning broadcasts for signal violations or crossings through voice guidance of pedestrian signals, including a system that induces pedestrians and drivers to comply with traffic signals. Moreover, it is possible to automate various traffic control systems by utilizing traffic-related IoT sensor nodes. One big advantage is that these data can be used as basic data for reinforcing traffic-safety facilities [30].

As shown in Table 4, we compared the differences and commonalities of public transport systems by continent, and we sought policy alternatives.
Table 4. Comparison of transportation system policies by continent.

| Continent | Policy                                                                 | Commonality                                                                 |
|-----------|------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| Asia      | In China, there is a policy to prohibit the use of motorcycles during morning rush hours. | Worldwide, urban areas are experiencing severe rush-hour traffic congestion. |
| Africa    | Implemented public transportation efficiency policy through Bus Rapid Transit (BRT) systems. | A lot of time and material costs are incurred due to traffic congestion during rush hour. |
| Europe    | In order to reduce the use of private vehicles during rush hour, public transport fares have been reduced. | It is assumed that individual traffic users behave rationally between public and private traffic. |
| Latin America | Improving individual labor performance by improving public transportation access to work. | Efficient public transportation can improve productivity and competitiveness nationally. |
| North America | Provides effective transfer system information to improve neighborhood accessibility. |                                                                                     |

Many studies concluded that extreme rush-hour traffic congestion occurs worldwide—Asia, the Americas, Australia, Africa, and Europe. Moreover, we can see that many countries are carrying out various studies and policies for efficient traffic control during rush hour. China, a representative country in Asia, has a policy of tightly controlling motorcycle operation to prevent traffic congestion and safety accidents during rush hour [31]. Several African countries have implemented Bus Rapid Transit Systems since 2007, to establish an effective public transport system and alleviate traffic congestion [32]. To increase neighborhood accessibility, North American countries focused on distributing platforms that provide traffic information to users [33]. Many European countries promoted legislation to prevent traffic congestion by revitalizing public transportation fare-reduction policies on the premise that individuals’ choice of transportation is quite rational [34]. Several countries in Latin America have suggested that improving access to work through revitalization of public transport can improve work performance [35].

2.4. Studies on Individual Behavior of Public Transportation Users

However, the implementation of a smart transportation system depends on not only the ICT but also the people using the system and their characteristics, such as switching and conversion behaviors. Most studies related to switching behavior in the traffic field have focused on categorizing each characteristic based on the presence or absence of the driver’s traffic information. It has been found that the number of routes perceived by the driver and the number of route changes are significantly related. In particular, the desire for switching using traffic information before departure appears to be much stronger than the desire to switch during driving. This is the consensus among several studies. The driver’s route selection differs considerably depending on the cause of road delay, degree of delay, local situation, past travel experience, and driver’s social and economic status. The driver’s initial desire to change appears much higher than the desire to change during driving. However, the higher the lag time due to traffic, the more sensitive the desire to change the driving route becomes. The experience of using existing routes will also have a significant influence on future route selection [36]. As such, research on switching behaviors in many transportation fields is more focused on a driver’s conversion sensitivity than the transfer decisions of public transportation. Owing to recent developments in navigation systems, there is quite a lot of research on this. In the case of individual drivers, the longer the experience of the driver, the greater is the confidence that driver has on the navigation system [37]. Moreover, the greater the traffic volume and congestion, and the greater the number of alternative traffic routes, and the greater is the effect of providing traffic information and the resulting switching behavior. That is, the content of traffic information provided by the navigation system and traffic broadcasts, delivery rate of information, and reliability of these media are different for different individuals. It can also be seen that the effects thereof are all different, depending on the point of contact with the information [38].
Transfer resistance theory can be used to best predict switching behavior. In a recent study, transfer resistance was classified into temporal elements, such as transfer waiting time and transfer walking time, and non-temporal elements of transfer convenience and psychological burden due to congestion [39]. Most studies have converted these non-temporal elements into time and then algorithms. For example, one study estimates the transfer resistance value that occurs when using public transport in Seoul using smart card data. The results show that there is an average transfer resistance of 11.24 min. In other words, it was confirmed that there is a difference in the transfer-resistance value, depending on the departure and destination points [40]. Moreover, facilities, such as plane distance, number of steps, transfer time, and escalators, are factors of transfer resistance. Transfer convenience was analyzed by converting vertical travel distance to plane distance [41]. A similar study calculated the transfer resistance value for passengers, using the transfer ticket gates of Seoul subway and airport railways, and estimated the transfer resistance between the heterogeneous subways to be approximately 5.35 min per transfer [42]. In this way, it was confirmed that there is transfer resistance even when transferring to the same mode of transportation; a higher resistance is generated when transferring to another mode of transportation. Previous studies have found the optimal Pareto value by considering the number of transfers and minimum arrival time, rather than by searching for K multiple routes [43]. However, when a public transport passenger selects a route, a multi-route search is necessary because the route is determined according to not only the minimum arrival time but also personal preference (minimum transfer or minimum walking) from among other alternative routes.

There are also studies on the conversion of public transportation users in the transit section. In particular, because the increase in traffic volume in large cities is considered a serious social phenomenon, there is a lot of interest in information on public transport systems, such as subways and buses [44]. Transfer facilities and environments have now become more efficient and convenient, but a considerable amount of media for such transfer information has been developed, which significantly affects the switching behavior of public transportation users. With the diversification of personal characteristics and the development of public transportation information, many studies have tried to explain individual behavioral models using statistical models, whereas others examine the characteristics of transfers between each mode of transportation by analyzing the marginal effects on probability. For example, one study that conducted individual behavior analysis on commuter traffic predicted changes in the user’s transportation choice when a new transportation mode was introduced [45]. To analyze how transfer affects route selection, user behavior was analyzed by using stated preference, revealing preference data for subway users. Moreover, a model equation for estimating the value of facilities, such as resting time in transportation facilities, number of transfers, transfer time, and escalators, was also established [46]. This study identified the changes in bus routes and the user’s transfer type in case of traffic congestion, and it suggested measures for formulating transportation policies, such as the establishment of a transfer system for transportation facilities in the future.

3. Materials and Methods

3.1. Experimental Settings

We used the traffic information of company T, which provides all smart equipment and software for public transportation in Seoul Metropolitan City. We targeted all bus combinations operating on eight stops (between Station-A and Station-H in Seoul Metropolitan Area) during peak hours (weekday commute) and other times (weekdays excluding commuting time and weekends). This combination consists of 5 branch buses (green), 6 trunk buses (blue), 1 midnight bus, 1 local bus, and 17 wide-area buses (red). We finally selected 4 lines out of a total of 30 buses, considering routes with the same fare and dispatch interval. We utilized bus data for the bypass section, to prevent traffic congestion during commute time, applied temporarily for six months, from May 2018. During this period, we analyzed the consumer-switching behavior between buses that bypass the section and buses that stay on the
existing route. We constructed the same environment, assuming three strict prerequisites in addition to the same fare and route. First, we chose an area with no external traffic environment factors or population changes, such as construction of new apartments on the route or change in subway lines. Second, we excluded situations wherein the departure and arrival points are different by making two stops the same in front of and behind the eight bus stops in the experimental section. Finally, we included six subway stations in the eight bus stops, allowing us to investigate the various transit effects of users.

The first of the three major investigations that we propose is a comparison of a bus combination that bypasses a section to minimize road congestion in a specific period (RC), and a bus combination that stays on the existing route (RS). The second is the impact of the traffic condition over time. In other words, during the peak commute time on weekdays, a large number of users want to use the bus but at other times—daytime on weekdays or most weekends—the road is less crowded. We define and classify the congestion time zone (C) and non-congestion time zone (NC) through time division by the social environment. The last category relates to the customer’s transfer status or the number of transfers. The customer group can be classified into the first boarding group (F0) and the first transfer (F1) to the fourth transfer (F4) customer, using the transit record information on the transportation cards. In the context of this study, we can infer that, in the case of the morning commute, the first boarding group consists of users who live within a 10-minute-walk radius from the station. Conversely, we can also assume that, the greater the number of transfers, the further the user lives from the station. However, in the case of evening commute, it is not reasonable to infer the first boarding or the number of boarding, because the starting point is a workplace rather than a living space. Based on these simple assumptions, we detect changes in bus routes and traffic conditions, using real-time traffic information from the BIS. Through this, we can empirically consider the impact on traffic consumer-switching behavior.

As mentioned earlier, we took four bus lines as the subjects of this study. The tag data of traffic users are organized as shown in Table 5. Specifically, the two trunk buses (Y108 and Y243) that changed routes start operating at 4 a.m. and end at 11 p.m. However, the service intervals are 5 and 11 min, respectively. We made another combination for trunk buses (X121 and X601) that retained the existing route. The second combination also has the same start and end times as the first combination, and the intervals are different at 5 and 11 min, respectively. Therefore, we can assume that the spatiotemporal environment for the combination of these two groups is virtually the same condition. We also checked the data from January to April 2018, before the route was changed, and found that the trends of user increase and decrease are quite similar between the two combinations (less than 0.2%), as shown in Table 6. Through this, we confirmed that there is no difference in the bus users between the two groups.

Table 5. Period before route change (four months, January–April, in 2018).

| Target (Green) Bus | Routes Stayed (RS) | Routes Changed (RC) | Comparison of Growth Rate |
|--------------------|--------------------|---------------------|--------------------------|
| Total # of Bus Users (1) | 1,425,601 | 1,423,476 | −0.15% |
| Weekdays | 1,140,476 | 1,139,507 | - |
| Morning Commute Time (2) | 129,637 | 129,341 | - |
| Ratio (2/1) | (11.37%) | (11.35%) | - |
| # of First Boarder (3) | 18,570 | 19,177 | - |
| Ratio (3/2) | (14.3%) | (14.8%) | 3.51% |
| # of Transfer Users (4) | 111,067 | 110,164 | - |
| Ratio (4/2) | (85.7%) | (85.2%) | 0.59% |
Table 6. Period after route change (six months, May–October, in 2018).

| Target (Green) Bus | Routes Stayed (RS) | Routes Changed (RC) | Comparison of Growth Rate |
|-------------------|-------------------|---------------------|---------------------------|
| Total # of Bus Users (1) | 2,059,203 | 2,097,016 | 1.84% |
| Weekdays | 1,647,363 | 1,685,174 | - |
| Morning Commute Time (2) | 187,201 | 206,123 | - |
| Ratio (2/1) | (9.09%) | (9.83%) | 8.14% |
| # of First Boarder (3) | 26,021 | 29,682 | - |
| Ratio (3/2) | (13.9%) | (14.4%) | 3.60% |
| # of Transfer Users (4) | 161,180 | 176,441 | - |
| Ratio (4/2) | (86.1%) | (85.6%) | -0.58% |

Through Tables 5 and 6, we can compare the two bus combinations before and after the target period and discover two singularities. We find that the ratio of users in the morning commute (7:00 to 9:00 a.m.) to the total number of users increased significantly in bus groups that changed routes, compared with bus groups that stayed on the routes (–0.18% → 8.14%).

For this study, we acquired, processed, and analyzed traffic data in accordance with the Personal Information Protection Act (PIPA) of Korea, and we complied with relevant laws and regulations. Personal information is all information about a living individual. PIPA strictly prohibits the inclusion of personally identifiable information. Moreover, researchers can take measures to de-identify personal information, such as an individual’s name or social security number. The criterion for marking the identifier must also be observed. However, identifiers that are essential for data use purposes can be used for research after de-identification measures. De-identification measures we used for this study include string fake-name processing, numeric total processing, data deletion, data categorization, and data masking. We performed data preprocessing and obtained basic demographic information from the traffic data to establish our hypothesis and verify it. To do this, we performed a quantitative test, using Oracle Virtual-Box, a virtual machine, and SAS University Edition, an open source statistical package.

3.2. Operational Definitions and Preprocessing

As shown in Table 7, we selected the data required for this study from the vast amount of data received from the transportation database and created new variables for analysis in conjunction with data preprocessing. Moreover, we described the prerequisites of our experiments and the operational definitions of variables for the environmental settings.

Table 7. Data generation and preprocessing through variable classification.

| Variable Classification | Data Generation and Preprocessing |
|-------------------------|-----------------------------------|
| Card ID                 | Unique # allocation after personal information protection (by case) |
| Date                    | Target period classification (0: target period, 1: previous period) |
| Pick-up time            | Time slot classification (0: congestion time, 1: non-congestion time) |
| Pick-up stop            | Labeling all the stops (from Station-A to H & Station-a to d) |
| Drip-off stop           | Same as above |
| # of transfers          | Number of transfers when boarding (F0: first boarding, F1 to F4: transit #) |
| Bus number              | Route change classification (RS: X121, X601, RC: Y108, Y243) |
| Duration time           | Generating residence time on bus (for calculating the degree of congestion) |

To set up the experimental environment for our empirical study, we show the bus route change as in Figure 4.
As shown, the total route is 4.4 km long northward during rush hour, starting from Station-A to Station-H. Although there is a dedicated bus lane on this line, the section is constantly blocked, especially during rush hour. For this reason, the Seoul Metropolitan Government allowed two blue bus routes to be changed temporarily from May to October 2018. In other words, this route changes the section from Station-C and goes back to the original section, Station-F, through Station-a to Station-d. With this, the distance increases by about 2.1 km to a total of 6.5 km, and the number of stations increases by 2, to a total of 10. Note that even though the distance has increased, the time can be shortened by an average of 8 min during the morning rush hour.

We previously set the morning commute (from 7:00 to 9:00 a.m.) and evening commute (from 6:00 to 8:00 p.m.) as the congestion time for explaining the experimental environment, data held, and secondary processing data. This corresponds to 4 of the total 19 hours of bus operation time (from 4:00 a.m. to 11:00 p.m.), which is approximately 20% of the total time. We set the time outside the congestion time zone as the non-congestion time zone. As previously stated, the proportion of public transportation use during the morning commute is around 11–12%. Contrary to popular belief, we can see that the bias on rush hour is not high. We also paid attention to the number of transfers to understand the user’s commuting tendency. We checked whether there is a previous boarding record when the user’s boarding tag was used and set it as F0 to F4. F0 denotes that the user boarded the bus for the first time, and F1 to F4 denote the number of transfers by the user from 1 to 4 times, respectively. For this study, we divided it into two categories: first-time passengers (F0) and transit users (FN).

### 3.3. Demographic Information

As shown in Table 8, the total number of parameters of the cluster that maintains the bus route for 6 months and the cluster that changes the route is 4,156,219 cases. The number of cases in the two groups is roughly the same: 2,059,203 and 2,097,016, respectively. Among them, the number of first passengers (F0) is 290,348, which is 14.1% of the total route maintenance group, and transfers (F1–F4) are 1,768,855 cases and 85.9% of the total. Conversely, the number of first passengers (F0) in the group that changes the route is 296,010 cases, which is 14.2% of this group, and in the case of the transfer (F1–F4), it is 1,801,006 cases or 85.8% of the route change group. We find that, during this period, passengers who board first are less likely to prefer the changed route (only 0.11% difference). Moreover, although there is a difference in parameters, we find that the difference in the ratio increases as the number of transit passengers increases. We hypothesize that passengers who board first are
people who reside near the bus station, and they have accurate information on when and which bus combination would be efficient to use. Despite the inconvenience of having to go to two more stations during peak hours, they use a fixed route to reduce the transport time and use existing routes during off-peak times. We consider this behavior to be a kind of consumer-switching behavior that is apparent in nearby residents based on accurate information on bus routes.

Table 8. Classification by bus combination and number of transfers.

| # of Transfers | Routes Stayed (RS) | Routes Changed (RC) | Comparison of Growth Rate |
|----------------|--------------------|---------------------|---------------------------|
|                | # of Users | %    | # of Users | %    |                      |
| Total          | 2,059,203 | 100.0 | 2,097,016 | 100.0 | -                    |
| First boarding (F0) | 290,348 | 14.1  | 296,010   | 14.2  | 0.11                 |
| Total transferring (FN) | 1,768,855 | 85.9  | 1,801,006 | 85.8  | -0.02                |
| 1st transfer (F1) | 1,517,633 | 73.7  | 1,544,001 | 73.6  | -0.10                |
| 2nd transfer (F2) | 201,802  | 9.8   | 196,192   | 9.4   | -4.53                |
| 3rd transfer (F3) | 35,006   | 1.7   | 39,843    | 1.9   | 11.77                |
| Over 4th transfer (F4) | 14,414  | 0.7   | 20,970    | 1.0   | 42.86                |

Table 9 shows the change in the combination of two buses with congestion time. During rush hour on weekdays, the rate of increase in route-change bus combinations is clear. Conversely, it can be seen that the rate of increase is rather low during non-congestion hours on weekdays and weekends. This is because, as in the previous case, people prefer conventional-route bus combinations that save more time during off-peak hours. Moreover, in terms of overall bus use level, the congestion period on weekdays is about 25–26% of the average on weekdays. Compared with the number of daily users on weekdays, Saturdays are about 75% of the weekday average, and Sundays are about 50% of the weekday average.

Table 9. Classification by bus combination and traffic condition.

| Traffic Condition | Routes Stayed (RS) | Routes Changed (RC) | Comparison of Growth Rate |
|-------------------|--------------------|---------------------|---------------------------|
|                   | # of Users | %    | # of Users | %    |                      |
| Weekdays total    | 1,647,363 | 79.9  | 1,685,174 | 80.3  | 0.50                 |
| Weekdays average (1 of 5 days) | 329,473 | -     | 337,035   | -     | -                    |
| Morning rush hour | 187,201  | 8.9   | 206,123   | 9.8   | 9.18                 |
| Evening rush hour | 205,836  | 10.1  | 234,090   | 11.2  | 9.82                 |
| Time outside commute | 1,254,326 | 60.9  | 1,242,961 | 59.3  | -2.69                |
| Weekend total     | 410,840  | 20.1  | 408,842   | 19.7  | -2.03                |
| Saturday          | 247,104  | 12.0  | 243,473   | 11.8  | -1.69                |
| Sunday            | 163,736  | 8.1   | 165,569   | 7.9   | -2.53                |

3.4. Hypotheses

As seen in previous studies, the ease of access to information affects the user’s switching behavior. We hypothesize that, using a real-time BIS and transportation card data, along with the theoretical basis, policy changes such as bus-route changes affect the behavior of transportation users. Changes in the traffic condition at different times, especially traffic jams in the city center that inevitably occur during rush hour, can induce traffic users to act based on their information. We set this choice as a change in consumer behavior based on whether they bypass a particular section and track changes in actual user behavior through empirical data. In other words, we classify the bus combination as a route change and compare the bus combination that stays on the same route (RS) and that bypasses by going on another route with less congestion (RC). At this time, as seen through descriptive analysis,
we assume that the behavior of public transportation users will change depending on the time, and we present the first hypothesis as follows.

**H1:** The group that does not change routes (RS) and the group that changes the routes (RC) have different effects on user increase or decrease, depending on the road environment classification (C: congestion/NC: non-congestion) according to time.

We hypothesize that there will be differences in switching behavior of users between passengers that board first and transit passengers, in addition to the division by time. In other words, we assume that the first boarding passengers are generally those who reside nearby, whereas the further the user’s residence is from the stop, the higher is the number of transfers. We hypothesize that people living near a bus stop (more first boarding passengers than transfers) are more likely to be exposed to information. It can be assumed that, in the case of heavy traffic congestion during rush hour, they will actively use the adjusted route, using their information. Conversely, during non-congestion times, they are more likely to stay on the existing route rather than the adjusted longer route. Therefore, our second hypothesis is as follows.

**H2:** The group that does not change routes (RS) and the group that changes the routes (RC) have different effects on user increase or decrease, depending on the transfer status (F0: first boarding/ FN: transferring).

To show that the effect of route adjustment varies with time for congestion and the number of transfers, we conduct an analysis of covariance (ANCOVA). Like regression analysis, ANCOVA allows us to see how several independent variables act on the dependent variable. Moreover, ANCOVA eliminates all effects of covariates that are not variables we want to identify. The formulation of the ANCOVA is as follows [47]:

$$\sigma_{XY} = \sum \frac{(X_i - \mu_X)(Y_i - \mu_Y)}{N(-1)}$$

The third hypothesis is the moderating effect of the previous two hypotheses. In other words, we create a cross table of existing/adjusted routes by mixing the time division of congestion/non-congestion and first boarding/transfer division. With this hybrid model, we create a combination of four cases and examine the effects of these combinations on traffic use. The four combinations are (1) congestion time–first boarding passengers, (2) congestion time–transfer users, (3) non-congestion time–first boarding passengers, and (4) non-congestion time–transfer users. To clarify the meaning of the third hypothesis for first boarding and transfer, we remove the evening commute data, where the significance of transfer is less directly related with the distance of residence. We also remove weekend data when the concept of commute time is ambiguous for fairness and convenience of analysis. In other words, we construct a new dataset, using only the morning time on weekdays and other weekday time data. The basic data for this new dataset are shown in Table 10.

**Table 10.** Hybrid model for traffic condition and number of transfers

| Division              | Routes Stayed (RS) | Routes Changed (RC) | Comparison of Growth Rate |
|-----------------------|--------------------|---------------------|---------------------------|
|                       | # of Users         | %                  | # of Users     | %    |                     |
| Grand total           | 1,441,527          | 100.0              | 1,449,084     | 100.0 | -                  |
| Congesting time (total)| 187,201            | 13.0               | 206,123       | 14.2  | 9.53               |
| First boarding (F0)   | 26,021             | 1.8                | 29,682        | 2.1   | 13.47              |
| Transferring (FN)     | 161,180            | 11.2               | 176,441       | 12.2  | 8.90               |
| Non-congesting time (total)| 1,254,326   | 87.1               | 1,242,961     | 85.8  | -1.42              |
| First boarding (F0)   | 201,946            | 13.9               | 191,416       | 13.2  | -5.71              |
| Transferring (FN)     | 1,052,380          | 73.0               | 1,051,545     | 72.6  | -0.60              |

As Table 10 shows, when only observing the difference in bus combination use during congestion/non-congestion times, the rate of change is 9.53% and –1.42%, respectively. However, with the addition of variables of first boarding passengers and transit passengers, the gap in the increment
changes markedly. Through these changes, we find that mixed variables enable a more sensitive and multifaceted analysis of the user’s switching behavior. We propose a mixed combination hypothesis through this descriptive analysis.

H3a: In the case of congestion time (C)–first boarding passengers (F0) combination, there is no difference in the increase or decrease in the use of the maintenance bus group (RS) and change bus group (RC).

H3b: In the case of congestion time (C)–transfer user (FN) combination, there is no difference in the increase or decrease in the use of the maintenance bus group (RS) and change bus group (RC).

H3c: In the case of non-congestion time (NC)–first boarding passengers (F0) combination, there is no difference in the increase or decrease in the use of the maintenance bus group (RS) and change bus group (RC).

H3d: In the case of non-congestion time (NC)–transfer user (FN) combination, there is no difference in the increase or decrease in the use of the maintenance bus group (RS) and change bus group (RC).

We divided the third hypothesis into four mixed regions. To test these hypotheses, we ran a \( t \)-test, comparing the differences between the two factors of each combination.

4. Results

4.1. Results of the First-Round Analysis

We classified Hypothesis 1 as a route change and compared the bus combination that stays on the existing route (RS) and the bus group that bypasses the route by going on another route with less congestion (RC). We hypothesize that the effects of the two bus combinations on the use of public transportation users differ depending on the traffic condition (C/NC) at different times. To test this, we perform the ANCOVA test, as shown in Table 11.

| Degree of Freedom | Sum of Square | Mean Square | F-Value | p-Value |
|-------------------|--------------|-------------|---------|---------|
| Model             | 51,461.32556 | 25,730.66278| 144.560 | <0.0001 |
| Error             | 2,059,201    | 366,524,324.12485 | -       | -       |
| Total             | 2,059,203    | 366,575,785.45044 | -       | -       |
| Parameter         | DF | Estimate | SE | T for H0 | p-value |
| Intercept         | 1  | 73.210000 | 0.44720 | 163.708 | <0.0001 |
| C                 | 1  | -47.740000 | 0.64520 | -73.993 | <0.0001 |
| NC                | 1  | -43.680000 | 0.75470 | -57.877 | <0.0001 |

C = congestion time zone; NC = non-congestion time zone (NC).

Hypothesis 1 has a secured model fit according to the F-test results (F = 144.560). With these results, we confirm that, when there is congestion, route adjustment has a significant effect on bus users. Similarly, we confirm that, even when there is non-congestion, route adjustment has a significant effect on bus users. As shown in the descriptive statistics, the overall users increase from 3.51% to 8.14% during weekday commutes, and the route-change group has a higher user preference than the route-maintenance group. Moreover, we confirm that there is a significant negative effect (−2.03%) during the non-congestion time and route-change period.

Through Hypothesis 2, we check for user characteristics according to the number of bus transfers, as well as the congestion/non-congestion period, depending on route change. We hypothesize that there will be differences in the user’s switching behavior between the first boarding passengers and transit passengers. Through this, we hypothesize that the two bus groups classified by route change will have different effects on public transportation users, depending on the number of transfers (F0/FN). The verification results of the covariance analysis are shown in Table 12.
Hypothesis 2 confirms the model fit according to the F-test results ($F = 134.728$). From these results, we can conclude that the difference between first boarding passengers and transit passengers has no effect on bus use. This shows a significant difference, with the difference between the two groups being 13.47% according to the descriptive statistics. However, this is because the descriptive statistics were surveyed only for the first boarding passengers during rush hour. In other words, because all users, including the number of users during non-congestion hours (~5.71%), were sampled, the effect was offset, which led to the confirmation of the second hypothesis.

4.2. Robustness Check as a Second-Round Analysis

We wanted to test the two hypotheses tested earlier more specifically because we believe that the second hypothesis resulted in different results as the effect on the transfer was offset. To this end, we verify the effects before offsetting using a subdivided model and establish four hybrid hypotheses. The parameter in Hypothesis 3a is a total of 55,703 first boarding passengers during rush hour. Among them, 26,021 users selected the route maintenance group, and 29,682 users selected the route change group. The target of the dependent variable is the increase or decrease in users of the two bus combinations belonging to this group. We compare each ratio for this and test the significance of the difference in the increase rate of users based on route change.

As shown in Table 13, Hypothesis 3a satisfies the assumption of equal variance with the F-test ($F = 1.14$). We refer to the pooled $t$-test, and the result rejects the hypothesis ($t < 0.001$). In other words, the combination of the two buses classified by route change shows a difference in the number of users during rush hour for first boarding passengers. In this case, assuming that there is congestion, the positive effect of RC preference is more evident than the offset effect, and the number of users increases. The parameters in Hypothesis 3b are congestion time and transit passengers, with a total of 337,621 cases. Of these, RS users are 161,180, and RC users are 176,441. Like H3a, the dependent variable is the increase or decrease in users of the two bus combinations. We compare each ratio for this and test the significance of the difference in the increase rate of users based on route change.

As shown in Table 14, Hypothesis 3b satisfies the assumption of equal variance with the F-test ($F = 1.09$). We refer to the pooled $t$-test and adopt the hypothesis for the test result ($t = -0.2214$). In other words, there is no statistically significant increase or decrease in usage between the congestion
time-transfer users of the two bus combinations classified by route change. The total parameters in Hypothesis 3c are 393,362 passengers, which is the number of first boarding passengers during the non-congestion time. Of these, 201,946 are RS users and 191,416 are RC users. As with the aforementioned hypothesis, the dependent variable is the increase or decrease in users of the two bus combinations. We compare each ratio for this and test the significance of the difference in the increase rate of users by route change.

Table 14. Two-sample t-test for H3b.

| Variance      | Degree of Freedom | t-Value  | p-Value |
|---------------|------------------|----------|---------|
| Pooled Equal  | 337,619          | −0.2214  | 0.142   |
| Satterthwaite Unequal | 336,674.0148    | −0.2213  | 0.142   |

As shown in Table 15, Hypothesis 3c satisfies the assumption of equal variance with the F-test (F = 0.95). We refer to the pooled t-test, and the test result rejects the hypothesis (t = −3.1145). In other words, there is a significant difference in the increase and decrease in the use of first boarding passengers during non-congestion time between the two bus combinations classified by route change. The t-value is negative, which means that the first boarding users increase the use of RS during non-congestion times. With these results, we again prove consumer-switching behavior by using information of nearby residents. Lastly, the total number of parameters in Hypothesis 3d is 2,103,925, which includes non-congestion time zone and transit users. Among these, 1,052,380 customers of the RS bus group maintain their route, and 1,051,545 of RC users change routes. As with the aforementioned hypothesis, the dependent variable is the increase or decrease in users of the two bus combinations. We compare each ratio for this and test the significance of the difference in the increase rate of users by route change.

Table 15. Two-sample t-test for H3c.

| Variance      | Degree of Freedom | t-Value  | p-Value |
|---------------|------------------|----------|---------|
| Pooled Equal  | 393,360          | −3.1145  | <0.0001 |
| Satterthwaite Unequal | 392,259.0707    | −3.1127  | <0.0001 |

As shown in Table 16, Hypothesis 3d satisfies the assumption of equal variance with the F-test (F = 1.00). We refer to the pooled t-test, and the hypothesis is adopted because of the test result (t = 0.5524). In other words, there is no difference in the use of non-congestion time-transfer users when comparing the two route changes.

Table 16. Two-sample t-test for H3d.

| Variance      | Degree of Freedom | t-Value  | p-Value |
|---------------|------------------|----------|---------|
| Pooled Equal  | 2,103,925        | 0.5524   | 0.241   |
| Satterthwaite Unequal | 2,098,031.5369  | 0.5521   | 0.240   |

| Variance      | Degree of Freedom | t-Value  | p-Value |
|---------------|------------------|----------|---------|
| Folded F 191,416 | 201,946          | 0.95     | 0.7443  |

| Variance      | Degree of Freedom | t-Value  | p-Value |
|---------------|------------------|----------|---------|
| Folded F 1,051,545 | 1,052,380        | 1.00     | 0.5829  |
5. Conclusions

In this study, we use the BIS and transportation card data to analyze the behavior of users on certain bus route sections. We empirically verified that the switching behavior of bus passengers is relevant to the information held by them. The practical implications of this study are as follows. We find that real-time BIS allows individuals to obtain real-time traffic information in the cloud environment. We also find that individuals can make decisions, using this real-time information. In particular, this study empirically verifies the effects of changing routes at certain times and sections from the Seoul Metropolitan Transportation Smart Card information, as well as public API information. Moreover, one of the three bifurcations presented by this study demonstrates the significant changes in routes that can be actually used by people during rush hour and non-congestion time. In other words, we find that, during rush hour, many people choose a less time-consuming route, even if there are more stops and the route is longer. We believe that these results can provide a theoretical basis for formulating policy measures on bus routes in future urban spaces. This study also has significant academic contributions. We empirically verify the relationship between transit switching behavior and information of people living close to bus stops. There are many optimization algorithms used by previous studies in the field of transportation, but there are also academic disciplines that study people’s behavioral patterns. This study confirms that switching behavior based on traffic information can change significantly depending on the user’s situation. In particular, passengers living near bus stops can quickly and easily obtain information about surrounding bus routes and change their behavior. Conversely, this switching behavior also appears during non-congestion time. Passengers use existing routes on shorter distances rather than the changed routes and enjoy the advantages of an uncrowded traffic condition. Third and lastly, we made use of the theoretical basis of this study to ensure that practically meaningful policies can be implemented through actual projects. We made it possible to continuously make route changes for the target buses, which were temporarily implemented due to traffic congestion (Y108 and Y243), through consultations with private bus companies and the government. Moreover, we added a service that posts the estimated arrival time to the major stops reflecting real-time traffic conditions in addition to the bus location and arrival time information through GPS, a service currently provided through the real-time aggregate status board of major bus stations.

This study has many limitations. First, six months is a very limited time to check the effectiveness of bus route adjustments. Moreover, it is difficult to make generalizations based on the results because they are limited to specific sections. In other words, it can be assumed that people’s behavior in areas that are not connected to the subway can change. Thus, the type of transfer should be studied in more detail. Second, owing to the nature of transportation card data, certain personal information could be collected. We expect that this specific information will lead to more diverse policy implications. Third, the dependent variable of this study is the increase and decrease in bus use. However, if we investigate the operating profits of bus companies or the bifurcated revenue structure of buses and subways, the actual practical contribution will be even greater. Fourth, we were not able to verify the electronic signals presented in the preceding study. While preparing this research, we wanted to see the overall flow of data, including a signal-based system for collecting public API information and a cloud system for storing it, but there was a lack of adequate data and time to study them. Based on these limitations, we make the following suggestions for future research. First, the regional limits should be broadened so that we can look at the number of cases in different regions and transit sections. We believe that we can explore a variety of policy alternatives and create new opportunities, such as personal information acquisition. Second, we believe that we will be able to produce more valuable research results by preparing various dependent variables in cooperation with various bus companies and other institutions, as well as existing independent variables. Third, we expect to be able to conduct various studies that can promote user convenience beyond issues such as privacy protection in academic research on individual movements. We believe that such data can also help us establish a transportation system that smart cities should have. Lastly, we hope to conduct empirical studies considering the larger picture, including links to cloud systems, in the future.
Author Contributions: Conceptualization, Z.L.; data curation, S.H. and J.K.; formal analysis, J.K.; investigation, S.H.; methodology, J.K.; project administration, Z.L.; software, S.H.; writing—original draft, Z.L.; writing—review and editing, J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chang, K.J.; Yoo, S.M. A Study on the Evaluation of Vehicle Operation Prior to Autonomous Vehicle Technology Deployment in Urban Area. Korea Contents Assoc. 2019, 19, 452–459.
2. Tang, C.S.; Veelenturf, L.P. The strategic role of logistics in the industry 4.0 era. Transp. Res. Part E Logist. Transp. Rev. 2019, 129, 1–11. [CrossRef]
3. Wan, C.; Yang, Z.; Zhang, D.; Yan, X.; Fan, S. Resilience in transportation systems: A systematic review and future directions. Transp. Rev. 2018, 38, 479–498. [CrossRef]
4. Mee-Young, L. A Study on Selected Station Analysis of AFC-Based Integrated Transit Network—Focused on Subway Transfer Stations in Seoul Metropolitan Area. Intel. Transp. Syst. 2018, 21, 1155–1161. [CrossRef]
5. Lee, M.-Y. Estimating Internal Transfer Trips Considering Subway Express Line—Focusing on Smart Card Data Based Network. Korean Soc. Civ. Eng. 2019, 38, 613–621.
6. Li, J.; Lv, Y.; Ma, J.; Ouyang, Q. Methodology for extracting potential customized bus routes based on bus smart card data. Energies 2018, 11, 2224. [CrossRef]
7. Kim, K.; Lee, I.; Min, J. Bus OD Matrix Estimation Using Smart Card Data. Korean Soc. Railw. 2018, 21, 1155–1161. [CrossRef]
8. Ding, K.; Jiang, P.; Su, S. RFID-enabled social manufacturing system for inter-enterprise monitoring and dispatching of integrated production and transportation tasks. Robot. Comput. Integr. Manuf. 2018, 49, 120–133. [CrossRef]
9. Zhao, H.; Anong, S.T.; Zhang, L. Understanding the impact of financial incentives on NFC mobile payment adoption. Int. J. Bank Mark. 2019, 37, 1296–1312. [CrossRef]
10. Xu, G.; Li, M.; Luo, L.; Chen, C.H.; Huang, G.Q. Cloud-based fleet management for prefabrication transportation. Entrep. Inf. Syst. 2019, 13, 87–106. [CrossRef]
11. Malokin, A.; Circella, G.; Mokhtarian, P.L. How do activities conducted while commuting influence mode choice? Using revealed preference models to inform public transportation advantage and autonomous vehicle scenarios. Transp. Res. Part A Policy Pract. 2019, 124, 82–114. [CrossRef]
12. Qu, S.R.; Li, J.; Shu, Y. Accurate vehicle location and tracking algorithms based on improved kernelized correlation motion model and Kalman filter in intelligent transport surveillance system. J. Ambient Intell. Humaniz. Comput. 2019, 1–10. [CrossRef]
13. Jia, R.; Jiang, P.; Liu, L.; Cui, L.; Shi, Y. Data driven congestion trends prediction of urban transportation. IEEE Internet Things J. 2018, 5, 581–591. [CrossRef]
14. Allen, J.; Farber, S. How time-use and transportation barriers limit on-campus participation of university students. Travel Behav. Soc. 2018, 13, 174–182. [CrossRef]
15. Yang, X.; Stewart, K.; Tang, L.; Xie, Z.; Li, Q. A review of GPS trajectories classification based on transportation mode. Sensors 2018, 18, 3741. [CrossRef] [PubMed]
16. Zhang, S.; Yao, Y.; Hu, J.; Zhao, Y.; Li, S.; Hu, J. Deep autoencoder neural networks for short-term traffic congestion prediction of transportation networks. Sensors 2019, 19, 2229. [CrossRef]
17. Lu, S.; Liu, Y. Evaluation system for the sustainable development of urban transportation and ecological environment based on SVM. J. Intell. Fuzzy Syst. 2018, 34, 831–838. [CrossRef]
18. Shen, L.; Stopher, P.R. A process for trip purpose imputation from Global Positioning System data. Transp. Res. Part C Emerg. Technol. 2013, 36, 261–267. [CrossRef]
19. Kusakabe, T.; Asakura, Y. Behavioural data mining of transit smart card data: A data fusion approach. Transp. Res. Part C Emerg. Technol. 2014, 46, 179–191. [CrossRef]
20. Fazayeli, S.; Eydi, A.; Kamalabadi, I.N. Location-routing problem in multimodal transportation network with time windows and fuzzy demands: Presenting a two-part genetic algorithm. Comput. Ind. Eng. 2018, 119, 233–246. [CrossRef]
21. Borucka, A. Three-state Markov model of using transport means. Bus. Logist. Mod. Manag. 2018, 3–19.
22. Hawkins, J.; Nurul Habib, K. Integrated models of land use and transportation for the autonomous vehicle revolution. *Transp. Rev.* 2019, 39, 66–83. [CrossRef]

23. Imeri, A.; Khadraoui, D. The security and traceability of shared information in the process of transportation of dangerous goods. In Proceedings of the 2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Paris, France, 26–28 February 2018; pp. 1–5.

24. Parulekar, W.R. Ubiquitous Computing Architecture, Applications and Challenges. *Int. J. Future Revolut. Comput. Sci. Commun. Eng.* 2018, 4, 694–696.

25. Caragliu, A.; Del Bo, C.F. Smart innovative cities: The impact of Smart City policies on urban innovation. *Technol. Forecast. Soc. Chang.* 2019, 142, 373–383. [CrossRef]

26. Caird, S.P.; Hallett, S.H. Towards evaluation design for smart city development. *J. Urban Des.* 2019, 24, 188–209. [CrossRef]

27. Arena, F.; Ticali, D. The development of autonomous driving vehicles in tomorrow’s smart cities mobility. *AIP Conf. Proc.* 2018, 2040, 140007.

28. Alam, M.; Moroni, D.; Pieri, G.; Tampucci, M.; Gomes, M.; Fonseca, J.; Leone, G.R. Real-time smart parking systems integration in distributed ITS for smart cities. *J. Adv. Transp.* 2018. [CrossRef]

29. Behrendt, F. Cycling the smart and sustainable city: Analyzing EC policy documents on internet of things, mobility and transport, and smart cities. *Sustainability* 2019, 11, 763. [CrossRef]

30. Branquinho, J.; Senna, C.; Zuquete, A. An Efficient and Secure Alert System for VANETs to Improve Crosswalks’ Security in Smart Cities. *Sensors* 2020, 20, 2473. [CrossRef]

31. Guo, Y.; Wang, J.; Peeta, S.; Anastasopoulos, P.C. Personal and societal impacts of motorcycle ban policy on motorcyclists’ home-to-work morning commute in China. *Travel Behav. Soc.* 2020, 19, 137–150. [CrossRef]

32. Walters, J. Overview of public transport policy developments in South Africa. *Res. Transp. Econ.* 2013, 39, 34–45. [CrossRef]

33. Guo, Y.; Peeta, S. Impacts of personalized accessibility information on residential location choice and travel behavior. *Travel Behav. Soc.* 2020, 19, 99–111. [CrossRef]

34. Bresson, G.; Dárgay, J.; Madre, J.L.; Pirotte, A. The main determinants of the demand for public transport: A comparative analysis of England and France using shrinkage estimators. *Transp. Res. Part A Policy Pract.* 2003, 37, 605–627. [CrossRef]

35. Hernandez, D.; Hansz, M.; Massobrio, R. Job accessibility through public transport and unemployment in Latin America: The case of Montevideo (Uruguay). *J. Transp. Geogr.* 2020, 85, 102742. [CrossRef]

36. Kang, A.S.; Jayaraman, K.; Soh, K.L.; Wong, W.P. Convenience, flexible service, and commute impedance as the predictors of drivers’ intention to switch and behavioral readiness to use public transport. *Transp. Res. Part F Traffic Psychol. Behav.* 2019, 62, 505–519. [CrossRef]

37. Cheung, E.; Bera, A.; Manocha, D. Efficient and safe vehicle navigation based on driver behavior classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1024–1031.

38. Ghaleb, F.A.; Zainal, A.; Rassam, M.A.; Saeed, F. Driving-situation-aware adaptive broadcasting rate scheme for vehicular ad hoc network. *J. Intell. Fuzzy Syst.* 2018, 35, 423–438. [CrossRef]

39. Dong, B.; Gu, Z.; Qiu, Q.; Liu, Y.; Ding, W.; Xing, F.; Hong, S. Electrochemical feature for chloride ion transportation in fly ash blended cementitious materials. *Constr. Build. Mater.* 2018, 161, 577–586. [CrossRef]

40. Inwoo, J.; Hyunwoo, N.; Chulmin, J. A schedule-based Public Transit Routing Algorithm for Finding K-shortest Paths Considering Transfer Penalties. *Intell. Transp. Syst.* 2018, 17, 72–86.

41. De Santis, R.; Montanari, R.; Vignali, G.; Bottani, E. An adapted ant colony optimization algorithm for the minimization of the travel distance of pickers in manual warehouses. *Eur. J. Oper. Res.* 2018, 267, 120–137. [CrossRef]

42. Sujeong, Y.; Jang-Ho, L. A Study on Route Choice Modeling in Metropolitan Urban Rail Network Considering Transfer Penalty and Angular Cost. *Korean Soc. Urban Railw.* 2018, 6, 120–137.

43. Kujala, R.; Weckström, C.; Mladenović, M.N.; Saramäki, J. Travel times and transfers in public transport: Comprehensive accessibility analysis based on Pareto-optimal journeys. *Comput. Environ. Urban Syst.* 2018, 67, 41–54. [CrossRef]

44. Han, Y.; Li, W.; Wei, S.; Zhang, T. Research on passenger’s travel mode choice behavior waiting at bus station based on sem-logit integration model. *Sustainability* 2018, 10, 996. [CrossRef]
45. Lupták, V.; Droždziel, P.; Stopka, O.; Stopková, M.; Rybicka, I. Approach methodology for comprehensive assessing the public passenger transport timetable performances at a regional scale. *Sustainability* 2019, 11, 3532. [CrossRef]

46. Belgiawan, P.F.; Dubernet, I.; Schmid, B.; Axhausen, K. Context-dependent models (CRRM, MuRRM, PRRM, RAM) versus a context-free model (MNL) in transportation studies: Comprehensive comparisons for Swiss and German SP and RP data sets. *Transp. A Transp. Sci.* 2019, 15, 1487–1521. [CrossRef]

47. Mangalathu, S.; Jeon, J.S.; DesRoches, R.; Padgett, J. Application of Bayesian methods to probabilistic seismic demand analyses of concrete box-girder bridges. *Geotech. Struct. Eng. Congr.* 2016, 1367–1379.

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).