

Can Trip Planner Log Files Analysis Help in Transit Service Planning?

*Martin Trépanier, Robert Chapleau, Bruno Allard
Ecole Polytechnique de Montreal*

Abstract

Transit trip planners are now found on most transit authority websites. This feature gives transit users a full itinerary from a point of origin to a destination. The web server on which the trip planner is installed usually stores usage logs on a daily basis. Log files contain data on origins, destinations, calculated paths, and other website entries. The main purpose of this article is to determine whether the analysis of trip planner log files can help to improve transit service by providing better knowledge on transit users. A website-oriented analysis and a transit-oriented analysis based on four years of observations on the Montreal Transit Commission website are presented. Results show that, even though not all transit users have access to the Internet or use the planner regularly, log files can be useful for identifying new locations to be accessed by a transit system, for better understanding user behaviors, and for guiding updates of the geographic information system (GIS) and the trip planner itself.

Introduction

Today, most transit authorities provide information to their users through Internet websites, which usually contain detailed information on routes, schedules, fares, and paratransit services. Some of them also provide a “trip planner” (Federal Transit Administration 2002). Trip planners are advanced web tools that create transit itineraries based on user input of a specific origin and destination pair. Of

course, these planners serve a dual function: In addition to providing information to users, they also provide the organization with information about users' mobility behaviors and trip patterns. Moreover, by examining the logs of the map reference search features of these websites, the subjacent geographical information system (GIS) can be enriched.

This article underlines the potential for the analysis and use of trip planners as tools in transit planning. A background overview of website usability and transit websites also describes the fundamentals that form the basis of this project's websites in order to support planning: the Totally Disaggregate Approach and Transportation Object-Oriented Modeling. Then, the article describes the trip planning process and the log file structure offered on the websites studied here. Some useful website usage statistics are provided (website-oriented analysis), followed by an in-depth examination of the space-time characteristics of the user's website declarations and possible resections with origin-destination survey data (transit-oriented analysis). There is also a section on possible bias and limitations of this research. Finally, the principal results of this research are given.

Background

Website Usability and Logging

It is usually difficult to adequately track user behaviors on a website because of the different "surfing paths" that can be followed by Internet users, depending on their needs (Piroli and Pitkow 1999). Some tools are designed to evaluate usability for general-purpose websites (Ivory and Hearst 2001), but a few authors have tried to use website logs to improve the website functions themselves, and not just the interface. Tan and Kumar (2002) have studied web navigational patterns in the case of web robots for a better knowledge of website usage. Smith and Ng (2003) have worked on a web page clustering method to evaluate the usability of websites. Usability refers to ease of use, access to website functions, and the extent to which user needs are met. Murphy et al. (2001) report that website log files can reveal valuable marketing information on customer habits and preferences. All these works are valuable for traditional "commercial" websites, but are not suited for application-specific websites like trip planners. Golob and Regan (2001) believe that in the transportation field, the Internet, with the help of adequate websites, can be a powerful survey tool for tracking people trips.

Transit Trip Planners

In public transportation, websites are seen as a natural extension of normal service activities and public relations tools. Efforts have been made to describe the elements that could be used to feed a transit information website for transit users. Lee et al. (1999) presented a web-based bus information system (WBI) implementation with data stored in a Microsoft Access database. Peng and Huang (2000) have also used Access to store route and transit networks, bus stops and time-point data, and their database is linked to a GIS and to network servers. Karimi et al. (2004) have developed a web interface to help the user to mount his or her own transit trip along with information. Maclean and Daily (2002) have made significant work about the measurement of the utility of a real-time information system using log file analysis, but this work does not refer to website usage.

The Federal Transit Administration (2002) reports that about 20 percent of American transit websites have a trip planner, but it does not rate these in terms of features. The Montreal Transit Commission (MTC) and the Laval Transit Commission (LTC) both have implemented transit trip planners on their websites (<http://www.stm.info> and <http://www.stl.laval.qc.ca>). At first glance, few of these trip planners have a true planning function such as those in Montreal and Laval, where users can specify any location on the territory (address, intersection, trip generator, station, and maps) and where they are not limited to lists of stations. In recent work (Trépanier et al. 2002), we have shown that a trip planner must rely on a comprehensive geographic information system (GIS) and the Totally Disaggregate Approach (TDA) to meet the needs of transit users. GIS are extremely useful and powerful for transportation modeling and operation (Arampatzis et al. 2004).

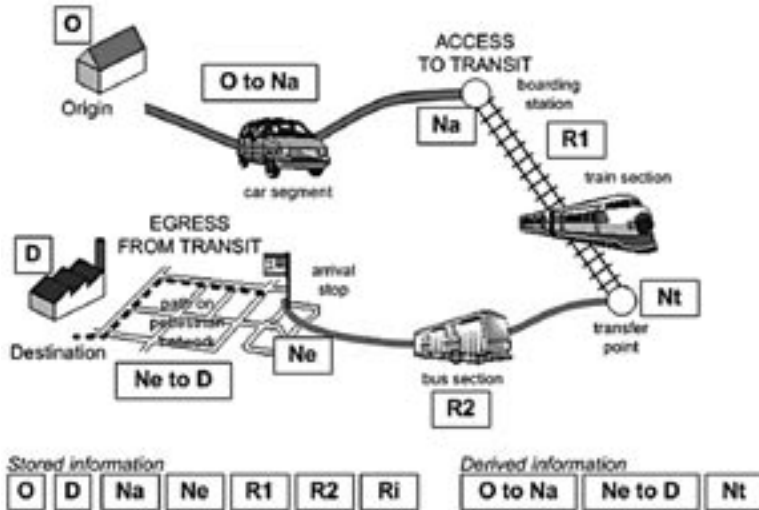
Totally Disaggregate Approach

The Totally Disaggregate Approach (TDA) is a transportation planning methodology used in the Montreal region since the beginning of the 1980s. The main principles of TDA are (Chapleau 1986):

- Every single piece of information from origin-destination surveys is kept in “tripfiles.” These flat files contain data on households (size, home location, number of cars), people (age, gender, car ownership) and trips (trip chain, origin and destination locations, junction points, mode, purpose, transit routes, and time of departure). The data are normalized and coded so that they can be retrieved rapidly and analyzed by software developed for the individual processing of trips. For the trip planners, simplified trip files are used in which information on origin and destination locations, date and

time, and transit routes is retained. The core information is the individual transit trip, as seen in Figure 1.

Figure 1. The Individual Trip in the Totally Dissagregate Approach



- Spatial districting, such as Traffic Analysis Zones (TAZ), is rejected for analysis in favor of the examination of individual trips, and trip extremities are geocoded at the most refined level of resolution. This is an essential feature of the trip planner because it enables users to specify precise origin and destination locations (such as addresses, intersections, monuments). Geographical information systems are thus essential at this point.

The use of the TDA for storage, validation, and analysis of large household origin-destination surveys constitutes the foundation of the data structure and the data itself for both the MTC and the LTC trip planners. In recent years, a new approach, the Transportation Object-Oriented Modeling (TOOM), has facilitated the integration of the TDA within multiple applications.

Transportation Object-Oriented Modeling

Transportation data processing generates several problems because it involves spatial components that are related together and have multiple time-dependent attributes. These problems have been partially reported by Goodchild (2000):

- **Temporality:** Transportation data always contain strong intrinsic temporal components (speed, time of movements), as well as extrinsic temporal components (date of effectiveness, time of collection).
- **Geography:** Transportation data are spatially referenced, without zones if general usage is the goal.
- **Structure:** Transportation data must be normalized so that they can be analyzed over time, by different software programs, and for different purposes.
- **Collection:** Transportation data are usually collected for a single purpose; far-sighted data collection can help to widen their range of use.

An object-oriented approach is well suited in this case because of its openness and its flexibility. Based on traditional OO approaches (computer science), the Transportation Object-Oriented Modeling is characterized by its special metaclasses of objects (dynamic, kinetic, static, and system), which are used to describe every transportation system (see Trépanier and Chapleau 2001 for more details). In this study, TOOM structures the database and the software environment in such a way as to link the Geographic Information System (GIS), the Operational Information System and the planner log files.

Methodology for Trip Planner Design and Implementation

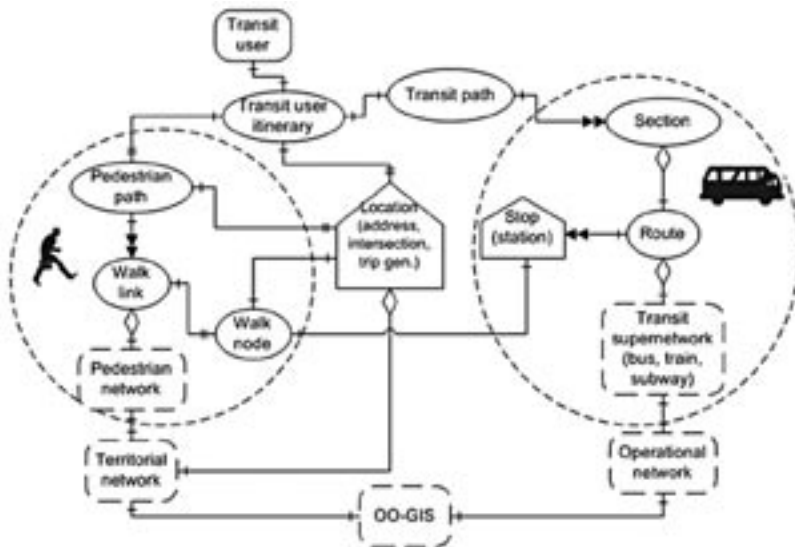
In this section, we will explain the design choices for the trip planner: methodology for design, trip planning process, and log file structure. These choices have helped the log file analysis because they are based on the same background elements (TDA and TOOM). The system architecture that is needed to obtain both the trip planner and the related log files are described in Chapleau et al. (1997) and Trépanier et al. (2002).

Methodology for Design

The methodology for both the construction and the operation of trip planners is a blend of the Totally Dissagregate Approach (TDA) and TOOM. The use of the TDA for the storage, validation, and analysis of large household origin-destination surveys constitutes the foundation of the data structure and the data itself for both the MTC and the LTC planners. TOOM always attempts to create a wide view of every system, even though the data for each object may not be available. Figure 2, for example, presents the general TOOM for transit trip planning. At the top of the model is the transit user, who uses two major systems to complete his

or her trip. On the left, the territorial network defines the pedestrian network, and includes streets and static objects (trip generators, addresses, intersections, postal codes). On the right, the operational network is composed of the bus, subway, and commuter train networks, and is seen as a single transit supernetwork by the user. The transit path is then defined as a sequence of route sections between two pedestrian paths (at the access and egress points, see Figure 1).

Figure 2. A TOOM for Transit Trip Planning



Trip Planning Process

To better understand the analyses presented below, let us describe the typical trip planning process on the MTC website (Figure 3). First, the user specifies his or her origin and destination locations in one of three ways: browsing through generator and stop lists; clicking on a map; or searching the database for a given intersection or address (the most common). Then, the user indicates the day and the time of the trip, and has the option of applying penalties on certain transit modes (subway or train).

The rest of the process is handled by the planning engine beneath the website. Access and egress are calculated using the pedestrian network. This determines the shortest walking path between the trip extremities and the transit network (which is reached through bus stops, train stations, or subway stations). Then, there is a path planning on the transit network (more than 50 paths can be calculated at a time, depending on the network complexity). A final selection of the best possible paths is made based

on the schedule, which is fully integrated into the paths. Figure 4 shows the kind of result that can be obtained on the MTC website: route selection, walking distances, path chronology, sequence of bus stops, and subway stations.

Figure 3. Trip Planning Process on MTC and LTC Websites

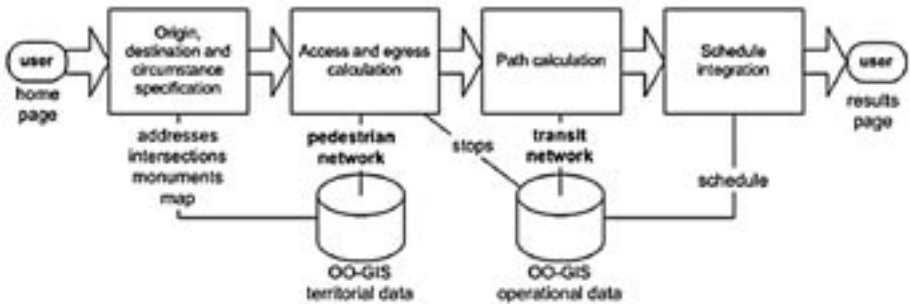


Figure 4. Trip Planning Results Page on the MTC Website

Calculated Itineraries

- Route(s) 208, 215, 2 (78 minutes)
- Route(s) 215, 2 (52 minutes)
- Route(s) 208, 215, 2 (44 (33 minutes)

Route(s) 208, 215, 2 (78 minutes)

Origin: prom. Hymen / prom. Sunshine (Dollard-des-Ormeaux)
 Destination: rue Rachel Est / rue Berri (Montréal)
 Day: 23/07/2004, Preferred departure time: 13h45
 Walk penalty: no, Metro active: yes, Train active: no
 Trip duration: 78 minutes

Dep. Time*	FROM	Walk / route	Arr. Time*	TO
13h45	prom. Hymen / prom. Sunshine (Dollard-des-Ormeaux)	Walk for 250 meters	13h45	Hymen / des Sources
13h45	Hymen / des Sources C-22020	208 Brunswick dr. Est	13h45	Brunswick / Deacon
13h44	Brunswick / Deacon C-22027	215 Henri-Bourassa dr. Est	14:15	Station Côte-Vertu
14:20	Station Côte-Vertu	Ligne orange dr. Henri-Bourassa	14:46	Station Mont-Royal
14:46	Station Mont-Royal	Walk for 350 meters	14:54	rue Rachel Est / rue Berri (Montréal)

*Approximate times. Please verify schedule with AUTORBUS system or with the Information Centre by calling 514 810-6100 -press 8 for English, option 7 (514 730-4636 + 8 + 7)
 Schedule patch 2004.06.21 - 2004.06.29
 Itinerary calculated on 03/07/2004

Log File Structure

Log files are kept for every transit trip planning. They typically contain the date and time of the visit, location of the origin and destination points (expressed in terms of identification, source and coordinates), parameters specified by the user, a summary result of the planning (route sequence), and the number of trip planning calculations made in a single visit. Only the last calculation of the visitor is retained in the log files. Table 1 shows a sample entry in the Laval’s website. At the MTC, there is also a full entry log file, containing all text entries and clicks made by users during their visits. Since there is only one file per log with normalized data, simple queries were made for the analysis.

Website-oriented Analysis

This section presents the principal results that can be obtained from website-oriented analysis. These analyses are intended to show the popularity of the trip planner function among users, to examine evolving trends in usage, and to improve the website functions themselves with the help of appropriate indicators.

Table 1. Sample Log File Entry for LTC Website

<i>Field</i>	<i>Sample Data</i>	<i>Description</i>
idsession	12P0KIUS5	Visitor’s web session ID
date	19-Oct-03 09:33:33	Date and time of visit
codeorig	FINT22999/24775	Code for origin location (FINT → intersection)
codedest	FGEN16015	Code for destination location (FGEN → trip generator)
orig	boul. Chomedey / boul. Cartier ouest <i>(Chomedey)</i></td>	Textual identification of the origin (for website display)
dest	métro Charlevoix <l>(Montréal)</l>	Textual identification of the destination (for website display)
xorig	286516	X coordinate at origin (UTM, meters)
yorig	5044373	Y coordinate at origin
xdest	611797	X coordinate at destination
ydest	5036849	Y coordinate at destination
result	[7606] 24 [1] 1002 [36] 1001 [37]	Result of the calculation (sequence of stops and routes)
nbcalc	2	Number of calculations for this user
c_jparc	di	Day for trip planning (here, di → Sunday)
c_hdep	12:00	Time of departure for trip planning
c_pmarc	n	Parameter for walk (n → user do not want to penalize walk)
c_ustrain	n	Parameter for train (n → user do not want to use train)

Statistics on Visits

Table 2 shows monthly statistics for the trip planning sub-website of the MTC website. A visit occurs when an Internet user accesses the website and performs a trip planning. As we can see, users usually calculate about two trips per visit. The number of clicks and the average time per visit are good indicators of the usability of the website. There is a click for every web request sent (change of web page). Between 2002 and 2003, these indicators are quite stable events, although there was a 57 percent increase in the number of visits.

Table 2. Monthly Statistics for the MTC Website (trip planner sub-website)

<i>Month</i>	<i>Calculations</i>	<i>Visits</i>	<i>Average Clicks per Calculation</i>	<i>Average Time per Visit(s)</i>
Jan 2000	9,158	4,970	n/a	n/a
Jan 2001	32,321	14,738	n/a	n/a
Jan 2002	62,355	27,625	5.41	241.16
Jan 2003	93,758	43,316	5.12	247.17
Jan 2004	129,609	62,289	n/a	n/a

The number of people using the trip planner is relatively small, compared to the total number of MTC users (there are about 300 million trips per year on the MTC network). There were 333,611 visitors to the entire MTC website during March 2002. Of this total, 27,625 (8.3%) used the trip planner.

Trends

The number of users of the MTC and LTC trip planners has been increasing steadily since these applications were launched in 2000. In 2001 alone, there was an increase of 221 percent for the MTC site. During the same period, CyberAtlas (2003) reported a 7.3 percent increase in the number of Internet users in general, and the average monthly web usage had increased from 8 hours and 17 minutes to 19 hours and 57 minutes.

Figure 5 shows the increase in the number of visitors between January 2000 and March 2002. The time series decomposition function starts at $x=0$ on January 1, 2000. In addition to the increase in the number of visits related to the growth of the Internet itself, two time-related effects can be observed. In a regular week, the number of trip planning calculations is 40 percent lower on the weekends. In

Figure 5. Number of Visitors to the MTC Website Planner, January–March 2002

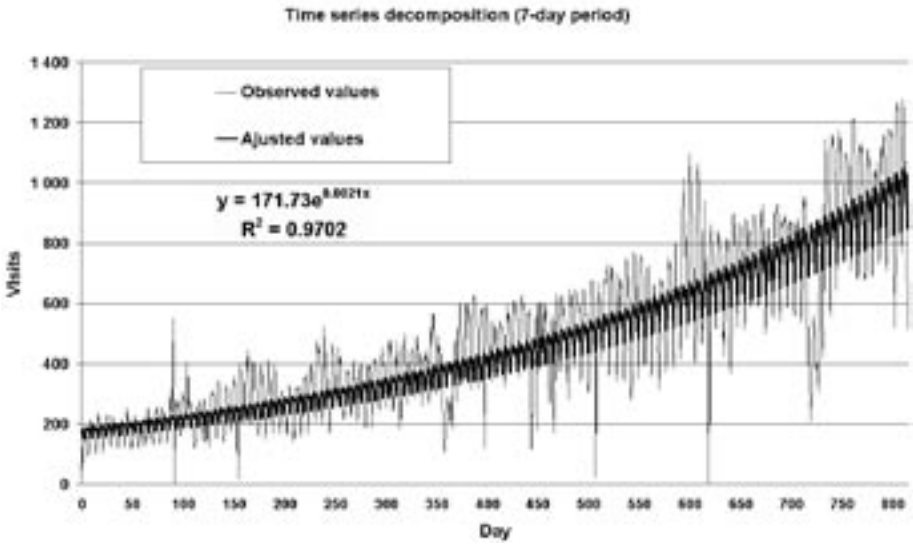


Figure 6. Average Number of Visitors per Day to the MTC Website Planner, January–May 2004

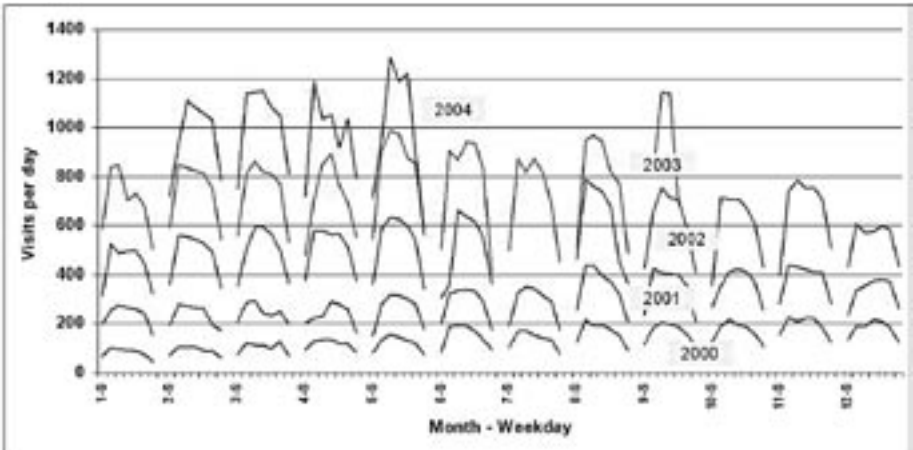
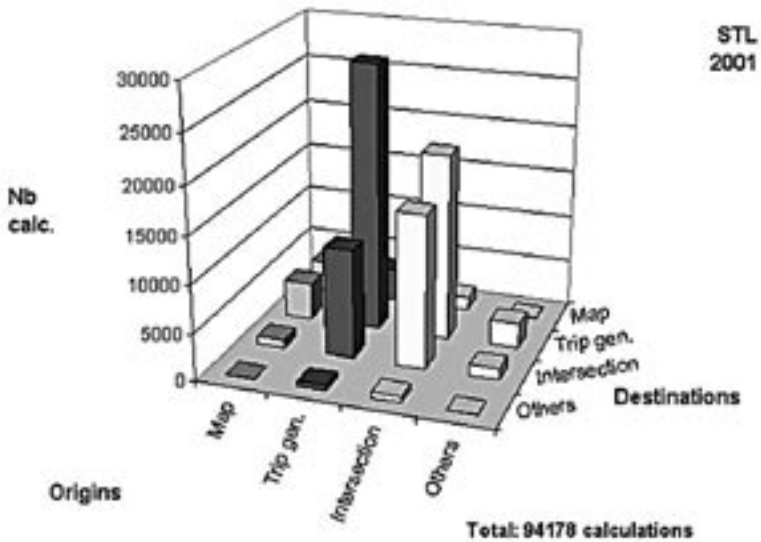


Figure 6, we observe that the peak period is on Monday and Tuesday. We can also see that there is a slight yearly peek during August and September, when most of the transit network changes are made, and students use the site to access new schools.

Location Specification

Also of interest is a look at the location specification mode itself to gain a better understanding of the use of the tools provided in the planner, and, when possible, to improve GIS data and the search engine. Figure 7 presents the distribution of location specification modes for the Laval website (2001). Note that there is no address search on this website. The study shows that the map-based search is much less used than the intersection or trip generator searches, which are more popular.

Figure 7. Location Specification Mode Distribution for the LTC Website, 2001



The same type of analysis has been conducted on the MTC website and shows that few visitors use only one type of location mode when they calculate their trip. Only 9.3 percent of all trip planning are from civic address to civic address. The number of calculations made with intersection pairs represents less than 5 percent of the total. The main finding is that users need a wide choice of location modes to adequately specify their origin and destination points. For this reason, these features were retained on the studied websites. Nevertheless, the preferences were taken into account in the design of the interface of the web pages.

Entry Log Analysis for GIS Updating

On the MTC website, the search functions have been thoroughly analyzed to search for the causes of data errors and typographical errors. A log file collects all keyboard entries and clicks on the planner sub-website. There were about 2,760,000 data inputs between June 2001 and March 2002. About 8.1 percent resulted in the display of an error message due to incorrect use of the interface by users, or because the desired location was not found (see Table 3). Of the 528,000 trip-planner visits made during this period, there was an abort-search rate of 3.9 percent which could be related to these error messages.

Table 3. Distribution of Entry Error Messages for the MTC Website, June 2001–March 2002

Location Type	Error Message	Error Type*	Count	% of Total
Address	Address not found	G	44,806	19.9%
Address	Unspecified civic number	I	15,284	6.8%
Address	Unspecified street name	I	2,115	0.9%
Address	Street not found	G	76,228	33.9%
Trip generator	Not found	G	21,444	9.5%
Intersection	Unspecified 1 st street	I	847	0.4%
Intersection	1 st street not found	G	21,458	9.5%
Intersection	Unspecified 2 nd street	I	7,257	3.2%
Intersection	2 nd street not found	G	19,065	8.5%
Intersection	Intersection not found	G	16,204	7.2%
Total			224,708	100%

* G = GIS-related, I = interface-related

Table 3 shows that a small proportion of the errors were due to incorrect use of the website interface by users. Missing fields, for example, are common, and one of a number of problems which can be easily corrected by changing the behaviors of the web pages. GIS errors, commonly features not found, are harder to tackle. A spelling error or language mismatch (both French and English are used on MTC website) can be the cause. Geographical corrections can help reduce missing intersection problems. For example, “virtual intersections” were coded at level intersection between freeways and streets, even though there is no physical intersection at these places. Based on most common errors, aliases were added for some street names, so that spelling variations would be accepted. With the help of these error entries and other inputs, more than 1,000 features are updated each year in the MTC’s GIS (e.g., street geometry, feature naming, new intersections, new address ranges).

Transit-Oriented Analysis

Although statistics on the number of visitors and their lengths of stay on the websites are useful for Internet designers, they do not provide information on the trip behavior of transit users. Let us now examine the log files to characterize the origin and destination locations, and also the use of the transit network.

Bias and Limitations of the Analysis

It is clear that the results of this study rely on the behaviors of only a portion of all transit users. A recent study (CEFRIO 2003) shows that about 60 percent of Canadians use the Internet on a regular basis. About 80 percent of them are between 15 and 25 years old, as are many of the transit users in the Montreal area. An internal survey of the Laval Transit Commission shows that about 70 percent of its users visit the authority’s website, and about 70 percent of these individuals use the trip planner. This translates into a trip planner usage of about 50 percent of all transit users in the Laval territory. The same survey reports that about 20 percent of them use the trip planner once a week, 45 percent a few times a month, and 20 percent at each schedule change (about once every three months). Finally, the underlying hypothesis is that website planner users do, in fact, make the trip that they calculated. This cannot be verified at the present time because no specific survey has been carried out for this purpose.

Main Destinations

In 2001, Montreal International Airport in the suburb of Dorval was the destination most often selected on the MTC website trip planner (4,514 trip planning calculations), followed by major subway terminal stations (Henri-Bourassa, Longueuil, Berri-UQAM, and Côte-Vertu). The relative difficulty in accessing the airport via the public transit network could be an explanation for the popularity of this destination. The map in Figure 8 shows all destinations related to more than 100 trips for 2001. Here, we can see some local concentrations in four places: A—the airport zone, B—the Saint Laurent industrial zone, C—the University of Montreal, and, of course, D—the Central Business District (CBD) zone.

On the Laval website, it is possible to obtain a path that leaves the LTC to go into the Montreal territory, and vice-versa. In Figure 9, the map shows the distribution of destinations over Laval's urbanized area for the whole of the year 2001. While the distribution is more spread out, there is a concentration in the CBD of Laval (A). On the Island of Montreal, calculations are made toward subway terminal stations, where users can easily find their way to their final destination (B—Henri-Bourassa and C—Montreal CBD stations).

Network Usage

Evaluation of transit network usage from individual trip declarations is a common feature of the TDA. The software MADITUC was used for this task. The following steps were carried out to obtain network load profiles and data on network components (MTC website):

- Individual trips are extracted from the website log file (pairs of origin and destination X-Y coordinates). Data are used “as is” and fully compatible to the TDA-based software.
- Each trip is simulated using MADITUC's trip simulator, which is basically the same as the website planner.
- The trips are “loaded” into the transit network objects to obtain the load profile.

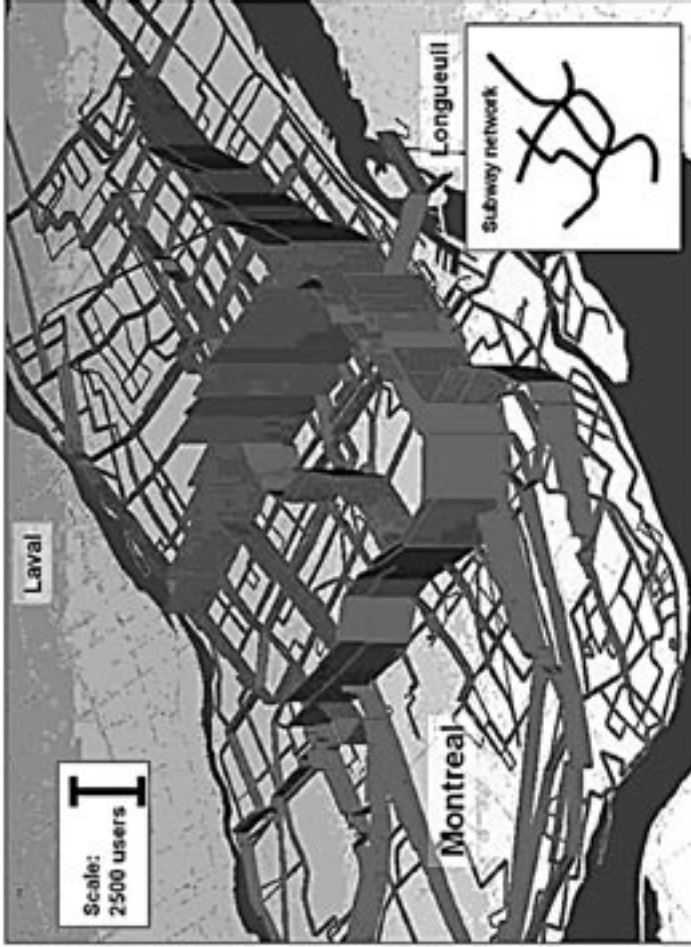
Figure 8. Map of Major Destinations of Planner Users, MTC Website



Figure 9. Map of Major Destinations of Planner Users, LTC Website



Figure 10. Transit Load Profile for Planner Users, January 2003, MTC Website



The load profile for the trips calculated on the MTC website in January 2003 is shown in Figure 10. It is significant on the subway network (enclosed map), which is considered normal for transit usage in Montreal. There is also an increased load on bus routes going to the western part of the Island of Montreal, compared to the normal load profile (not shown).

Table 4 presents statistics on transit network usage by users of the trip planner (based on calculated trips) and regular MTC transit users (data from 1998 household survey). The average total travel time and the average in-vehicle time are higher for the trip planner user. This would suggest that the trip planner is used to plan longer trips. However, the average access time (time to walk from origin to nearest bus stop or station) is lower for the trip planner, possibly because the locations specified in the planner are mostly intersections (many bus stops are located directly at intersections). The proportion of subway users is higher for the trip planner due to the concentration of destinations at the CBD.

Table 4. Transit Network Load Statistics for Website and Regular Transit Users

<i>Attribute</i>	<i>Trip Planner Users, Jan. 2003, STM</i>	<i>STM Transit Users (1998 survey)</i>
No. of trips loaded	43,297	35,533
Average total travel time (min)	49.99	46.28
Average access time (min)	10.05	11.71
Average waiting time (min)	17.57	15.88
Average in-vehicle time (min)	19.51	16.59
Average no. of transit routes per trip	2.02	1.84
Prop. of subway users (%)	63.0	57.6

Comparison with Origin-Destination Survey Matrices

Even though the websites do not collect any information on trip attributes, such as purpose or trip mode habits, we can compare the O-D matrices obtained from the website trip planner to those obtained from the 1998 O-D survey, to better characterize the website users. The goal is to know if the behavior of the website user is different from the regular transit user.

Table 5. Correlations Between Survey and Website Matrices

<i>O-D Matrices for Trip Calculation on STM Site</i>	<i>WEB2000</i>	<i>WEB2001</i>
STM website, 2000 (WEB2000)	1.000	0.982
STM website, 2001 (WEB2001)	0.982	1.000
1998 survey O-D matrices		
All trips, Montréal residents (TOT98MTL)	0.694	0.689
All trips, all residents (TOT98TOT)	0.707	0.701
Transit trips, Montréal residents (TC98MTL)	0.790	0.825
Transit trips, all residents (TC98TOT)	0.795	0.830
Work trips, Montréal residents (TRAV98MTL)	0.797	0.812
Work trips, all residents (TRAV98TOT)	0.822	0.834
Study trips, Montréal residents (ETUD98MTL)	0.586	0.581
Study trips, all residents (ETUD98TOT)	0.592	0.587
Transit trips, work trips, Montréal residents (TCTRAV98MTL)	0.565	0.610
Transit trips, work trips, all residents (TCTRAV98TOT)	0.571	0.616
Transit trips, study trips, Montréal residents (TCETUD98MTL)	0.683	0.709
Transit trips, study trips, all residents (TCETUD98TOT)	0.688	0.713

In this case, we defined 41X41 O-D matrices. Table 5 presents the correlation coefficient (R at a 95% confidence level) between these matrices, obtained by single-cell couple comparison. First, it shows a strong correlation between 2000 and 2001 for the website, confirming that the planner user's behavior remains constant, in spite of a 221 percent increase between these two years. Next, it shows that the correlation is stronger for transit-trip and work-trip matrices. This would identify the main characteristics of the Internet users, who are, we suppose, mostly transit users.

A quick look at Figure 11 demonstrates the similarity between the website usage matrix and the matrix developed from the transit trips in the O-D survey (the cells are shaded at proportional scales). The upper right-hand part of matrix A is the airport zone, and shows the importance of this zone for Internet users. Our hypothesis is that the website captures users who are not interviewed in O-D surveys, such as tourists and nonresident business travelers. In Figure 12, a comparison based on the importance of the matrix diagonal and CBD destinations confirms our previous analysis. The pattern for transit trips (TC98/MTL) is similar to the pattern of users' website declarations (WEB2001). The 2001 matrix to better fit the household survey data conducted three years before.

Departure Time

The distribution of the departure time is also a measure that can be used to compare website users' behaviors with behaviors from household survey data. Figure 13 shows the distribution of departure times for three categories: calculated trips on the STL website in 2000, declared trips by Laval residents (1998 survey), and declared trips by STL users (1998 survey). Results show that the proportion of departure times in the morning is higher for the trip planner than for the declared trips. This would suggest that the great proportion of planned trips are home-based (homeplace as origin).

Figure 11. Matrix Pattern for MTC Website 2001 and 1998 Survey (transit trips)

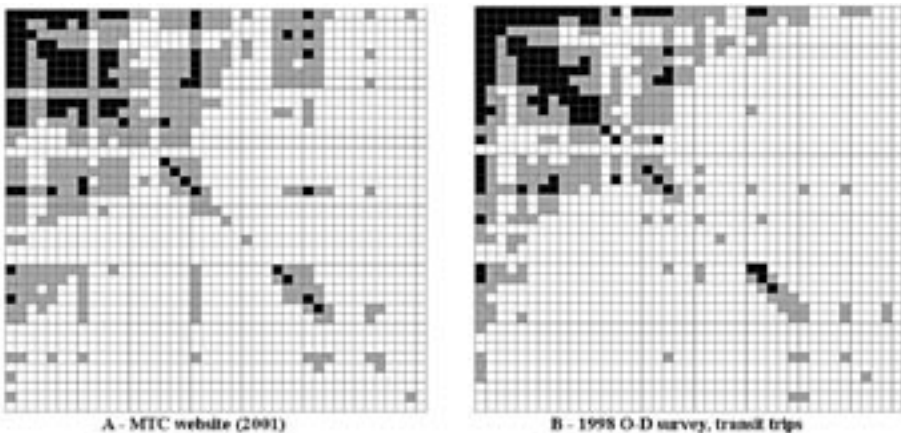


Figure 12. Matrix Comparison Based on Diagonal and CBD Destinations (MTC website)

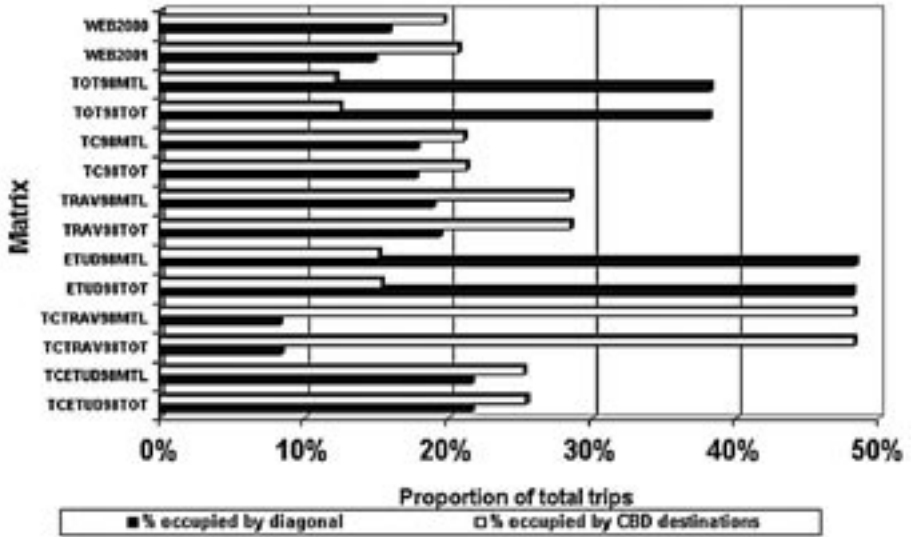
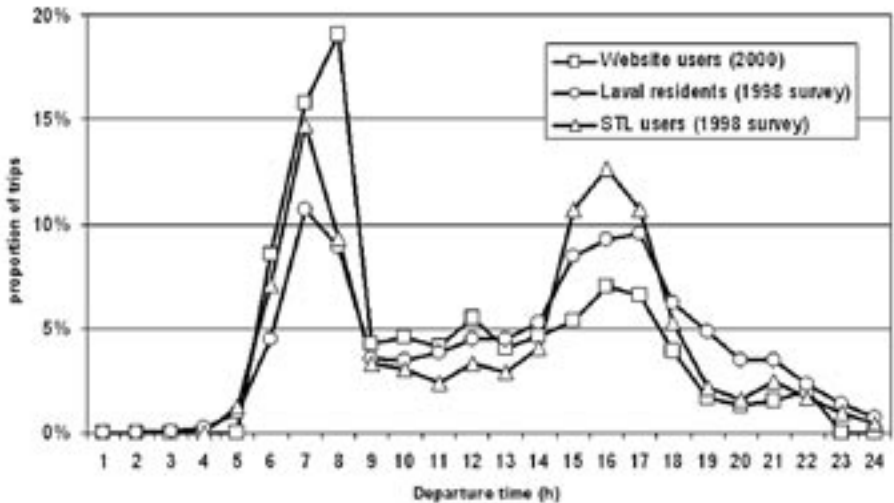


Figure 13. Distribution of Departure Time for Declared and Calculated Trips



Conclusions

The aim of this article was to determine whether trip planner log file analysis could help to improve the transit service, and, if so, how. The question cannot be completely resolved until we find methods to overcome, or at least evaluate, the following bias and limitations that were identified:

- Transit users do not all have access to the Internet; thus, they do not all visit the transit authority website.
- Visitors to the website do not all use the trip planner, although, in the case of Laval, more than 70 percent of the visitors have used it.
- The trips that have been calculated are only “planned” trips, and users can, of course, choose not to make the trip.

Nevertheless, it seems that relevant findings can be gleaned from the huge amount of data collected in the trip planner log files. In this project, examination of the log files of the transit trip planners of the Montreal Transit Commission and the Laval Transit Commission websites has led to the following observations:

- The growth in usage of this kind of website is faster than the growth of Internet usage itself, showing that there is an interest in the availability of transit trip planners on the Web.
- The Internet user usually performs more than one calculation during a visit. The tool can be accessed whenever it is needed.
- The trip planner is used to access destinations that are more difficult to reach by public transit. Such a tool is useful when the transit network is complex (several transfers) or when the areas are not well serviced.
- The user of a trip planner is typically a regular commuter and a worker. The trip planner also can be accessed and used by other regular users and by tourists relatively easily.

In addition to these observations, trip planner log files were found to be useful in some specific cases. The MTC had a better knowledge of its most popular destinations. The trip planner entry log helps to improve the underlying GIS and the website interface (the design of the new 2004 STL website has been based on findings from this project). Often, trip planners are used to access newly built structures or new developments; the GIS must, in this case, be updated as soon as possible.

This article has shown some potential usages of trip planner log files. Actually, with the help of the Totally Disaggregate Approach and appropriate data structures,

similar user behavior analysis are performed on other individual data, such as log files obtained from the usage of smart cards in transit systems.

Acknowledgments

The authors thank the Montreal Transit Commission, the Laval Transit Commission, and the Natural Sciences and Engineering Research Council of Canada (NSERC) for their supporting of this research project.

References

- Arampatzis, G., C.T. Kiranoudis, P. Scaloubacas, and D. Assimacopoulos. 2004. A GIS-based support system for planning urban transportation policies. *European Journal of Operational Research* 152: 465–475.
- CEFRIO (Centre francophone d'informatisation des organisations). 2003. *Netendances 2003, Utilisation d'Internet au Québec*.
- Chapleau, R. 1986. Transit network analysis and evaluation with a Totally Disaggregate Approach. *World Conference on Transportation Research Proceedings*, Vancouver.
- Chapleau, R., B. Allard, and M. Trépanier. 1997. Transit path calculation supported by a special GIS-transit information system. *Transportation Research Record* 1521: 98–111.
- CyberAtlas. 2003. <http://cyberatlas.internet.com/>.
- Federal Transit Administration. 2002. *Effective use of transit websites: A synthesis of transit practice*. Transit Cooperative Research Program (TCRP) 43, Washington.
- Golob, T. F., and A. C. Regan. 2001. Impacts of information technology on personal travel and commercial vehicle operations: research challenges and opportunities. *Transportation Research Part C* (9): 87–121
- Goodchild, M. F. 2000. GIS and transportation: Status and challenges. *Geoinformatica* 4, 2: 127–139.
- Ivory, M. Y., and M. A. Hearst. 2001. The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys* 33, 4: 470–516.
- Karimi, H. A., R. Peachavanish, and J. Peng. 2004. Finding optimal bus service routes: Internet-based methodology to serve transit patrons. *Journal of Computing in Civil Engineering*, April: 83–91.
- Maclean, S.D., and D. J. Daily. 2002. Measuring the utility of a real-time information system. *Proceedings of the IEEE 5th Conference on Intelligent Transportation Systems*, 846–850.
- Lee, J. S., J. A. Baumgartner, and J. K. Tschango. 1999. A web-based bus information system. *ESRI User Conference Proceedings*, no.494.

- Murphy, J., C. F. Hofacker, M. Bennett. 2001. Website-generated Market-research Data. *Cornell Hotel and Restaurant Administration Quarterly*, February: 82–91.
- Peng, Z.-R., and R. Huang. 2000. Design and development of interactive trip planning for web-based transit information system. *Transportation Research Part C* 8: 409–425.
- Piroli, P.L.T., and J. E. Pitkow. 1999. Distributions of surfers' paths through the World Wide Web: Empirical characterizations. *World Wide Web* 2: 29–45.
- Smith, K.A., and Alan Ng. 2003. Web page clustering using a self-organizing map of user navigation patterns. *Decision Support Systems* 35: 245–256.
- Tan, P.-M., and V. Kumar. 2002. Discovery of web robot sessions based on their navigational patterns. *Data Mining and Knowledge Discovery* 6: 9–35.
- Trépanier, M., R. Chapleau, and B. Allard. 2002. A transit user information system for transit itinerary calculation on the web. *Journal of Public Transportation* 5, 3: 13–32.
- Trépanier, M., and R. Chapleau. 2001. Linking transit operational data to road network with a transportation Object-Oriented GIS. *Urban and Regional Information Systems Association Journal* 13, 2: 23–27.

About the Authors

MARTIN TRÉPANIÉRIE (mtrepanier@polymtl.ca) is professor of industrial engineering at the Ecole Polytechnique de Montreal. His research is mainly related to logistics, information systems, object-oriented modeling, GIS development, and Internet applications. He worked along with the MADITUC group in this research project.

ROBERT CHAPLEAU (rchapleau@polymtl.ca) is professor of civil engineering (transportation planning) and founder-director of the MADITUC group, Civil Engineering Department, Ecole Polytechnique de Montreal. He developed the Totally Disaggregate Approach in transportation and is participating in several research projects in the Montreal area.

BRUNO ALLARD (ballard@polymtl.ca) is research associate in the MADITUC group, Civil Engineering Department, Ecole Polytechnique de Montreal. He is an experienced system developer and manager, with applications in origin-destination surveys, transit operation, and GIS-T-based websites.