Missed Connections: Quantifying and Optimizing Multi-modal Interconnectivity in Cities

Cathal Coffey  
cathal.coffey@nuim.ie  
National Centre for Geocomputation  
National University of Ireland Maynooth  
Maynooth, Co. Kildare, Ireland

Rahul Nair  
rahul.nair@ie.ibm.com  
Smarter Cities Technology Centre  
IBM Research - Ireland  
IBM Technology Campus, Dublin 15, Ireland

Fabio Pinelli  
fabiopin@ie.ibm.com  
Smarter Cities Technology Centre  
IBM Research - Ireland  
IBM Technology Campus, Dublin 15, Ireland

Alexei Pozdnoukhov  
alexei.pozdnoukhov@nuim.ie  
National Centre for Geocomputation  
National University of Ireland Maynooth  
Maynooth, Co. Kildare, Ireland

Francesco Calabrese  
fcalabre@ie.ibm.com  
Smarter Cities Technology Centre  
IBM Research - Ireland  
IBM Technology Campus, Dublin 15, Ireland

ABSTRACT
We present a methodology to measure multi-modal interconnectivity between different transportation modes that operate in a city. The interconnectivity of an urban network represents how well different services integrate to offer seamless transportation options to users. On the supply side, we leverage open data sources that cities provide to accurately model the services that cities offer. On the demand side, we account for myopic user behavior through the use of journey planners and travel demand estimates. The reciprocal interaction between supply and demand is then used to characterize interconnectivity. In a multi-modal setting, we present different measures that can be employed to understand shortcomings in connectivity across the network. Using these metrics, improvements in service schedules are proposed using an optimization model that seeks to perturb existing schedules to improve transit connectivity. Real-world data from Washington, D.C. is used to demonstrate presented measures and optimal schedule perturbation for one route are presented which result in a 19% reduction in delays.

1. INTRODUCTION
We present a methodology to measure multi-modal interconnectivity between different transportation modes that operate in a city. The interconnectivity of an urban network represents how well different services integrate to offer seamless transportation options to users. The interconnectivity measures are then utilized to optimize service schedules, such that transport services are better coordinated. Civic authorities, faced with growing urban populations, environmental problems, and shrinking operational budgets, seek to offer better mobility services using less public resources. This challenge in improving supply has to be considered with citizens’ needs, since the two are inextricably intertwined. Urban transportation systems are complex networks, where several modes, often operated by different entities, may not be planned to provide a globally optimum service as a whole, but rather a local solution, where each stakeholder may benefit. The purpose of this paper is therefore to study such multi-modal networks at the city level to characterize the nature of interconnectivity. By identifying hierarchies and opportunities for improvements at the system level, the proposed work aims to arm civic authorities and operators to proactively address missed connections at the urban level.

The research is facilitated by the availability of open data sources such as GTFS (General Transit Feed Specification) data; which describes general transit networks, as well as AVL (Automatic Vehicle Location) data from transit vehicles that offer a real-time picture of the system. At a coarse level, information on the network structure provides a great opportunity to investigate methodological approaches

Categories and Subject Descriptors
G.1.6 [Optimization]: Constrained optimization, Linear programming

General Terms
Design, Experimentation, Algorithms

Keywords
Optimization, multi-modal, transit, schedule, route

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
ACM SIGSPATIAL IWCTS’12, November 6, 2012. Redondo Beach, CA, USA
Copyright (c) 2012 ACM ISBN 978-1-4503-1693-4/12/11 ...$15.00.
to measure the performance of public transit systems and proposes improvement and optimizations. Many recent studies provide a variety of analysis focusing on a particular transportation mode, measuring its performance in terms of reliability and accessibility. An example is given in [14] where the authors study the accessibility of bus transport system using GPS data. Other works as [13] where the authors use GIS data to develop isochrones on the basis of bus route timetables and street network data. The use of operational data to support strategic decision making in transport has been pursued recently where delay costs are used to justify highway investments [15].

This work addresses the gap on the use of operational data, such as AVL and schedules, for multi-modal network analysis. In contrast to previous studies, this paper also takes into account the supply and demand processes that govern system performance, i.e. the system is evaluated on how well it meets demand. The multi-modality of the system and the citizens’ demand represent key aspects to perform an accurate evaluation of the reliability of the system and its quality. Analysis of a single mode can provide insights on how the mode performs. However, typical urban journeys may require more than one mode, thus unimodal analysis represents a partial view of the system. For instance, a typical commuter journey schedules one part of the trip by bus and the next one by metro. If the two modes are not well interconnected the efficiency of both systems is wasted. Several questions arise when considering a multi-modal network for public transit. At the city level, how much time do people spend waiting, on average, when transferring between modes of transit? What are simple ways that different operators can coordinate to better serve the traveling public?

A quantitative approach to measuring the interconnectivity properties of a transport network is needed. For example, measuring the amount of time that people spent waiting for a change of transportation mode can be used to evaluate if the multi-modal system presents a good level of interconnectivity between the different transportation modes. It also highlights where and when the infrastructure presents bottlenecks and specific problem w.r.t the travel demand. The transit system can be optimized in different ways: adding new lines, new stops/stations but all of these are costly solutions and some of them can have a considerable impact on the urban environment. Perturbations of the temporal schedule, instead, can produce considerable results while keeping the operational costs and the environment impact the same.

The paper makes the following contributions. We propose a formal definition to measure interconnectivity of the multi-modal transit network based on the traffic volume passing through these nodes at a specific time. Then, we formally define the optimization problem through a perturbation of the time schedule in order to minimize the amount of waiting time along a specific line of a transit service. Finally, we demonstrate the proposed methodology on a case study of Washington D.C. The paper makes methodological contributions by sampling approximately 10% of all trips from aggregate travel patterns and generating travel itineraries that support high fidelity analysis. The research also demonstrates how open data and open tools such as OpenTripPlanner (OTP), can be leveraged to provide additional insights into city functions. By relying on open standards, it is hoped that the results can be replicated in other cities as well.

2. PROBLEM DEFINITION

In this section, we provide formal definitions of the problems and discuss a systematic methodology to evaluate the level of interconnection between the different transportation modes in terms of the waiting time that people spend in each station during their journey. We then propose a schedule optimization problem in order to reduce the waiting time due to interconnections with other transportation modes or vehicles.

2.1 Interconnectivity measures

We assume that a good interconnectivity level among transportation modes in a multi-modal infrastructure corresponds to low waiting time during a change of mode or vehicle. In the following we propose two formal indexes that measure the amount of waiting time at each node $n$ of the network $G(N,E)$. We define the total waiting time for each node $n \in N$ given a set of queries $Q$ performed on a typical weekday as:

$$\text{Total Waiting Time}(n) = \sum_{q \in Q} w_q \times n_q$$

where $w_q$ represents the time a person must wait at node $n$ while following query $q$. Similarly $n_q$ is an estimate for the number of people following query $q$ and therefore are passing through node $n$ because of query $q$. This measure highlights particular nodes of the network that suffer from considerable waiting time and are thus possible bottlenecks for the entire system. Moreover the interconnectivity on a specific node $n$ needs to be averaged by the real usage of this node. For this reason we define the average waiting time of a node $n$ given a set of queries $Q$ performed on a typical weekday as follow:

$$\text{Average Waiting Time}(n) = \frac{\sum_{q \in Q} w_q \times n_q}{\sum_{q \in Q} n_q}$$

In this way, we can measure systematically the response of each node of the network with respect to its relative usage during a typical weekday. As previously mentioned, a high value corresponds to low performance of the node in terms of its cooperation level between transportation operators. Moreover, it is worth noticing that we can associate to the waiting time a cost, this give us an idea of the important role played by the synergy between the transportation modes w.r.t to the real usage when we deploy a complex system such as the public transit. In addition, the distribution of waiting times at nodes, and the transfer modes that define the waiting times are also studied. A deeper investigation of these measures is provided in the case study where we demonstrate how to use them in a practical problem and what we can highlight through their usage.

2.2 Schedule optimization

Given estimates of transfer delays at various points on the network, we present a linear optimization model that seeks to perturb existing schedules such that transfer waiting times are minimized. The aim of the model is to alleviate systematic delays that occur at transit facilities. Since public transit networks, especially multi-modal ones, evolve over a long period of time, operators often lack the comprehensive picture on how interconnected services are used by the
traveling public. This schedule perturbation model therefore serves as a feedback loop in the overall transit planning process [4], where accurate estimates of transfer delays serve as key performance measures.

There is a sizable literature on schedule generation, and optimization for public transit networks (see [9] for a review). The review also highlights the need for multi-modal considerations in transit planning and the importance of service coordination. We focus on previous work that addresses where schedules of different services are better aligned to minimize transfer times. Several studies have looked at coordinating services [8, 7, 3] where the aim is to coordinate arrivals of different services at a stop. [1] considered the problem of minimizing waiting times when travel time between stops is random and services have fixed headways. Several other works that study coordination have employed meta heuristic approaches to solve the general problem [12, 6, 5, 11, 10] using methods such as genetic algorithms, tabu search, and local search. The approach presented herein is similar in spirit to [2] who present a Lagrangian-based heuristic that seeks to optimize one line at a time.

Our approach differs from the previously studied model in the following ways. Our focus is in achieving small perturbations in the schedule times to reduce transfer penalties instead of redesigning the transit network or schedule. This is a practical way for operators to improve services, given resource constraints and existing travel time constraints. The model also employs detailed transfer volume data at the trip level, such that time-of-day variations are included and trip-to-trip transfer volumes are known. This level of disaggregation affords more fidelity in schedule adjustments that have previously not been feasible.

The proposed optimization model focuses on one line and seeks a local improvement in the dispatch time at the start point of the trip, such that ensuing transfers to and from this service experience minimal waiting time needed to complete the transfer. The model is considered strategic, in that operational issues such as uncertainties in travel times and variability in demand are not considered. The existing schedule is assumed to have adequate slack built in the travel times to absorb these uncertainties. The model keeps these slacks and existing travel time, but only seeks and optimal temporal shift, should one be available. The aim therefore is to remedy particularly egregious transfer delays, as identified by the interconnectivity measures.

Given (a) the existing schedule of a transit service that is being improved, (b) the schedules of a set of services from which users transfer to and from the service being optimized, (c) the volume and times of transfers based on journey plans, we seek an optimal temporal shift such that transfer waiting times are minimized. The model considers only minor deviations from the existing schedule, where minor is defined as being within the time headway of the service (time headway is the temporal difference between two successive runs of the same route), this headway is not necessarily constant it can vary depending on the time and day of the model. The model is defined on a set of nodes $N$ indexed by $i$, which are the stops for the service $S$. The service $S$ includes several runs, indexed by $p$. The schedule for $p$ is denoted by $t_{pi}$ which is the time run $p$ services node $i$. Denote $h_{pi}$ as the average time headway for the route during the hour of run $p$. At each node $i$ there are transfer opportunities to and from other services, which are defined by the set $Q$. The set $Q$ also has two subsets $Q_i^-$ and $Q_i^+$ indexed by $q$ (where the subscripts $+/-$ indicate if passengers join or leave service $S$). The transfer volume between two services is denoted by $C_{pqi}$ as the number of people who transfer from service $q, q \in Q_i^-$ to run $p$ at stop $i$. Similarly, $C_{pqi}^+$ denotes the number of people who transfer from $p$ to $q, q \in Q_i^-$ at node $i$. These quantities are shown in Figure 1.

![Figure 1: Model of service network showing transfers and associated notation](image1.png)

![Figure 2: Relative intensity of zones as trip origins (left) and destinations (right)](image2.png)

Define two indicator variables

$$
\xi_{pqi} = \begin{cases} 
1 & \text{if } C_{pqi}^- > 0 \\
0 & \text{otherwise}
\end{cases} \quad \xi_{pqi}^+ = \begin{cases} 
1 & \text{if } C_{pqi}^+ > 0 \\
0 & \text{otherwise}
\end{cases}
$$

(3)

Define $\gamma$ as the minimum time required to make the transfer (this can be relaxed to be node specific or service-pair specific). The decision variables are denoted by $\delta_p$ which is the perturbation in time to the schedule of run $p$, and $w_{pqi}$ which is the waiting time in passenger-time units for transferring from $p$ to $q$ at node $i$. With these definitions, the schedule perturbation model can be expressed as follows.

$$
\min_{\delta_p} \sum_{p \in S} \sum_{i \in N} \sum_{q \in Q} w_{pqi}
$$

subject to

$$
w_{pqi} = \sum_{q \in Q_i^-} (t_{qi} - (t_{pi} + \delta_p)) C_{pqi}^- + \sum_{q \in Q_i^+} ((t_{qi} + \delta_p) - t_{qi}) C_{pqi}^+
$$

$$
\forall p \in S, q \in Q, i \in N
$$

(5)
from repositories such as GTFS-data-exchange. GTFS and OSM data are already available for many cities.

It is very important because it enables us (as future work) to optimize multi-modal networks. 2) OTP takes as input two open formats: the Google Transit Feed Specification (GTFS) and Open Street Map data. This is very important because it enables us to manipulate the transit network we need to be able to manipulate the transit schedules themselves and then rerun our analysis. The GTFS format enables us to do this, it’s a human readable open format so it’s trivial to for example shift the starting times of any service. More complicated modifications such as adding a completely new service or removing a stop from an existing service are also relatively easy to accomplish, these modifications are however not so easy to justify or validate and so we leave this as potential future work.

It is important to generate a sufficiently large collection of multi-modal journeys such that you get a realistic sample of how the real transit network of a city is utilized by a population. OTP queries require a minimum of three parameters: a starting location, a destination location and a time. We set OTP to optimize for minimum arrival time and we constraint the maximum number of transfers to 5.

We use OpenTripPlanner\(^1\) (OTP) to generate multi-modal journeys for a city. We choose OTP for a number of reasons. 1) It runs as a local server so there is no time-out or restriction in terms of the number of queries we can make. 2) OTP takes as input two open formats: the Google Transit Feed Specification (GTFS) and Open Street Map data. This is very important because it enables us (as future work) to manipulate the transit network of any city for which this data (and an origin destination matrix) is available. High quality GTFS and OSM data are already available for many cities from repositories such as GTFS-data-exchange\(^2\) and Cloud-Made\(^3\). 3) In order to evaluate our modifications to the transit network we need to be able to manipulate the transit schedules themselves and then rerun our analysis. The GTFS format enables us to do this, it’s a human readable open format so it’s trivial to for example shift the starting times of any service. More complicated modifications such as adding a completely new service or removing a stop from an existing service are also relatively easy to accomplish, these modifications are however not so easy to justify or validate and so we leave this as potential future work.

In order to generate a realistic collection of queries we must respect the true spatial and temporal distributions of transit queries. To do this we select our start and destination locations using an origin/destination matrix for the city. We also use a temporal distribution for the transit network that captures the morning and evening rush hours.

So we need four pieces of information to do our analysis. 1) An origin/destination matrix which captures the number of people flowing from every region of the city to every other region. 2) A temporal distribution for transit usage which captures the number of people traveling with respect to time. 3) An accurate GTFS file for the city which represents the entire transit network structure and timetabling. 4) OSM data for the city which provides real network distances and directionality. Once we have these 4 pieces, we simply setup a local OTP server and bombard it with queries that respect our spatial and temporal distributions.

3. DATA ACQUISITION

We use OpenTripPlanner\(^1\) (OTP) to generate multi-modal journeys for a city. We choose OTP for a number of reasons.

1) It runs as a local server so there is no time-out or restriction in terms of the number of queries we can make. 2) OTP takes as input two open formats: the Google Transit Feed Specification (GTFS) and Open Street Map data. This is very important because it enables us (as future work) to manipulate the transit network of any city for which this data (and an origin destination matrix) is available. High quality GTFS and OSM data are already available for many cities from repositories such as GTFS-data-exchange\(^2\) and Cloud-Made\(^3\). 3) In order to evaluate our modifications to the transit network we need to be able to manipulate the transit schedules themselves and then rerun our analysis. The GTFS format enables us to do this, it’s a human readable open format so it’s trivial to for example shift the starting times of any service. More complicated modifications such as adding a completely new service or removing a stop from an existing service are also relatively easy to accomplish, these modifications are however not so easy to justify or validate and so we leave this as potential future work.

It is important to generate a sufficiently large collection of multi-modal journeys such that you get a realistic sample of how the real transit network of a city is utilized by a population. OTP queries require a minimum of three parameters: a starting location, a destination location and a time. We set OTP to optimize for minimum arrival time and we constraint the maximum number of transfers to 5. For an exhaustive list of optional OTP parameters please see the API\(^4\).

In order to generate a realistic collection of queries we must respect the true spatial and temporal distributions of transit queries. To do this we select our start and destination locations using an origin/destination matrix for the city. We also use a temporal distribution for the transit network that captures the morning and evening rush hours.

So we need four pieces of information to do our analysis. 1) An origin/destination matrix which captures the number of people flowing from every region of the city to every other region. 2) A temporal distribution for transit usage which captures the number of people traveling with respect to time. 3) An accurate GTFS file for the city which represents the entire transit network structure and timetabling. 4) OSM data for the city which provides real network distances and directionality. Once we have these 4 pieces, we simply setup a local OTP server and bombard it with queries that respect our spatial and temporal distributions.

4. CASE STUDY

The presented method for quantifying interconnectivity and optimizing multi-modal networks is applied to Washington, D.C. The city is served by the Washington Metropolitan Area Transit Authority (WMATA) along with several smaller transit agencies. WMATA operates a heavy rail system, called Metrorail, consisting of five primary lines, and an extensive bus network, called Metrobus. The entire transit network consists of 311 routes that serve an estimated 1.07 million trips per day.

Using the transit schedule information from GTFS, the OSM network, and OTP, a sampling methodology is devised to sample approximately 10% of the daily transit flows. The transit flow patterns are estimated from the origin-destination matrix estimated by the regional planning organization, the Washington Metropolitan Council of Governments. The OD matrix available shows daily estimated transit patterns in the region between 3,574 zones with 2.1 million OD flows. For the purposes of this study, a subset within the service area of the WMATA transit network is considered. This subset encompasses 2,899 zones and is shown in Figure 2 with relative intensity of each zone as origin and destination. The figures show the movement of people from outlying suburbs to the city center. The OD pattern represents a total of 1.07 million trips. For analyzing multi-modal connectivity, a spatial sample of approximately 10% is constructed from this overall demand data. Random origins and destinations within each zone are selected to constitute trips performed. To establish the temporal distribution of the sample, a temporal distribution is assumed based on known mobility patterns [15].

While the 10% sample is intended to provide an accurate description of the demand and supply processes that occur on an urban multi-modal network, there are some limitations that must be noted. Primarily, due to lack of data, the directionality of travel was only available as a daily aggregate. This means that time-of-day impacts, beyond the aggregate temporal distribution could not be accounted for. For certain OD pairs that experience highly directional flow, it implies that our estimates of interconnectivity are conservative. In practice, other data sources, such as crowdsourced journey planner queries, could be utilized to infer directionality. Secondly, there are several alternative modes of travel available for users in the Washington region. These
include bicycle-sharing schemes such as Capital Bikeshare\(^5\) and car sharing services such as Zip Car and Car2Go\(^6\) which offers one-way car sharing. While these systems contribute to the overall multi-modal basket of travel opportunities, these were not included in the present study, since robust journey planners are currently not available. Lastly, there are operational considerations that are inadequately represented within OTP. Use of a personal automobile to access transit stations is not currently supported. A large number of transit stations within the periphery have park-and-ride facilities that have not been modeled. As a result, approximated 14% returned infeasible trips for given OD pairs. These have been excluded from the subsequent analysis.

### 4.1 Measures for Washington, D.C.

The measures introduced in Sec.2.1 are evaluated for Washington D.C. Figure 3 show the overall spatial distribution of the total waiting times, a measure of interconnectivity. The rail system which forms the backbone of the network with high-volume of passengers contributes significantly to the overall interconnectivity measure. Figure 4 disaggregates the total waiting time by different mode combinations, showing the hierarchy in modal connections. While the high-frequency rail services offer minimal delays when transferring to other rail services, it is more expensive to transfer from rail to bus than between any other transit mode pair. While this is to be expected, on account of rail systems having a dedicated right-of-way, the comparison of rail-to-bus and bus-to-rail shows the asymmetry of modal interconnectivity.

Figure 5 shows the distribution of waiting times at selected nodes on the network. The plot shows the invariance of the waiting time distribution w.r.t. volume and the number of transfers that give rise to the distribution. This exponential distribution of waiting times also hold across different modes. Figure 6 shows the distribution of resulting trips based on distance bands and their modes as a relative weight of total travel distance. It is intended that this plot describe the interplay between transit modes therefore only trips which contain at least one transfer are included.

### 4.2 Optimizing a Route

To demonstrate the schedule perturbation model, a route with high associated transfer costs was chosen from the network-wide analysis of interconnectivity. The bus route 38B intersects with metrorail stops and offers several transfer opportunities. The route has 62 trips daily and accounts for 1173.14 passenger-hours of delay daily. The major transfer stops service high-volume interchanges to the metrorail system. The optimization model, implemented in MATLAB with IBM ILOG Cplex as the linear solver, seeks to perturb the schedule of Route 38B, based on interconnectivity analysis and estimated transfer volumes. The model was solved in 1.3 seconds, yielding the optimal schedule. Since the program works on a single line, the model can be scaled to the entire system by optimizing one line at a time.

Figure 7 shows the recommended schedule perturbation in comparison with the existing schedule. As a result of the perturbation, the modified schedule yields a total delay of 940.95 passenger-hours, a 19% improvement over the original schedule. It should be noted that these savings are conditional on the sampled transfers being realized. While

\(^5\)http://capitalbikeshare.com/
\(^6\)http://www.car2go.com/washingtondc/en/concept/
city-wide features were demonstrated. There is a hierarchy in mode transfers that is asymmetric. Certain attributes, such as the waiting time distribution at nodes, is invariant w.r.t. volume. Perturbation of the schedule on one service showed reduction of transfer delays by 19%, under the sampled transfer volumes. This work can be extended in several dimensions. Primarily, instead of sampling multi-modal trips, crowd-sourced journey plans and users as probes to provide trajectory level data, could serve to provide a richer data source and potentially drive real-time strategies to coordinate services. Normalized metrics, such as the average waiting time, could be measured in different cities to evaluate the impact of their size and service offerings. The optimization problem can be extended to the entire system to satisfy global objectives of minimal system-wide waiting time.

6. ACKNOWLEDGMENT

The authors would like to thank Ron Kirby, Ron Malone, and Mary Martchouk at the Metropolitan Washington Council of Governments for sharing data on transit demand for the Washington, D.C. region. C. Coffey acknowledges support from the Strategic Research Cluster grant (07/SRC/I1168) by Science Foundation Ireland, and an IBM PhD Fellowship award.

References


