Public Transport Systems' Connectivity: Spatiotemporal Analysis and Failure Detection

Yuval Hadassa,*, Riccardo Rossib, Massimiliano Gastaldib, Gregorio Geccheleb

*aBar-Ilan University, Department of Management, Max ve-Anna Webb Street, Ramat Gan 5290002, Israel
bUniversity of Padova, Department of Civil, Environmental and Architectural Engineering, Via Marzolo 9, Padova 35131, Italy

Abstract

Public Transport (PT) plays a major role in passenger flow as an affordable and efficient mode contributing for sustainable transportation by way of traffic congestion and air pollution reduction. Those advantages are impaired if the PT system does not provide a continuous accessibility and connectivity for all prospect passengers. Hence, it is imperative to assess the performance of PT systems based on the system's spatial and temporal properties. For the failure detection, three connectivity indicators are being used: a) transportation network coverage (direct and indirect); and b) stop transfer potential. These indicators are used for the identification of connectivity issues and flaws. Each indicator provides the means to identify the causes in terms of network coverage, routes structure and coverage, stops locations, frequencies, and transfers synchronization. A case study of Dolo area, which is part of Veneto region (Italy), is introduced. The analysis is focused on the hospital connectivity. The current PT system is analyzed, followed by identifying connectivity failures, and improvements recommendations. Results show that connectivity to the hospital by PT is characterized by long ingress and egress distances, low frequencies, and lack of fast and efficient transfers. The local authorities can easily use the tool to pinpoint stops to be relocated, as well as time-tables change, all in order to increase the connectivity by PT to the hospital.

* Corresponding author. Tel: +972 3 5317009, Fax: +972 4 9590565.
E-mail address: yuval.hadas@biu.ac.il

Keywords: public transport; connectivity, spatiotemporal analysis; failure detection; case study
1. Introduction

Public Transport (PT) plays a major role in passenger mobility and contribute for sustainable transportation, by way of traffic congestion and air pollution reduction if it is an affordable and efficient mode. PT systems must provide a continuous accessibility and connectivity for passengers, otherwise these advantages cannot be achieved. To enhance decision making in view of better accessibility and to allow comparison among PT systems over time, the assessment of PT systems connectivity is fundamental. In this context the assessment is defined as the ability to extract and analyze data in an automated and recurring process, since PT systems consist of several physical features (roads, railways, routes, stops), represented by a complex network of spatial and temporal data with millions of entities (Ceder, 2007; Vuchic, 2005). Any PT analysis should refer to the four availability factors of PT systems, as described by the Transit Capacity and Quality of Service Manual (Kittelson & Associates. et al., 2003). 1) spatial – where the service is provided, 2) temporal – when the service is provided, 3) information – how to use the service, and 4) capacity – space available for the passenger. For PT connectivity analysis the first two are particularly important, since connectivity has both spatial (routes coverage, stops locations, transfer availability, etc.), and temporal (waiting time, travel time, transfer time, etc.) components. For general network analysis several connectivity measures were developed and are commonly employed (Black, 2003; Rodrigue et al., 2006). Among the others are connectivity and strong connectivity of graphs (Ahuja et al., 1993); the cyclomatic number, which is essentially a measure of the number of circuits in a graph; the alpha index, which is the ratio between the number of existing circuits and the maximum of circuits possible. For transportation networks other connectivity measures were proposed (Mishra et al., 2012), such as the longest shortest path of a network (which is the longest distance possibly traveled among all shortest paths in a network); the degree of a node, which can take the form of the number of arcs connected or the form of the sum of shortest paths to all other nodes (Black, 2003); the ratio between the network- based shortest path and the direct line between node pairs. Public transport network models are far more complex than general and other transportation networks: arcs represent roads and routes, whereas nodes represent intersections and stops. The definition of connectivity measures has to consider this complexity, accounting for the routes as well as timetables, access, transfer, etc.. Vuchic (2005) presented a set of measures such as transfer permutation (for routes sharing the same stop or station); network complexity (ratio of arcs and nodes); line overlapping; directness of service, etc.. O’Sullivan, Morrison et al. (2000) introduced an isochronic approach for modeling PT systems based on timetables, which is best used for a small number of origins. A similar approach, which uses a schedule-based, Dijkstra shortest-path algorithm, was suggested by Lei and Church (2010). Connectivity measures that integrate demand forecast and transfers classification were developed by Hadas and Ceder (2010) and Hadas and Ranjitkar (2012). Other models require more complex datasets, such as demand, demographic properties, transportation zones, attractions, etc. As an example, Wu and Hine’s model measures changes in bus service accessibility (Wu and Hine, 2003), Currie’s model quantifies spatial gaps in PT supply based on social needs (Currie, 2010), Mamun et al.’s method defines PT opportunity space (Mamun et al., 2013). When analyzing these complex datasets, GIS-based approaches may be helpful to integrate land use (e.g., activity locations) and PT network information (e.g., network structure and service frequencies). The SNAMUTS model (spatial network analysis for multimodal urban transport systems) developed by Curtis and Scheurer (2010) provides a network accessibility index composed of the following components: degree centrality; closeness centrality; contour catchment; congested speed ratio; nodal betweenness; and connectivity. Most of the assessment models here presented rely heavily on multiple sources of data, which require extensive efforts and time consuming processes to be obtained, extracted and analyzed, with different PT software systems. For transportation agencies and decision makers the availability of unified frameworks for assessing PT networks based on the minimal data required, namely, PT and the underlying infrastructures (road network, rail network, ferry, seaways, etc.) is important.

This work adopted the unified methodology for extracting, storing and analyzing PT data, first presented by Hadas (2013). The approach is based on Google Transit feeds (Google Transit 2010) which provides an easy to use PT data source, any transportation layers, and an origin-destination estimation matrix. Road network layers are also relatively easy to acquire, whether commercially from NAVTEQ (NAVTEQ, 2012), TeleAtlas (TeleAtlas, 2010), freely from OpenStreetMap project (OpenStreetMap, 2012), or other online sources, such as the US National Transportation Atlas Database (The Bureau of Transportation Statistics, 2013), or by digitizing the networks from published maps.
In this paper three connectivity indicators were used to identify PT connectivity issues and flaws: transportation network coverage (direct and indirect) and origin-destination stop transfer potential. Spatiotemporal properties of connectivity measures were analyzed and compared, obtaining useful information for decision makers.

This work is organized as follows. Section 2 presents the detailed formulation of the data required, connectivity indicators, and the detection of failures methodology. A case study, analyzing the accessibility and connectivity to a regional hospital in the Veneto region (Italy), is introduced in section 3. Conclusions and further research are presented in section 4.

2. Public Transport Spatiotemporal Analysis Model

The proposed detection of connectivity failures of a PT system is based on three steps: a) data acquisition, b) spatiotemporal connectivity indicators calculation, and c) sensitivity analysis.

The spatiotemporal analysis is aimed to reflect the public transport components contribution to the connectivity level between origin and destination. Each trip can be characterized by: 1) Ingress and egress to (and from) the system, which are carried-out by a different mode (such as walking, driving a car, or cycling). 2) Trips' temporal properties (ride time and transfer time), and 3) Trip's spatial properties (transfer walking distance). The objectives set were the development of a data-independent model that can be easily adapted to different systems and data sources, and to provide an operational decision tool for the policy maker.

2.1. Data acquisition

Three data sources are required. For the PT network a GTFS (Google Transit 2010) structured data is preferred, as it provides in a standard form the location of stops, the sequence of stops for each route, and the time-tables. Many PT operators, as well as transit authorities provide GTFS based data for the public and researchers. For the transportation network, any GIS based layer is sufficient, as the proposed model uses the transportation layers for visualization purposes, as origin-destination paths are unnecessary for the model. The last data source is an origin-destination demand matrix.

2.2. Spatiotemporal Connectivity Indicators

The connectivity indicators are a revised form of two connectivity indicators, previously developed by one of the authors, namely road coverage level indicator and stop-transfer potential indicator (Hadas, 2013). These revised indicators are: a1) destination-direct stop coverage level, a2) destination-indirect stop coverage level, and b) origin-destination stop-transfer potential. Figure 1(a) illustrates indicator a1, Figure 1(b) illustrates indicator b, while indicator a2 is a combined indicator of Figure 1(a) and (b), as a transfer is required to reach (indirectly) a destination from origin. The main enhancements with regards to the original connectivity indicators, is as follows: a) all indicators are correlated to stops, which represent the PT system, and not the road network. b) all indicators are origin or destination oriented, those, can assess the connectivity of an origin or destination set, not the general connectivity of the network. c) the destination-indirect stop coverage level indicator is a new indicator that has the same dimension of the destination direct stop coverage level indicator, hence they are comparable (which is not the case with the transfer potential indicator. d) the connectivity indictors are being calculated for the zone and municipality level. e) sensitivity analysis is carried out.

![Figure 1. Public transport connectivity indicators](image-url)
2.2.1. Destination-direct stop coverage level indicator

The destination-direct stop level indicator is the aggregate departures from a stop arriving at a set of destinations without any transfers.

$$SC_{o,d} = \sum_{r|i_o, d\in N(r)} f_r$$  \hspace{1cm} (1)$$

Where $o$ is a stop, $d$ is a set of destination stops in a network $G$, $f_r$ the frequency of route $r$, derived from the time-tables, and $N(r)$ is a set of nodes traversed by route $r$. The higher the frequency at a stop, the higher the coverage.

2.2.2. Origin-Destination stop-transfer potential indicator

The origin-destination stop-transfer potential is the average number of potential transfers from a routes departing from a set of origins and routes arriving at a set of destinations. A transfer is possible only if a given maximal walking distance and maximal waiting time are not exceeded.

Transfers between routes are a common practice in modern PT networks, even though they detract from the convenience and smoothness of trips. Since a passenger is apt to associate a high service level with ease of transfer, assessing the transfer potential of a PT system is crucial. Since transfers are attributes of both space and time (Hadas and Ranjitkar, 2012), the assessment of transfer potential is based on the possible departures within a specified time window and a walking distance, as illustrated in Figure 1(b). The following equations formally define the potential.

$$xp^{r_1, r_2, x}_{s_1, s_2, o, d} = \begin{cases} 
\frac{d(s_1, s_2)}{ws} \leq T_{r_2}^{s_2,x} - T_{r_1}^{s_1,x} \leq \Delta T_{\text{max}} & |d(s_1, s_2)| \leq d_{\text{max}} \\
\text{otherwise} & sq(o, r_1) < sq(s_1, r_1), sq(d, r_2) > sq(s_2, r_2), x \in P(r_1), y \in P(r_2) 
\end{cases}$$  \hspace{1cm} (2)$$

$$XP_{o,d} = \sum_{x} \frac{\sum_{r_1 \in S(r_1)} \sum_{r_2 \in S(r_2)} \sum_{x_1, x_2} xp^{r_1, r_2, x}_{s_1, s_2, o, d}}{|x|} \hspace{1cm} (3)$$

where $r_1, r_2$ are two routes, $x, y$ trips of route $r_1, r_2$ respectively; $s_1, s_2$ are two stops; $o, d$ are origin and destination sets respectively; $sq(s, r)$ is the sequence of stop $s$ in route $r$; $S(r)$ is the set of stops for route $r$; $T_{r_1}^{s_1,x}$ is the arrival or departure time of route $r_1$ at stop $s_1$ and trip $x$; $ws$ is the walking speed; $P(r)$ the trip set of route $r$; $\Delta T_{\text{max}}$ is the maximal walking and waiting for a transfer; $d(s, t)$ is the distance between stop $s$ and stop $t$; and $d_{\text{max}}$ is the maximal walking distance between two stops.

Equation (2) calculates the possibility of a transfer from trip $x$ of route $r_1$ (that departs from stops $o$) at stop $s_1$ to all other routes (that will arrive at stops $d$) within a maximal walking distance. Such a transfer is possible if the time between arrival and departure is not smaller than the walking time and not larger than the maximal waiting time. Equation (3) aggregates all possible transfers per trip $x$ and then averages for the stop, that connects stops $o$ and stops $d$. Thus, for an average stop-transfer potential of 3, a passenger alighting will have 3 transfers available.

2.2.3. Destination-indirect stop coverage level indicator

The destination-indirect stop level indicator is the aggregate departures from a stop arriving at a set of destinations with a transfer, with maximum wait and walk. Equation (4) is a revised version of equation (1), taking into account a possible transfer (equation (5)), for a feasible multi-leg trip.
2.2.4. Aggregated indicators

Based on an estimated origin-destination demand matrix, it is possible to calculate the aggregated and weighted indicator for different zone levels (specific origin to aggregated destinations, aggregated origins to specific destination etc.).

2.3. Sensitivity Analysis

In order to provide the policy maker with a decision tool that detect PT connectivity failures, a sensitivity analysis approach is selected. Altering the values of the key spatiotemporal variables, directly effect the indicators output. As those key variables reflect the PT system's level of service, it is possible to identify connectivity and accessibility problems resulting from the system's spatiotemporal structure. Those key variables are the ingress and egress distance, the transfer waiting time, and the transfer walking distance.

For ingress and egress, let \( O_{i,j} \) and \( D_{i,j} \) be origins and destinations sets respectively, from origin \( i \) (or to destination \( i \)), with increasing distance from the origin (or destination such as set \( j+1 \) completely contains set \( j \) (denote as \( o \) and \( d \) in equations (1)-(3)). For example, Figure 3 illustrates a set of 10 stops, with stops 611, 612 being the destination. Hence it is possible to construct the following 3 sets: \( D_{1} = \{ 611, 612 \} \), \( D_{2} = \{ 611, 612, 394, 395, 435, 436 \} \), \( D_{3} = \{ 611, 612, 394, 395, 435, 436, 370, 371, 372, 373 \} \), each provides additional connectivity options, with lower level of service. Hence, it is possible to define the tuple \( s = \{ j, D_{\max}, T_{\max} \} \) as a scenario related to the three key spatiotemporal variables. Furthermore, let \( CI_{s} \) be the value of on the above mentioned connectivity indicators as a function of scenario \( s \). If \( CI_{j+1} \geq CI_{j} \), then it is fair to assume that the additional reduction with level of service significantly increases connectivity, hence a detailed spatiotemporal analysis is required, based on the changed variable. If \( D_{\max} \) is the cause, then relocation of connecting stops at the transfer area is required. On the other hand, if \( T_{\max} \) is the cause, then frequency increase or time-table synchronization are required.

3. Case Study

The paper analyzed the PT connectivity to a regional hospital in the city of Dolo (Province of Venice, Italy). The area of study is the hospital service area which includes several municipalities in the Province of Venice (Table 1). A population of 12,7850 inhabitants is served by the hospital (Italian National Institute of Statistics, 2001).

The main transport facilities (roads and railways), for both goods and people, are mainly used by traffic from Padova to Venice and vice versa. At the time of the evaluation, about 20% of commuter trips were made using existing public transport system which consisted of buses travelling along the main road connecting the two cities. Some details concerning the extension of primary roads (arterial and collector roads), and secondary roads (local roads) are reported in Table 1.

Figure 2 shows the bus network serving the study area, with a detailed representation of bus stops (circles). For the analysis, a typical mid-week day was selected (Tuesday) for the morning peak (7AM-9AM). Each bus stop size represents the destination-direct stop coverage indicator level, where the destination set is the set of all stops, meaning the indicator reflects the area's overall connectivity. It is evident that most of the PT flow is on the main road connecting Padova and Venice, with lesser connectivity throughout the case study area.

\[
\overline{SC}_{a,d} = \sum_{j \in N(r)} f_{r} \quad \text{equation (4)}
\]

\[
xp = \sum_{x} xp_{r_{1},r_{2},o,d} \quad \text{equation (5)}
\]
Table 1. Characteristics of the area of study.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Campagna Lupia</td>
<td>6,950</td>
<td>79</td>
<td>28.89</td>
<td>47.57</td>
<td>19.27</td>
<td>27</td>
</tr>
<tr>
<td>Campolongo Maggiore</td>
<td>10,350</td>
<td>440</td>
<td>25.65</td>
<td>80.48</td>
<td>22.01</td>
<td>34</td>
</tr>
<tr>
<td>Camponogara</td>
<td>12,950</td>
<td>606</td>
<td>23.00</td>
<td>57.76</td>
<td>22.13</td>
<td>52</td>
</tr>
<tr>
<td>Dolo</td>
<td>15,000</td>
<td>621</td>
<td>40.38</td>
<td>76.63</td>
<td>19.57</td>
<td>46</td>
</tr>
<tr>
<td>Fieso d‘Artico</td>
<td>7,750</td>
<td>1,228</td>
<td>10.40</td>
<td>33.32</td>
<td>5.06</td>
<td>14</td>
</tr>
<tr>
<td>Mira</td>
<td>38,550</td>
<td>390</td>
<td>85.22</td>
<td>186.44</td>
<td>41.49</td>
<td>102</td>
</tr>
<tr>
<td>Pianiga</td>
<td>12,000</td>
<td>598</td>
<td>34.43</td>
<td>60.29</td>
<td>9.52</td>
<td>23</td>
</tr>
<tr>
<td>Stra</td>
<td>7,600</td>
<td>861</td>
<td>16.12</td>
<td>35.89</td>
<td>9.03</td>
<td>21</td>
</tr>
<tr>
<td>Vigonovo</td>
<td>9,900</td>
<td>775</td>
<td>15.14</td>
<td>67.17</td>
<td>9.58</td>
<td>18</td>
</tr>
</tbody>
</table>

As the objective of the case study is the connectivity analysis to the regional hospital, only destination sets were constructed as follows. Set 1: the bus stops located at the hospital entrance \((D_1 = \{611,612\})\). Set 2: the addition of the bus stops located to the north of the hospital \((D_2 = \{611,612,394,395,435,436\})\). These stops are approximately 500 meters of walking distance, with relatively easy access. Set 3: the addition of the bus stops located to the south of the hospital \((D_3 = \{611,612,394,395,435,436,370,371,372,373\})\). Walking from these bus stops to the hospital requires street crossing. Set 4: all bus tops of the region. Figure 3 presents the various bus stops locations with relation to the hospital and road network.
The development of the O-D matrix was based on the transportation zones' census data (Italian National Institute of Statistics, 2001), as it was assumed that all inhabitants are prospect visitors to the hospital. Each bus stop was assigned with the nearest zone's population, as well as the ingress distance. All stops within a 100 meters were treated as a stop cluster, and the same population and distance were used. This procedure was used to reflect the potential to reach all destinations from all nearby stops, including the inbound and outbound routes.

3.1. Stop coverage and ingress analysis

Examining the destination-direct stop coverage (Table 2) reveals that a direct access to the hospital from most of the municipalities is almost nonexistent. The column "departures" represents the number average departures from 7AM to 9AM, and the "ingress distance" is the average distance from the zones' centroids to the stops. It is evident that connectivity increases if destination set 2 is used, hence for lower level of service (extra walking distance), better connectivity is achieved. Furthermore, from Figure 4 it is clear that destination set 2 has better coverage than set 1, with particular emphasis on Pianiga, which does not have access to the hospital bus stops.
Table 2. Weekday (Tuesday) morning peak (7AM-9AM) destination-direct bus stop coverage indicator.

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Destination set</th>
<th>Direct stop coverage</th>
<th>Ingress distance [meters]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Campagna Lupia</td>
<td>0.3</td>
<td>2.3</td>
<td>2.6</td>
</tr>
<tr>
<td>Campolongo Maggiore</td>
<td>0.8</td>
<td>1.5</td>
<td>1.7</td>
</tr>
<tr>
<td>Camponogara</td>
<td>0.3</td>
<td>1.9</td>
<td>2.0</td>
</tr>
<tr>
<td>Dolo</td>
<td>1.7</td>
<td>7.1</td>
<td>7.3</td>
</tr>
<tr>
<td>Fieso d'Artico</td>
<td>14.6</td>
<td>14.6</td>
<td>14.6</td>
</tr>
<tr>
<td>Fosso'</td>
<td>2.1</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Mira</td>
<td>2.7</td>
<td>7.6</td>
<td>7.6</td>
</tr>
<tr>
<td>Pianiga</td>
<td>5.0</td>
<td>5.0</td>
<td>5.9</td>
</tr>
<tr>
<td>Stra</td>
<td>8.6</td>
<td>9.1</td>
<td>9.1</td>
</tr>
<tr>
<td>Vigonovo</td>
<td>3.3</td>
<td>4.7</td>
<td>4.7</td>
</tr>
</tbody>
</table>

The departures indicator confirms the structure of the existing PT service: it has been designed to accommodate the demand along the Padova-Venice route, rather than to serve demand related to Dolo hospital. This is well visible by the values of “departures” with reference to both destination set 1 and 2 (sets of bus stops along the regional road). The highest values of these indicators (see column 2) are those associated with the municipalities located along the regional road connecting Padova and Venice (Fiesso d’Artico, Strà, Mira, and Dolo). Furthermore, the “ingress distance” for some municipalities (i.e. Fiesso d’Artico) is extremely high: with reference to the existing bus stops in its area, this municipality is the best connected to Dolo (departures) but the ingress distances to the PT system are very long. This reflects that the bus stops are not effectively distributed over the municipality area.

3.2. Transfer potential analysis

The results of the weighted average transfer potential for each municipality for 8 scenarios, clearly show that inhabitants wishing to reach the hospital by a multi-leg trip, will suffer a long waiting time, as well as a long walking distance in order to perform transfers (Figure 5).
3.3. **Indirect stop coverage analysis**

The analysis of the indirect stop coverage provides similar conclusions, to obtain a higher connectivity, lower level of service is required (Figure 6).

4. **Conclusions and further research**

When combining the direct and indirect coverage level indicators (for destination set 2), a range of 2.5-12.5 departures per hour are calculated, meaning that for some municipalities the trip to the hospital by PT is not a high quality option, as it requires long ingress and egress walking, long waiting time and transfer time.
The model that was developed is an easy-to-use tool enabling decision-makers to analyze the connectivity of PT networks and to detect connectivity flaws.

Performing what-if analysis, by varying ingress, egress, and transfer distances, as well as waiting time, enables the decision-maker to measure the level of service impact on the connectivity.

The GIS tool assists with the identification of the exact location of spatial-related flaws, such as inefficient stop location, and time-table synchronization points.

With reference to the study case, future directions of the research will be devoted to:

- Extensive sensitivity analysis.
- Evaluation of other points of attraction (school district)
- A more detailed description of the walking paths considering the spatial properties of the transfers (street crossings), as suggested by Hadas and Ranjitkar (2012).
- Inclusion of demographic data.

Acknowledgements

The authors would like to thank the Department of Transportation, Province of Venice and Consuelo Bolzonella (ATraKI S.r.l.) for their kind support.

References

NAVTEQ. 2012. What is NAVTEQ Map Data.