Temporary User-Centred Networks for Transport Systems

António A. Nunes [a,*], Teresa Galvão Dias [a], Chris Zegras [b], João Falcão e Cunha [a]

[a] INESC TEC, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
[b] Massachusetts Institute of Technology, Department of Urban Studies and Planning, 77 Massachusetts Avenue, Cambridge, MA 02138
[*] Corresponding author. Tel.: +351 22 508 1400; Fax.: +351 22 508 14 40; E-mail address: antonio.nunes@fe.up.pt

Keywords: spatiotemporal data; intelligent transport systems; social networks; collaboration

Abstract

The concept of Temporary User-Centred Networks (TUNs) for transport systems is introduced. Affinity in these networks is defined as the time-specific degree of equivalence between travel patterns of users in the system. TUNs reveal latent social structures typically invisible to their users, enabling circumstantial collaboration opportunities amongst them. To make TUNs explicit we quantify affinity as a combination of two measures: journey similarity and journey substitutability. In the urban public transport domain, TUNs enable the diffusion of knowledge across the system in real-time. This can assist passengers adjusting travel decisions to their preferences and objectives according to service status. An enriched Automated Fare Collection (AFC) system dataset is used to demonstrate the market potential of TUNs in the urban public transport domain.

List of Figure captions

Fig. 1. Journey similarity (SIM) and journey substitutability (SUB).
Fig. 2. Journeys with SIM and SUB scores.
Fig. 3. Journey paths with maximum SUB at shorter substitute portion.
Fig. 4. SUB according to distance between location points within SPODs for various levels of substitution potential.
Fig. 5. Journey transaction throughout the day.
Fig. 6. Timeline of active journeys.
Fig. 7. Active passenger journeys at 10:00.
Fig. 8. Aggregate results.
Fig. 9. Example passenger’s TUN evolution.
Fig. 10. Journey paths of passenger and highly *relevant* peers at 10:00.

List of Table captions

Table 1
Relevant peers for an example passenger journey

Table 2
Analysis of TUNs across time instants at various *relevance* threshold levels

List of Keywords

spatiotemporal data
intelligent transport systems
social networks
collaboration

Temporary User-Centred Networks for Transport Systems

António A. Nunes [a,*], Teresa Galvão Dias [a], Chris Zegras [b], João Falcão e Cunha [a]

[a] INESC TEC, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
[b] Massachusetts Institute of Technology, Department of Urban Studies and Planning, 77 Massachusetts Avenue, Cambridge, MA 02138
[*] Corresponding author. Tel.: +351 22 508 1400; Fax.: +351 22 508 14 40; E-mail address: antonio.nunes@fe.up.pt

Keywords: spatiotemporal data; intelligent transport systems; social networks; collaboration

1. Introduction

Developments in information and communication technologies (ICT) have paved the way for new forms of collaboration across diverse service domains (Greengard 2011). This evolution has allowed for various kinds of underlying interpersonal affinity to materialise in actual ties between users. Personal ties such as friendships or professional relationships are tangible social structures replicated in ICT-based social networking services. Yet affinity also exists in more abstract forms, with resultant ties increasingly facilitated by the ongoing evolution in ICTs. For example, interpersonal affinity may be a shared interest or ideology revealed across a social media platform, such as a recommendation or rating system. Whatever form affinity takes, the resultant ties between users have already enabled unprecedented collaboration-based value creating processes by harnessing knowledge distributed amongst users themselves (O'Reilly 2005). This can be viewed as a step beyond value "co-creation", in which suppliers engage customers as a co-creator of value; instead, ICT-based social network services represent a form of consumer co-creation of service value (e.g., Payne et al. 2008).

This trend has emerged across service industries, from retail (e.g., reviews on Amazon.com, ratings on eBay) to tourism (e.g., reviews on TripAdvisor.com), to health (e.g., crowd creation on CrowdMed), and to media (e.g., crowd creation on Wikipedia, film ratings on IMDb), just to name a few. The users involved, which may be consumers, service providers, or members of the public in general, create value in a variety of forms, including decision-making support with review and rating platforms, and problem-solving know-how compiled in crowd creation systems. In the transport systems domain specifically, applications already exist that leverage interpersonal affinity as the shared consumption of a particular service, aggregating reviews to assist travel choices. E-hailing applications that feature driver and passenger ratings (e.g., Uber) illustrate this well, by letting consumers and service providers each review their experiences of the other. Still, much potential for collaboration in transport systems remains to be explored. In particular, drawing greater value from interpersonal affinity in transport systems most likely must account for the inherent spatiotemporal dynamics.

This paper introduces and operationalizes the concept of Temporary User-Centred Networks (TUNs) in which *affinity* is defined as the time-specific degree of equivalence between travel patterns of users in a transport system. TUNs aim to identify circumstantial user-based collaboration opportunities to facilitate the diffusion of knowledge spread across a system in real-time. TUNs have the potential to produce benefits ranging from improved travel experiences to the reduction of carbon emissions. For example, urban public transport TUNs may increase the visibility of service status for passengers, permitting them to adjust travel decisions in real-time to better fit their preferences and goals. In private motorised transport TUNs could facilitate dynamic routing choices by car users based on the distribution of real-time traffic information. TUNs may also help urban cycling beginners identify experienced peers with equivalent travel patterns, to become their training buddies. TUNs may even provide an entirely new approach for the assignment and routing of Demand Responsive Transport (DRT).

TUNs materialise spatiotemporally dependent ties between users, using two *affinity* measures that are tangible and quantifiable. TUNs differ from existing ICT-based social networking services where the spatial element is typically absent or, in fewer cases, stationary (for example, to connect guests at a local attraction or event). TUNs generate opportunities to leverage knowledge that is distributed across the transport system at any time through collaboration between users. This may expand the knowledge users have of the transport system by increasing the type, availability and timeliness of information at their disposal. This will ultimately help improve travel experiences and satisfaction, of clear relevance to users, agencies, and policy makers. This paper demonstrates the concept of TUNs and evaluates their potential, based on the proposed *affinity* measures, using an urban public transport system as a case study.

2. Leveraging distributed knowledge in transport

In recent years, attempts have been made to capitalize on increasingly ubiquitous personal communication, computing and sensing devices (e.g., the smartphone) to leverage the knowledge distributed amongst transport system users. Transport agencies, for example, are ever more present in social media services (Austin 2010; Gault et al. 2014). This facilitates engagement with citizens and enables the distribution of timely information (Bregman 2012). Yet these social media services tend to be top-down (i.e., from supplier to user), and structured according to forms of interpersonal affinity poorly related to short-term travel patterns (Cho et al. 2011), so it becomes difficult to source and filter information and make it reach the target audiences in real-time (Nunes et al. 2011). This limitation has prompted recent studies on harvesting transport-related information from social media using text-mining techniques (Carvalho et al. 2010; Gal-Tzur, Grant-Muller, Kuflik, et al. 2014; Gal-Tzur, Grant-Muller, Minkov, et al. 2014) and on developing smartphone applications that spatially structure and aggregate crowdsourced data from travellers in a transport system, such as Waze, Moovit (Olson 2014), and Tiramisu (Steinfeld et al. 2011; Steinfeld et al. 2013). At best, these data-centred approaches only superficially reflect affinity between users in the transport system, thus providing limited ability to proactively filter and forward information to them based on their travel patterns.

So in spite of these efforts, user collaboration-based value creation processes have yet to fulfil their potential in passenger transport. In public transport, some evidence has emerged of companies and agencies taking a co-creation approach to service enhancement. Gebauer et al. (2010), for example, argue that the Swiss federal railway operator has relatively recently moved towards a value co-creator approach in interacting with its customers in problem-solving, co-designing, among other activities. In the USA, many public transport agencies, partly due to financial constraints, have also taken what one could call a value co-creation approach with users, releasing real-time data for use by third-party developers to create real-time information services for users (e.g., Brakewood et al. 2015). The relatively modest amount of co-creation activity in this realm may be partly attributable to the fact that public transport services tend to be monopoly-provided or, at least, oligopolistic (Evans 1991). Adding to this institutional setting, the spatiotemporal nature of transport-related information makes it highly contextual and transient (Wolfson & Xu 2010). Leveraging distributed knowledge through collaboration is therefore a challenge that remains to be addressed. For example, information about a traffic incident is only relevant to those who are geographically affected by it, and only over a short time period until it gets resolved.

Still the user-based value creation potential of distributed knowledge in a transport system has been revealed across a range of studies. The proliferation of smartphones has rapidly increased the spatial and temporal resolution, quality, speed and communicability of user-generated information on mobility systems, even in the most traditionally data-sparse settings (e.g., Zegras et al. 2015). Research on the information needs of travellers has highlighted the importance of early warnings regarding unscheduled disruptions and information about safety and comfort-related aspects associated with travel alternatives (Caulfield & O'Mahony 2007; Chorus et al. 2006; Chorus et al. 2007; Windmiller et al. 2014). Those types of information have the potential to raise travel satisfaction by managing travel time expectations (Li 2003), and allow travellers to adjust travel choices according to their preferences and needs (Costa et al. 2012; Lathia et al. 2013). Whilst it may be unfeasible for transport agencies to provide such information in real-time, it can be sourced from travellers that are scattered across the transport system, who are the first to observe most events. It has also been shown that the availability of information obtained from others has potential to raise the efficiency and utility drawn from travel choices, particularly in relation to non-recurrent travel behaviours (Iryo et al. 2012).

3. Conceptualising relevance

TUNs represent affinity as spatiotemporally dependent ties between users in a transport system. The concept of TUNs offers a solution to the problem of selecting, for any user, other users (peers) with the highest collaboration potential within the system at a specific instant. The creation of a TUN therefore requires the identification and ranking of a user's *relevant* peers in real-time. In this context, *relevance* is an indicator of *affinity* strength between peers. This problem shares some basic characteristics to that of searching for information on the Internet. Kleinberg (1999) highlighted that the inherent subjectivity of the notion of relevance, which determines the ranking of search results and underpins search quality, made the problem of discovering pages from a given query challenging. TUNs hinge upon a similarly crucial concept of relevance but, by virtue of timeliness, the focus shifts from the actual information to the ranking of users who generate that information. Subjectivity, however, remains, as does the challenge.

Clustering methods have often been applied to problems dealing with movement patterns in spatiotemporal data (Gudmundsson et al. 2008), in particular to the automatic detection of traffic congestion events (Anbaroglu et al. 2014; Li et al. 2007). A trajectory clustering method, however, is insufficient for identifying a user's relevant peers, in terms of those whose travel patterns have either similar or alternative characteristics, or a combination of both. An underlying premise of a TUN is that a user will likely derive greater benefits from collaborating with peers that either have similar behaviours, or experience feasible alternatives to achieve the same travel goal. This implies that a user will be relevant to another user travelling simultaneously if at least a part of their journeys is bounded by two common locations, irrespective of their paths. This notion of relevance renders trajectory clustering methods unfeasible as two journeys may not share origins nor destinations and follow different paths, and still remain significantly relevant to each other. To better capture relevance, we propose two measures of *affinity*, defined in the following section.

4. Affinity as spatiotemporal relevance

Affinity in the context of TUNs is the time-specific degree of equivalence between travel patterns. Pattern here generalises two types of instances that differ in relation to their time frame: travel history, relating to past travel behaviours; and travel intentions, which are future travel behaviours that have somehow been expressed by travellers. Travel history may be captured, for example, through fare card usage (Pelletier et al. 2011), GPS traces (e.g., Zheng et al. 2008), or mobile network data (e.g., Calabrese et al. 2013; Chen et al. 2014; Järv et al. 2014). Intentions, which are tentative, may be obtained or inferred, for example, from usage of journey planning applications and/or from appointments detailed in electronic agendas. Both travel history and intentions are useful to predict future behaviours of travellers. This provides the necessary information for the creation of TUNs based on the identification and ranking of a user's relevant peers in real-time. The *relevance* score between two users simultaneously in the system is proposed as the sum of two *affinity* measures, called journey *similarity* (SIM) and journey *substitutability* (SUB), which we define to be mutually exclusive over a given portion of a journey path.

4.1. Journey similarity (SIM)

SIM represents the portion of simultaneous journey paths of two users that share spatial characteristics (Fig. 1, left). SIM measures the degree of similarity between travel patterns. In public transport, similar journey paths are those of two users travelling along a common portion of the same route and in the same direction. In the context of private transport, whether by motorised vehicle or bicycle, similar journey paths are those of two users who share the same routing between two locations. The proposed SIM measure is asymmetrical. It looks at the portion of a journey path that is contained by another similar journey path, and therefore it is a function of the overall length of the first. For example, the path of user 1 may be perfectly similar to the path of user 2 but the converse may not hold. This would be the case where the path of user 1 is longer and contains the path of user 2. In public transport, for example, passenger 1 would have the information about the entire length of the journey of passenger 2, whereas passenger 2 would only be partially relevant to passenger 1.

The SIM between users in a transport system is represented by a square matrix ***S***, the dimensions of which are given by the number of users simultaneously in the system. The entries of matrix ***S*** at a given moment *t* are determined by Eq. (1). The SIM of one user to another equals the common length of their journey paths divided by the total length of the first. Hence matrix ***S*** is asymmetric. The SIM of two journey paths is a ratio that varies between 0 (dissimilar) to 1 (perfectly similar). Lengths can be expressed in actual distance, Euclidean distance, or in journey segments, depending on the available data, purpose, and domain of application.

$$s_{ij} = \frac{c_{ij}}{l_i} \tag{1}$$

Where:

i is the journey path of the user
j is the journey path of a peer

s_{ij} is the SIM of j to i
c_{ij} is the common length between i and j
l_i is the length of i

Place Fig. 1 about here.

4.2. Journey substitutability (SUB)

SUB represents the portion of simultaneous journey paths of two users where SIM is 0, but they serve at least two of the same potential origin-destination (O-D) pairs (Fig. 1, right). SUB measures the degree of path substitutability. In public transport, substitute journey paths are those of two users travelling along a different route or direction but having at least two shared potential origins or destinations (SPOD). In the context of private transport, substitute journey paths are alternative routings between two SPODs. The proposed SUB measure is also asymmetrical, for the same reasoning as for the SIM measure as described above. Simultaneous journey paths for two users may present a combination of SIM and SUB, however portions with a SIM score cannot have a SUB score and vice-versa. In other words, we define the *affinity* measures to be mutually exclusive over a given portion of a journey path, as illustrated by Fig. 2: user 2's first journey stage is similar to a portion of user 1's single stage journey; user 2's second journey stage is a substitute of a downstream portion of user 1's journey.

Place Fig. 2 about here.

The SUB between users in a transport system is represented by a square matrix $\boldsymbol{B}$, the dimensions of which are given by the number of users simultaneously in the system. Calculating the entries of the SUB matrix $\boldsymbol{B}$ at a given moment t is more complex, for two reasons.

One reason for SUB's additional complexity is the need to define the concept of shared potential origin or destination (SPOD), best illustrated with an example. Consider alternative routes for a public transport passenger in two different modes, bus and train. The bus departs from a stop adjacent to the train station. The departing locations, while not the same, are near enough to be considered a shared origin, so both routes are true substitutes. This instance can be generalised for private transport since alternative routes may not overlap yet still have nearby passing points (i.e., SPODs). The extent to which alternative routes actually *share* a potential origin or destination, depends upon the proximity of their respective Os or Ds. We operationalize proximity with the notion of *vicinity*, akin to a catchment area. Vicinity is a fixed parameter representing a cut-off distance within which two points may be considered to represent a SPOD. Nonetheless, the real substitutability of alternative routes decreases as the distances between points in the SPOD grow, reducing the degree to which the points *share* origin or destination potential and, thus, reducing the substitutability of the routes. So, even if two points are in the *vicinity* of each other, the resulting SUB of route alternatives is inversely proportional to the distance between the points. For that reason, the proposed SUB measure incorporates a penalty factor, a function of distance among points within the *vicinity*.

The second source of additional complexity in calculating SUB is that substitute journey paths may have more than two SPODs. In such cases, various combinations have to be considered, because considering just the first and the last SPODs between two journey paths could underestimate SUB. This calculation is sensitive to the penalty factor, described above. Thus, SUB is influenced by a trade-off between lengths of the substitute portions and penalty factors (see Fig. 3).

Place Fig. 3 about here.

The entries of matrix ***B*** at a given moment *t* are determined by Eq. (2):

$$b_{ij}: Max\left[\frac{p_{ij}}{l_i} * \left[1 - \frac{\left(d_{O_{ij}} + d_{D_{ij}}\right)}{2k}\right]\right] \quad subject\ to\ d_{O_{ij}} \leq k, d_{D_{ij}} \leq k \qquad (2)$$

Where:

- i is the journey path of the user
- j is the journey path of a peer
- b_{ij} is the SUB of j to i
- p_{ij} is the substitute portion between i and j
- l_i is the length of i
- $d_{O_{ij}}$ is the distance between origins
- $d_{D_{ij}}$ is the distance between destinations
- k is the vicinity parameter

The SUB between users is given by a ratio multiplied by the penalty factor. The ratio is the substitute portion of the users' journey paths divided by the total length of the first. Hence matrix ***B*** is also asymmetric. The ratio depicts substitution potential, the portion of a journey that the can be substituted by the alternative. The penalty factor is the sum of distances between the location points at the SPODs of the substitute portion, divided by *vicinity* doubled and subtracted from 1. The penalty factor tends to 0 as the distances between location points at the SPODs get close to the set *vicinity* parameter, *k*, and equals 1 when those distances are null. This accounts for the reduction of the convenience of an alternative due to the necessity to detour. The multiplication in Eq. (2) ensures that the penalty is applied proportionally to the substitution potential of journey paths. When the penalty factor equals 1, no penalty is applied. Fig. 4 illustrates how SUB varies according to the distances between location points for various levels of substitution potential.

Place Fig. 4 about here.

Considering the first and the last SPODs between two journey paths may not return the highest value for SUB. The mathematical maximisation shown in Eq. (2) avoids underestimating the true SUB by considering all candidate pairs of SPODs between two journeys that may yield the optimum SUB. The SUB of two journey paths varies between 0 (not substitute) to 1 (perfect substitute). As in SIM, lengths can be expressed in actual distance, Euclidean distance, or in journey segments, depending on the available data, purpose, and domain of application.

4.3. Relevance

Relevance in TUNs provides an indicator of the equivalence between simultaneous journey paths of two users. Relevance estimates the degree to which simultaneous travel patterns of two users either share or have alternative spatial characteristics. The *relevance* score between users in a transport system is represented by a square matrix $\boldsymbol{R}$, the dimensions of which are given by the number of users simultaneously in the system. The *relevance* score is obtained for a given moment *t* as the sum of the *affinity* measures SIM and SUB, calculated separately, as shown in Eq. (3). Being the sum of asymmetric matrices $\boldsymbol{S}$ and $\boldsymbol{B}$, $\boldsymbol{R}$ is asymmetric too. Since SIM and SUB are mutually exclusive over a given portion of a journey path, the *relevance* score will range between 0 (not relevant) to 1 (perfectly relevant).

$$\boldsymbol{R} = \boldsymbol{S} + \boldsymbol{B} \tag{3}$$

Where:

$\boldsymbol{R}$ is the relevance matrix
$\boldsymbol{S}$ is the SIM matrix
$\boldsymbol{B}$ is the SUB matrix

Relevance is fundamental for creating TUNs. In fact, for users in a transport system, TUNs are created based on the highest *relevance* scores. Different cut-off rules and their parameters may be defined for this. A cut-off rule may be a minimum *relevance* score or a maximum number of users ranked by their *relevance*, or a combination of both. The score and number of users are the parameters to be defined. The magnitude of a TUN will reduce with the strictness of the cut-off rules. For example, a TUN for a user may be defined to include all other users with *relevance* higher than 0.5, but limited to the top ranking 100 users. This process selects the spatiotemporally most relevant peers to be included in a TUN, and must be continuously updated to consider the inherent dynamics of the transport system.

The *relevance* score reveals potentially useful, yet latent, social structures based on travel patterns. SIM instances may sometimes be visible over time to commuters, those who undertake regular or oft-repeated trips typically between home and work or school. Commuters may notice, for example, familiar faces in public transport. But many other SIM instances are not recurrent and most are likely unnoticeable. SUB is a different case, more likely invisible to commuters because the alternative routes do not share space, only potential origins and destinations. Hence, the combined *relevance* score represents fluid socialities (Büscher et al. 2011) based on actual, yet largely invisible affinity, enabling collaboration opportunities, and facilitating user engagement. Note, however, that the quantification of *affinity* through measuring *relevance* will not suffice; ties and communication channels between users must also be created to fulfil the potential of TUNs. Therefore, materialising engagement requires social ICT-based applications that either enable information sharing or reveal other collaboration opportunities between highly relevant users in real-time, whilst preserving their safety and privacy (e.g., Nunes et al. 2013). Such collaboration between users may be voluntary or automatic, depending on whether they actively provide information or passively allow information extraction from sensors attached to them, like accelerometer data from travellers' smartphones.

5. Case study: TUNs in urban public transport

We illustrate the process of creating TUNs and their respective potential using the case of urban public transport. Public transport has clear relevance because, typically, many passengers simultaneously populate a public transport system at any moment in time. Despite this commonality, these passengers have individual journey purposes and plans, making the behavioural dynamics of these systems complex. Users often have numerous alternatives available, ranging from different modes to different routes for the same mode, and even to various scheduled services on the same route. Since passengers may rely on more than one route to reach their destinations, a large number of available combinations for their travel plans exist.

This complexity contains a rich source of knowledge distributed across the system. Each passenger has limited visibility of the system as a whole, but a unique perspective of her chosen travel alternative that could be useful for the travel decisions of others. A passenger may know, for instance, that a specific vehicle is crowded, a traffic incident is about to delay a service, a route is particularly scenic, or a great street performer is playing at a certain station. A TUN aims to facilitate the diffusion of that collective knowledge in real-time, using personal mobile devices. Using spatiotemporal *affinity* to identify target user audiences, TUNs enable the aggregation and relay of information via users' personal mobile devices. The utility of potential information exchange within a TUN depends on, for example, whether it allows users to adjust travel decisions to better fit their preferences and goals according to service characteristics learned from peers in the system.

The potential for harnessing this collective knowledge is confirmed in light of graph theory. Random graphs consist of a set of N nodes and E edges, in which a pair of nodes has a probability p of being connected by an edge. Take a public transport system where passengers are the nodes, and the connections in all simultaneous TUNs are the edges of a random graph. Paul Erdős and Alfréd Rényi (1960) found the interesting property that when the average degree $<k>$ (number of edges connected to a node) reaches one, a great cluster will almost certainly appear in the graph. Barabási (2003) nicely illustrates this property with an example of gossip spreading at a party; if each person knows at least one other guest then everybody will rapidly learn the gossip. Unlike those exchanging hot gossip at a party, however, all passengers of an urban public transport system will not likely be engaging in some sort of information exchange using their personal mobile devices. Still, if only a small portion engages in such exchange, as long as the average degree $<k>$ of the resulting graph reaches one, knowledge will likely spread. And $<k>$ is indeed typically higher, as illustrated in our case study.

The analysis of complex networks, in particular social networks, has received considerable attention in recent years, reflecting the impact of social networks in the current society, and the availability of large datasets. New concepts and measures have emerged from these developments. Experimental results have been obtained for complex networks that include communication networks, science collaboration, and ecological systems (Albert & Barabási 2002; Barabási 2007). Different network models with roots in graph theory have been studied, particularly random graphs and scale-free models. Random graphs are used in several fields, namely as benchmarks for modelling and empirical studies, since complex networks with unknown organisational principles are often similar to random graphs.

A relevant discovery in random graph theory established that if a certain fraction of nodes were removed from a graph, its connectivity would disappear. This result strongly relates to percolation theory, in which a graph will likely collapse when a certain fraction of nodes or edges fail (Stauffer & Aharony 2003). This raises the question of how many nodes *(N)* or edges *(E)* are necessary for maintaining a cohesive graph. In the context of our work, this number is of great interest, providing insight into the minimum number of passengers required in the transport network to create significant TUNs. Section 5.5

describes the experimental analysis carried out on several networks obtained from our case study.

5.1. The Andante system

We use an enriched dataset from *Andante*, an Automated Fare Collection (AFC) system that covers bus, metro, and railway services in the metropolitan area of Porto (Portugal). *Andante*'s fare cards can be used across all service operators. All participating services operate under an entry-only configuration, which means that fare cards are only read at the beginning of each journey stage. A journey stage corresponds to a transaction record in the dataset. A complete journey has one or more stages, which start every time a passenger enters a new route or vehicle of the same route. Being an entry-only system, additional logic is required for estimating the destination of those journeys stages (Barry et al. 2009).

Porto's main bus service operator, *Sociedade de Transportes Colectivos do Porto, SA* (STCP), provided the data for this case study. STCP runs the vast majority of routes within the city and in the surrounding metropolitan area. The available dataset contains transaction records from April 2010. At that time the average daily number of journey transactions was approximately 330,000 on weekdays and 130,000 on weekends, distributed across 67 bus routes and 2,365 bus stops. Approximately 38% of journey transactions in STCP bus services related to *Andante* fare cards. Each transaction record contains several data attributes, including the travel card serial number, the transaction timestamp, the route, direction of travel, bus stop code, and the vehicle on which the transaction took place.

The dataset was previously enriched with the inferred destinations and time of arrival of passenger journeys stage using a method described by Nunes et al. (2015, in preprint). This method is based on two key assumptions introduced to the literature by Barry et al. (2002). These consist fundamentally of considering that the origin of a journey is the most likely destination of the previous journey (i.e., continuity), and that the first origin of the day is the most likely destination of the last journey of the day (i.e., circularity). Furthermore, the method used four spatial validation rules, the last two being newly introduced and owing to specificities found in the *Andante* system. The spatial validation rules test the key assumptions at disaggregate level to raise the quality of inference results. The first rule evaluates if the origin and inferred destination of journey are reasonably apart. The second rule evaluates if the destination of a journey is feasibly within walking distance of the origin of the following journey.

The third rule was newly introduced and relates specifically to entry-only systems with distance-based fare structures. It evaluates if the journey destination falls within the allowable travel bounds of the respective fare. The fourth rule was also newly introduced and relates specifically to time-based entry-only systems. If during a journey a passenger taps the travel card on a reader to read from its screen the time left for travelling, this rule checks if the recorded location is upstream from the inferred journey destination. The first two spatial validation rules were prolific in the identification of false positives to raise the accuracy if inference results, and the last two rules supported the validity of the key assumptions for the vast majority of journeys in the case study. Accurately inferring destinations is vital for the reliability of TUNs. The method also enhanced the data with a flag indicating if a transaction relates to a new journey or to an additional stage of a previous journey.

5.2. The case study data subset

Part of the enriched AFC system dataset was used to demonstrate the potential of creating TUNs in real-time. The TUNs users are passengers. We chose a specific date and time from the enriched AFC system dataset to simulate an instant in real-time, and then repeated 10-minute increments to capture the evolution of TUNs. In this case, we chose Wednesday, 7th April 2010 at 10:00 as the initial time instant, with subsequent increments of 10:10, 10:20, and 10:30. The date represents a typical mid-week working day and a regular school day, with a total number of journey transactions close to the weekday average for the month. The time chosen represents the inter-peak period, with the lowest level of journey transactions for that particular day (Fig. 5). This aims to provide the worst-case scenario for the period of interest bounded by the morning and evening peak times.

Place Fig. 5 about here.

Ongoing journeys are simulated for each time instant, considering all journeys that have at least one stage that started in the previous 30 minutes and have not been completed. This parameter is marginally conservative since the average duration of journeys using bus as the main mode of transport in Porto is 34 minutes (Pinho & Vilares 2009). In real-time, this would also require predicting destinations from the travel history of passengers because ongoing journeys are not completed. But passengers who are about to start their journeys should also be included, to benefit from information that allows them to adjust travel plans according to service status. These are travel intentions, which in real-time either must be captured from aforementioned sources, or predicted from passengers' travel histories. Journey intentions are simulated as those starting within 15 minutes of each time instant. The aggregate of ongoing journeys and journey intentions will hereafter be referred to as active journeys (Fig. 6). To summarise, the data selected for the simulation of real-time consists of the enriched AFC system transaction records that represent ongoing journeys and journey intentions.

Place Fig. 6 about here.

Place Fig. 7 about here.

Fig. 7 presents a snapshot of active journeys in the data subset at 10:00 over a map of the *Andante* system. The points represent bus stops in their geographic location, sized proportionally to the number of departures and arrivals. The lines represent passenger journeys between the respective origin and destination stop pairs. The size and colour of lines depict the number of passenger journeys between each pair. The *Andante* system is divided into travel zones. Those relevant to the present case study are labelled and shaded. As expected, the zone corresponding to Porto's city centre, C1, concentrates a major share of passenger journeys. Not only does this area have higher activity levels and urban density, it is also where most STCP routes pass by.

5.3. Creating TUNs

TUNs were created for all active passengers for the initial time instant at 10:00. This required separate calculation of the SIM matrix $\boldsymbol{S}$ and the SUB matrix $\boldsymbol{B}$ as per Eq. (1) and

Eq. (2) respectively, and adding the values to obtain the *relevance* matrix ***R*** as in Eq. (3). The same process was repeated for the subsequent time instants at 10:10, 10:20, and 10:30. The dimension of the resulting square *relevance* matrices varied slightly, ranging between 2,739 and 3,012 for the inter-peak time instants considered. This number represents simultaneous active passengers in the enriched AFC system dataset, for whom the destination could be successfully inferred using the method described by Nunes et al. (2015, in preprint). Given that only 38% of STCP transactions in April 2010 were *Andante*, transaction information on the complete set of journeys would expectedly yield *relevance* matrices of considerably greater dimensions. Likewise, availability of a multi-modal dataset (e.g., including trains) would further increase the number of simultaneous active passengers.

5.4. Case study results

Fig. 8 summarises the estimated market potential of TUNs based on the initial time instant at 10:00. There were 56 bus routes in operation served by a total of 331 buses, with 1389 bus stops as the origin or destination of 4060 active passenger journey stages. We set the *relevance* threshold at 50%, meaning that the TUN of a passenger includes all peers with *relevance* of at least 50%. We expressed lengths in journey segments, the span between two bus stops. We set the *vicinity* parameter to a Euclidean distance of 80 meters, representing an approximate walk catchment area of 1 minute at 4.8km/h.

Place Fig. 8 about here.

At 10:00, each passenger had on average 39 peers in their network. Of these, 18 had similarity and 20 had substitutability, with at least 50% of SIM and SUB respectively. If considering the entire set of journeys, not just those captured by *Andante*, the number of routes and buses in operation would likely remain unchanged, whereas active bus stops could be slightly higher. However, the average number of relevant peers in each TUN would certainly increase, because we only use *Andante* users as candidates for each passenger's TUN. Regarding the sensitivity of the *relevance* threshold, if raised to 60% the average number of peers per TUN in the data subset at 10:00 would drop from 39 to 29, to 23 for 70%, to 18 for 80%, to 14 for 90%, or to 12 for 100% (perfectly relevant peers) (Fig. 8). Stricter thresholds reduce the magnitude of resulting TUNs.

Place Table 1 about here.

These results provide insight into the market potential of TUNs. But an understanding of their spatiotemporally dynamic nature requires deeper insight. To illustrate, we present the results from a single typical passenger journey. The selected passenger journey consisted of two stages using two bus routes, and a total of 22 journey segments, 18 in the first stage and 4 in the second stage. Table 1 lists the relevant peers included in the TUN in each time instant considered, their SIM, SUB, and *relevance* score. Each peer is identified via a two-letter label for anonymity and ease of reading.

Place Fig. 9 about here.

The number of relevant peers in this TUN was slightly below average, ranging from 23 to 26 across the time instants considered. Some peers were purely similar, others were purely substitutes, but the majority was a combination of both. Fig. 9 shows the evolution of this

TUN. Each node is a relevant peer, labelled accordingly, and the edges are the connections in their own TUNs. The nodes colour coded: white to represent peers that joined the TUN in that time instant, grey for those that remain from a previous time instant, and black for those who left because their journeys were no longer active. The sizes of nodes evolve dynamically with the number of connections to other nodes within the TUN, meaning that the largest nodes have the most connections in common with the chosen passenger in a specific time instant. The edges are coloured based on *relevance* between peers, with darker meaning higher. This representation highlights the temporal volatility of TUNs.

Fig. 10 provides a geographical representation of the selected passenger journey, shown in red, and of the travel peers with highest *relevance* at 10:00: solid journey paths represent predominantly similar peers; dotted journey paths represent predominantly substitute peers. This representation provides geographic evidence of the validity of the method for creating TUNs. It clearly shows the time-specific likeness of journey patterns from automatically identified peers.

Place Fig. 10 about here.

5.5. Discussion of market potential

The market potential of TUNs is promising. Although the case study only used the *Andante* subset and focused on a period with a relatively small number of journey transactions, the average number of passengers in each TUN is considerable. Even assuming that only 10% of the average of 39 peers per TUN would engage in some sort of information exchange using their mobile devices, the aforementioned property of Erdős and Rényi (1960) would still apply. This means that large clusters should emerge. Fig. 9 exhibited this clustering potential with the density of connections between peers of the selected passenger. Hence, information may travel quickly across TUNs. Furthermore, the results indicate a good balance between predominantly similar and substitute peers, meaning that passengers may effectively have access to information about their selected journey paths inasmuch as feasible alternatives.

Generally, when a small fraction of nodes are removed from a graph, only small clusters break from the main cluster, known as giant component (G). However, at a critical threshold, p_c (the percolation threshold), the giant component G gets fragmented into small clusters. For large random graphs, it is suggested that $p_c = 1/<k>$ (Newman 2010). Graphs with a large $<k>$ can withstand the loss of many of its nodes while keeping their main connectivity. A graph loses its connectivity when the number of failed nodes reaches $N - [N * p_c]$. Hence, in order to assess the reliability of the network we can define the expected minimum number of nodes to keep the network connected as the threshold value $[N * p_c] + 1$ (Li et al. 2015).

Analysis of the structure of TUNs obtained for this case study, at various levels of the cut-off *relevance* threshold and at different time instants, demonstrates their structure to be resilient to node removal. Table 2 shows that the average degree $<k>$ is consistently high, even at high levels of the cut-off threshold. Additionally, the percolation threshold obtained is relatively low and the expected minimum number of nodes to keep connectivity is significantly lower than the number of nodes found across TUNs. These empirical results show that the TUNs obtained are robust and strongly connected. It is observed that, at a specific time instant, raising the cut-off *relevance* threshold causes the number of edges *(E)* to drop much faster than the number of nodes *(N)*. As Albert and Barabási (2002) point out, the topology of the network can determine the robustness of the network. Since the removal of a node implies removing all of its edges, node removal inflicts more damage than edge removal.

Place Table 2 about here.

Generalising these results across passenger transport, we suspect that TUNs can effectively uncover an invisible layer of social engagement, which facilitates collaboration-based value creating processes. TUNs offer opportunities for circumstantial user collaboration across transport domains, overcoming limitations of existing ICT-based social networking services to deal with spatiotemporally dynamic information. The sensitivity analysis of the *relevance* threshold provides evidence that those opportunities are greater with larger numbers of simultaneous users in a transport system. This is due not only to the breadth of information distributed, but also to the chance of applying stricter *relevance* thresholds in larger systems and still obtaining adequate numbers of connections per TUN.

While the concept of TUNs is applicable across passenger transport domains, some of our previous work has been focussing on the development of an ICT-based application for mobile devices called *Journata* that enables collaborative exchanges of information among passengers (Nunes et al. 2011; Nunes et al. 2013). The application is specific to the urban public transport domain. It resembles an ICT-based social networking service whereby users interact with peers in their TUNs by rating aspects relating to their journeys (e.g., crowding of a specific station or vehicle) and sharing comments about transport services (e.g., reporting incidents that cause delays) in real-time. The application aims to increase the breadth and timeliness of information that is available to public transport users beyond what is already provided by transport operators, helping to adjust travel decisions in real-time. The significant market potential of TUNs supports the feasibility of such an application to generate and distribute significant amounts of relevant information.

6. Conclusions

We introduce and operationalize the concept of TUNs for transport systems that materialise *affinity* as spatiotemporal *relevance* between users. TUNs aim to overcome limitations of existing social media that have inhibited user collaboration-based value creation processes in passenger transport domains. These networks introduce the ability to identify relevant audiences for transport-related information, dynamically and in real-time. TUNs deal with the transient nature of such information, and enable its timely distribution to realise the potential of distributed knowledge in a transport system. The literature review reveals that information generated by others may influence the utility of travel choice alternatives. Yet much information, such as crowding on public transport, remains hard to know in advance. Hence the development of ICT applications for personal mobile devices based on TUNs will support the dissemination of new knowledge amongst travellers to help them improve their travel choices. Furthermore, those applications may improve experience by letting travellers tailor travel plans to their individual preferences.

The case study of bus users in Porto demonstrates the market potential of TUNs in the urban public transport domain. It shows that, even with a small subset of passenger journey data and under conservative assumptions, passengers have large TUNs of highly relevant peers. Therefore, potential exists for information to travel quickly across networks. The case study also reveals a good balance between predominantly similar and substitute peers; yet this also relates to density of the public transport system. Denser systems are likely to increase the portion of substitute peers. Sensitivity analysis of the *relevance* threshold indicates that the circumstantial user collaboration opportunities in various transport domains will grow with the numbers of simultaneous users in a system.

Three streams of future work will follow. The first is continuing development of algorithms for creating TUNs in real-time. The second stream is to improve the ICT application for personal mobile devices described, which incorporates the concept of TUNs and applies it to the urban public transport domain. This will include tests with users in a real environment. Furthermore, a more comprehensive rollout of the application as a pilot study will enable the analysis of the distribution of information flows, and a better understanding of the market potential and usefulness of TUNs. The last stream of future work is to consider additional case studies, in order to evaluate the transferability of TUNs to other public transport systems and even to other transport domains.

Acknowledgements

This work was supported by Fundação para a Ciência e Tecnologia (FCT) through a doctoral fellowship of the MIT Portugal Program. The authors thank STCP for providing the AFC system data necessary for this work.

References

Albert, R. & Barabási, A.-L., 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1), pp.47–97.

Anbaroglu, B., Heydecker, B. & Cheng, T., 2014. Spatio-temporal clustering for non-recurrent traffic congestion detection on urban road networks. *Transportation Research Part C: Emerging Technologies*, 48, pp.47–65.

Austin, J., 2010. Use of Social Networking to promote public transport and sustainable travel. *Local Transport Today, ITS Review*.

Barabási, A., 2003. *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*, New York: Plume.

Barabási, A.-L., 2007. The Architecture of Complexity. *IEEE Control Systems Magazine*, 27(4), pp.33–42.

Barry, J.J. et al., 2002. Origin and Destination Estimation in New York City with Automated Fare System Data. *Transportation Research Record: Journal of the Transportation Research Board*, 1817, pp.183–187.

Barry, J.J., Freimer, R. & Slavin, H., 2009. Use of Entry-Only Automatic Fare Collection Data to Estimate Linked Transit Trips in New York City. *Transportation Research Record: Journal of the Transportation Research Board*, 2112, pp.53–61.

Brakewood, C. et al., 2015. An Analysis of Commuter Rail Real-Time Information in Boston. *Journal of Public Transportation*, 18(1), pp.1–20.

Bregman, S., 2012. *TCRP Synthesis 99: Uses of Social Media in Public Transportation*, Washington, DC.

Büscher, M., Urry, J. & Witchger, K., 2011. *Mobile Methods*, Abingdon: Routledge.

Calabrese, F. et al., 2013. Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies*, 26, pp.301–313.

Carvalho, S., Sarmento, L. & Rossetti, R.J.F., 2010. Real-Time Sensing of Traffic Information in Twitter Messages. In *4th Workshop on Artificial Transportation Systems and Simulation (ATSS @ IEEE ITSC 2010)*.

Caulfield, B. & O'Mahony, M., 2007. An Examination of the Public Transport Information Requirements of Users. *IEEE Transactions on Intelligent Transportation Systems*, 8(1), pp.21–30.

Chen, C., Bian, L. & Ma, J., 2014. From traces to trajectories: How well can we guess activity locations from mobile phone traces? *Transportation Research Part C: Emerging Technologies*, 46, pp.326–337.

Cho, E., Myers, S.A. & Leskovec, J., 2011. Friendship and Mobility: User Movement In Location-Based Social Networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. pp. 1082–1090.

Chorus, C., Molin, E. & Van Wee, B., 2006. Use and Effects of Advanced Traveller Information Services (ATIS): A Review of the Literature. *Transport Reviews*, 26(2), pp.127–149.

Chorus, C.G. et al., 2007. Travelers' Need for Information in Traffic and Transit: Results from a Web Survey. *Journal of Intelligent Transportation Systems*, 11(2), pp.57–67.

Costa, P.M. et al., 2012. Cloud2Bubble: Enhancing Quality of Experience in Mobile Cloud Computing Settings. A Framework for System Design and Development in Smart Environments. In *MCS 2012 The Third International Workshop on Mobile Cloud Computing & Services*.

Erdős, P. & Rényi, A., 1960. On the Evolution of Random Graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5, pp.17–61.

Evans, A.W., 1991. Are urban bus services natural monopolies? *Transportation*, 18(2), pp.131–150.

Gal-Tzur, A., Grant-Muller, S.M., Minkov, E., et al., 2014. The Impact of Social Media Usage on Transport Policy: Issues, Challenges and Recommendations. *Procedia - Social and Behavioral Sciences*, 111, pp.937–946.

Gal-Tzur, A., Grant-Muller, S.M., Kuflik, T., et al., 2014. The potential of social media in delivering transport policy goals. *Transport Policy*, 32, pp.115–123.

Gault, P. et al., 2014. You'll Never Ride Alone: The Role of Social Media in Supporting the Bus Passenger Experience. In *2014 Ethnographic Praxis in Industry Conference Proceedings*. pp. 199–212.

Greengard, S., 2011. Following the Crowd. *Communications of the ACM*, 54(2), pp.20–22.

Gudmundsson, J., Laube, P. & Wolle, T., 2008. Movement Patterns in Spatio-temporal Data. *Encyclopedia of GIS*, pp.726–732.

Iryo, T., Yamabe, K. & Asakura, Y., 2012. Dynamics of information generation and transmissions through a social network in non-recurrent transport behaviour. *Transportation Research Part C: Emerging Technologies*, 20, pp.236–251.

Järv, O., Ahas, R. & Witlox, F., 2014. Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies*, 38, pp.122–135.

Kleinberg, J.M., 1999. Authoritative Sources in a Hyperlinked Environment. *Journal of the ACM*, 46(5), pp.604–632.

Lathia, N. et al., 2013. Individuals among commuters: Building personalised transport information services from fare collection systems. *Pervasive and Mobile Computing*, 9, pp.643–664.

Li, D. et al., 2015. Network reliability analysis based on percolation theory. *Reliability Engineering & System Safety*, 142, pp.556–562.

Li, X. et al., 2007. Traffic Density-Based Discovery of Hot Routes in Road Networks. *Advances in Spatial and Temporal Databases*, 4605, pp.441–459.

Li, Y., 2003. Evaluating the Urban Commute Experience: A Time Perception Approach. *Journal of Public Transportation*, 6(4), pp.41–67.

Newman, M.E.J., 2010. *Networks: An Introduction*, Oxford: Oxford University Press.

Nunes, A.A. et al., 2011. Using social networks for exchanging valuable real time public transport information among travellers. In *2011 IEEE Conference on Commerce and Enterprise Computing*. pp. 365–370.

Nunes, A.A., Galvão Dias, T. & Falcão e Cunha, J., 2015 (in preprint). Passenger Journey Destination Estimation from Automated Fare Collection System Data using Spatial Validation. *IEEE Transactions on Intelligent Transportation Systems*.

Nunes, A.A., Gonçalves, T. & Galvão, T., 2013. A Prototype for Public Transport Service Co-creation Using Social Media: Results from Usability Testing. In *International Conference on Exploring Service Science 1.3*.

O'Reilly, T., 2005. What Is Web 2.0. *O'Reilly Network*.

Olson, P., 2014. Why Google's Waze Is Trading User Data With Local Governments. *Forbes*.

Payne, A.F., Storbacka, K. & Frow, P., 2008. Managing the co-creation of value. *Journal of the Academy of Marketing Science*, 36, pp.83–96.

Pelletier, M.-P., Trépanier, M. & Morency, C., 2011. Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19, pp.557–568.

Pinho, P. & Vilares, M., 2009. *A Avaliação das Grandes Obras Públicas - O Caso do Metro do Porto*, Porto: FEUP Edições.

Stauffer, D. & Aharony, A., 2003. *Introduction To Percolation Theory*, London: Taylor & Francis.

Steinfeld, A. et al., 2013. Co-Producing Value through Public Transit Information Services. In J. C. Spohrer & L. E. Freund, eds. *Advances in the Human Side of Service Engineering*. Boca Raton: CRC Press, pp. 231–240.

Steinfeld, A. et al., 2011. Mobile Transit Information from Universal Design and Crowdsourcing. *Transportation Research Record: Journal of the Transportation Research Board*, 2217, pp.95–102.

Windmiller, S., Hennessy, T. & Watkins, K.E., 2014. Accessibility of Communication Technology and the Rider Experience: Case Study of Saint Louis, Missouri, Metro. *Transportation Research Board*, 2415, pp.118–126.

Wolfson, O. & Xu, B., 2010. Spatio-temporal Databases in Urban Transportation. *Bulletin of the Technical Committee on Data Engineering*, 33(2), pp.18–25.

Zegras, P.C. et al., 2015. Tracing a path to knowledge? Indicative user impacts of introducing a public transport map in Dhaka, Bangladesh. *Cambridge Journal of Regions, Economy and Society*, 8(1), pp.113–129.

Zheng, Y. et al., 2008. Understanding Mobility Based on GPS Data. In *UbiComp '08 Proceedings of the 10th international conference on Ubiquitous computing*. pp. 312–321.

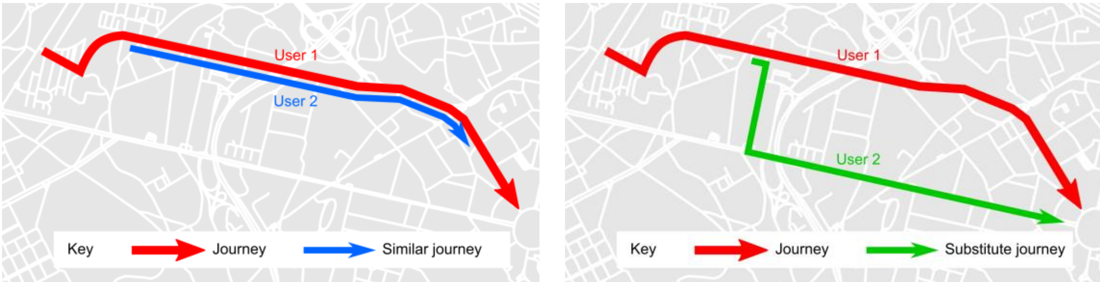

Fig. 1. Journey similarity (SIM) and journey substitutability (SUB).

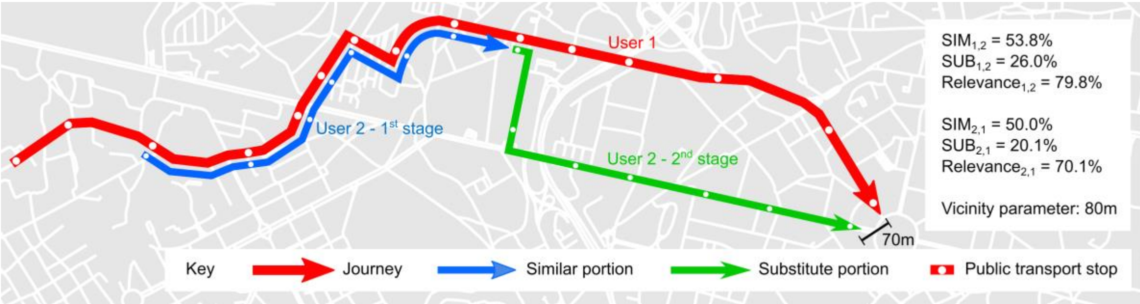

Fig. 2. Journeys with SIM and SUB scores.

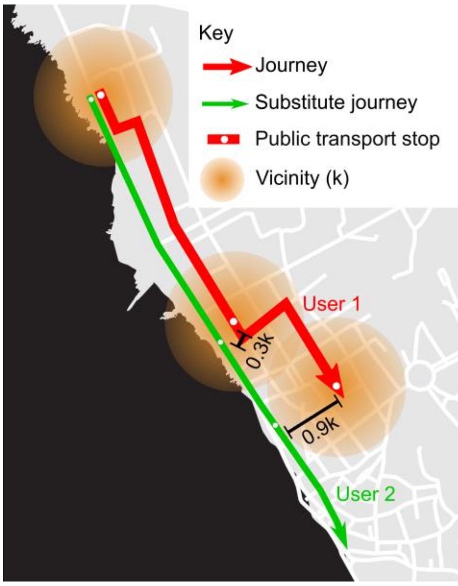

Fig. 3. Journey paths with maximum SUB at shorter substitute portion.

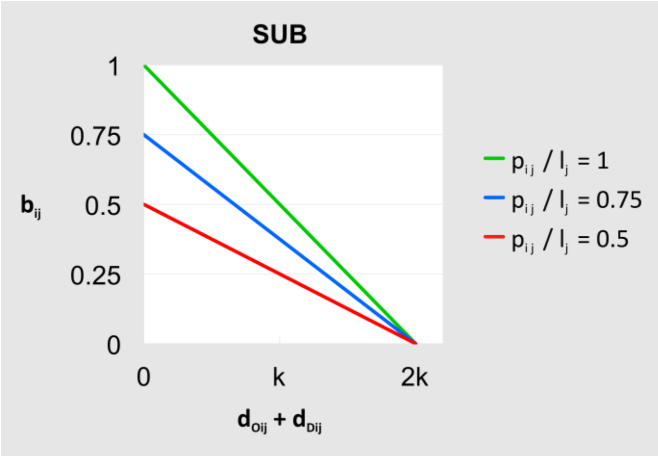

Fig. 4. SUB according to distance between location points within SPODs for various levels of substitution potential.

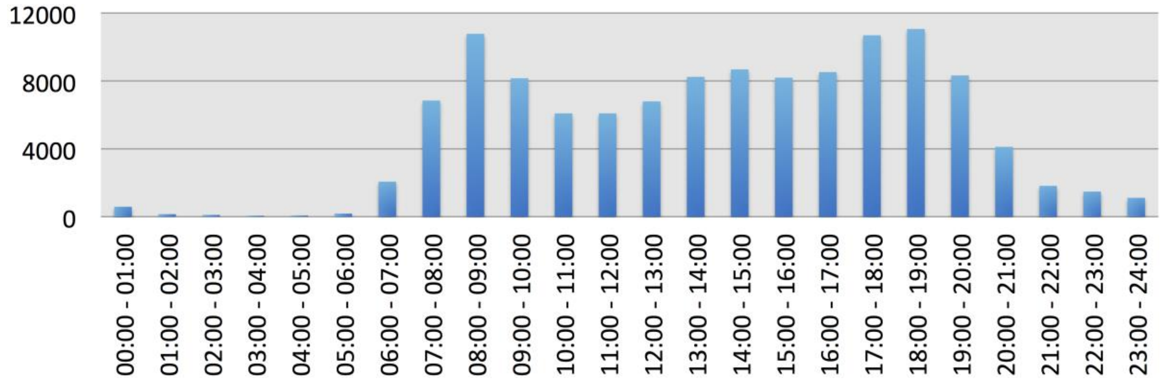

Fig. 5. Journey transaction throughout the day.

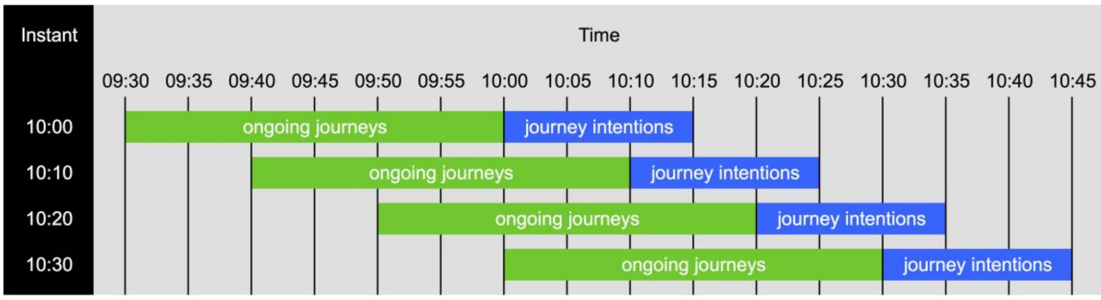

Fig. 6. Timeline of active journeys.

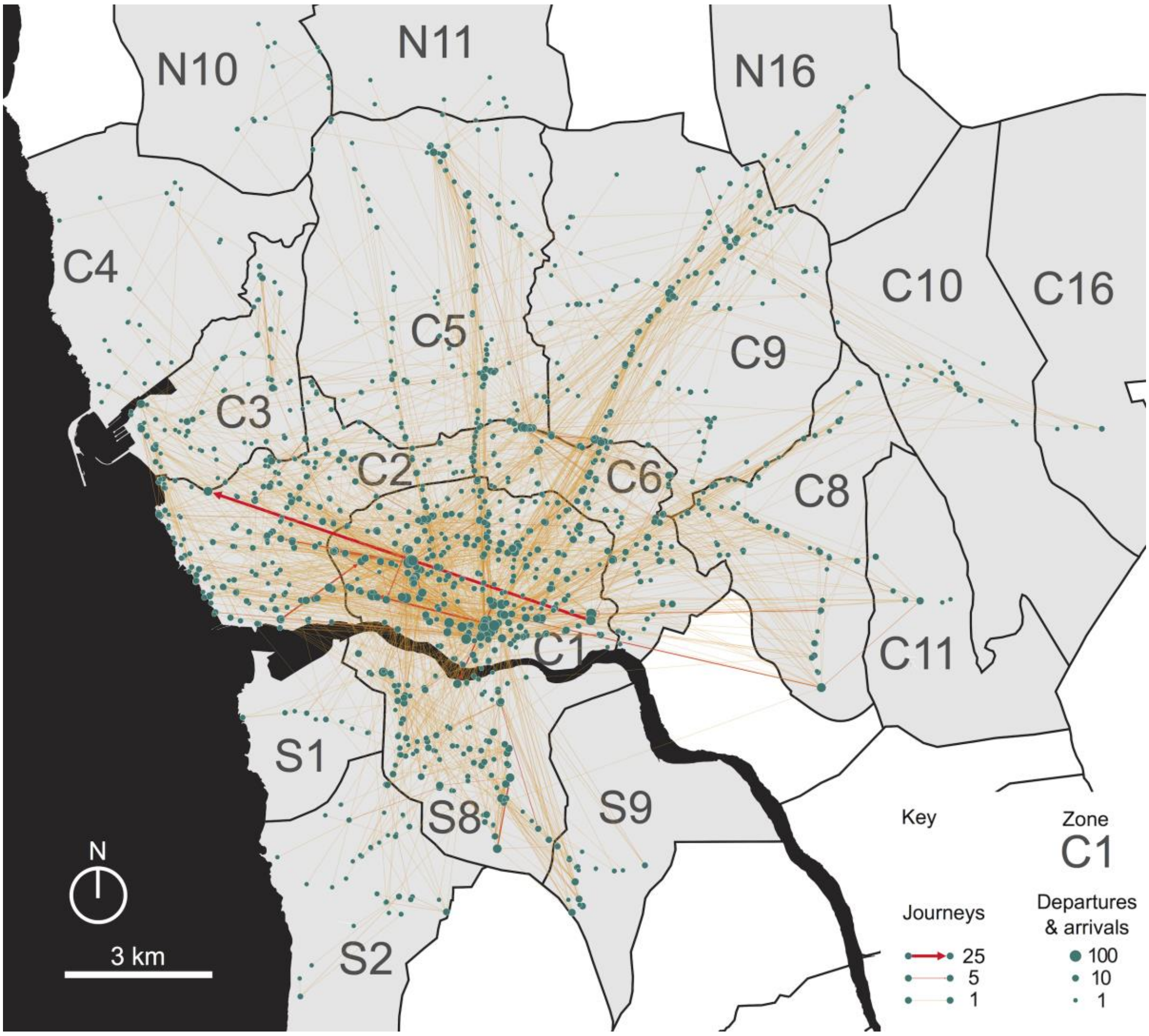

Fig. 7. Active passenger journeys at 10:00.

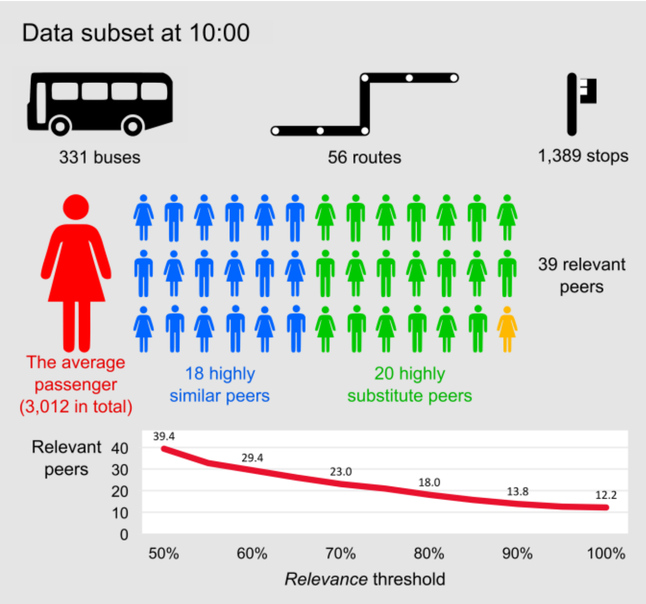

Fig. 8. Aggregate results.

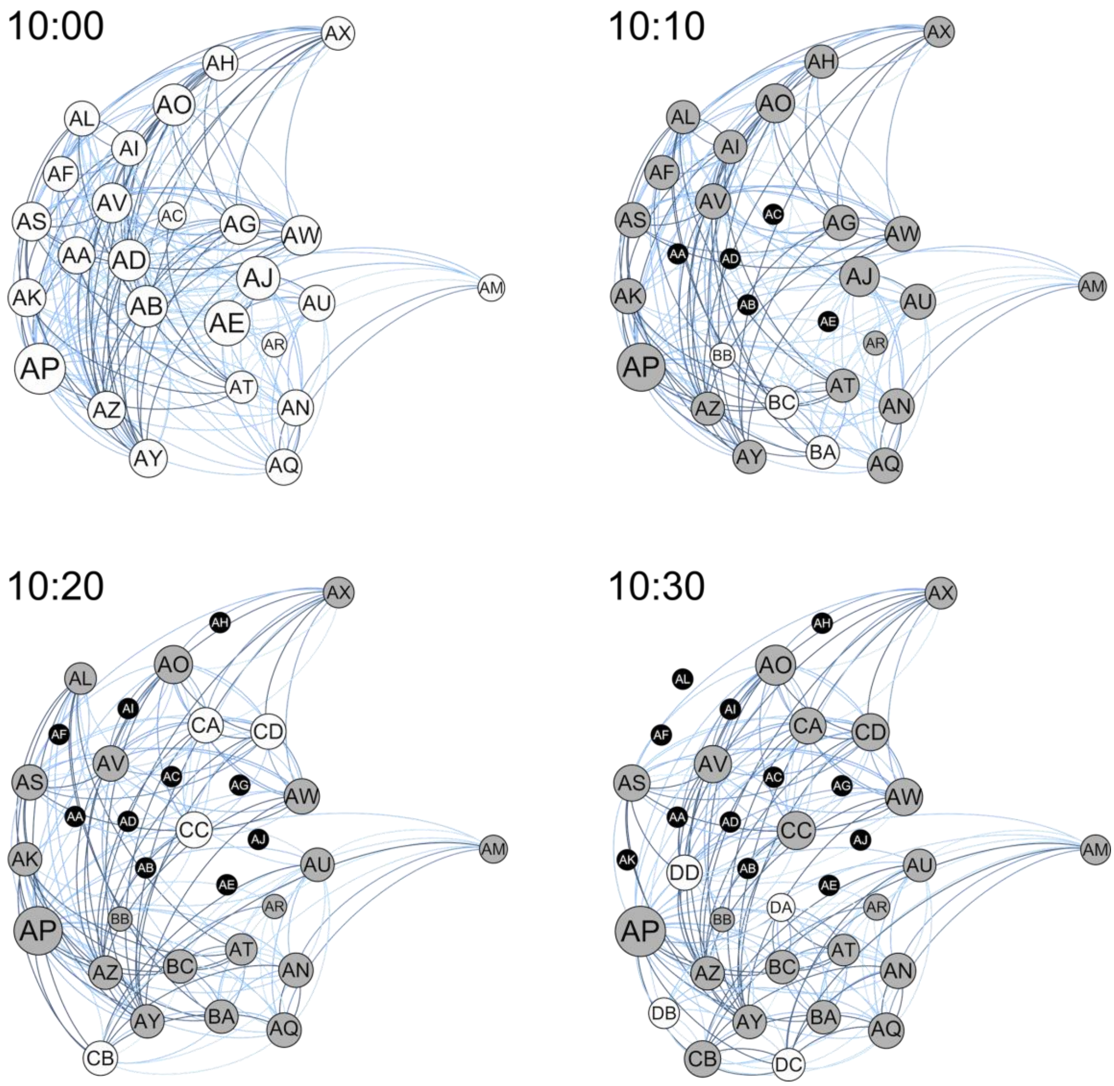

Fig. 9. Example passenger's TUN evolution.

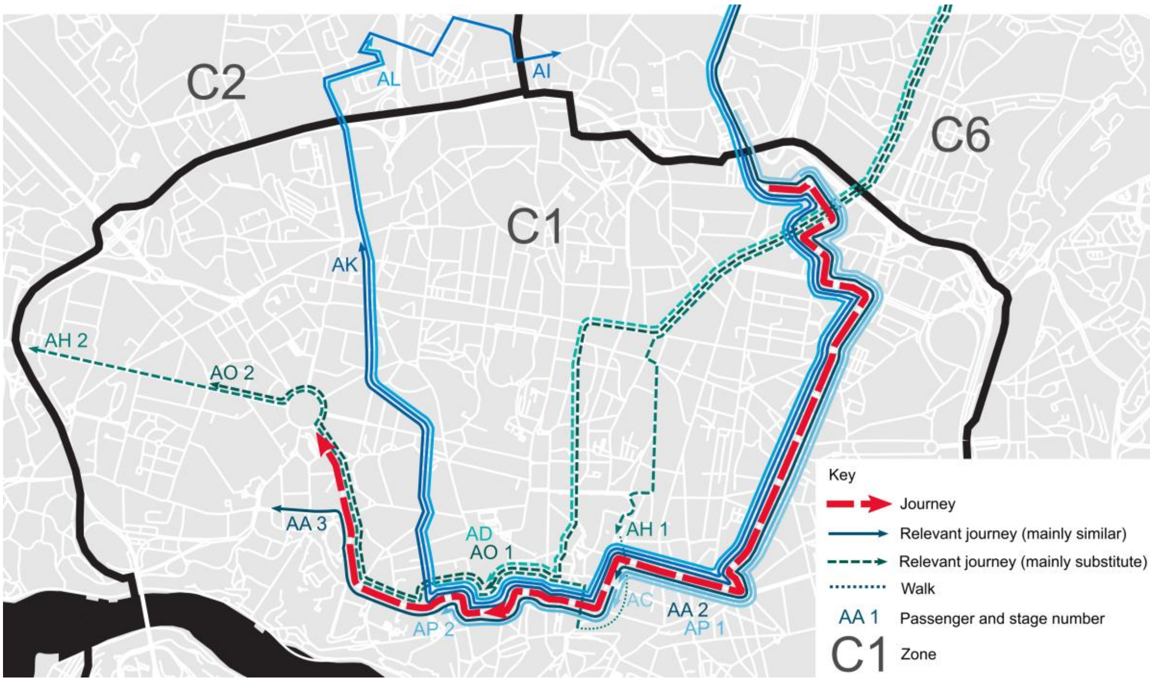

Fig. 10. Journey paths of passenger and highly *relevant* peers at 10:00.

Table 1
Relevant peers for an example passenger journey

Passenger	**SIM**	**SUB**	**Relevance**	**10:00**	**10:10**	**10:20**	**10:30**
AA	0.636	0.227	0.863	✓	×	×	×
AB	0	0.532	0.532	✓	×	×	×
AC	0.682	0	0.682	✓	×	×	×
AD	0	0.631	0.631	✓	×	×	×
AE	0	0.545	0.545	✓	×	×	×
AF	0	0.513	0.513	✓	✓	×	×
AG	0	0.532	0.532	✓	✓	×	×
AH	0	0.665	0.665	✓	✓	×	×
AI	0.818	0.045	0.863	✓	✓	×	×
AJ	0	0.5	0.5	✓	✓	×	×
AK	0.818	0.045	0.863	✓	✓	✓	×
AL	0.818	0.045	0.863	✓	✓	✓	×
AM	0	0.545	0.545	✓	✓	✓	✓
AN	0.545	0.045	0.59	✓	✓	✓	✓
AO	0	0.665	0.665	✓	✓	✓	✓
AP	0.591	0.227	0.818	✓	✓	✓	✓
AQ	0.545	0.045	0.59	✓	✓	✓	✓
AR	0	0.5	0.5	✓	✓	✓	✓
AS	0	0.565	0.565	✓	✓	✓	✓
AT	0.5	0	0.5	✓	✓	✓	✓
AU	0	0.591	0.591	✓	✓	✓	✓
AV	0	0.565	0.565	✓	✓	✓	✓
AW	0	0.532	0.532	✓	✓	✓	✓
AX	0	0.565	0.565	✓	✓	✓	✓
AY	0	0.532	0.532	✓	✓	✓	✓
AZ	0	0.532	0.532	✓	✓	✓	✓
BA	0.545	0	0.545	×	✓	✓	✓
BB	0	0.513	0.513	×	✓	✓	✓
BC	0.545	0	0.545	×	✓	✓	✓
CA	0	0.532	0.532	×	×	✓	✓
CB	0.5	0	0.5	×	×	✓	✓
CC	0	0.665	0.665	×	×	✓	✓
CD	0	0.532	0.532	×	×	✓	✓
DA	0.5	0.045	0.545	×	×	×	✓
DB	0.182	0.318	0.5	×	×	×	✓
DC	0	0.5	0.5	×	×	×	✓
DD	0	0.565	0.565	×	×	×	✓
			Total	26	24	23	25

Table 2
Analysis of TUNs across time instants at various *relevance* threshold levels

10:00	***Relevance* threshold**		**0%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
	Number of nodes	N	3010	3003	2940	2862	2727	2528	2347
	Number of edges	E	477207	118738	88608	69363	54358	41605	36668
	Average degree	$\langle k \rangle$	158.54	39.54	30.14	24.24	19.93	16.46	15.62
	Maximum degree	$max\ k$	775	232	217	210	210	210	210
	Percolation threshold	$p_c = 1 / \langle k \rangle$	0.01	0.03	0.03	0.04	0.05	0.06	0.06
	Min. No. of nodes to keep G	$n * p_c + 1$	20	77	99	119	138	155	151
10:10	***Relevance* threshold**		**0%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
	Number of nodes	N	2938	2929	2875	2800	2671	2460	2284
	Number of edges	E	466058	112612	83594	64579	51243	38757	34233
	Average degree	$\langle k \rangle$	158.63	38.45	29.08	23.06	19.18	15.75	14.99
	Maximum degree	$max\ k$	749	238	214	186	160	142	142
	Percolation threshold	$p_c = 1 / \langle k \rangle$	0.01	0.03	0.03	0.04	0.05	0.06	0.07
	Min. No. of nodes to keep G	$n * p_c + 1$	20	77	100	122	140	157	153
10:20	***Relevance* threshold**		**0%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
	Number of nodes	N	2847	2836	2792	2710	2581	2366	2172
	Number of edges	E	424487	102203	75529	57250	45319	34137	30210
	Average degree	$\langle k \rangle$	149.10	36.04	27.05	21.13	17.56	14.43	13.91
	Maximum degree	$max\ k$	725	237	209	138	132	132	132
	Percolation threshold	$p_c = 1 / \langle k \rangle$	0.01	0.03	0.04	0.05	0.06	0.07	0.07
	Min. No. of nodes to keep G	$n * p_c + 1$	20	80	104	129	148	165	157
10:30	***Relevance* threshold**		**0%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
	Number of nodes	N	2736	2724	2678	2590	2458	2265	2086
	Number of edges	E	377461	94576	69978	52729	41685	31353	27723
	Average degree	$\langle k \rangle$	137.96	34.72	26.13	20.36	16.96	13.84	13.29
	Maximum degree	$max\ k$	679	223	198	168	153	140	140
	Percolation threshold	$p_c = 1 / \langle k \rangle$	0.01	0.03	0.04	0.05	0.06	0.07	0.08
	Min. No. of nodes to keep G	$n * p_c + 1$	21	79	103	128	146	165	158