A Computational Approach for Integration of Soft Mobility and Regular Public Transports

Norberto Henriques
Centro de Informática e Sistemas
da Universidade de Coimbra, Portugal
Email: norberto@dei.uc.pt

Carlos Bento
Centro de Informática e Sistemas
da Universidade de Coimbra, Portugal
Email: bento@dei.uc.pt

***Abstract*— During a regular commuter's day, one may be faced with several situations where choosing a soft mobility option (e.g. walking, cycling) offers more advantages than a regular public transport (e.g. bus, subway). The main issue is to evaluate when such advantages exist.**

In this paper we analyze the Lisbon bus network on face of the evaluation of when this integration is interesting. We present its availability and range (in terms of distance and time) in order to highlight the areas where the commuter may benefit from a soft mobility option. The final goal is to use this information to determine which are the best locations for providing soft mobility means of transportation, as well as to advise a commuter about the best travelling option. We argue that this approach is useful for similar bus networks in other cities.

***Index Terms*— soft mobility, intelligent transportation systems, urban computing, Global Positioning System (GPS) traces analysis, user profile**

I. Introduction

Nowadays, soft mobility is becoming more than a simple way of life. The continuous growth of private vehicle usage on urban areas has achieved a unsustainable level and the public transportation system is nowhere near to suit everyone's needs.

The concept of soft mobility (or slow traffic, as it is also referred[14]) is commonly understood as any non-motorized transport (human powered mobility) [10]. Nonetheless, on the context of our work, we are also considering small electric transports, such as electrical bicycles, as soft mobility transports. These new motorized vehicles provide the same advantages of their non-motorized counterparts (low space profile and low direct impact on the environment) and provide the often necessary motivation for using a soft mobility transport: it enables the commuter to travel with a considerable low level of human effort.

In order to make the public transportation more attractive for the commuters, there are several actions that can be taken.

Amsterdam, which is a common reference in terms of transportation systems, holds a set of future scenarios for interesting transport innovations [3]. Nonetheless, although those innovations look very promising on a mid/long-term goal, we believe that the current technology and information can provide other interesting solutions on a short-term basis.

N. Henriques is with Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Leiria, Portugal and with the Centro de Informática e Sistemas da Universidade de Coimbra, Portugal

From the commuters' point of view, the first thing to understand is the availability of public transports on a given location and time. Moreover, on such cases where the public transport may not suit the commuter's needs, choosing a soft mobility mean can in fact represent a valid choice [14].

In the next sections, we will present our approach to bridge traditional transport systems with soft mobility means of transportation.

II. Data Description

In this section we take a closer look to the available data about the urban space of Lisbon.

A. "Carris" buses

Our research is based primarily on mobility data related to Lisbon "Carris" [1] buses. Figure 1 illustrates the distribution of bus stops on the metropolitan area in Lisbon.

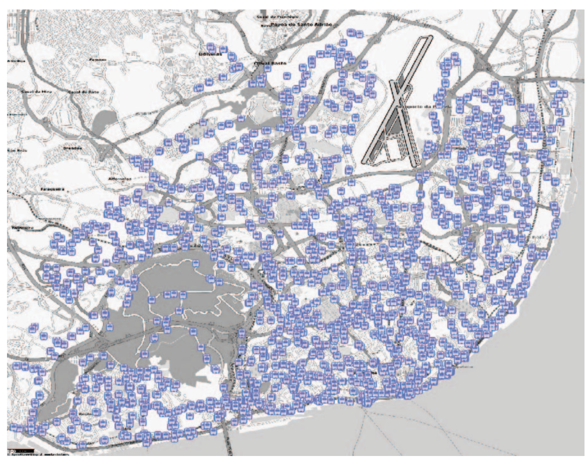

Fig. 1. Bus stops in Lisbon.

This data gives us an overall view of the availability of bus stops throughout the urban area.

Data on "Carris" buses also comprises location tracking for each operating bus. For this study, we are focusing on the particular points where the bus arrives to a bus stop, due to the fact that we want to study not only where but also when the bus is available.

During its normal course, a device mounted onboard of a bus records a set of data each time the bus stops. This data includes identification of the bus (bus number, line), event's timestamp, bus stop identification number (id), information about the trip itself (when it started, direction, trip variant).

Each bus stop is also geo-referenced with its GPS coordinates. By cross-referencing the bus stop id (registered by the onboard device) with this information, we are able to rebuild the bus trip's route using the bus stops as waypoints.

The ticketing information for the "Carris" buses is also available, although we are not using it on this study. Each ticket validation is recorded individually.

B. TMN mobile operator

Another important dataset that we are using on our research is the information about the GSM cells from the TMN (Telecomunicações Móveis Nacionais) mobile operator[1].

Each data record is acquired on an hourly basis and includes, among other information, the number of call attempts (on site and handovers), blocked calls, congestion time (time during which the cell had all its channels busy), successful call attempts and the number of dropped calls.

We used this data as a clue on the distribution of people along the city of Lisbon.

C. ASTER Global Digital Elevation Model

The Ministry of Economy, Trade and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA) released, back in June 2009, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) [5]. This elevation model is freely distributed for research purposes and provides a highly reliable dataset with a resolution of about 30 meters per pixel. The model's elevation accuracy standard deviation is within the [7m,14m[range.

The purpose for this data is to provide an elevation measurement for calculating the terrain slope. This is particularly important for the user profiling that we will discuss later in section V.

III. Determining bus availability

As seen on section II-A, the data about "Carris" buses that we are using comprises the GPS tracking of each operating vehicle (bus). Based on this information, we are able to extract several indicators about the availability of buses not only in the space, but also in time dimension.

A bus line may include more than one bus operating at a given time.

The data set we used refers to all the bus arrivals at bus stops on the week between May 26, 2010 and May 31, 2010. This represents a total of 1.701.134 arrivals, distributed across 64.589 bus trips.

[1]TMN has a national market share of 43 to 45%, which is even higher in Lisbon. This provides us a representative information about the actual usage of GSM mobile devices within the city of Lisbon.

A. Bus arrivals

Figure 2 shows the number of bus arrivals at each bus stop for all the week. The lighter blue (and larger) circles represent the bus stops where there is a higher rate of bus arrivals. As expected, the main streets are the ones with more frequent buses.

As our purpose was to identify the areas of Lisbon that are provided with less buses, we did not consider the direction of each bus.

In order to more clearly identify the bus availability on the urban area of Lisbon, we modeled its map with a grid of 500x500m cells. Each cell represents the sum of arrivals at each bus stop contained on its area, as seen on Figure 2. The bus availability is represented from yellow to red, where yellow represents a lower availability, whereas red represents a higher availability. This gives us an overall view of the location-based accessibility of the bus network[9].

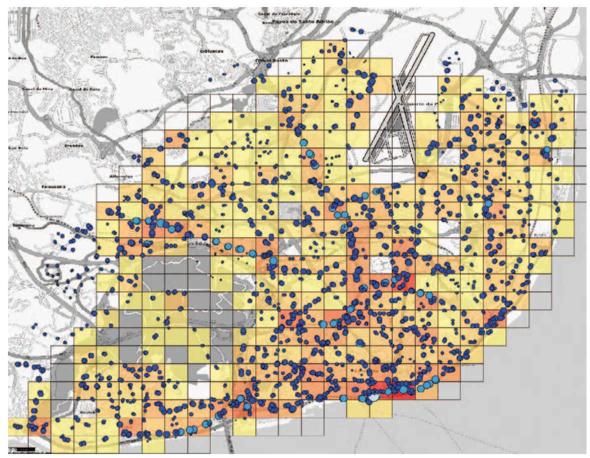

Fig. 2. Bus arrivals in each bus stop during one week: 500x500m cells.

In Figure 2 we can see several cells with a limited number of bus stops, but with a great number of bus arrivals. This means that, although there is the need for a high availability of buses on these areas (thus anticipating a high number of commuters), there will be many individuals that will need to travel by their own means a considerable distance in order to reach the designated bus stops.

We also analyzed the time between bus arrivals at a bus stop, in order to calculate the mean waiting time. Figure 3 shows the mean time (in seconds) between arrivals for the top thirty bus stops with more frequent buses.

This figure also shows the mean of bus arrivals per bus stop, aggregated per hour, for the complete week. We can clearly see two peaks on [6h, 9h[and [16h, 20h[. These two periods correspond to the rush hours, hence the reinforcement of the bus availability.

Based on all the previous data analysis, we are able to determine the areas of Lisbon where the availability of buses is on its lowest level. Figure 4 represents those areas, according to Table I, which specifies the lower and upper value for the bus arrivals on each cell area.

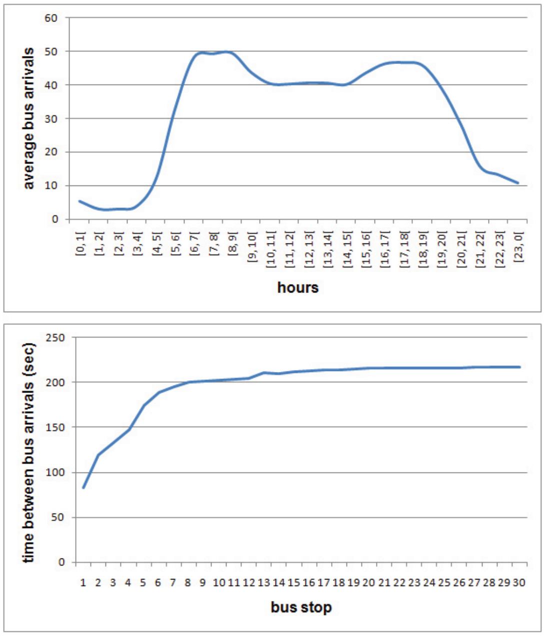

Fig. 3. a) time between arrivals for the top 30 bus stops, b) bus arrivals per hour.

TABLE I

RANGES REPRESENTED ON FIGURE 4

	Arrivals (per week)	
color	**lower value**	**upper value**
black	0	1
red	1	300
orange	300	600
yellow	600	900

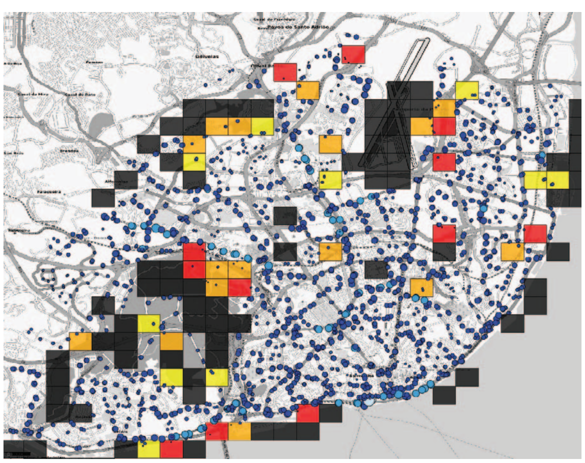

Fig. 4. Areas of Lisbon with the lowest bus availability.

B. GSM antenna load

We determined the number of attempted connections for each GSM antenna in order to represent its load. In Figure 5 the red diamonds represents the most loaded antennas, which overlap the 500x500m cells referred on Figure 2. The blue circles represent the bus stops with the highest number of bus visits.

Although we don't have enough data to irrefutably prove it, we consider that a higher number of connection attempts is translated to a higher crowded area.

As we expected, the areas with the highest number of bus visits match the areas where the GSM antennas register a higher number of connection attempts.

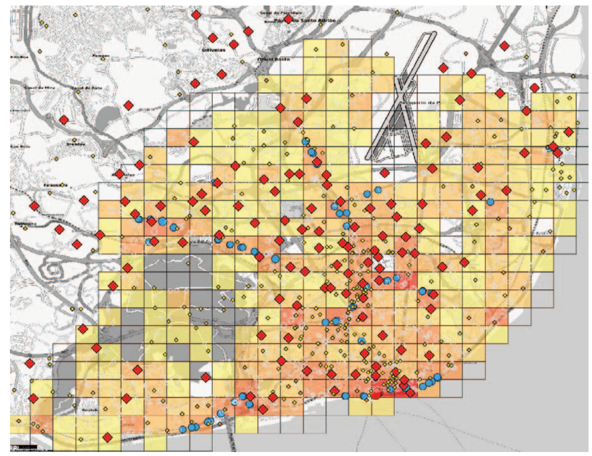

Fig. 5. Antenna load compared to the bus availability.

This exploratory analysis can assist us, for instance, to determine the potential locations for bicycle stations. Placing those bicycle stations within the areas represented on Figure 4 could provide the necessary means for commuters to either travel to their usual destination, or at least to the nearest bus stop.

C. Bus lines

During the course of our study, we also took a closer look to a number of statistics about each bus line. Our first task at this point was to find the bus lines with the highest number of trips, as well as the bus lines that feature lower travel speeds.

There are several events that affect the normal traffic behavior of a bus (e.g. traffic jams, bus breakdown). By using the "real-time" GPS data, we were able to research the dynamic accessibility[9] of the bus network, as opposed to the static nature of a typical timetable defined by the operator.

As described earlier on Section II-A, the provided information about the bus trips has a record for each time the bus stops. We executed a series of SQL queries on a Postgres database engine with the PostGIS extension[6], in order to rebuild each bus route.

Figure 6 shows the bus lines with the highest number of trips (red represents the highest number of trips, whereas yellow represents the lowest).

Figure 7 shows the traces of the slowest bus lines (total trip mean speed less than 14 Km/h). The bus lines represented in red are the slowest ones within the selected subset, in dashed-blue are the medium ones and in dashed-green are the fastest ones. The 500x500m cells that include those traces are highlighted on the right of Figure 7.

Fig. 6. Bus lines with the highest number of trips.

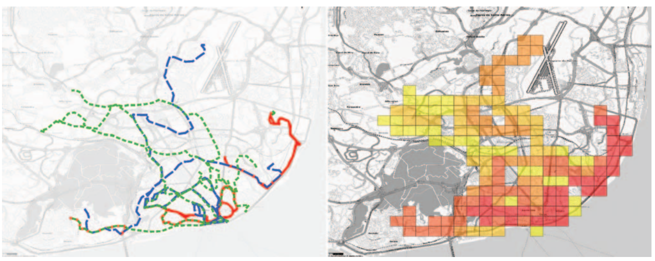

Fig. 7. Slowest bus lines and the corresponding 500x500m cells.

The speed is described as the mean trip speed of all the trips that were recorded during the period we are studying, for each bus line. This helps us to identify the bus lines where the mean trip speed may eventually lead a commuter to choose a soft mobility mean, such as a shared bicycle.

IV. Introducing Soft Mobility In The Bus Network

The next step in our analysis was to determine where a soft mobility option presents itself as a better option as opposed to the bus.

In our experimental approach we began by choosing one of the slowest bus lines. We selected the bus line 742, which goes from the Bairro Madre Deus/Escola to Pólo Universitário da Ajuda. This represents a trip of around 14.5km (13km considering the Euclidean distance between bus stops) and it is the 17th slowest bus line on our calculated rank.

The mean speed of the trips on this bus line is about 12km/h. This value was determined by averaging the speed for all the trips of this bus line (limited to the trip direction mentioned before). The calculation of the mean speed includes the waiting time in each bus stop.

Using the ASTER digital elevation maps[5], we determined the elevation for each bus stop included in this bus line. Figure 8 represents the elevation profile of this bus line along the entire trip. Table II shows the distribution of the main slope ranges along the trip's route.

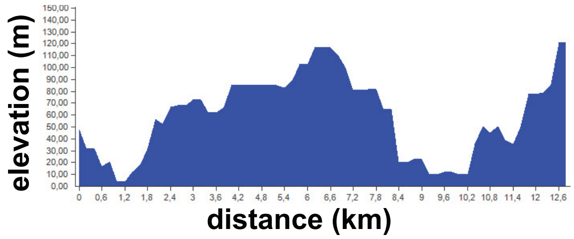

Fig. 8. Selected bus line's elevation profile

TABLE II
Slope along the trip

slope gradient	distance	% of distance
downhill (-3 to -2%)	0.31 km	2.36%
downhill (-2 to -1%)	0.50 km	3.82%
downhill (-1 to 0%)	4.40 km	33.75%
flat	0 km	0%
uphill (0 to 1%)	7.30 km	56.00%
uphill (1 to 2%)	0.53 km	4.06%
uphill (2 to 3%)	0 km	0%

According to Zhan and Wan [14], the usual bicycle speed on an urban context is about 10 to 15 km/h. The mean speed for the bus line we selected is within this range.

Parkin and Rotheram[8] also present a very interesting study on the speed and characteristics of cyclists and which are the main factors that influence those characteristics. Their model suggest that, on the flat, the mean speed of cyclists is 21.6km/h. For each additional 1% of uphill gradient, the mean speed is reduced by 1.44km/h, whereas an additional 1% of downhill (negative) increases speed by 0.86km/h.

Considering this model and the mean of each slope gradient for the selected bus line, we calculated the mean speed for each slope gradient, as seen on Table III.

TABLE III
Mean speed per slope gradient

slope gradient	speed
downhill (-3 to -2%)	23.75 km/h
downhill (-2 to -1%)	22.89 km/h
downhill (-1 to -0%)	22.03 km/h
flat	0 km/h
uphill (0 to 1%)	20,88 km/h
uphill (1 to 2%)	19.44 km/h
uphill (2 to 3%)	0 km/h

Next we calculated the weighted mean of the speed of the cyclist. The percentage of total distance in Table II represent the contribution of each slope gradient for the overall mean speed. The final result for this measurement is 21.35 km/h.

According to the model proposed by Parkin and Rotheram, and considering the mean speed on the selected bus line, there is clearly an advantage on the soft mobility alternative. There are, of course, several questions to be addressed when using, for example, a bicycle (e.g. weather conditions, items do carry, traffic safety, wind resistance). Nonetheless, from the strict point of the travelling speed, bicycles provide a possible alternative to the bus network.

In order to compare these results with real GPS data, we collected a set of GPS tracks generated by bicycle commuters. It was not possible to recruit volunteers for explicitly recording GPS data, so we took in consideration a few community based sites that rely on the concept of social networks for sharing this kind of data. We used the Wikiloc website[2] for this purpose, because it is one of the few social network sites that retain the original timestamps of the uploaded GPS tracks.

We selected 15 GPS tracks with 10 to 30 km in length. In order to obtain accurate elevation measurements on these GPS tracks, we also applied the ASTER GDEM[5] elevation model.

We calculated the mean speed for each slope gradient of all the 15 GPS tracks, considering the actual recorded speed. Table IV shows these results.

TABLE IV
RESULTS REAL GPS DATA

slope gradient	speed	distance	% of distance
downhill (-3 to -2%)	—	0 km	0%
downhill (-2 to -1%)	9.94 km/h	0.03 km	0.10%
downhill (-1 to -0%)	16.73 km/h	14.79 km	49.99%
flat	7.37 km/h	1 km	3.38%
uphill (0 to 1%)	14.95 km/h	13.71 km	46.4%
uphill (1 to 2%)	17.66 km/h	0.05 km	0.17%
uphill (2 to 3%)	—	0 km	0%

Next we calculated the weighted mean of the speed, based on the percentage of total distance in Table II for each slope gradient. The result was 15.32 km/h. This result is a bit far from the 21.35 km/h predicted by the model, although is is within the speed range suggested by Zhan and Wan. Moreover, the mean speed is still above the mean speed of the bus, therefore the bicycle option is still a possible alternative.

We should take in consideration the fact that there we have no knowledge about the behavior and motivation of the cyclists that anonymously provided the GPS tracks. We expect that during the evolution of our research we will be able to use real GPS data acquired specifically for this purpose. Nonetheless, the available data at the present time enabled us to demonstrate the proof of concept.

V. Future Work: Modeling the Commuter's Mobility

One of the purposes of our study is to identify the potential for creating a multimodal interface between public transportations (the bus network, in our study) and soft mobility transports. Nonetheless, in order to achieve that, the human variable cannot be discarded.

Our envision incorporates the use of a personal travel assistant that takes in consideration not only the information about the available means of transportation in real-time, but the commuter's physical strength. This physical strength will be translated to the effective range of the commuter based on his soft mobility choice.

A. Energy expenditure

The estimate of energy expenditure has been studied in the past [7][12]. The common approaches use specific devices, such as the ActiGraph activity monitor range [4].

Nowadays, smart handheld devices are widespread. These devices are commonly equipped with accelerometers and GPS receivers.

Our approach is based on simple handheld devices, such as a smartphone. Although these devices may provide data that is less precise than the data generated by activity monitors, they are less intrusive and do not require the commuter to carry an additional device. Our goal is not to determine the exact energy expenditure of a walking or cycling journey, but to determine the commuter's behavior during that journey.

B. Determining commuter's range

By analyzing the mobility data features (speed, rhythm [11], acceleration, location, among others) against the terrain topology [5], we will be able to create a user profile. This user profile will allow us to determine how far can a commuter go (on foot or by bicycle) at an acceptable effort level.

Figure 9 shows a prototype application designed for Android smartphones that will capture in real-time the mobility data from the commuter. Due to the high power consumption associated with the intensive use of motion and GPS sensors on handheld devices, this application will take into account several power-aware techniques that minimize the battery usage [13].

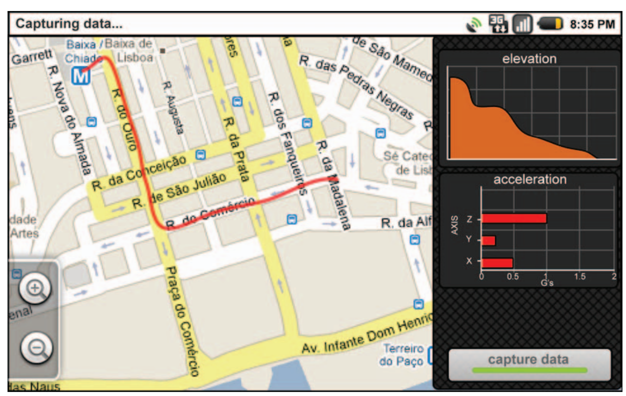

Fig. 9. Capturing GPS and accelerometer data.

So, considering the following information:

1) the commuter's location
2) where the commuter wants to go
3) the bus stops and bicycle stations nearby the start and destination locations
4) the statistical information about the nearby bus stops that are part of a bus route that includes the destination area

5) the commuter's profile
6) the land topology

The application will be able to determine the best choice for the commuter in terms of means of transportation. If this implies choosing a soft mobility option, the application will provide an optimized route as well.

C. Commuter's dynamics within the city

In order to anticipate the distribution of bicycles on a bicycle station network, it will be necessary to analyze the dynamics of the commuters within the city. Determining the flow of individuals on several periods of the day will help do establish the needed quantity of bicycles in each station, and eventually the need to move them from station to station on key moments of the day.

The ticketing information can be used to determine how many commuters enter a bus on a particular bus stop. Unfortunately there is not any information about the number of commuters that leave a bus. This presents itself as a challenge for our future work, to infer how many people leave a bus on a particular bus stop.

VI. Conclusions

We strongly believe that integrating the information about the availability of public transportation with the profiling of commuters can benefit all the users of that public transportation system. Not only it can encourage the use of soft mobility means, it also can provide the proper information for distributing bicycle stations.

Our study shows that although Lisbon has a large set of bus stops, there are several places where the bus availability is not as regular as desired. It also shows that in some of the areas with frequent bus arrivals, there is a considerable amount of streets in a 500m radius that do not have bus stops. Furthermore, our study highlights the bus lines which would benefit from a shared bicycle network, based on the mean speed along those bus lines.

Finally, we demonstrated that a cyclist will be able to travel along a bus line faster than the bus itself, based on real data.

References

[1] CARRIS - Transportes Públicos de Lisboa. http://www.carris.pt/.

[2] Wikiloc. http://www.wikiloc.com/.

[3] Federico Casalegno and Dave Chiu. White paper on sustainable transportation. Technical report, MIT Mobile Experience Laboratory, 2008.

[4] Scott Crouter, James Churilla, and David Bassett. Estimating energy expenditure using accelerometers. *European Journal of Applied Physiology*, 98:601–612, 2006.

[5] Earth Remote Sensing Data Analysis Center (ERSDAC). Aster global digital elevation model (aster gdem), 06 2010.

[6] Stephan Holl and Hans Plum. PostGIS. *GeoInformatics*, 03/2009:34–36, April 2009.

[7] Alberto E. Minetti, Christian Moia, Giulio S. Roi, Davide Susta, and Guido Ferretti. Energy cost of walking and running at extreme uphill and downhill slopes. *Journal of Applied Physiology*, 93(3):1039–1046, 2002.

[8] John Parkin and Jonathon Rotheram. Design speeds and acceleration characteristics of bicycle traffic for use in planning, design and appraisal. *Transport Policy*, 17(5):335–341, September 2010.

[9] F. Pinelli, A. Hou, F. Calabrese, M. Nanni, C. Zegras, and C. Ratti. Space and time-dependant bus accessibility: a case study in rome. In *IEEE International Conference on Intelligent Transportation Systems*, Sr. Louis, Mussouri, USA, October 2009.

[10] Rosa La Rocca. Soft mobility and urban transformation. *Tema. Journal of Land Use, Mobility and Environment*, 2(0), 2010.

[11] K. Van Laerhoven, D. Kilian, and B. Schiele. Using Rhythm Awareness in Long-Term Activity Recognition. In *Proceedings of the eleventh International Symposium on Wearable Computers (ISWC 2008)*. IEEE Press, 2008.

[12] David Gordon Wilson. *Bicycling Science 3rd edition*. The MIT Press, 2004.

[13] Moustafa Youssef, Mohamed Amir Yosef, and Mohamed N. El-Derini. Gac: Energy-efficient hybrid gps-accelerometer-compass gsm localization. *CoRR*, abs/1004.3174, 2010.

[14] Wang Zhang and Wan Jun. Research on the development mode of slow traffic system in cities based on low-carbon concept. In *Proceedings of the 7th International Conference on Innovation & Management*, 2010.