Visualizing Public Transport Systems: State-of-the-Art and Future Challenges

Massimo De Marchi
University of Illinois at Chicago, USA

Abstract
People every day rely on Public Transport Systems for their daily commutes or for their journeys through new cities. The availability and usability of information that are relevant to the traveller are very important for the system to be easy to use. However, due to the increasing complexity of public transportation networks and the needs of different travellers, designing effective methods to visualize and explore these systems is highly challenging. In this state-of-the-art report I surveyed the literature related to public transport visualization and I interviewed two experts from the field in order to provide a systematic overview of the state-of-the-art techniques to visualize public transport data. I describe what kind of datasets are commonly used as well as what are the common tasks that a visualization tool should support. At the end I identify challenges in the area that need further research.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation

1. Introduction
Transportation is the backbone of our civilization and cities invest a lot of money for maintenance and improvements of public transport infrastructures. Research in this field is essential to develop sustainable cities and several factors need to be taken into consideration. From an economical perspective, cities administrations look for solutions to reduce the cost of the service and to improve its efficiency. From a social perspective, a Public Transport System (hereafter, PTS) ensures that all citizen are able to travel. From an environmental perspective PTS allows to save more energy compared to private transport.

Many opportunities exist for the visualization community to help to solve PTS problems. On one hand the network need to be easy to use for the travelers, therefore suitable visualizations should provide them with the information they need to plan their journeys. On the other hand, urban planners and transportation researchers look for patterns in the PTS usage that could give insights on how to improve the overall system efficiency.

Most of the techniques present in the literature seems to focus more on visualizations that help to understand the system usage and to improve urban planning, rather than to improve the usability and accessibility of the system for the travelers. In fact few works have been found that address visualizations that aid wayfinding (see section 3.1) and this is a very important aspect for the overall quality of the PTS. A study conducted by the Israeli Ministry of Transport identified system legibility as one of the greatest barriers to use the PTS [WS05].

Most of the people plan their journeys using mobile applications such as Google Maps or using printed maps that they can find nearby the stations. The problem with printed maps is that they lack of temporal information and often they are not so easy to navigate. Using Google Maps the user has few control over the route selection parameters, and is not able to see all the possible solutions. Although Google Maps provides the most common solutions, this does not enable the exploration of the systems.

Research in PTS visualization resembles techniques and concepts from many fields, such as cartography, psychology and graph theory. For this reason I cited related work of these field which theories and techniques are still valid for public transport visualization, because they can be viewed as a contextualization of such techniques.

In the literature many work on PTS data visualization and visual analytics have been conducted, but finding the appropriate visual encoding for the needs of different users is still
challenging due to several factors, the first of which is the high and growing complexity of these networks.

In this report I survey state-of-the-art techniques for visualizing PTS data. I also integrate information coming from the interviews of a transportation researcher and an expert of visual analytics of movement.

In section 2 I briefly describe what kind of data are used in public transport visualization and how it is gathered. In section 3 I give an overview of the theoretical aspects behind research in public transportation and in section 4 I list some of the common tasks that a visualization tool need to support. In section 5 I give an overview of what are the most common techniques used in the field and in section 6 I list some of the problems that seems to be more exciting to tackle. Finally in section 7 I draw some conclusions.

2. Transportation data

Today more and more the transport authorities are monitoring their infrastructures at an increasingly unprecedented level of detail and they are making it available to the public. This opens the way for research in visualizing and analyzing these data to improve the PTS. These data can include geospatial, temporal and categorical information. In order to understand the domain a common terminology is defined below and will be used in this document.

- **Transportation network**: directed graph structure where node are lines stops and edges are available lines that connect nodes.
- **Transit route**: a sequence of nodes and edges with two locations that are defined as the departure location and the arrival location (alternatively, origin and destination).
- **Transit line**: a Public Transportation Service offered by a certain transport mode (e.g. bus, trains, etc).
- **Trip**: an individual ride on a certain road, therefore it is characterized by a temporal information.
- **Transfer**: a connection between one or more routes.
- **Journey**: a passenger travel from an origin to a destination within the Public Transport System.

The PTS data is available in various formats, and one of the most popular is **GTFS** (Google Transit Feed Specification) [gtf] which is rapidly becoming a standard for publishing Public Transit feed data. These data are available offline and many transport authorities have started to provide also real-time online data collected by vehicles GPS. Many transport authorities such as Singapore public transport system provide also Passenger RFID card data which allow to track behaviors of travelers.

An important and widely used source of geospatial data is **Openstreet Map** [ope]. Google Maps [goob] is also a well known source of geospatial data. These services allow to get both raster tiles and vector tiles with a custom amount of features.

Jakob Eriksson, a public transport researcher that I have interviewed, says that the most common kind of data that he used in his project are GPS traces. This holds also for Gennady Andrienko, an expert of visual analytics of movement, who mentioned also other kind of data coming from other contexts, since visualization of movement data is a broad subject. As reported by the researchers, these data are gathered using open APIs, collaborating with owners of the data, and even by manually installing tracking device on vehicles. The respondents reported also that often data is aggregated and several techniques can be used to analyze it, depending on the context; examples are inferring the geometry and topology of the underlying road map from large GPS trace collections, and "skeletonization", which produces a discrete map-like graph from density estimates.

3. Theoretical foundations

In this section I describe some theoretical concepts that are important when it comes to design visualization tools for public transportation. A great introduction to some of these concepts is provided by Colin Ware [War12].

3.1. Wayfinding theory

Everyday we use public transportation to move from an origin to a certain destination. To help people move through an environment we need to think of how people build an understanding of an environment over time and how they use this understanding to seek information. One aspect of this problem is called wayfinding, or alternatively pathfinding. The term was first used by the architect Kevin Lynch [Lyn60] referring to maps, street numbers and directional signs as "way-finding" devices. Muhlhause [Muh06] explained that wayfinding requires more that maps and street numbers, namely it involves understanding the spatial organization of a place. Wojciechowicz [WS05] described wayfinding as the process of collecting information from an environment to know where we are relative to where we want to go to and how to get there. Therefore a concept that is directly linked to wayfinding theory is spatial cognitive maps.

3.2. Spatial Cognitive Maps

Spatial cognitive maps is the term used to refer to human internal representation of an environment. Siegel and White [SW75] suggested that the formation of a cognitive map is the result of a three stages process. According to Siegel et al. in the first stage we learn information about key landmarks of the environment we are exploring; this kind of knowledge is also referred to as declarative knowledge. In the second stage we develop routing instructions from one landmark to another; the knowledge that we build in this second step is both procedural knowledge and topological knowledge since we develop this information by mean of sequential instructions and connections between different locations. The final step of this map formation process is the actual development...
of a cognitive map that let us also to estimate distances between places and navigate from one place to one another even if we do not have the exact routing information. But in their work Siegel et al. ignored the importance of map technologies, in fact as proved by Thorndyke and Hayes-Roth [THR82], people can easily build cognitive maps directly from actual maps.

Another interesting insight comes from the study of Colle and Reid [CR98] that suggest the importance of overview maps since the formation of cognitive maps is easier when the viewer can see everything at once.

Darken et al. [DAA98] argue that perspective views are not quite effective to support the generation of mental maps since there are evidences that terrain features are not encoded in memory as three-dimensional structures; this is an important information to take into consideration when choosing the visual encoding for a geospatial visualization tool.

Kosslyn [Kos87] proposed a theory of spatial knowledge based on two kind of knowledge, namely categorical and coordinate. With these two term Kosslyn encoded two ways in which humans build their cognitive maps. The former is due two the actual experience of going through an environment. The latter is due to visual imagery obtained with an overview.

3.3. Legibility measures

The study of Woyciechowicz [WS05] has identified system legibility as a barrier to use the PTS. To compare the quality of different network solutions, transportation researchers developed a set of measures to evaluate the legibility of a network.

Their work is based on some insights from previous work in wayfinding and cognitive map theories. In particular they identify wayfinding as the combination of three kind of knowledge: survey knowledge, procedural knowledge and landmark knowledge. They also give definitions of these terms, namely survey knowledge is the ability of using a cognitive map, procedural knowledge is the ability to follow directions and landmark knowledge is the ability of building paths through landmarks.

The set of measures developed by the authors are based on how easy is to exploit these three kind of knowledge within a certain transportation network. Here I briefly sketch the most important:

- **Pattern repetition**: if some pattern is present in the network, that make it easier to navigate the network.
- **Simplicity**: it makes more easy for the user to exploit survey knowledge since less overlapping occurs between different lines.
- **Use of arterial road**: arterial road are more familiar to passengers and therefore they help to better exploit procedural knowledge.
- **Joining of transit hubs to landmarks**: this way the transit system becomes the link that connects different activity centers and this make it more easy for a passenger to navigate the transit system.

These measures could be important when studying a PTS to decide what visual encoding would better fit the considered scenario. For example we can imagine a preliminary study of the system using these measures to highlights certain characteristics of the system, and therefore it can give important insights on what is the most effective visual encoding.

3.4. Landmarks

Landmarks are significant physical, built, or culturally defined objects that stand out from their surroundings and therefore help locating the geographic position [Go99].

In the literature there are many approaches to develop formal models for the automatic extraction of landmarks from existing GIS dataset. Raubal et al. [RW02] proposed measures to formally specify the landmark saliency of a feature base on its visual, semantic, and structural attraction. These measures are used to extract landmarks from a dataset. Elias [Eli03] and Elias et al. [EB05] works described knowledge discovery and data mining methods for landmarks extraction. These approaches are limited to the identification of buildings as landmarks, but also other terrain features such as parks and bridges could by identified as landmarks [ES02].

4. Tasks Analysis

When designing a visualization of Public Transport data, it is important to determine what tasks it needs to support. The tasks that the visualization has to support depend also on the end user. I identified two kind of users, namely the traveler and the transportation researcher.

4.1. Traveller

The root of all the tasks that a visualization targeted for a traveler should support is **wayfinding**. When it comes to wayfinding every person has different capabilities and a visualization tool needs to take this into account.

Travelers every day use the PTS for their journeys through the city, therefore they need to find a path from an origin to a destination. However, we need to take into consideration other factors such as user needs; for example a tourist might want the most enjoyable path, but a worker would probably prefer the journey that takes the least amount of time.

I have interviewed Jakob Eriksson, a transportation researcher that now is working on a project that has to do with improving access to transit, both in terms of on-demand routing of vehicles, and in terms of better navigation support for transit users. An ideal result for his research would be a dramatic improvement in the ease of use and efficiency of transit services and a tool to support navigation, such as an ambient
display on a wall [PS06], which allows passers-by to plan a real-time route through the transit network at a glance. He also mentioned Gnuplot [gnu] and Google Earth [gooa] as common tools that he uses to make sense of the data. The visualizations described in section 5, enables the following tasks:

- Find the best journey in terms of time efficiency from the departure location to the arrival location.
- Find the journeys with the minimum waiting time at the departure and/or transfer stations (some users might prefer to take a longer ride rather than waiting more time at the station).
- Find the most "enjoyable" path from the departure station to the arrival station (the term enjoyable has not a unique measure and depends on the context [QSA14]).
- Show all the possible journeys from the departure point within a certain amount of time.

4.2. Transportation researcher

Transportation researchers would like to explore several aspects of a PTS, such as the degree of connectivity of the network, the travel efficiency of a city, etc. Exploiting these information researchers are able to understand which areas of the city are less developed, find out inefficient usage of the PTS, understand what facilities are lacking and make predictions on the behaviors of the system.

I have interviewed Gennady Andrienko, an expert of visual analytics of movement, and he reported that a predictive and descriptive model for mobility would be an ideal result for his research. This reflect the fact that transportation researchers need tools able to predict the behaviors of travelers and of the system.

A list of the common tasks that a visualization tool that aim to help transportation researchers needs to support has been obtained analyzing the tasks that the tools in section 5 support.

- Given a departure location, show reachable destinations within a given time span (this is important to answer questions regarding the accessibility of a city, for example "how long does it take to get to the closest hospital?").
- Show possible routes that connects the origin location to the destination.
- Show and compare travel time of different routes.
- Display useful information along the routes, that are mobility factors such as waiting time at the transfer station, riding time on a vehicle, transfer time to walk to the next transfer stop, etc.
- Show accessibility to the primary activity centers of the city.

5. Survey

Several possible techniques are used to visualize PTS data, and all of these techniques follow one or more of the these approaches:

- **Filtering**: remove data that is not relevant to support the desired task.
- **Aggregation**: used very often especially to avoid visual clutter due to the high degree of over-plotting.
- **Distortion**: used to improve maps readability and a core part of focus+context approaches.
- **Abstraction**: a form of simplification that improves readability and makes information more understandable.
- **Animation**: used to highlight changes in data and convey movements, useful for example when giving user information about vehicles positions and directions.

I conducted this survey taking into consideration articles published on the main venues and other articles that are cited by these articles. I am also including in the survey articles that describes techniques used in other fields, such as visualization of graphs or visualization of movement data, since these techniques could be easily applied to the PTS domain. Based on these articles I have highlighted the main techniques used in Public Transport Visualization and I have subdivided them in the following five categories.

5.1. Isochrone Maps

Isonchone maps [iso] are traditional visual representations used in transportation to show regions of equal travel time from a given departure location. Isochrone maps usually are visually encoded using colored contours and can be easily overlaid on geographical maps. This technique is used in the literature [OMS00] [WFNH15] to represent accessibility information at a coarse grain, in fact is very difficult to use this technique to estimate precise travel time. Moreover isochrone maps cannot reveal other mobility factors and therefore are often used in combination with other techniques, such as in the work of Zeng et al [ZFA*14].

Figure 1: Isochrone visual representation of travel time from the work of Zeng et al.

5.2. Schematic Maps

Schematic maps [iso] are the outcome of a process of simplification and/or distortion. A pioneer in this field is the work of Beck [Gar94] that proposed a redesigned map of the London underground that today is widely adopted to display metro...
Many works in the literature that aim to produce easy-to-use schematic maps have been inspired by Beck’s work.

Figure 2: H.C. Beck’s 1946 Underground map

Clark [Cla77] argued that travel time through a network is more important than travel distance. He proposed a schematic map in which space represents travel time, however route planning based on his map is challenging due to the large amount of overlapping lines. The work of Wang et al. [WC11] introduces a focus+context technique that allows interactive exploration of metro maps in small displays.

Figure 3: Stockholm metro map rendered with the Wang et al. method.

The outcome of focus+context methods are maps that highlight user routing information while partially showing related information. This method is effective in orienting the user but does not work if the user wants to explore different possibilities. Usually focus+context maps are implemented using distortion techniques such as fish-eye lens. Haertt [HS11] argued that distortion should preferably take place in those areas where the transport network is sparse and proposed a graph-based approach to compute a new spatial mapping that better uses the available space.

Figure 4: Smart distortion using Haubert et al. method.

5.3. Visual Analytics of Movement

Analysis of movement is currently a hot research topic in visual analytics and a wide variety of tools for analyzing movement data have been developed in recent years. Andrienko et al. [AA12] listed three visualization categories for movement data, namely direction depiction, summarization and pattern extraction. Zhong et al. [ZWZA12] and Goncalves et al. [GAM13] described common techniques that are used to visualize movement data, that are static maps, space-time cubes, animated maps, and small multiples. The work Nguyen et al. [NDBT12] address the problem of representing the temporal information on the routes, namely visually represents the different trips that occurs along the day in the same route; the solution that they come up with is to use a space-time cube in which two of the three axes represents the spatial information and the third axis gives the temporal information.

Figure 5: Normalized bus trip visualized on a space-time cube (Nguyen et al.).

This visual encoding is used to represent a normalized bus trip that is a bus trip obtained with a shape-based aggregation of all the bus trips of a given route. By picking a time on the third axis the user can infer all the other information. An interesting work that exploits animated maps is TRAVIC
[BBS], a visualization tool that shows live world-wide vehicle movements using smooth animations. The focus of this work is to show many transportation modes, since usually animated maps are limited to one transportation mode (e.g. trains, bus, etc.).

Figure 6: A screenshot of the TRAVIC tool showing vehicle movements of the Chicago PTS.

Another interesting work that uses animated maps is "Shangai Metro Flow" by Nagel et al. [NG]. The purpose of this work is to encourage users to reflect on their perception of the environment and one thing to notice is the visualization aesthetic which is very pleasant. Visual aesthetic is very important especially if the visualization is targeted for the general public.

Figure 7: A screenshot of the "Shangai Metro Flow" visualization (Nagel et al.).

Many approaches combine more than one technique to show movement data from different perspectives such as in the work of Janetzko et al. [JJDK13] where trajectory abstraction on a map is applied and a simplified graph visualization with coherent topology is also provided. A small-multiple technique is used also to show changes in movement of the observed subjects.

Another common visual encoding for geospatial movement data are flow maps. Flow maps are thematic maps used to display the movement of objects (e.g. vehicles, goods, etc) between different geographic regions. A pioneer of this technique is Tobler [Tob87] but in his work visual clutter easily occurred when applied to large datasets. An effective method to overcome this issue is edge bundling. Phan et al. [PXYH05] proposed the Flow Map Layout method to automatically cluster nodes into a tree hierarchical structure and then bundle neighboring flow lines to present the general flow trend. Verbeek et al. [BSV11] proposed a spiral-tree-based method to compute crossing-free flows that improves the flow maps readability.

Figure 8: Motion analysis of lions (Janetzko et al.).

Figure 9: Motion analysis of lions (Janetzko et al.).

An interesting overview on time-oriented data is provided by Aigner et al. [AMM’08].
5.4. Visual encoding of landmarks

Landmarks are important to help the user to orient within a map and to support user wayfinding process, therefore choosing the most effective visual encoding for landmarks is very important. Many salient features such as lakes, parks and bridges are already identified by their shapes and colors, and their visual encoding is straightforward. In the literature most of the works focus on the visual encoding of buildings. Buildings can be very helpful in supporting user wayfinding process since it is very common for people to identify different areas of a city through building landmarks, but finding the right visual encoding for building is not as straightforward as for lakes and parks.

Lee et al. [LKPM01] used cutouts from photographs and put them directly on the map to illustrate individual facade landmarks. Dollner et al. [DBNK05] following the work of Lee et all proposed a variation using non-photorealistic rendering techniques and this technique is often used in tourist maps. MacEachren [Mac86] discussed about iconic representation of landmarks and highlighted how pictorial representations are better than geometric abstract markers. Bruyas et al. [BLBP98] added that the symbol does not have to be too detailed and Klippel [Kli03] highlighted that logos are a useful candidate to depict trademarks shop landmarks. Elias et al. [EPK05] described a solution for choosing the right visual encoding depending on the level of zoom scale.

5.5. Three-dimensional visualizations

Three-dimensional visualization research is still in its preliminary stages in public transportation systems. Few works have been found in the literature that employs three-dimensional techniques to visualize transportation data. I have already mention in section 5.3 a three-dimensional technique [NDBT12] that uses space-cubes to show spatiotemporal information of buses. Kapler [KW05] developed a technique for displaying and tracking events, objects and activities within a combined temporal and geospatial display. Pina et al. [PCS12] developed a semantic visualization for 3D urban environments, illustrating a scenario in which the user decides which bus line to take using the visualization tool.
6. Future Challenges and Opportunities

The techniques surveyed in section 5 demonstrate the state of the art in visualizing public transportation systems. Nevertheless, research in this area is still in its early stage and many open problems and challenges need to be addressed in the future. In the following I give a list of interesting challenges that I think are worth to further explore.

6.1. Visualization of movement data

Visualization of movement data is a hot topic in the visualization community, yet no definitive solutions have been found that address these problems. An effective technique to visualize movement data should be able to effectively convey direction and speed. One of the major issues is the visualization of movement of thousand objects, such as bus trips in a city; finding an effective way to convey movements of all these objects is difficult and often depends on the purpose of the visualization tool.

6.2. Visualization of overview maps

Overview maps are effective to help the user to build a cognitive map of the environment [CR98]. The problem with overview map is the high degree of clutter usually due to the great amount of information to display. Several approaches such as filtering and aggregation can be used to solve the problem. Ultimately, being able to display a transit network in an overview map could help the user to explore different choices that he has at hand and thus can satisfy many user needs.

6.3. Three-dimensional visualization

Three-dimensional visualization of public transport data is another hot topic due to the advent of 3D maps and the availability of these data in vector format. The possibilities of research in this topic are many and very few works have addressed it. Possible research questions to be answered are:

When is appropriate to use them? How to optimize such visualizations for speed and portability? Performances are often an issue in these kind of representations.

6.4. Effective animations on real-time data

Another open problem is the visualization of real-time vehicles data. The problem relies on the fact the most of transit APIs today update bus positions about every minute, and therefore some interpolation is needed to animate vehicles on a map. A very rough approach is to update bus position only when there is a new update, but this cause the vehicle to jump from the old position to the new position and makes it difficult to estimate vehicles movement directions; moreover, the vehicle position is out of date until a new update comes.

New visualizations should study how to effectively interpolate bus updates, which is not an easy to solve problem.

6.5. Visualization of massive dataset

Often PTS datasets of cities present huge amount of data. All these amount of data exposes visualizations to performance issues, especially if they are web applications. Solutions available on the web employ server side preprocessing of data to reduce the amount of work that need to be done by the client. Nevertheless, the amount of information that could be displayed is still limited and guidelines on how to deal with this problem can be very useful.

7. Conclusion

Visualizing Public Transportation Systems is important to help to understand their usage in order to fix issues, to improve transportation efficiency and ultimately, make the system easy to use for travelers.

I have described what data is commonly use, GTFS being the standard and most commonly used format. I have overviewed also the main theoretical concepts that can help decide which is the most effective visual encoding based on the task. I have identified the task that visualization tools for public transportation data should support. These tasks are based on an identification of two kind of user, namely the traveler and the transportation researcher. I have reported many state-of-the-art techniques categorizing them in five main categories. All these techniques follow common methodologies such as aggregation/filtering of data, simplification and abstraction. Finally I have identified some promising research topics that can be worth to explore.

References

Massimo De Marchi / Visualizing Public Transport Systems


[BB98] Bas H., Broel P., Storandt S.: Travic: A visualization client for public transit data. 6


[Gar94] Garland K.: Mr beek’s underground map. 4


[NG] Nagel T., Gross B.: Shanghai metro flow-multiple perspectives into a subway system. 6


[SW75] Siegel A. W., White S. H.: The development of spatial representations of large-scale environments. Advances in child development and behavior 10 (1975), 9. 2


[WS05] Woyciechowicz A., Shliselberg R.: Wayfinding in public transportation. *Transportation Research Record: Journal of the Transportation Research Board* 1903, 1 (2005), 35–42. 1, 2, 3

[ZFA*14] Zeng W., Fu C.-W., Arisona S. M., Erath A., Qu H.: Visualizing mobility of public transportation system. 4