Atvis: A New Transit Visualization System

Jiaxuan Pang[1], Charles Tian[2], Yan Huang[3], Bill Buckles[4], Arash Mirzaei[5]

[1,3,4] University of North Texas
Email: tjpjx2012@gmail.com, huangyan@unt.edu, bbuckles@unt.edu

[2]University of Texas at Austin
Email: c.y.t@me.com

[5]North Central Texas Council of Governments
Email: amirzaei@nctcog.org

1. Introduction

Due to urbanization and an increase in private vehicles, most cities are becoming extremely crowded. Moreover, with rising awareness of environmental conservation, people are looking for more economical and eco-friendly ways to travel. As of 2011, there were 7,100 public transportation providers in United States [American Public Transportation Association, 2013]. These providers range from single-vehicle response servers to large multimodal systems. The largest amongst them, MTA New York City Transit, served 3.3 billion trips covering 12.2 billion miles in 2011. Cumulatively, bus transportation accounts for more than half of the 10.3 billion total trips taken and 36% of passenger miles covered for the year, making it the most popular form of public transportation. As a result, appropriate bus route optimization and stop allocation is crucial in helping public transportation providers adapt to the needs of people in rapidly developing cities.

Data visualization provides a medium through which decision makers can observe systems and draw conclusions. Agrawala (2002) presents a variety of visualization methods for map and route data and organized these methods by their level of abstraction, e.g. one-dimensional route map, standard road map, and physical models. More details provide more fidelity, but do not necessarily offer greater usability. A well-designed visualization system identifies the appropriate level of abstraction to provide maximal usability.

In this paper, we introduce a new abstract visualization system, *Atvis*[1], which uses a more appropriate level of abstraction for presenting bus route/traffic data compared to traditional systems. Map-based geographic visualization systems such as GIS, while commonly used, often struggle with the issue of readability, especially when large-scale data is presented [Mashima, Kobourov, & Hu, 2013]. Often times, these map-based systems are too detailed, thus distracting decision makers from observing individual and systematic traffic patterns.

Atvis proposes a system of arc and bar diagram representations of bus route data. It extends upon the arc diagram visualization method, used by Martin Wattenberg to look for patterns in strings and music [Martin, 2002], to represent bus ridership levels and distance between bus stops. Additionally, the system implements a bar diagram which offers more detailed information regarding the flow of bus passengers. *Atvis* places bus routes on a line map, and uses these diagrams to display key traffic statistics.

2. *Atvis*: A New Transit Visualization System

---

[1] http://hpproliant.cse.unt.edu/nctcog/

2.1 Goals and Objectives

The primary objective of *Atvis* is to provide decision makers with an intuitive representation of bus data in order to explore spatial and temporal relationships and observe systematic traffic patterns. Specifically, one can observe visually:

(1) Distance between stops
(2) Passenger allocation throughout the system for any time interval
(3) Boarding and alighting at each stop, individually and cumulatively at any time interval

We first explored traditional visualization methods using map-based data presentation. However, this adds a high level of geographical detail, which creates more distraction than information. Figure 1 summarizes our process in transitioning from a GIS based model to abstract data visualization.

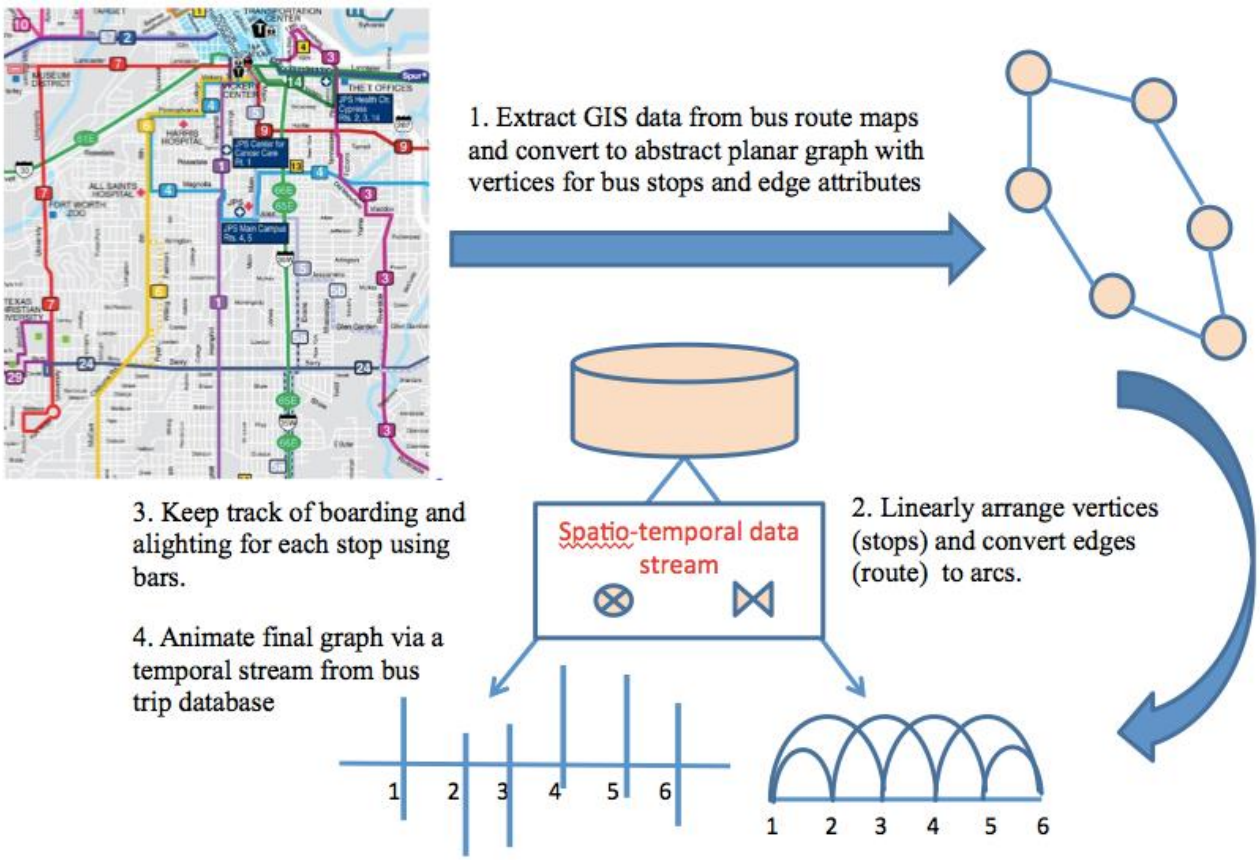

Figure 1. Geographical to Abstract Visualization

2.2 Data Description

The bus dataset we used for visualization was provided by the North Central Texas Council of Governments (NCTCOG). It contains 1,827 records from November 30, 1999, detailing bus trips on Route 2 in Fort Worth, Texas. Each record represents a single stop made on a single bus trip along the route (from the first to last stop). Figure 2 is a map of Route 2 provided by NCTCOG.

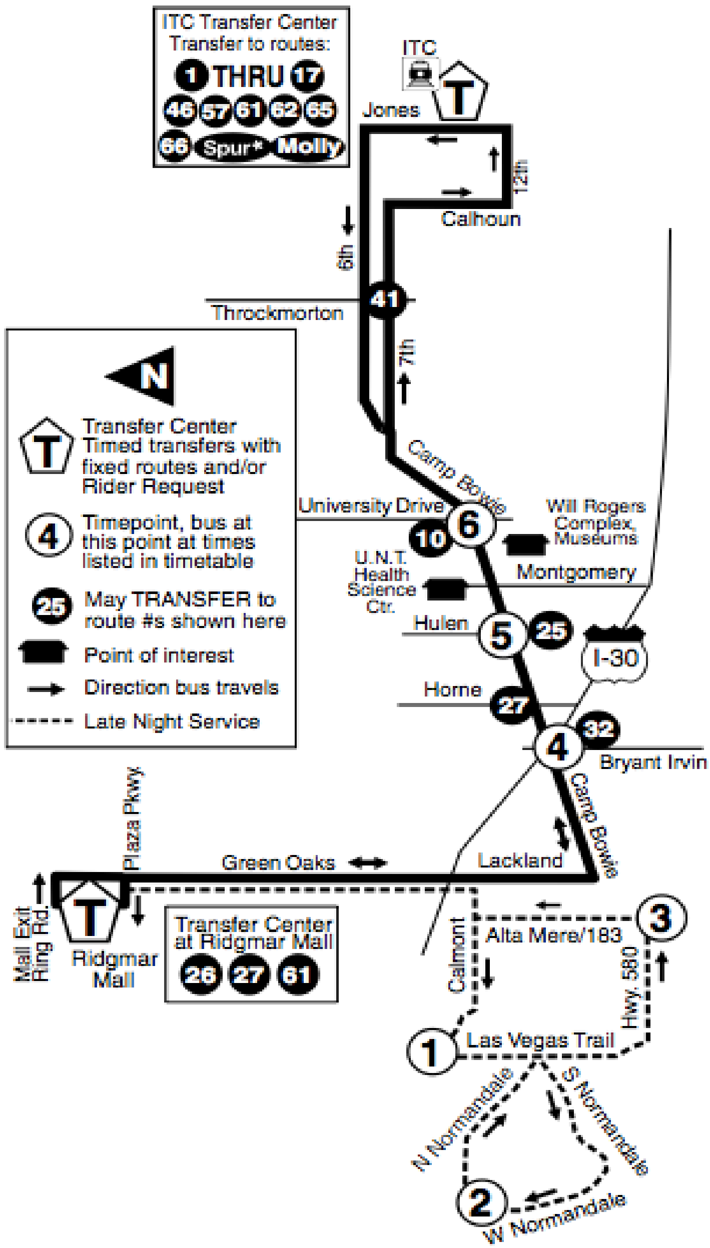

Figure 2. Map of Route 2

Table 1 offers a detailed description of the NCTCOG bus data.

Field Name	Data Type	Field Description
ID	Integer	Unique identifier for each row of data
Route_name	String	Name of the bus route being visualized
Trip_ID	Integer	Unique identifier for each bus trip

Stoporder	Integer	Unique identifier for each stop, based on order of all stops along the route
Latitude	Double	Latitude coordinate the bus stopped at
Longitude	Double	Longitude coordinate the bus stopped at
Arrival_time	Integer	Time in seconds from midnight to when the bus arrived at the "Stoporder" stop, on November 30, 1999
Departure_time	Integer	Time in seconds from midnight to when the bus departed from the "Stoporder" stop, on November 30, 1999
Dwell_time	Integer	Time in seconds that the bus dwelled at the "Stoporder" stop
Board	Integer	Number of passengers who boarded the bus at the "Stoporder" stop
Alight	Integer	Number of passengers who alighted the bus at the "Stoporder" stop

Table 1. NCTCOG Data Description

2.3 *Atvis* System Design

Atvis allows users to view and analyze trends and information for customized bus routes and datasets. Our system offers the following three major advantages:

(1) Multidimensional data profiles in a two-dimensional format
(2) Seamless integration of spatial and temporal information
(3) Combination of local and global information in one view

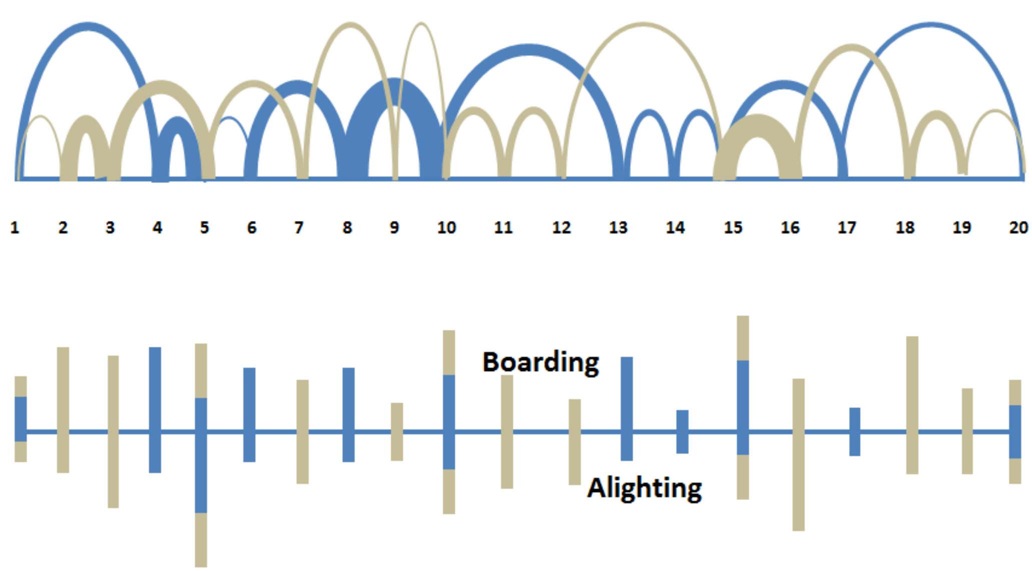

Figure 3. *Atvis* Final Visualization Model

Figure 3 depicts the *Atvis* final visualization model, using sample data for two routes consisting of a total of 20 stops. The endpoints of each arc in the arc diagram (upper graph) represent stops along the visualized route. These stops are arranged corresponding to the route's stop order. For multiple routes, it is unlikely that a single linear order will be consistent with all routes' stop order. In that case, a partial order will be created among the stops instead. The height of each arc represents the distance between the two stops, while the thickness of each arc represents ridership data. The bar diagram (lower graph) consists of bars which correspond to stops in the arc diagram. The component of each bar above the axis represents the number of boarded passengers, while the component of each bar below the axis represents the number of alighted passengers. The combined height of these two components measures the total passenger flow at each stop.

Table 2 summarizes how data is represented through the arc and bar diagrams.

Arc Legend		**Bar Legend**	
Integer	Bus stop order	**Integer**	Bus stop order
Arc Color	Bus route	**Bar Color**	Bus route
Arc Thickness	Number of passengers on the bus	**Bar Above Axis**	Number of passengers boarding
Arc Height	Distance between stops	**Bar Below Axis**	Number of passengers alighting

Table 2. Bar and Arc Legend Description

Atvis visualizations progress through time in Accumulative or Real-Time fashion and can be slowed down, sped up, paused, and resumed by the user. Changes in passenger flow are reflected when the system clock reaches the time that they occurred. Users can specify start and end times for data visualization, allowing for more straightforward analysis of bus traffic based on time of day.

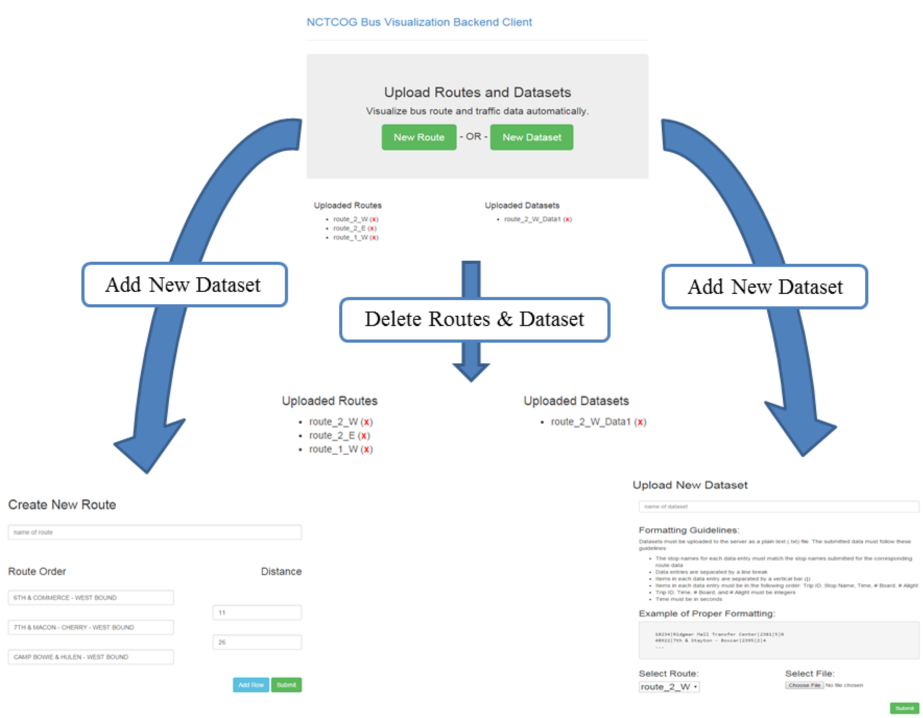

Figure 4. *Atvis* Backend Data Management System

2.3.1 Backend System

Our backend system[2] allows users to manage, define, and upload bus routes and corresponding datasets to be visualized. Figure 4 describes the flow of route and trip data submission.

The *Atvis* backend system supports the creation of multiple bus routes and multiple bus stops within each bus route. Users must provide the name of each stop, as well as the distance in miles between stops. Figure 5 is a screenshot of the route creation system.

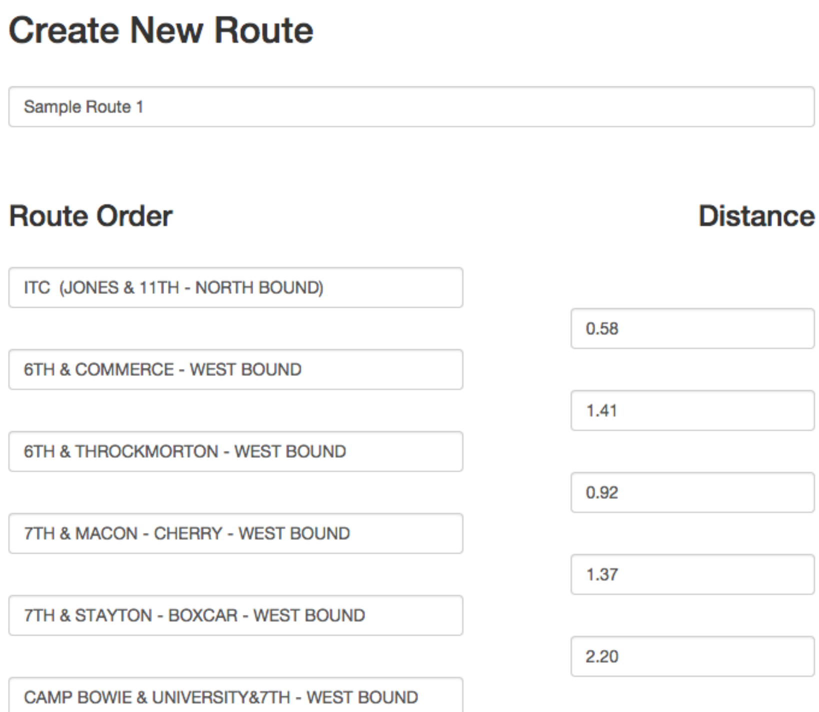

Figure 5. Backend Route Creation System

Datasets are uploaded to the server as a plain text file. Users must also indicate the route with which their dataset corresponds. The data must be formatted to comply with the following requirements:

(1) Stop names for each data entry match stop names for the corresponding route
(2) Data entries are separated by a line break
(4) Items in each data entry are separated by a vertical bar (|)
(5) Items in each data entry are in the following order: Trip ID, Stop Name, Time, # Board, # Alight
(6) Trip ID, Time, # Board, and # Alight must be integers
(7) Time must be in seconds

Figure 6 offers an example of how the dataset should be properly formatted.

[2] http://hpproliant.cse.unt.edu/nctcog/backend/

```
18234|Ridgmar Mall Transfer Center|2381|5|0
48922|7th & Stayton – Boxcar|2395|2|4
...
```

Figure 6. Example of Proper Dataset Format

2.3.2 Frontend System

Atvis can visualize multiple routes and datasets which are defined in our backend system. Thus, we offer a menu that allows users to select the data they wish to display, as shown in Figure 7.

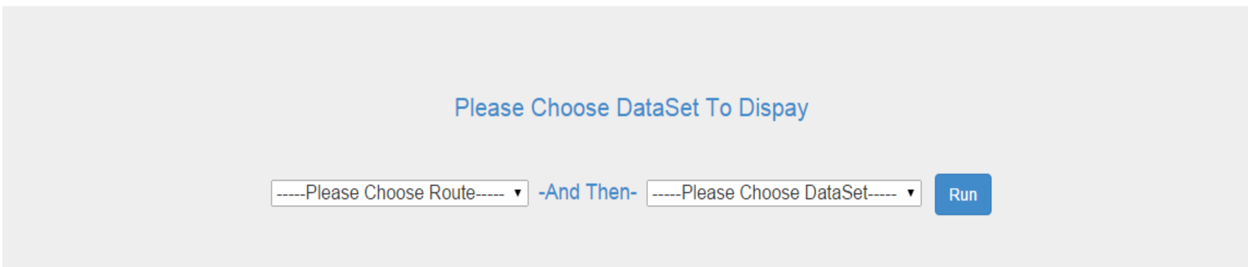

Figure 7. *Atvis* Frontend Data Selection

Figure 8 is a screenshot of the *Atvis* visualization system. Our original model and final implementation differed in a few key ways:

(1) Our final implementation only allows for the visualization of one route and dataset at a time

(2) We add interfaces which allow users to control the simulation timing and visualization methodologies (described in greater detail in section 2.4)

(3) We add a timeline which allows users to specify the starting/ending time of the simulation, and view the current simulator time

(4) Users can mouse over both the arc and bar diagrams to view more detailed ridership information at each stop

(5) Users can zoom in/out of the arc and bar diagrams in order to view ridership and passenger flow statistics in greater/lesser detail

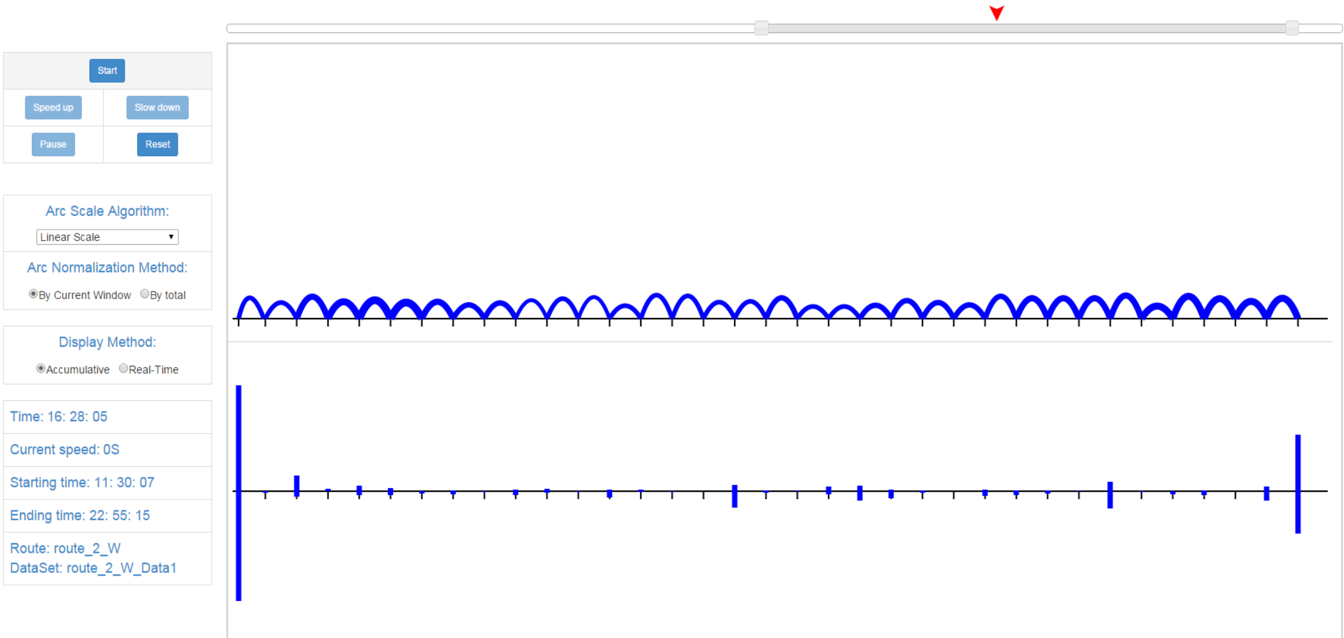

Figure 8. *Atvis* Visualization System

2.4 *Atvis* Visualization Methodologies

Atvis supports a variety of visualization options, which display different types of information using different algorithms and calculation methods. These options are as follow:

2.4.1 Display Method

Atvis offers Accumulative and Real-Time options for displaying route and traffic data. The Accumulative method displays the ridership situation from the start of the simulation, up until the current time. The Real-Time method presents the ridership situation at exactly the current time in the simulation.

2.4.2 Arc Normalization Method

Atvis normalizes ridership values as arc thickness using two different methods: By Current Window and By Total. These two methods differ based on the scope of data that is presented.

Using the Accumulative display option:

Both methods calculate a normalization factor to be used in determining arc thickness, then apply this factor to the Arc Scaling Algorithm (described in section 2.4.3). The By Current Window method does this based on the range of data from the start time to the current simulator time, while the By Total method does this based on the range of the entire dataset.

Using the Real-Time display option:

The By Current Window method determines the arc thickness normalization factor based on the data at the current simulator time. Using the By Total method while displaying data in real time produces a visualization that lacks detail; as a result, we have disabled this option from the *Atvis* system.

2.4.3 Arc Scaling Algorithm

The arc scaling algorithm controls the method in which ridership values are presented in the arc graph. *Atvis* offers three different algorithms for scaling data in this view:

(1) Linear Scale – The base scaling function utilized by the following two scaling algorithms. A

scale is determined for the dataset by linearly distributing thickness values throughout the range of data.

(2) Square Root Scale – Takes the square root of all ridership values, and then applies the Linear Scale to them.

(3) Relative Magnitude Scale – Takes the square root of the largest ridership value, then uses the Linear Scale to scale the rest of the values relative to this value.

3. Discussion/Conclusion

Atvis allows a transit planner to visually observe bus data, pinpoint individual stops and transitions, and come up with research hypotheses. The following is a list of questions that can be observed or answered by the *Atvis* system:

1. Which stop has the highest boarding/alighting level at 8:11 A.M?
2. How many people ride the bus between stop 2 and stop 3 from 9 A.M. to 10 A.M?
3. Which two bus stops have the most ridership on Saturday?
4. Does the distance between two stops have any correlation with their ridership levels during rush hours?
5. Do any stops have abnormal boarding and alighting ratios for a given time period?

Atvis excels at visualizing transit data at an abstract level. Our focus on ridership and inter-stop relationships creates a visualization that emphasizes key bus/route statistics and removes the distraction of unwanted geographical data, allowing for more direct observation of transit patterns. Our innovative use of the arc diagram in conjunction with a bidirectional bar graph offers a holistic representation of transit data in a single view.

Atvis is currently is not implemented fully for multiple routes and datasets. Future implementation of this functionality can be useful to transit planners. A simple extension is to add the relationship between trips headed towards opposite bounds, but along the same geographical route. A general extension is to add multi-route using partial order and multi-dataset visualization to allow users to compare ridership and passenger flow through different bus systems at the same simulator time.

Atvis can also be improved by expanding data presentation to allow zooming in to individual bus trips. Doing so allows *Atvis* to not only help analysts determine trends in bus data for the overall system, but also offer a breakdown of ridership and passenger flow information for each trip. This feature would allow users to analyze data from a different perspective, while maintaining the other advantages of our visualization.

5. References

2013 Public Transportation Fact Book. 64th ed. Washington, DC: American Public Transportation Association, 2013. Online.

Agrawala, Maneesh "Visualizing Route Maps." (Doctoral Dissertation) (2002) Stanford University.

Daisuke Mashima, Stephen G. Kobourov, and Yifan Hu. "Visualizing Dynamic Data with Maps." Visualization and Computer Graphics, IEEE Transactions on 18.9 (2012): 1424-1437.

Wattenberg, Martin. "Arc Diagrams: Visualizing Structure in Strings." *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on*. IEEE, 2002.