Virtual City
A Heterogeneous System Model of an Intelligent Road Navigation System Incorporating Data Mining Concepts

by

Mike Kofi Okyere
School of Informatics
Indiana University - Bloomington
300 N. Jordan Avenue
Bloomington, IN 47405. USA
mokyere@indiana.edu

Faculty Advisor: Edward A. Lee, PhD.
Mentor: Yang Zhao

Abstract

Generating satisfactory directions for route guidance is a challenging task, because the effectiveness and advantage of particular routes depend on various road specifications and characteristics. Current route planners, such as MapQuest.com, Yahoo.com and Maps.com, present only limited route options to the drivers of personal or commercial vehicles. Such directions are based on static evaluation criteria and do not consider real-time information related to traffic conditions, road construction, road inclination or weather conditions. This paper proposes a more efficient approach to route finding by applying data mining concepts in order to create a user customized route guide. The Intelligent Road Navigation System consists of a dynamic data warehouse that contains real-time road information, ranging from road name and length to road inclination and traffic density. The most efficient route is calculated by an extension of the Dijkstra's Shortest Path Algorithm to obtain driving directions that focus on minimizing either travel time, gasoline usage, driving mileage or a combination of all.

1. Introduction

In today's world one of the most critical social and economical tribulations is traffic congestion. As the population grows so will its travel needs, therefore it is crucial to develop an intelligent route planning system that reduces the time spent traveling. In addition the steady increase of the traffic volume and the decrease of infrastructure expansions will ultimately reinforce the need for a more reliable and efficient route navigation system.

Traffic jams, and particularly unexpected congestion, have a very strong impact on travelers' lives. Motorists and transit riders would like to know what they need to expect when traveling. The possession of precise real-time data, regarding road and highway performance as well as characteristics, changes the experience and length of a commutation, since the provided information presents the motorist the ability to control the time spent on journeys. The awareness of the approximate travel time enables the driver to, either postpone the departure time, change the route, destination or even cancel the trip should the travel length cause scheduling or interest conflicts. Receiving advanced notice on major construction projects via the radio is nowadays very common, yet integrating that type of information into route planning leads to a more accurate travel time calculation. The employment of such data decrease the overall congestion occurrences and time at the road construction sites by supplying alternative routes that make use of the existing infrastructure. Not only are dynamic route planning features crucial for the general public but also for commercial freight carriers, since their reliability decreases due to the arrival delays caused by traffic and road conditions. Many companies are experiencing decreased profits, due to the rapid growth in overtime cost due to the traffic density. "Costs also increase from an inability to schedule work for their vehicles over the complete workday, as the companies lengthen expected delivery times just to ensure that they don't have to pay overtime."[1]

Throughout this research project an Intelligent Road Navigation System has been created that is linked to a dynamic data warehouse, which could revolutionize commuting. The dynamic route planner differs from current inventions, by placing a greater emphasis on real-time data, such as traffic conditions, road constructions, road inclination and weather conditions. The most efficient route is calculation by accessing the given real-time data that is then processed by a Ptolemy II model that employ an extension of the Dijkstra's "Shortest Path Algorithm" on the dynamic data warehouse.

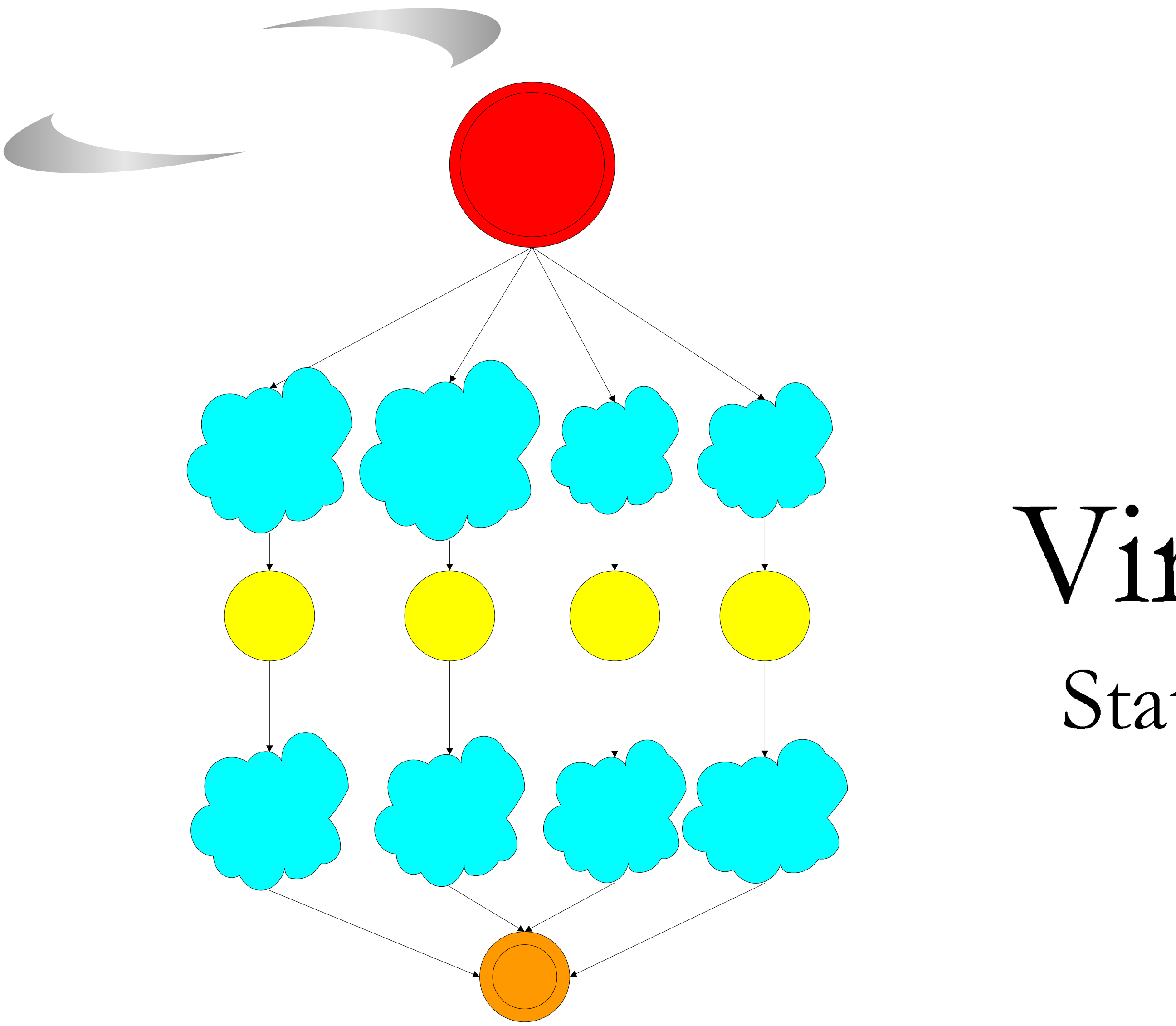
2. Disadvantages of Contemporary Route Planners

Current route planning applications, such as MapQuest.com, Yahoo.com, and Maps.com, provide limited approaches to decrease travel times and distances. Most of these route guides make use of the shortest path algorithm and generally use static data, such as road length and a constant travel speed, to calculate the shortest or "fastest" route. In order to significantly reduce the cost associated with congestion and to drastically increase the capacity of goods and people that can pass through the existing infrastructure, one can not exclusively rely on static road information. In the case of Downtown New York City traffic, local commuters naturally assume that the time of day is intertwined with the travel speed and adjust their departure in accordance to their earlier experiences. Yet it would be very beneficial for everyone, who desires to travel through New York City to have the same information as visitors as local residents do.

3. Intelligent Route Navigation System

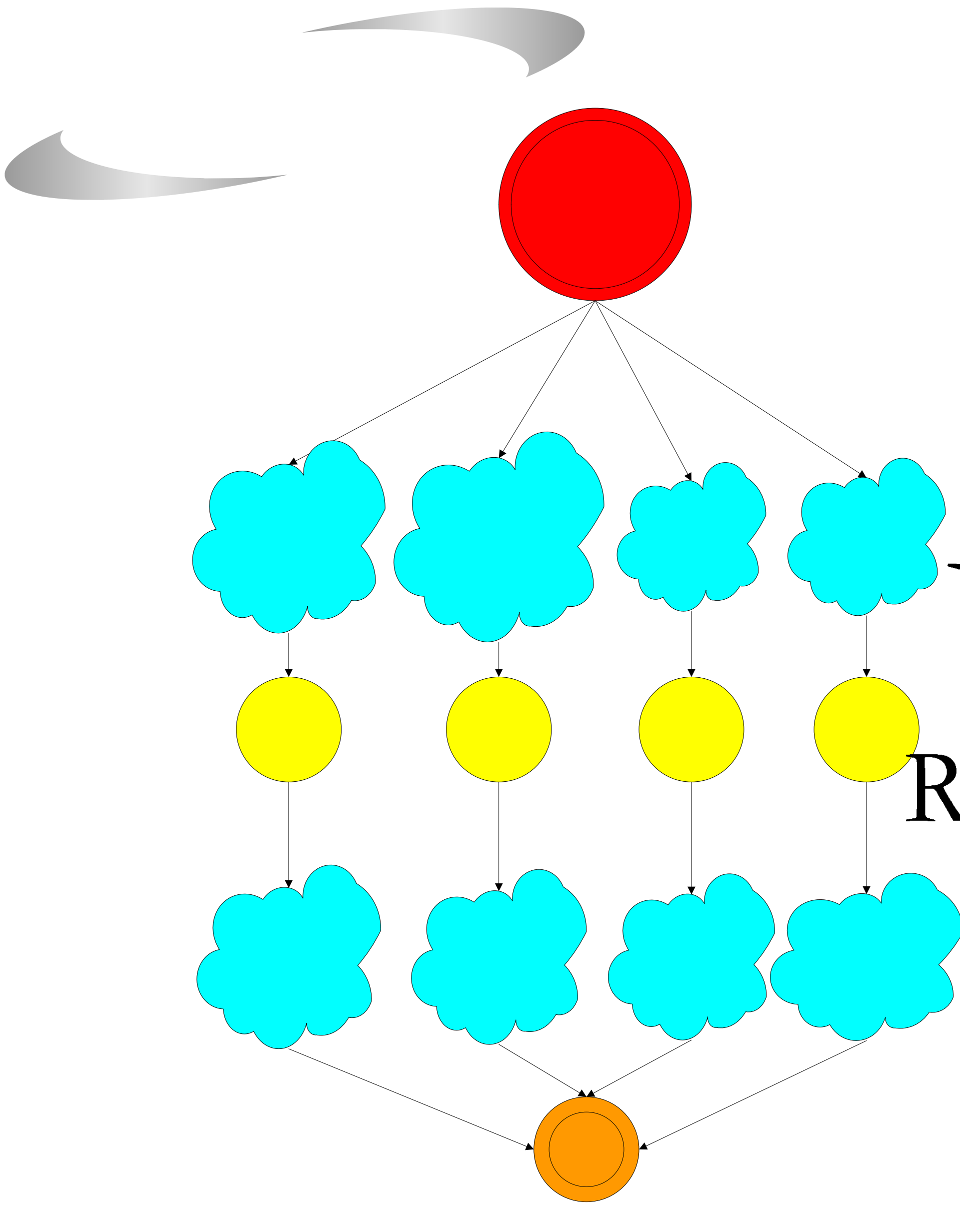
The important features of intelligent route guides are their ability to take into account dynamic road characteristics and information such as congestions, construction sites, weather conditions and alternative routes. Such information could be retrieved through advanced traffic management systems, such as Intelligent Transportation System (ITS), which are currently used in law enforcement, in order to detect congestions and unusual traffic activities. Through the wide usage of such navigation systems, roadways could operate at peak volume levels and in the rare cases of traffic overflow the navigation guide would transfer the overload to other routes. The average delay time would over time be significantly decreased, leading to a reduction of the total amount of travelers caught in congestion, as well as the duration of slight congestions when they occurs.

Possessing the ability to choose ultimately improves the traveling experience and the commuter's attitude towards traveling in general, and since many drivers choose to avoid backups, traffic jam occurrences will over time decline. Although this could increase the traffic behavior and volume the Intelligent Route Navigation System can retrieve patterns within the old and current data in order to recreate traffic system balance.



Static Data			
	Travel Time(min)	***Total Length(miles)***	***Average Speed (mph)***
Route A	100	95	57
Route B	107	105	59
Route C	110	110	60
Route D	118	100	51

Figure 1



Virtual

Real-Time R

Real Time Data			
	Travel Time(min)	***Total Length(miles)***	***Average Speed (mph)***
Route A	122	95	47
Route B	118	105	53
Route C	124	110	53
Route D	123	100	49

Figure 2

Highwa
Speed L
70mph

Even those who do not possess the option of postponing their trips will benefit from the change in travel perception, since there will be an overall lower traffic density in contrast to the one that they has to face in the pre-intelligent navigation era.

As an example, Figure 1 and Figure 2 show the main differences between the static and dynamic route planning. Figure 1, which portrays a static road map that mainly consists of road lengths and speed limits. In this figure Route A appears to be the fastest connection between Work and Home, in accordance to the given speed limits and road lengths. Those two values are used to calculate the total travel time, using the formula shown in Figure 3.

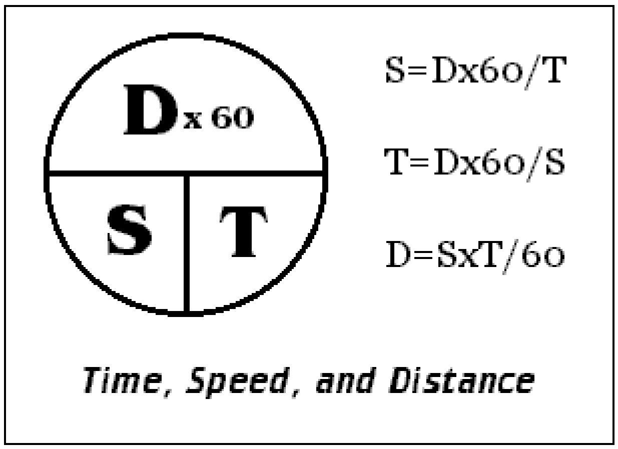


Figure 3

When considering the various traffic conditions that one will encounter when traveling from Work to Home, we arrive at the conclusion that Route B must be the faster connection between the two points. In order to arrive at this conclusion, we factor in the speed reduction that certain traffic behaviors cause. Therefore the updated formula will be as seen in Figure 4.

$$T = [D*60*SR]/S$$

SF = Speed Reduction Factor

Figure 4

The effects of the road congestion make traveling on Route A 15 minutes longer, when comparing it to Route B. Yet the static route planner suggests taking Route A, since in their calculation omitted dynamic data leading to the conclusion that the travel time from Work to Home would take 100 minutes, since none of the speed reducing factor of travel were considered.

4. Intelligent Route Navigation System

The Intelligent Route Navigation System makes use of static and dynamic road information to calculate the most efficient route between two user-specified points. Ptolemy II is used as the main platform, where components are aggregated and connected to create the dynamic route finder. The Intelligent Route Navigation System integrates a data warehouse, a digraph constructor and a component that specializes in finding the shortest path between two points by applying an extension of Dijkstra's Shortest Path Algorithm.

4.1 Data Warehouse

The data warehouse collects data from various resources to then systematize and organize the findings in a logical and accessible formation. The aspiration is to provide one dependable and uniform assembled view of the collected data, rather than a collection of multiple scattered files. The strength of the data warehouse is its organization and delivery of data in support of route planning decision-making process. The designed data warehouse is the physical implementation of a multidimensional decision support model which presents the route planner with the information it requires to determine the most efficient routes in accordance to the user-specified request

Possessing more background information that can be retrieved through the data warehouse will enhance decision making and increase driver confidence in the navigation system, as well as decreasing the overall road congestion on the current infrastructure, due to the data's time and assessment accuracy. The ultimate outcome therefore is the opportunity to further analyze trends and traffic patterns and to predict possible traffic behavior using historical data. The data warehouse that is being used for the navigation system will consist of a collection of incorporated, road condition-oriented databases designed to support the decision support system function, giving each unit of data its relevance at some point in time and therefore ensuring open access to the same data at the same level of extraction. This feature will consequently eliminate conflicting analytical results and concerns over the source and quality of data, creating an ideal infrastructure and travel environment. The complete database Model can be found in Appendix A.

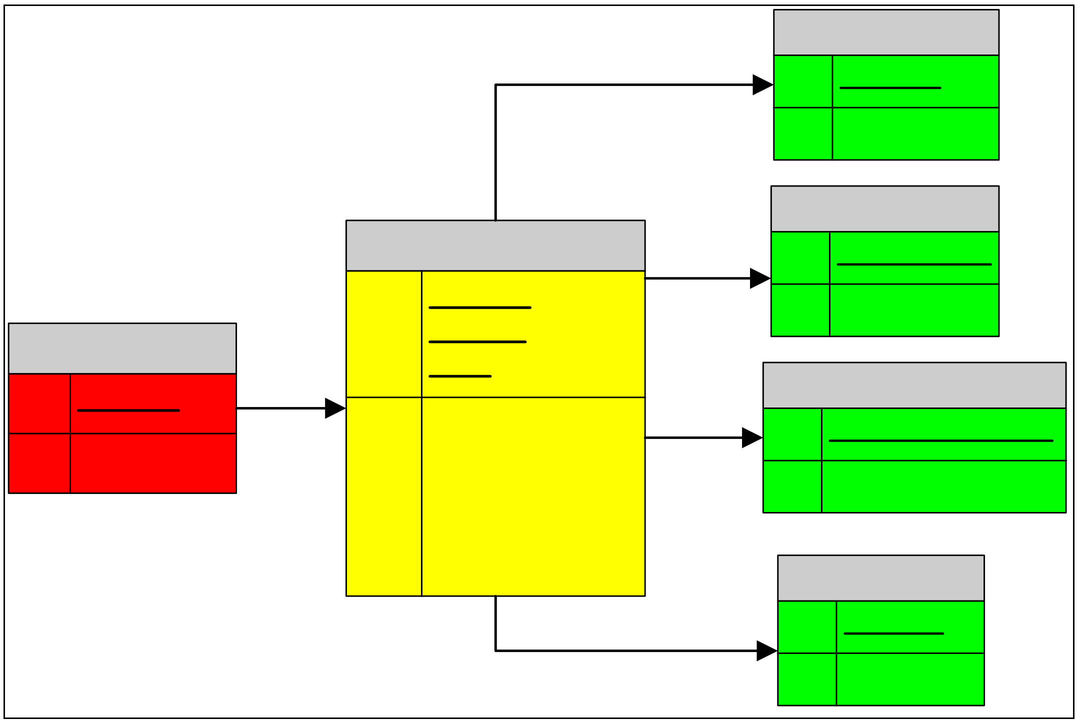


Figure 5

A prototype data warehouse has been implemented to assess these ideas. Figure 5 shows the relationships amongst the various entities. The system considers road behaviors such as traffic, construction sites, and road inclination when calculating the most efficient route between two locations. The database that consists of roadway characteristics could also be linked to vehicle characteristics, such as the cargo weight and type for trucks, in order to return a more accurate route in accordance to the vehicle performance. A sampling scheme to select vehicles and days could provide representative data for any geographic region, or vehicle or driver population and expand the traveling and transportation options as well as reduce the overall time and travel expense.

4.2 GraphBuilder Actor

The GraphBuilder actor accepts a single input from the DatabaseReader Actor, which sends a set of results that are being retrieved by executing an inputted SQL query

Roads

string on the created data warehouse, which is specified by the databaseURL parameter as can be seen in Figure 6.

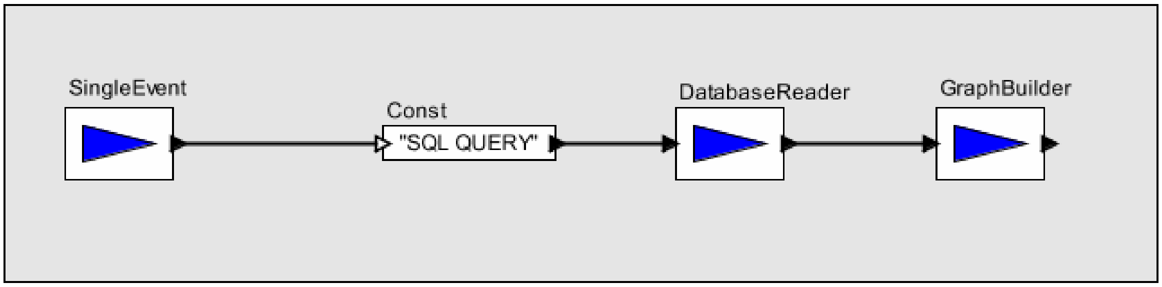
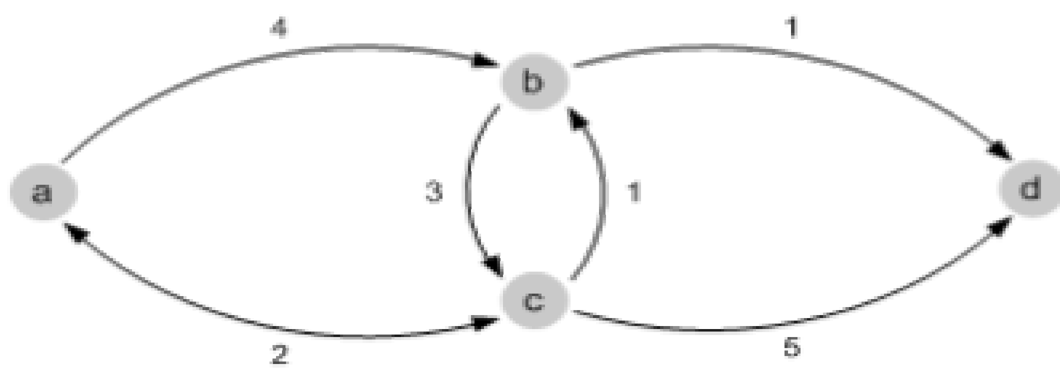


Figure 6

The inputted result set has a three column format, consisting of start point, end point and a weight. The set represents a directed graph (digraph). The concept of a digraph is extremely useful, with the finite set of elements called vertices or nodes and finite set of directed arcs that pairs of vertices that carry a certain weight. In this system, the weights represent the cost of going down a particular arc



Graph G = (V, E):
V = {a, b, c, d}
E = {(a, b, 4), (a, c, 2), (b, c, 3), (b, d, 1), (c, a, 2), (c, b, 1), (c, d, 5)}

Figure 7

The main purpose of using the digraph is to represent a network in this scenario consisting of roads. In such a graph the vertices represent the road intersections and the edges relate to the possible routes between the intersections as can be seen in Figure 7. In this setting there are several possible cost metrics: To compute travel distance, the digraph's weights represent the distance between road intersections. Should the user be interested in receiving a quota on travel time, using static data, then the weights on the edges would represent the calculated time considering the speed limit and the road length.

4.3 Dijkstra's Shortest Path Algorithm Actor

The ShortestPath actor receives two inputs, one being a constant, which consist of a string array that describes the user's current location and desired destination. Figure 8 shows how the input is being used to then create a start and end node that are being used to search the directed graph, which is being provided by the GraphBuilder actor.

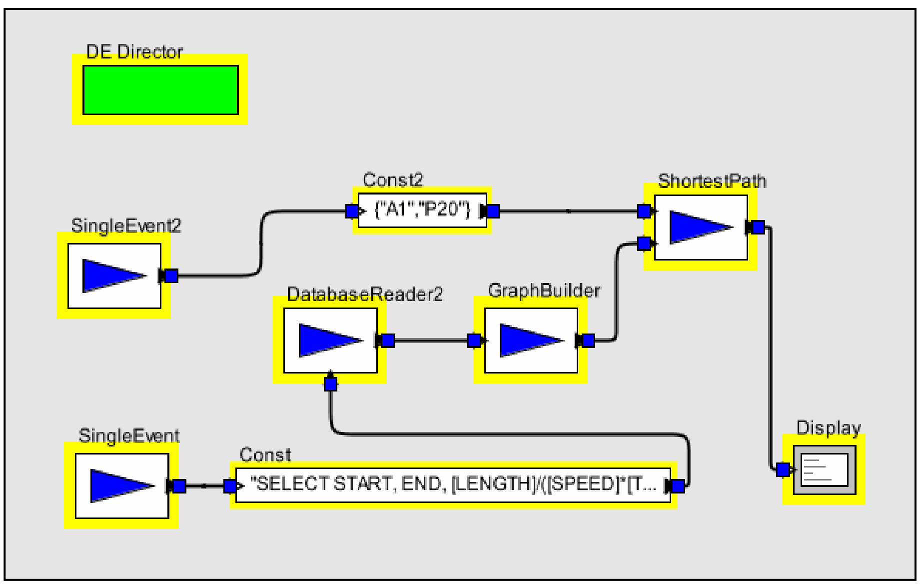


Figure 8

The Shortest Path actor initially verifies the validity of the input nodes and graph, in order to search the directed graph to find the shortest path. The search takes place by starting at the start node and then finding the neighboring node with the lowest weight, after successfully completing that task, it stores the value and adds the next lowest weight from the next edge it chooses. Through this approach the algorithm branches out by selecting specific edges that lead to new vertices until it reaches the final destination as shown in Figure 9.

Shortest Path from (a) to (d) is calculated as follows:

1. Init: d(a) = 0, d(b) = INF., d(c) = INF., d(d) = INF.
2. 0+ [a,b] = 4 < INF. [set distance from (a) → (b) = 4]
   0+ [a,c] = 2 < INF. [set distance from (a) → (c) = 2]
   [Pick [c]]
3. 2 + [b,c] = 3 < 4 [set distance from (a) → (b) = 3]
   2 + [b,d] = 7 < INF. [set distance from (a) → (d) = 7]
   [Pick [b]]
4. 3+ [b,d] = 4 [set distance from (a) → (d) = 4]

The algorithm stops, since the shortest path has been found.

Shortest Path: (a) → (c) → (b) → (d)

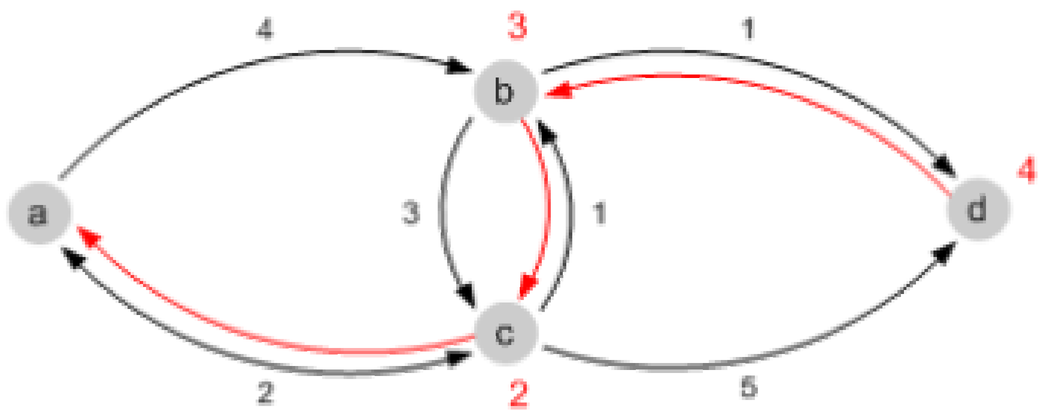


Figure 9

5. The Intelligent Route Navigation System (IRNS) – Sample Run

In order to start the route retrieval, the user inputs two locations on the map that is shown in Figure 12, in this case "A1" and "P20". The Intelligent Route Navigation System calculates the most efficient route between the two locations, taking into consideration traffic behavior and construction sites, which both affect the travel speed. To receive the proper digraph, we input the Query seen in Figure 10 into the DatabaseReader actor, which then outputs the result set showing the travel times for all road segments ranked from shortest to longest.

```
SELECT START, END,
[LENGTH]/([SPEED]*[TrafficFactor]*[ConstructionFactor])*60 AS TrafficTime

FROM Construcion INNER JOIN (Traffic INNER JOIN RoadRules ON
Traffic.TrafficTypeID = RoadRules.TRAFFIC) ON
Construcion.ConstructionSiteTypeID = RoadRules.CONSTRUCTION;"
```

Figure 10

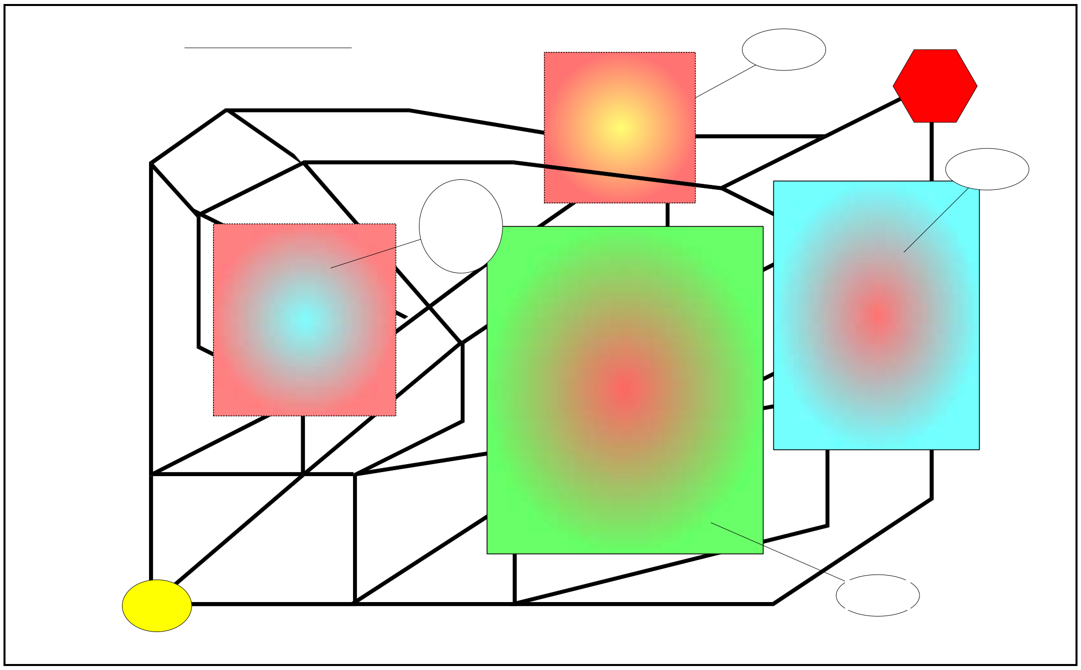


Figure 11

The GraphBuilder actor is the recipient of these sets and converts them into directed graph (digraph). After the successful construction of the graph the Shortest Path actor then verifies the validity of the input nodes and graph, in order to search the directed graph for the shortest route.

Figure 11 shows the least time-consuming route, using dynamic data, such as road construction sites and city and highway traffic. In contrast to Figure 12 Figure 13 shows the suggested route using static information, mainly being road length and constant speeds. The routes fundamentally differ in direction and road and highway usage.

Street

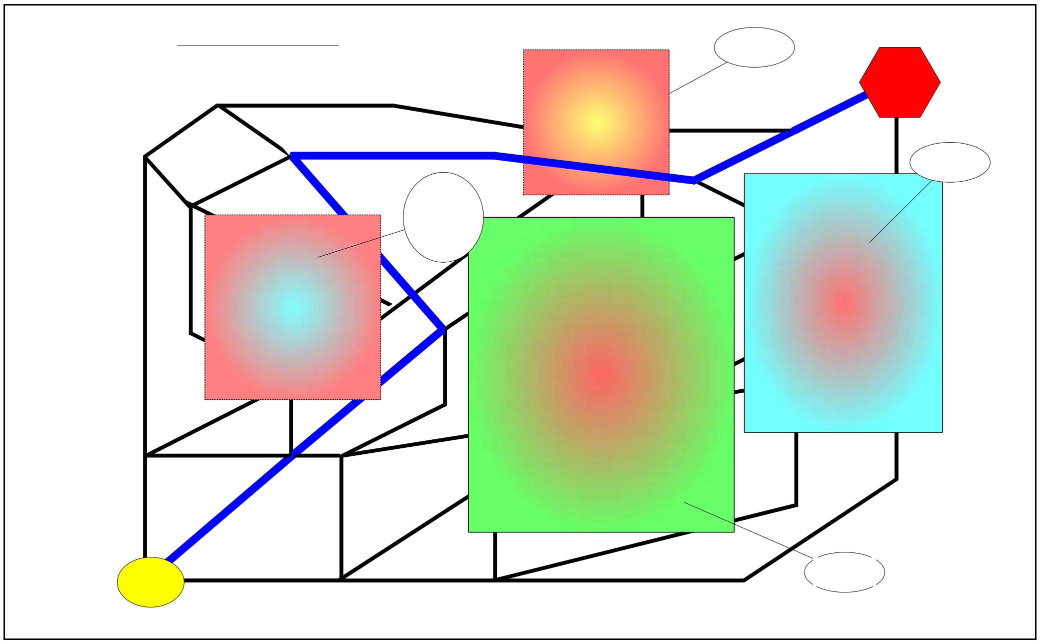


Figure 12

Stree

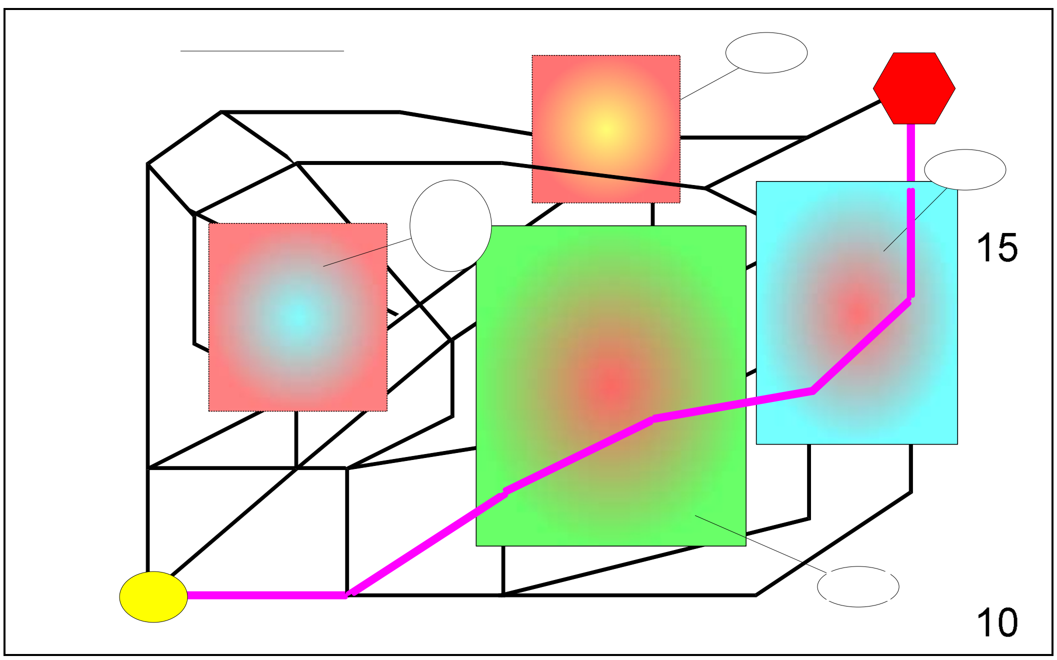


Figure 13

5

Conclusion

The Intelligent Route Navigation System helped us find the most efficient route between two locations using the gathered data that has been stored in a dynamic data warehouse. The user can pick the type of information that he/she would like to receive through the dynamic route finding system, such as minimum travel time, lowest gas usage, shortest distance, highest average speed etc. The different options are calculated using SQL queries on the data warehouse and can easily be expanded. The current data warehouse still lacks depth, yet the one of the future goals will be the gathering of real-world road information in order to expand the data warehouse, and to test the existing system.

Acknowledgement

Thanks to my parents, Dr. Kofi and Monika Okyere

Prof. Edward A. Lee

Yang Zhao

Steve Neuendorffer, Xiaojun Liu, Daniel Cuadrado & Vinny Blakes

Superb-IT

Reference:

[1] Seattle Wide-Area Information for Travelers (SWIFT) Consumer Acceptance Study, SAIC (Trombly, Wetherby & Dixson), October 19, 1998.