Representation, Visualization and Querying of Sea Turtle Migrations Using the MLPQ Constraint Database System

SEMERE WOLDEMARIAM and PETER Z. REVESZ
Department of Computer Science and Engineering
University of Nebraska-Lincoln
Lincoln, NE 68588-0115
USA
revesz@cse.unl.edu http://cse.unl.edu/~revesz

Abstract: - Animal migration is one of the areas biologists actively investigate in order to gain a better understanding of animal behavior. In general, the representation, visualization and querying of the observed animal migration data poses many challenges. This paper describes a novel approach to solving these challenges using the MLPQ constraint database system. Our modeling of the animal migration problem and specific software implementation using the MLPQ constraint database system provides an easy to understand and use scientific tool for biologists as is demonstrated on a test case of migration data for Atlantic Ocean sea turtles. The migration data include the position, latitude and longitude of the sea turtles as reported to satellites at certain time intervals. The migration observational data are interpolated and converted into a continuous movement data that are represented within a constraint database that can also animate and query the movement of the sea turtles.

Key-Words: -Animation, constraint databases, migration, moving objects, spatio-temporal interpolation, turtle.

1 Introduction

Animal migration is one of the most observed phenomena in biology because it is an important part of animal behavior [1]. Archie Carr [2] mentions an incident regarding sea turtles told to him by a captain of a turtle fishermen's boat in the Cayman Islands. According to the captain, a boat filled with sea turtles captured in their feeding area around northern Nicaragua, was shipped to Florida. Around the coast of Florida the boat capsized. The turtles were marked, customarily with the fishermen's initials, and to the surprise of the fishermen the turtles got back to where they were captured initially.

More recently, Lohmann, et al. [1] have studied hatchling and migrating juvenile loggerhead (*Caretta caretta*) and green (*Chelonia mydas*) sea turtles. These juvenile turtles originate from the east coast of the USA. They show strong homing behavior, that is, they usually return to their feeding sites after migration.

The representation, visualization and querying of the animal migration data, and in general moving objects data, currently still poses many challenges [5]. This paper describes a novel approach to solving these challenges using the MLPQ constraint database system [6], which was developed at the University of Nebraska–Lincoln. It is an advanced implementation of the constraint database model that was proposed by Kanellakis et al. [3].

This rest of this paper is organized as follows. Section 2 describes the original data source and our data curation method. Section 3 outlines a constraint database solution. Section 4 presents the main results and analysis. Finally, Section 5 presents some discussion of the results and gives some conclusions.

2 Data Source and Curation

We obtained Atlantic Ocean sea turtle migration data from Prof. Lohmann at the University of North Carolina at Chapel Hill [1]. The data used are the latitude and longitude of the boundaries of the US states [7], the geographical span of the United States [8] and the tracked latitude and longitude location of three sea turtles during their migration in the Atlantic Ocean. The United States map is used to give a perspective of the movement of the sea turtles relative to the US coast in the Atlantic Ocean.

The data from Dr. Lohmann includes the location of three sea turtles tracked over the time while the turtles navigate the Atlantic Ocean. For convenience, the three turtles are named here as,

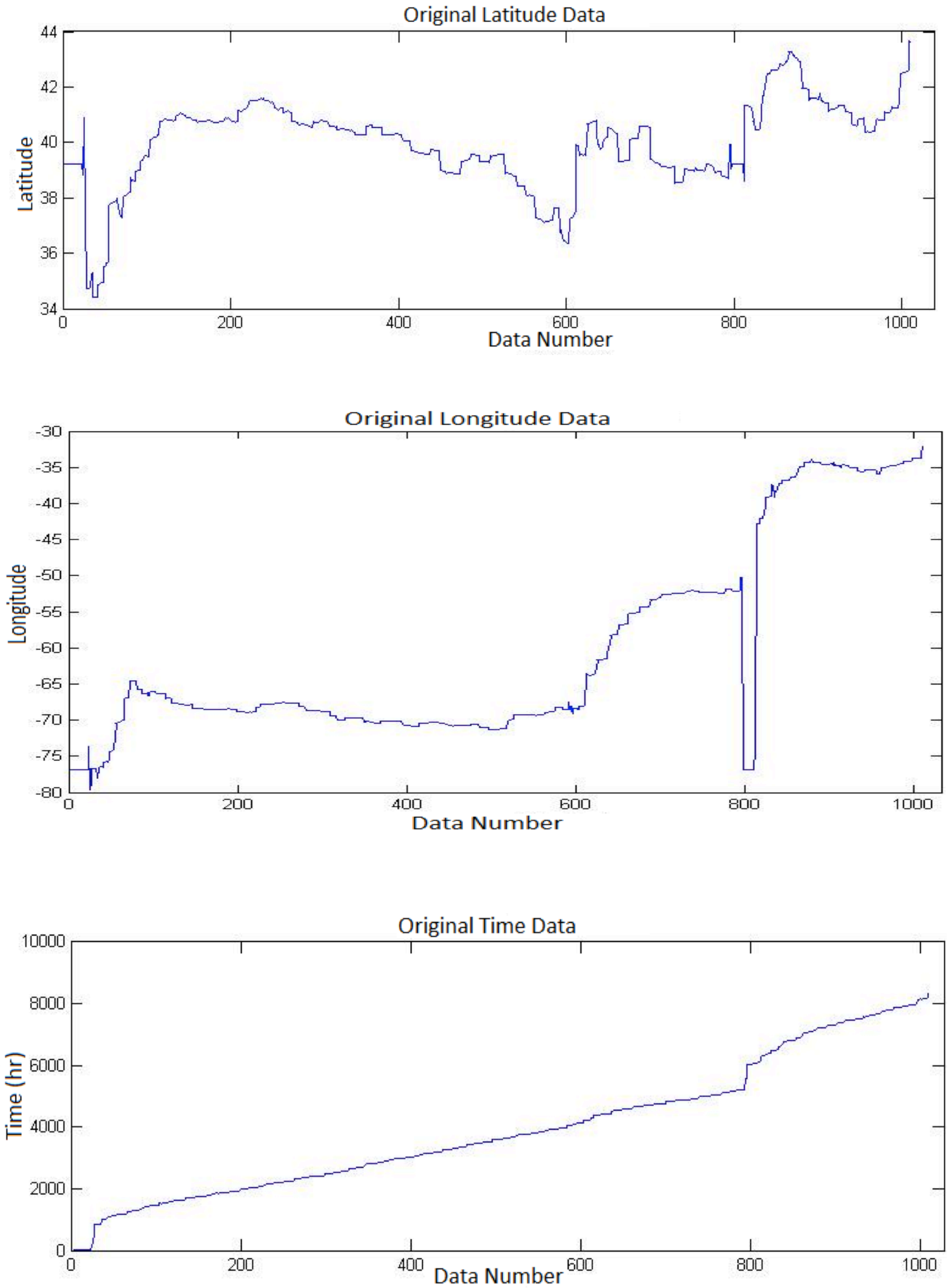

Fig. 1. The original data of sea turtle - for Irving - location, in latitude and longitude, and time in seconds.

Dexter, Irving and Nala. The turtles were tracked at different times. Each location is specified by the combination of latitude and longitude coordinates. An electronic device is attached to a sea turtle. This device communicates with a satellite and feeds back the location of the turtle to a device at a data collecting station [1].

The data of the sea turtle migration is subject to some irregularities. The cause of the irregularities of the data is not the focus here, but it is important to minimize them. The original data for Irving is shown in Fig. 1.

The time was recorded in a date format as "yyyy/mm/dd hh:MM." A Java program is used to convert it into duration of time in seconds from the start of recording. The irregularity of the original data of time is not hard to see in Fig. 1. Time should increase linearly. In Fig. 1, the time is more or less linear except around data point 800. It can be seen the time graph jumps up vertically. Time data taken

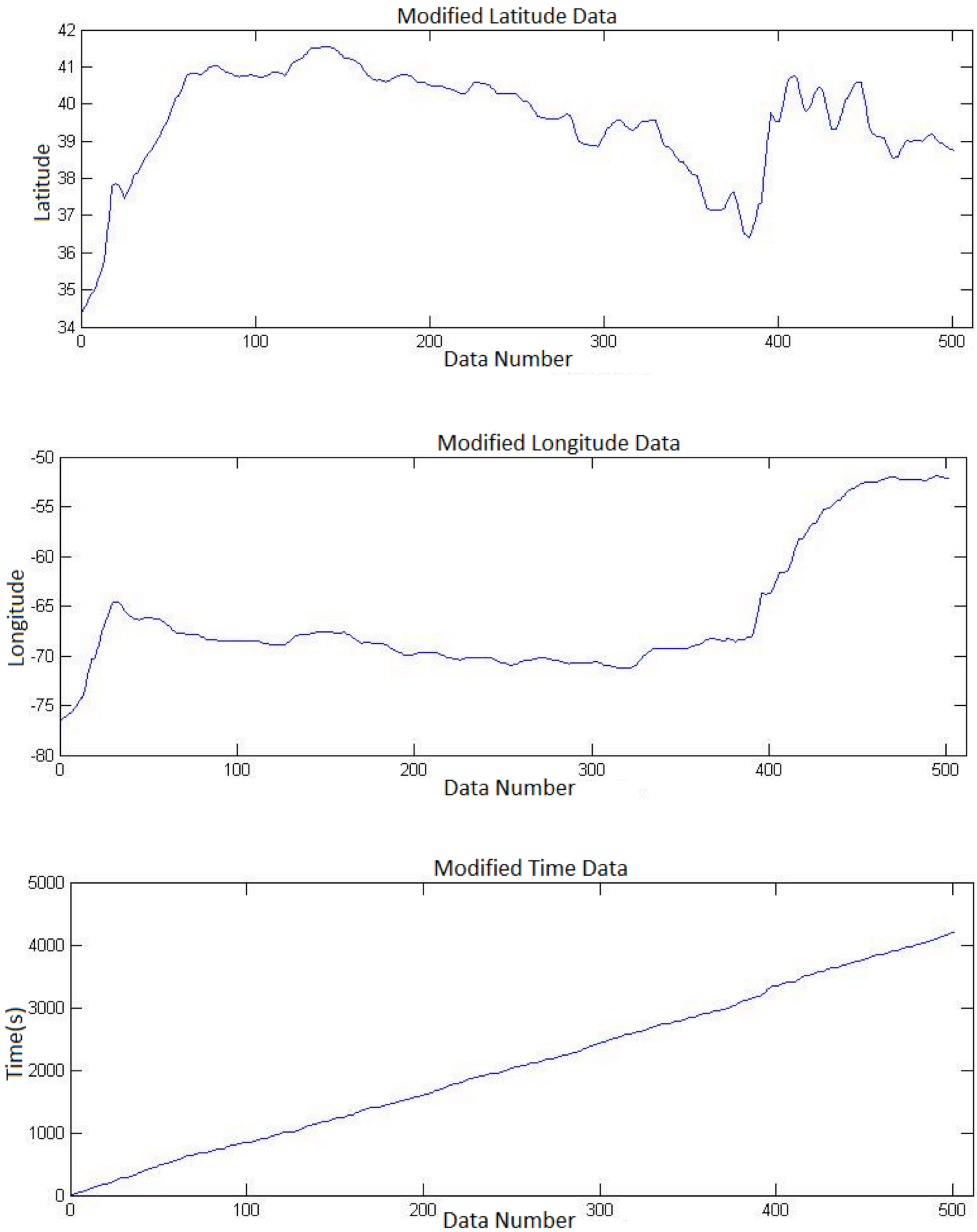

Fig. 2. Modified data of the original data of sea turtle - for Irving - location, in latitude and longitude, and time in seconds.

at a regular pace, which is what almost all time counters do, cannot jump instantly. Therefore it is fair to assume the time data at this point is aberrant.

Similarly, the graphs of location, both longitude and latitude, show some vertical lines. Vertical line on those graphs would mean that the turtle is at two different latitudes, or longitudes, at the same time. That is not possible; hence, the location data points at which the location graphs have vertical slope indicate aberrant data.

2.1 Minimizing Data Aberration

The aberrant data points on the original data cause the visualization of the migration of the sea turtles, in the Atlantic Ocean, using the MLPQ system (on any system for that matter), to be discontinuous and unrealistic. Without changing the general essence of original data, the data irregularities can be minimized. One way to minimize the irregularities is to smooth out the deviant data points. Here, for the purpose of smoothing the data, each data point was replaced by the average of the five neighboring data points. Moreover, the data from the beginning and end of the data recording were trimmed out so as to focus on the steady data. The smoothing and the trimming of the data were done using a Java program. Fig. 2 shows the smoothed and out trimmed out version of the original data from Fig. 1.

```
US - Notepad
File  Edit  Format  View  Help
NorthDakota(id,x,y)  :- id=18, y+4.457952468x>=-384.85639140766477, y+13.06442577x<=-1
NorthDakota(id,x,y)  :- id=19, y+0.001404212638x>=45.798399257773312, y+3.162548765x<=
NorthDakota(id,x,y)  :- id=20, y-23.97561x>=2367.4498780481158, y+6.145786x<=-547.9879

Nebraska(id,x,y)  :- id=1, y-0.07734876x>=50.327023896365475, y+0.1101804124x<=32.0986
Nebraska(id,x,y)  :- id=2, y+1.131446x<=-67.092593731039585, y+2.406721241x>=-189.6450
Nebraska(id,x,y)  :- id=3, y-2.305825x>=265.60182864075534, y+0.09615385x<=33.35615192
Nebraska(id,x,y)  :- id=4, y+0.5683183x>=-12.356307732735637, y+1.044061x<=-58.3653641
Nebraska(id,x,y)  :- id=5, y+1.670854x<=-118.74633065326255, y+20.70833333x>=-1948.407
Nebraska(id,x,y)  :- id=6, y+2.173796x<=-167.26116408477745, y+0.3323169x<=10.18060968
Nebraska(id,x,y)  :- id=7, y+1.830853563x<=-134.23565896634256, y+3.353603604x>=-280.4
Nebraska(id,x,y)  :- id=8, y+1.488619x<=-101.23200182094594, y+2.793494x>=-226.8133861
Nebraska(id,x,y)  :- id=9, y+3.323615x<=-277.5446174928191, y+12.07273x>=-1116.2931236
Nebraska(id,x,y)  :- id=10, y+0.717973005x>=-26.980558205067581, y+1.206921413x<=-74.0
Nebraska(id,x,y)  :- id=11, y+0.9286771x<=-48.512180822727515, y-0.04835762148x<=44.93
Nebraska(id,x,y)  :- id=12, y+0.9629612x<=-51.621275131781204, y-0.6772147x<=109.11268
Nebraska(id,x,y)  :- id=13, y+0.001048061x>=40.89364473723618, y+0.9971552628x<=-60.75
Nebraska(id,x,y)  :- id=14, y+0.5897959x<=-15.051820408166961, y+0.3149918x<=11.918833
Nebraska(id,x,y)  :- id=15, y-0.1096529x<=51.19340637449551, y-0.04835762148x>=44.9382
Nebraska(id,x,y)  :- id=16, y+0.3088128x<=12.809575455206282, y+0.4523809524x>=-1.1131
Nebraska(id,x,y)  :- id=17, y+0.6818811x>=-23.357152697096787, y-5.569536x>=579.397018
Nebraska(id,x,y)  :- id=18, y+0.1101804124x>=32.098644394330428, y+0.278974359x<=15.69
Nebraska(id,x,y)  :- id=19, y+0.6465061656x<=-19.938675866119091, y+0.2268314x<=20.759
Nebraska(id,x,y)  :- id=20, y+767.0769x>=-78239.659646122323, y-0.812569429x>=122.9255
Nebraska(id,x,y)  :- id=21, y-0.812569429x<=122.92554143245927, y-0.6772147x>=109.1126
Nebraska(id,x,y)  :- id=22, y+0.5839300758x<=-14.476979570345271, y+0.9629612x>=-51.62
Nebraska(id,x,y)  :- id=23, y+0.717973005x<=-26.980558205067588, y-0.07734876x<=50.327
Nebraska(id,x,y)  :- id=24, y-0.5354377x>=95.642334604813726, y+1.257943925x<=-80.9133
Nebraska(id,x,y)  :- id=25, y+0.9981534162x<=-60.861555013225789, x>=-104.054300, y+0.
Nebraska(id,x,y)  :- id=26, y-0.5624753424x>=98.401546175956653, y+0.0007204222x<=42.9
Nebraska(id,x,y)  :- id=27, y-0.7801837x<=119.20344488189201, y-0.2657269642x>=68.7877
Nebraska(id,x,y)  :- id=28, y-0.2657269642x<=68.787711041502888, y-0.1974417165x>=62.0
Nebraska(id,x,y)  :- id=29, y+1.429118774x<=-96.105989272013517, y+0.1012433392x<=33.0
Nebraska(id,x,y)  :- id=30, y+3.089256198x>=-255.47337553716935, y-1.413907285x>=178.4
Nebraska(id,x,y)  :- id=31, y+0.6465061656x>=-19.938675866119091, y+0.9272237x<=-47.06
Nebraska(id,x,y)  :- id=32, y+3.089256198x<=-255.47337553716937, y+0.6818811x<=-23.357
Nebraska(id,x,y)  :- id=33, y+0.6339805825x<=-18.981418058256267, y+1.488619x>=-101.23
Nebraska(id,x,y)  :- id=34, y+1.034648x<=-57.494779092008159, y+1.431837426x>=-95.7673
Nebraska(id,x,y)  :- id=35, y+1.431837426x<=-95.767320237089365, y+2.173796x>=-167.261
```

Fig. 3. Snippet of geographic constraint database of the US in MLPQ system for Irving

US_irving - Notepad

File Edit Format View Help

```
irvingLonLat138830a(id,x,y,t)  :- id=1,

 x - 0.01287t >= -76.48180,
 x - 0.01287t <= -76.23,
 y - 0.00795t >= 34.43,
 y - 0.00795t <= 34.68,
 t >= 5.00, t <= 16.20.

irvingLonLat138830a(id,x,y,t)  :- id=1,

 x - 0.01395t >= -76.33760,
 x - 0.01395t <= -76.09,
 y - 0.00880t >= 34.51,
 y - 0.00880t <= 34.76,
 t >= 16.20, t <= 27.20.

irvingLonLat138830a(id,x,y,t)  :- id=1,

 x - 0.01231t >= -76.18420,
 x - 0.01231t <= -75.93,
 y - 0.00941t >= 34.61,
 y - 0.00941t <= 34.86,
 t >= 27.20, t <= 38.00.

irvingLonLat138830a(id,x,y,t)  :- id=1,

 x - 0.01211t >= -76.05120,
 x - 0.01211t <= -75.80,
 y - 0.00948t >= 34.71,
 y - 0.00948t <= 34.96,
 t >= 38.00, t <= 48.80.

irvingLonLat138830a(id,x,y,t)  :- id=1,

 x - 0.01671t >= -75.92040,
 x - 0.01671t <= -75.67,
 y - 0.00907t >= 34.82,
 y - 0.00907t <= 35.07,
 t >= 48.80, t <= 59.80.
```

Fig. 4. Snippet of moving object constraint database in MLPQ system for Irving.

3 A Constraint Database Solution

At this point, the data is raw data with its irregularities minimized. However, in order to be able to visualize the sea turtles movement through the Atlantic Ocean, using the MLPQ system, it is necessary to represent the data in the form of constraint database. In MLPQ system, the constraint equations must be linear.

In order to help visualize the sea turtles' location, in the Atlantic Ocean, in relation to the US land mass, the latitude and longitude data of boundaries of all states from [7] and [8] were converted into geographical constraint database of MLPQ system. A Matlab program, which is a slight modification of an earlier Matlab program of the second author, is used to convert raw data from [7] and [8] into MLPQ geographical constraint database data. A snippet of this geographical constraint database is given in Fig. 3.

As mentioned before, the location of a sea turtle at a time is defined by its longitude and latitude. This means the location of the sea turtle changes as either the longitude or the latitude change, or perhaps they both change at the same time.

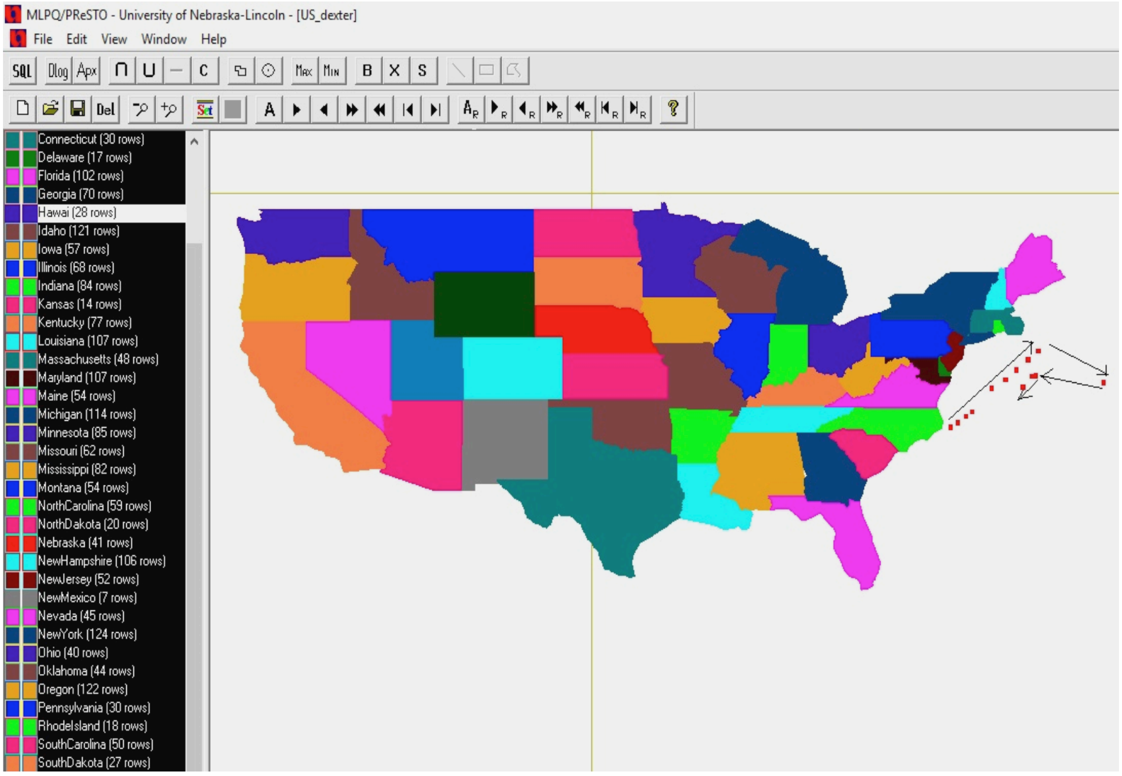

Fig. 5. A series of screen shots of a running MLPQ system with US' geographical constraint database and Dexter's moving object constraint database.

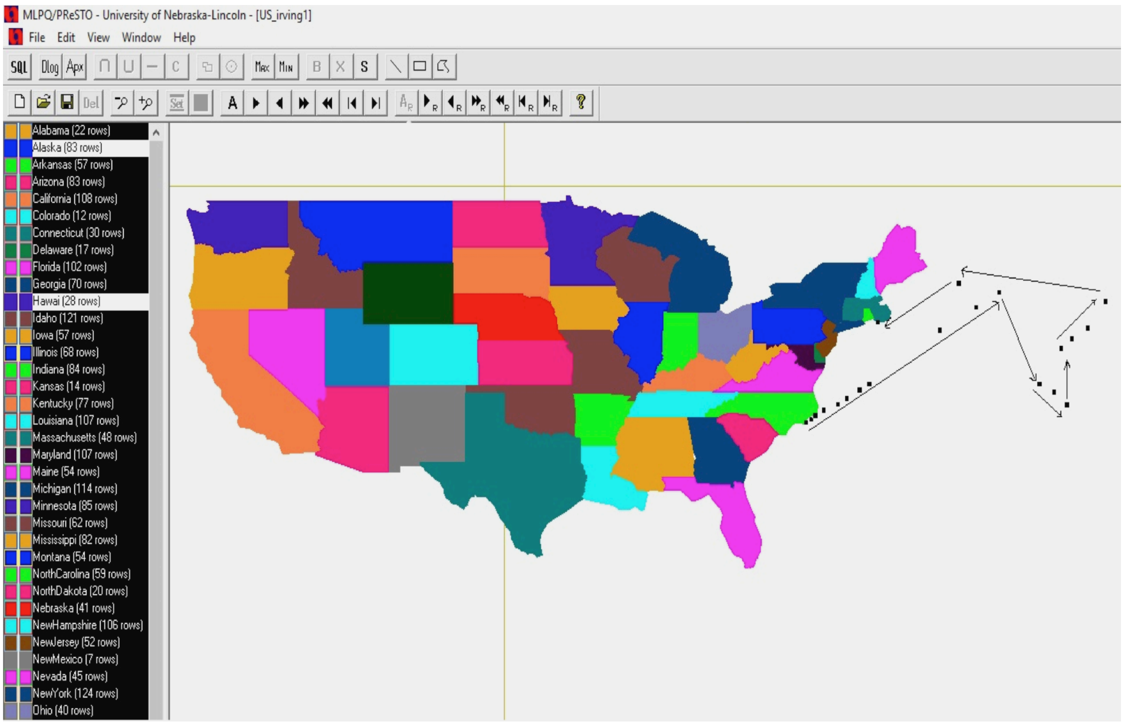

Fig. 6 A series of screen shots of a running MLPQ system with US' geographical constraint database and Irving's moving object constraint database.

Therefore, the constraint database is represented by the linear variation of the latitude in time and the linear variation of longitude in time. Because there is large quantity of data points, Java program is used to convert the large data into a constraint database of MLPQ system for each sea turtle. A snippet of the large MLPQ's constraint database for Irving is given in Fig. 4.

4 Results and Analysis

Once the constraint database for MLPQ system is prepared, running it in MLPQ is what ensues in order to visualize the movement of the sea turtles in the Atlantic Ocean. There is the US map where each state is delineated using latitude and longitude. A sea turtle is represented by square box, which is out of scale.

Fig. 5 is the series screen shots of running Dexter's moving object and US' geographical databases on MLPQ system. The screen shots were taken at about 10 steps, in a fairly constant manner. That means the denser the boxes are the more time Dexter took there. The arrows also indicate the direction of motion of Dexter. Fig. 6 also shows, like Fig. 5, the series of screen shots of MLPQ system running the geographical and a moving object constraint databases for Irving. Each box represents the position of Irving in the Atlantic Ocean in about 15 steps in the MLPQ system. The density of the boxes represents the speed, inversely, and the arrows represent the direction of Irving. Finally, Fig. 7 is like Figures 5 and 6 with the sea turtle being Nala. The shots were taken at about 15 steps of MLPQ system. Since each step is fairly equal, the box density indicates how fast Nala was moving, while the arrows show the direction.

5 Discussion and Conclusions

The data collection of the location of Dexter seems to start when the sea turtle was at the coast of North Carolina. After spending some time in the Atlantic Ocean, it seems to head back to the North Carolina coast. Irving started from the coast of the North

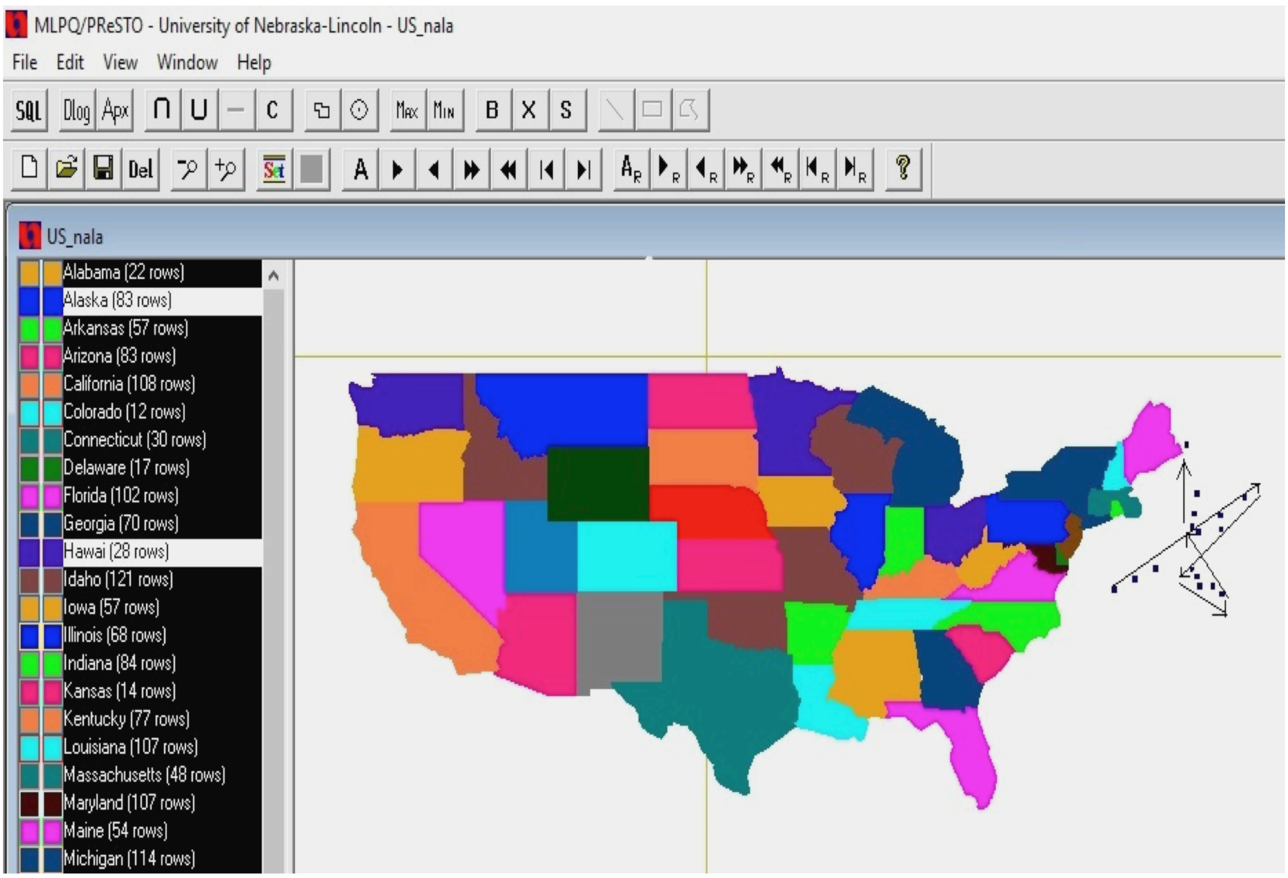

Fig. 7. A series of screen shots of a running MLPQ system with US' geographical constraint database and Nala's moving object constraint database.

Carolina. It swam in a relatively wide area in the Atlantic Ocean. The recording of its location ends when it get near the coast of Rhode Island. Nala also was recorded near North Carolina coast and the recording ended near Maine. Dexter's migration shown by the MLPQ system seems to be consistent with the hypothesis that the sea turtles would return to their original place after migrating deep in the Atlantic Ocean. Irvin and Nala, even though, they seem to return to a different US state, in the MLPQ system, it is important at this point to remember that the data used was trimmed at its beginning and its ending of the recording during the irregularity minimization process. So there is still a possibility that the turtles have returned to the same state. Moreover, in relation to the spread of the distance they travelled in the Atlantic Ocean, the distance between the origin and the destination is not too far. The fact that the MLPQ system, developed and being upgraded at the University of Nebraska-Lincoln, can serve as a tool to visualize the vast migration of sea turtles conveniently, shows its potential as an important tool for researchers.

When running the geographical constraint database and the moving object database of the three sea turtles in the MLPQ system, there exist some random movements of the boxes which seem to be out of sync from what would be normally expected the box to move to. Moreover, at times, there exist double boxes for one turtle at one time. It is fair to assume that these drawbacks originate from the data itself rather than from MLPQ system because first, when a simulated data is converted into a moving object constraint database and is run in the system, then there is no experience of abnormal movement. This is shown with simulation of a projectile motion and an inclined plane.

The second reason is understood by analyzing the given data of the location of the sea turtles. Looking at Fig. 1, it can be seen that there are vertical lines on the graphs of the latitude and longitude. Vertical lines on these graphs mean that the sea turtles exist at different latitude or longitude at the same time. To minimize such data aberration, averaging of neighbor data points was done. Even though, this method helps minimize such data aberration, it can't be expected to solve the problem completely.

References:

[1] L. Avens, and K. J. Lohmann, Navigation and Seasonal Migratory Orientation in Juvenile Sea Turtles. *Journal of Experimental Biology,* Vol. 207, 2004, pp. 1771-1778.

[2] A. Carr, *The Windward Road: Adventures of a Naturalist on Remote Caribbean Shores*. University of Florida Press, 2013.

[3] P. C. Kanellakis, G. M. Kuper, and P. Z. Revesz, Constraint Query Languages, *Journal of Computer and System Sciences*, Vol. 51, No. 1, 1995, pp. 26-52.

[4] E. J. Milner-Gulland, J. M. Fryxell, and A. R. E. Sinclair, *Animal Migration: A Synthesis*, Oxford University Press, 2011.

[5] P. Z. Revesz, *Introduction to Databases: From Biological to Spatio-Temporal*, Springer, 2010.

[6] P. Z. Revesz, R. Chen, P. Kanjamala, Y. Li, Y. Liu, and Y. Wang, The MLPQ/GIS Constraint Database System, *Proceedings of the ACM-SIGMOD International Conference on Management of Data*, ACM Press, 2000, pp. 601.

[7] M. William, *States*, Available at: http://econym.org.uk/gmap/states.xml

[8] D. B. Zwiefelhofer, Find Latitude and Longitude. Available at: http://www.findlatitudeandlongitude.com/