CHAPTER 3

MAP DIGITIZING

3.1 Introduction

This chapter deals with the route map digitizing process. The type of data collected, the plotting of the trajectory, digitizing the map into regular intervals of distance are discussed here. The data collected depends on the type of application. The GPS units can collect a variety of data including latitude, longitude, velocity, features and attributes. However, for our purposes, the time, latitude, longitude data are sufficient.

Digitizing the route map refers to the division of a route into a sequence of ordered segments with equal distance. The ends of each interval are generally associated with a node along the route or map. The algorithm that was developed was coded using Matlab language.

3.2 Description of data collection

In order to illustrate the process of developing the procedure, the Toms Creek A bus route will be used as the test route. The data was collected along this route using the GeoExplorer II receiver in a car. Four trips were made along the route and the data was collected at intervals of one second. The data collected using the GPS units was then processed as explained in Chapter 2. A text file containing the coordinates of the GPS data points is shown in Appendix A. An additional trip was made to collect the data at 5-second interval.

An alternate route was chosen and four trips were made along that route in the car to collect similar GPS data. Data was collected in these trips to compare the digitizing procedure for different routes.

3.2.1 Bus trajectory from GPS data after processing

The fit of the trajectory from the GPS data onto a DRG map is shown in Figure 3.1. The trajectory was found to overlap with the bus route. The picture was obtained by plotting the coordinates of the GPS data points from one of the files superimposed on a map of Blacksburg. Figure 3.2 shows the same plot for a section of the Tom's Creek A road and the GPS data points along the route.

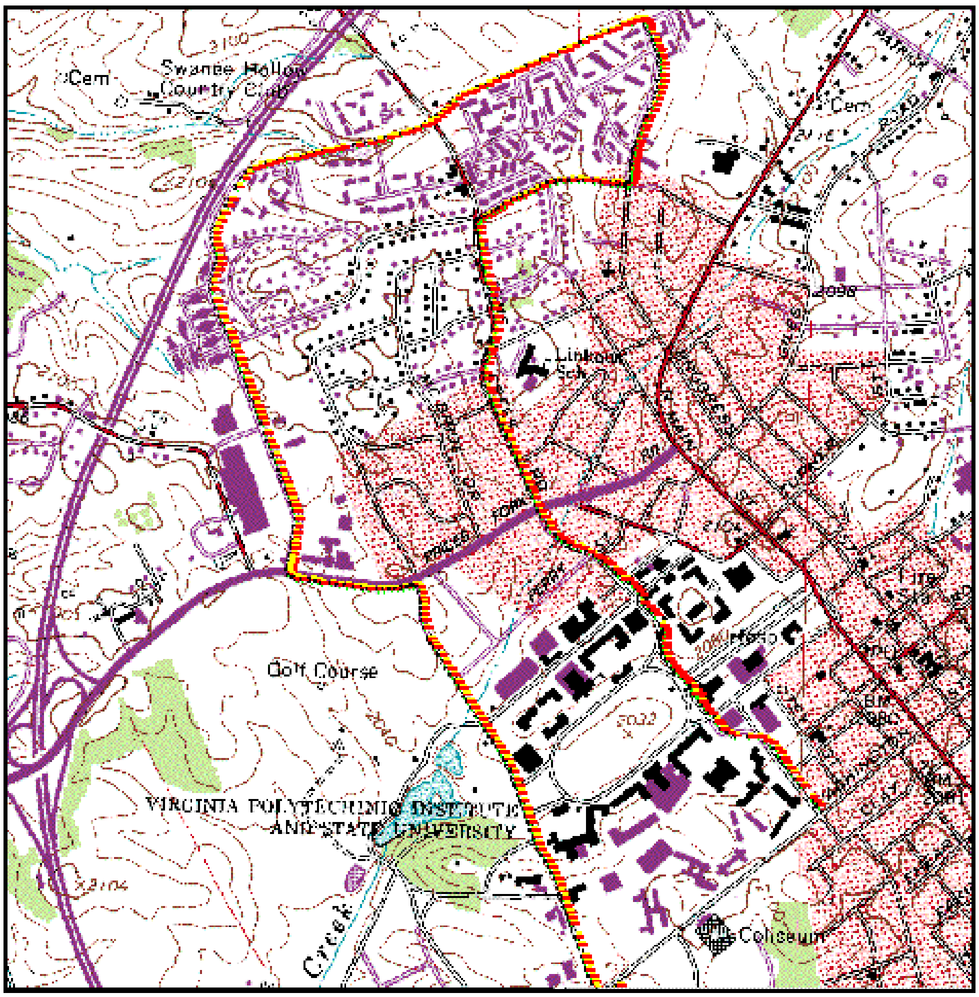

Figure 3.1: Sample plot of Tom's Creek A bus route (marked in orange) superimposed on a AutoCAD map

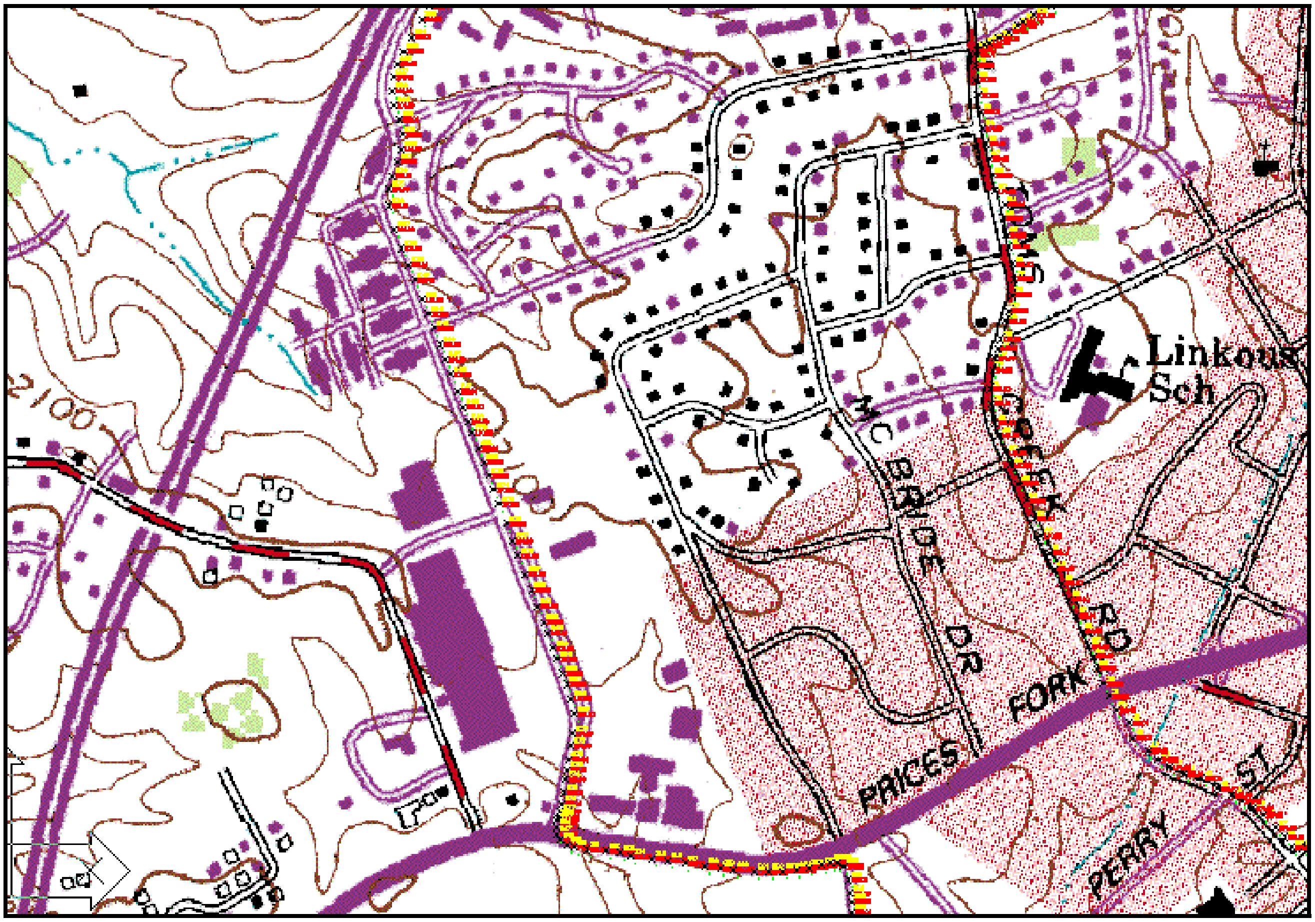

Figure 3.2: A section of the picture showing GPS data points along Tom's Creek Road

The plot shown in Figure 3.3 is a section of the route map obtained using the GPS data points. As shown in the plot, the distances between two consecutive points are not uniform, indicating the speed change of the vehicle along the route. The GPS data points that are located very close to each other indicate a slower speed of the vehicle. The larger intervals denote that the vehicle is traveling at higher speed.

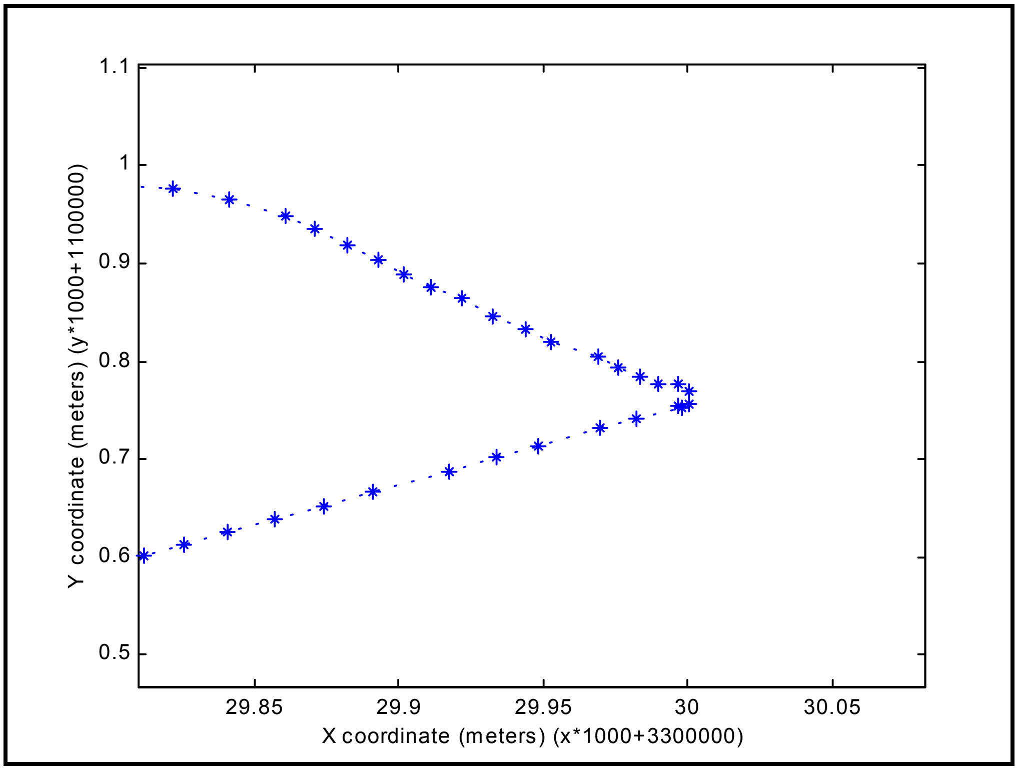

Figure 3.3: Section of plot of GPS data points

.

3.3 Construction of route based on GPS data

For bus arrival time prediction, it is useful to use links that are short and uniform in length. Short links are useful because, the knowledge of the actual location of a bus within a link becomes insignificant to the prediction accuracy. Besides, it is unlikely that two adjacent bus stops would ever reside in the same link. Uniform links are easier to obtain and since the length of the links are known, the travel time can be obtained easily. In some bus arrival time prediction procedures, the arrival times are calculated by summing up the travel times associated with all the links that occur between the final destination and the current location of the vehicle. In such procedures the map plays a very important role. Travel time between two links could be preprocessed and stored in a database. The database will have a set of links and corresponding travel times associated with each link.

3.4 Procedure for constructing route map with link-node representation

The procedure describes the method to obtain the digitized route map from the base map. The first node in the digitized route map is the first GPS data point in the base map. The next node is computed using the following basic conditions.

(1) The coordinate of the node always lies on a segment joining two points in the base map. This ensures that the new coordinate is not far away from any point on the base map.
(2) The coordinate of any node is always at a prefixed distance or interval from the previously obtained node in the digitized map.

The newly obtained is used to obtain the next node and so on. The last node is the last GPS data point in the base file.

A Matlab code was written to convert the raw data into data points at equal distance. This is done through eight steps as detailed below.

(1) The file Gps_Data1.txt was opened and the initializations take place.
(2) The first point in the raw data file is taken as the initial point in the new digitized route map.
(3) The distance from an initial point to the subsequent points is calculated until the interval where the new point lies is obtained. (See next section)
(4) The coordinate of the new point is obtained by interpolation between the coordinates of the end points of the interval in which it lies.
(5) The newly obtained point is used to obtain the next point by repeating steps (3), (4).
(6) The process is stopped when the distance between the latest point and the last point in the file is less than the desired interval.
(7) The last point in the new file is the last point in the Gps_Data1.txt file.
(8) The file containing the new points is used to plot the new map digitized at desired intervals of distance.

The steps (3) and (4) are explained in detail in the sections below.

3.4.1 Step (3)-Obtaining the interval for new point

This obtains the interval in which the next point in the digitized map would lie. Suppose the desired interval for the digitized map is **x** meters.

The distance from the initial point, **i** to each of the subsequent points is calculated. Let d(i,j) be the distance between i and j, j=i, i+1,………, p, q… Let n be the new point. Then n must fall on the link (p,q) if d(i,p) <=x <= d(i,q).

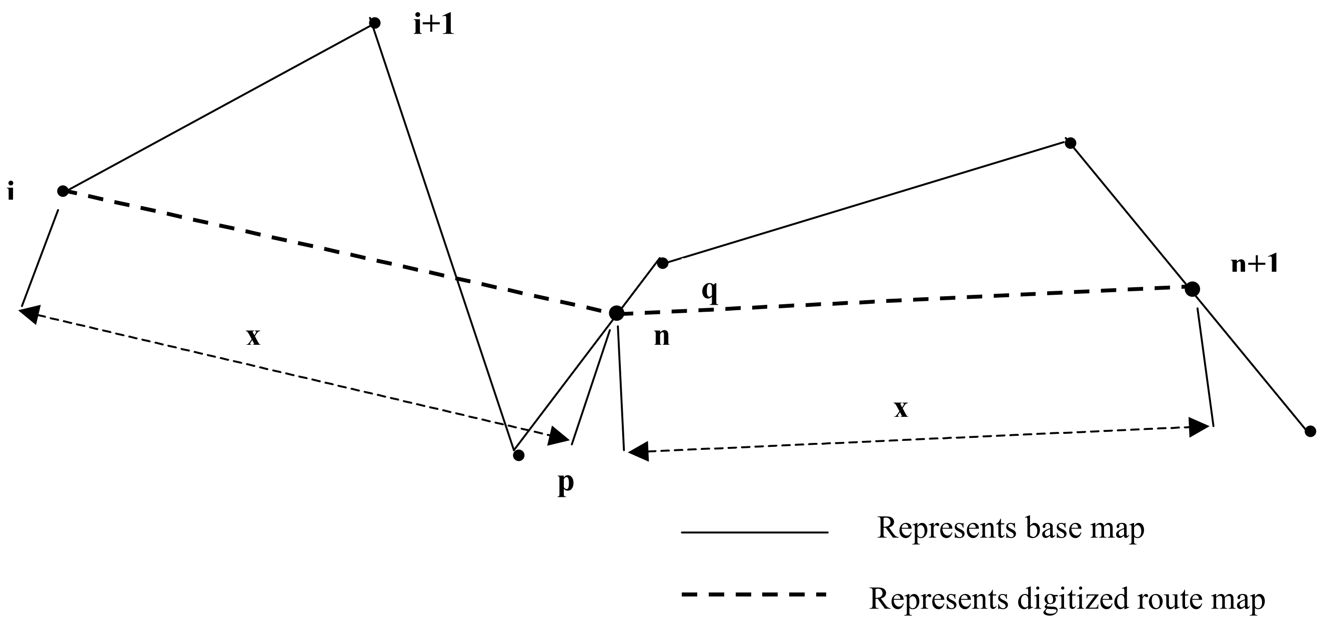

Figure 3.4: Obtaining the new point n

Once a new point is determined, the procedure is repeated after resetting the new point to **i**. The procedure is terminated when ultimately there is no point at a distance **x** from the latest **i**. The last point in the base map is also taken as the last node in the digitized route map even though the distance is less than **x**.

3.4.2 Step (4)- Coordinate of new point

Once the link where the new point is located is identified, the coordinate of the new point needs to be calculated. The following procedure obtains the coordinates of the point. The coordinate of the point, **n** shown in Figure 3.5 can be calculated in terms of **p** and **q**.

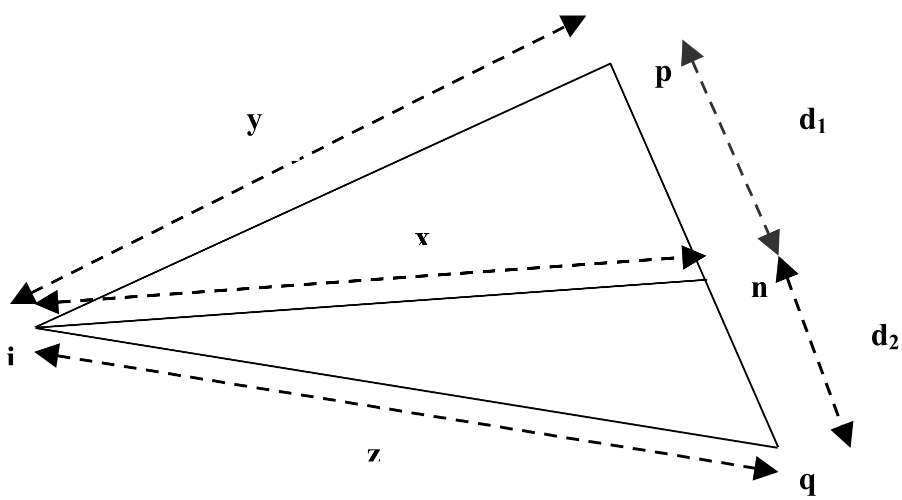

Figure 3.5: Obtaining the coordinate of new point n

Let (X_1, Y_1) be coordinates of **i**, (X_2, Y_2) be the coordinates of **p** and (X_3, Y_3) be the coordinates of **q**. The coordinates of **n**, (X_4, Y_4) have to estimated. The (X_1, Y_1), (X_2, Y_2), $(X_3, Y3)$ are known from the GPS data file. The distances **x**, **y**, **z** can be estimated since from the coordinates of **i**, **p**, **q**.

Calculation:

angle1 =$\angle$**ijk** = $\cos^{-1}$((**y**^2 + **d**^2 –**z**^2)/(2***y*****d**))

angle2 =$\angle$**ikj** = $\cos^{-1}$((**z**^2 + **d**^2 –**y**^2)/(2***z*****d**))

$\mathbf{d_1}$ = ((2***y***cos(**angle1**))+(((2***y***cos(**angle1**))^2-4*(**y**^2-**x**^2))^0.5)/2

$\mathbf{d_2}$ **= ik-d1**

Case (1)

If ($\mathbf{d_1}$>$\mathbf{d_2}$) then **r** = $\mathbf{d_1}$/$\mathbf{d_2}$

Then the Coordinates of new point **d** are:

$X_4 = (r^* X_3+1^* X_2)/(1+r)$

$Y_4 = (r^* Y_3+1^* Y_2)/(1+r)$

Case (2)

If ($\mathbf{d_2}$>$\mathbf{d_1}$) then **r** = $\mathbf{d_2}$/$\mathbf{d_1}$

Then the Coordinates of new point **d** are:

$X_4 = (r^* X_2+1^* X_3)/(1+r)$

$Y_4 = (r^* Y_2+1^* Y_3)/(1+r)$

A program was written in Matlab to carry out the above procedure to yield the final digitized map of any desired interval. The algorithm for the program, which is based on the above procedure, is as follows.

Step 1: Initialization:

Label the time sequence data as {**p[0]**, **p[1]**, …,**p[T]**}

Where **p** refers to the point and the number in the brackets refers to the time when it was recorded.

Set **N** = {}, and t = 1, Desired Interval = **x**

Set **i** = p[t-1]; **p** = p[t]; **q** = p[t+1]

Step 2: Compute **d(i, p)** and **d(i,q)** (Euclidian Distance between two points)

Step 3: if D(**i**, **j**) > **x** and D(**i**, **k**) < **x** then

n= GetNew(**i, p, q**)

N = {N,n }

i = **n**; **p** = **q**;

Go to Step 2

Else

increment t by 1

if t > T then

go to step 4

else

p= **q**; **q** = **p**[t+2]

Go to Step 2

Step 4: N = {N, p[T]}

Stop

GetNew is a function that returns a new point as described in section 3.4.1 and 3.4.2 and which satisfies the following two conditions:

1) **n** must be on the line segment [**p**,**q**]

2) **d(i, n**) = **x**

3.4.3 Digitized route map

The link-node representation of Tom's Creek A bus route obtained using the GPS data and described in the preceding sections is shown in Figure 3.6. In this route map, the link length is chosen to be 20 meters. When compared with the base map shown in Figure 3.3, the digitized map has points at larger but intervals. Figure 3.7 shows a section of the digitized route map and the base map. It can be seen that all the nodes in the digitized route map lie on the base map. The X axis and Y axis are not set to the same scale in the following figures.

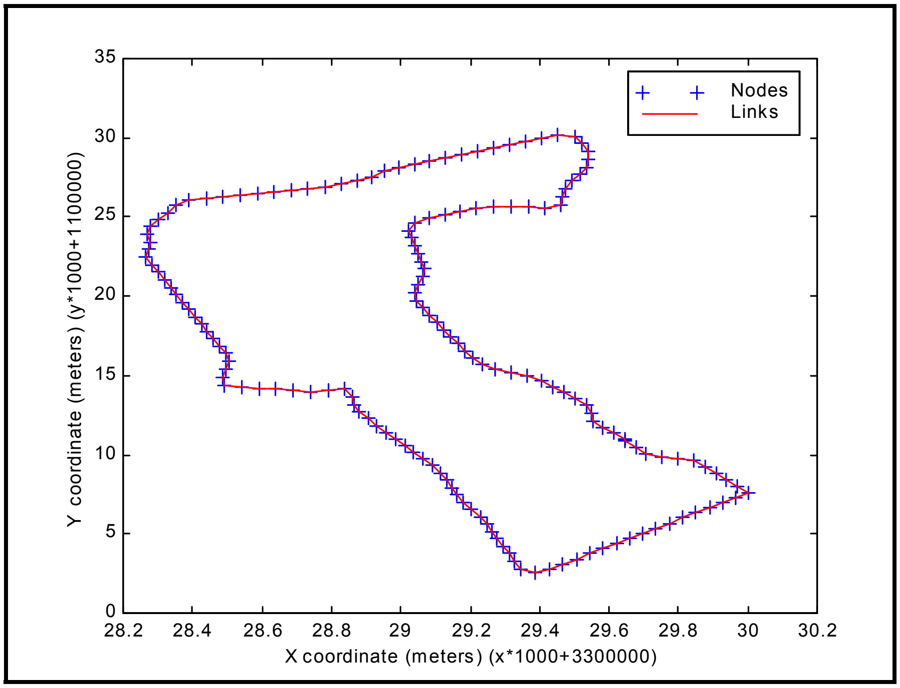

Figure 3.6: Digitized route map of Tom's Creek A route.

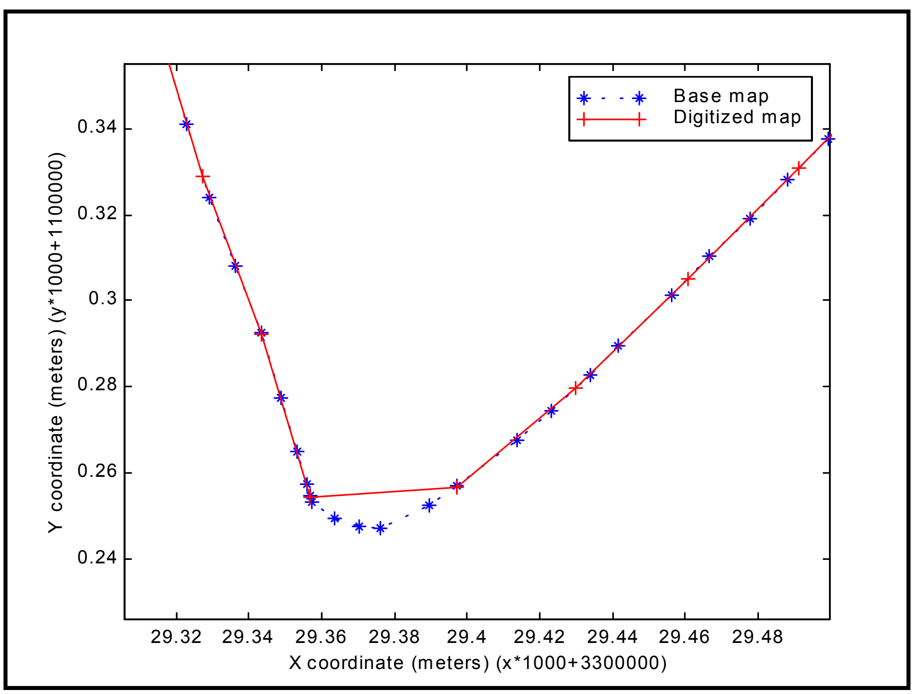

Figure 3.7: Comparisons of plots

3.4.4 Link and node numbering

Each link in the digitized map was assigned a number for easier reference. The links of the route are ordered such that one can trace the vehicle trajectory by the sequence of links. The current location of the bus and the destination of the bus can be associated with link numbers. The same section of the road can be represented by two different links if a bus is traveling along that section more than once during a trip. In addition, each link number can be provided with a unique travel time based on data collected over a period of time. This segmentation of the whole route eases the procedure of prediction and also reduces computation time.

The output text file was captured and is shown in Figure 3.8. The first two columns indicate the X coordinate and Y coordinate of the point and the last column indicates the node number. The label of a link is the same as the number assigned to the node at the start of the link.

sample - Notepad

File Edit Search Help

3329654.20	1101078.90	1.00
3329647.50	1101086.40	2.00
3329641.20	1101094.10	3.00
3329634.90	1101101.90	4.00
3329628.80	1101109.80	5.00
3329622.70	1101117.70	6.00
3329616.70	1101125.70	7.00
3329610.80	1101133.80	8.00
3329604.90	1101141.90	9.00
3329599.00	1101150.00	10.00
3329593.10	1101158.10	11.00
3329587.30	1101166.20	12.00
3329581.30	1101174.20	13.00
3329577.70	1101183.50	14.00
3329573.60	1101192.70	15.00
3329568.20	1101201.10	16.00
3329562.80	1101209.50	17.00
3329556.50	1101217.30	18.00
3329548.00	1101222.60	19.00
3329544.10	1101231.80	20.00
3329545.00	1101241.70	21.00
3329545.90	1101251.70	22.00
3329546.70	1101261.60	23.00
3329547.60	1101271.60	24.00
3329547.70	1101281.60	25.00
3329546.30	1101291.50	26.00
3329543.10	1101301.00	27.00
3329537.70	1101309.40	28.00
3329532.60	1101318.00	29.00
3329527.10	1101326.30	30.00
3329518.50	1101331.50	31.00

Figure 3.8: Output text file showing node coordinates and node numbers.

3.5 Location of Stops

The location of stops on the Tom's Creek A route was obtained logging GPS location data only at stops. The position of the stops can be represented on the Y-axis with their distance from the origin. The stops are plotted on the route map and shown in Figure 3.9.

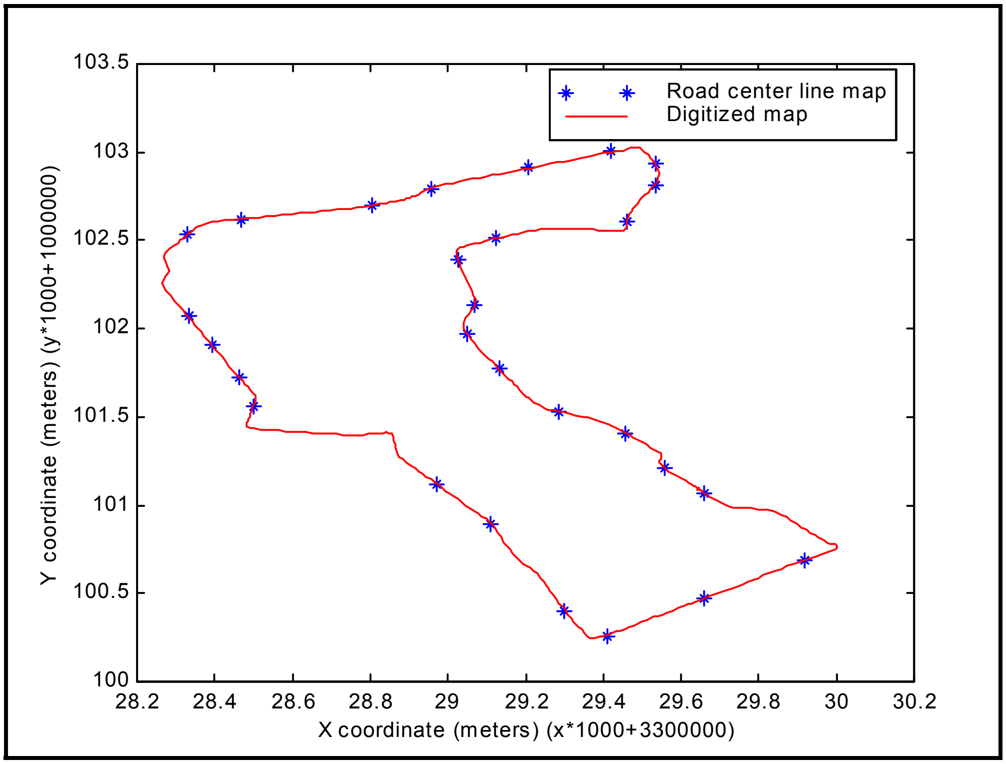

Figure 3.9: Location of Stops on Tom's Creek A bus route

3.5.1 Time-distance diagram

The two-dimensional plots of the bus route trajectories are converted into a time-distance diagram with time on the X-axis and distance on the Y-axis. In a time-distance diagram, the vehicle trajectory has a single dimension, distance, which is the distance of the vehicle from the origin. The distances of the stops from the origin are projected on the Y-axis as shown in Figure 3.10.

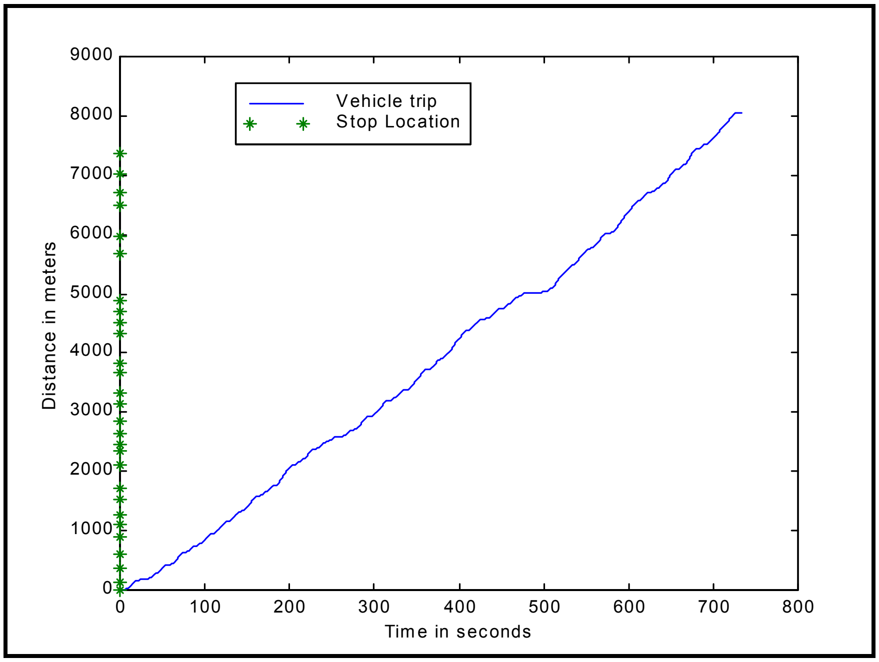

Figure 3.10: Time-distance plot of vehicle trip and stops