

NEW APPROACHES TO DISTANCE BASED CHARGING

Jakub Rajnoch
Rapp Trans UK Ltd
6-8 Market Place, Reading, RG1 2EG, UK
+44 (0) 118 925 5448, Jakub.Rajnoch@rapp.uk.com

Distance based charging (DBC) is an electronic fee collection system, where the total fee is calculated on the basis of the distance travelled. The specification for DBC does not limit the technology that could be used to implement such a system. With respect to the existing DBC schemes, the technologies commonly used can be of any combination of the GNSS, microwave or odometer/tachograph solutions. The process of determining the driven distance also varies from solution to solution. While such a process can be very straightforward and reliable for the microwave and odometer/tachograph solutions, the GNSS solution can sometimes provide inaccurate data. These inaccuracies usually relate to the location data of the vehicle or road network representation; following the matching process, these inaccuracies may result in the incorrect journey and distance identification. The aim of this paper is to suggest a new approach to DBC, which could reduce or remove the errors in the journey and distance calculations that are typically associated with GNSS solutions.

Keywords: Distance based charging, GNSS, matching process, distance calculation

INTRODUCTION

Distance based charging (DBC) is a type of electronic fee collection system (EFC) where the fee is based on the distance travelled. Individual technologies used in such a system may lead to individual processes to determine the distance. While such a process may be very simple and reliable for the system based on microwave (distance calculated as the sum of the lengths of segments defined or bounded by the gantries) or odometer technologies (distance read from the odometer or tachograph), it may require complex data analysis and processing if the system is based on the Global Navigation Satellite System (GNSS).

APPROACHES TO DBC BASED ON GNSS

The approach taken in the majority of existing GNSS-based charging systems is to match the location data of the vehicle against a geometric representation of the road network in order to determine the journey the vehicle has travelled. The vehicle location data is usually obtained during the journey, via a GPS unit mounted in the vehicle. This data is used in the matching process to identify the elements of the road network that the user travels on. The geometric representation of the road network consists of a set of charge objects, defined through the geometric apparatus, describing the real elements of the road network that are subject to a charge. Therefore the result of the matching process shall be a sequence of the charge objects (i.e. roads or sections of roads) as travelled through during the time of the vehicle's journey. The resulting distance is calculated as a sum of lengths associated with each individual charge object identified during the matching process.

The following criteria may be used in order to identify the approaches to the GNSS-based DBC systems:

- The charge object definition
- Matching process
- Distance calculation

The Charge Object definition

The charge (or the fee) is imposed on the basis of user's usage of the elements of the road network (the charge objects). By the definition (7), the charge object is a virtual shape on the surface of the road network given by the co-ordinates of a set of points and other attributes (7). In principle, the charge object or set of charge objects could represent a single road, section of the road, selected or all roads in a given area. The various types of charge objects have been identified as follows:

- The zone is a polygon-shaped area defined by the set of points and the straight lines between the zone points in the sequence, including the line between the last and the first zone points. The zone usually represents an area covering roads subject to charge.
- The corridor consists of at least one segment and a sequence of at least two corridor points. The segment is a diagonal polygon-shaped area of such a proportion to correlate the subject to charge. The corridor shall have at least one entry and one exit point. The corridor usually represents a section of the road subject to charge (usually spans the road between two junctions).
- The virtual gantry is a set of points; first two of which mark the posts of the virtual gantry. The gantry consists of two adjacent quadrilaterals, being adjacent in the line representing the virtual gantry. The virtual gantry usually represents a gantry on the road, where passage through incurs a charge.

Matching process

The nature of individual approaches to the matching process, designed according the environment where shall the charging scheme operate, enables the distance based charging to be used in either urban or inter-urban environment.

The inter-urban environment does not constrain the technology employed or the matching process implemented, as long as all the areas where driving incurs a charge are captured in the map or other geometric representation of the road network, used in the matching process. On the contrary, the urban environment, with its limiting factors represents a much larger challenge. The built-up characteristics of the urban environment may cause several inaccuracies to occur in the determination of the vehicle location. As a consequence of this there may be further implications in determining the journey and distance driven during the matching process.

The approaches taken differ depending on the charge object definition; the approaches include the following:

- Points/Curves Matching Approach which uses the charge object defined as a set of curves and points (known as a digital map). This approach may be divided into the following sub-approaches:

- Point-to-point method uses the matching algorithm to identify the closest node or shape point of the road network to the location fix. This method is very easy to implement, however very sensitive to how accurate the road network has been digitised.
- Point-to-arc method uses the matching algorithm to identify location fix and arcs in order to find the closest. This method gives better results than the point-to-point method, however may give incorrect results in some cases, e.g. very dense road network.
- Arc-to-arc method takes into account a set of location fixes grouped in the shape of the arc to be matched with arc of the road network. Such an approach is very sensitive to outliers and relies on point-to-point matching method, which therefore may be slow in performance.
- Polygons Matching Approach which uses the charge object defined as a set of polygons stored as a sequence of its co-ordinates. The polygon usually complies with the zone or the corridor definition in (7), i.e. closed shape dividing the area into inner and outer. This approach may be divided into the following sub-approaches:
 - Segment-based method uses the matching algorithm to identify the polygon which has been intersected by at least one line connecting the location fixes in the sequence according to the time received. The polygon definition is equivalent to a corridor in (7), although it is usually rather smaller size.
 - Corridor-based method uses the same matching algorithm as the segment-based method. The corridor complies with the definition in (7).
 - Concentric-circles method uses the matching algorithm to identify the circles which have been intersected by at least one line connecting the location fixes in the sequence during the time period involved. The concentric circles could be viewed as the equivalent of several zones overlaid.
 - Grid-shape method uses the matching algorithm to identify the grids (square shaped areas overlaying the road network) which have been intersected by at least one line connecting the location fixes in the sequence according to the time received. The grids may be understood as equivalent to several close adjacent zones.

The purpose of the matching process is to determine whether the vehicle has entered, used or exited a segment of the road network that is subject to a charge, given the location data of the vehicle during the journey.

Distance calculation

There is not a big variety of approaches to calculating the distance calculation. The definition of the charge objects covers the assignment of the distance to each of them. Therefore once the journey is identified, i.e. the sequence of the charge objects, the distance is calculated as a sum of lengths of charge objects visited during the journey.

The nature of the location data obtained via GNSS (N.B. this can only be achieved using a GPS device at the current time) could be quite unstable especially with regard to the built environment. If the data received is inaccurate, errors may be caused

during the matching process. Therefore, this can result in the incorrect identification of charge objects. A consequence of this situation might be that the incorrect charge is levied with the user.

NEW APPROACH

The majority of the recent GNSS-based DBC systems use the matching techniques. Regardless of the complexity of the charge objects definition, it is always the distance driven within individual charge objects that defines the charge. The knowledge of the journey may play a significant role in the interim processing, but in the end it is the distance that decides the total charge. The aim of this approach is to remove potential inaccuracies arising from the determination of the distance, caused by initial inaccuracies of the location data and the matching process. The way of doing so is fairly simple and straightforward; there is no matching process involved.

Although there is no intention to encompass the matching process in this approach, the outcome of the approach will include the algorithm for the distance driven as well as the detection of the road types defined according the speed limits. By removing the matching process from the determination of the distance, the following advantages may apply:

- No need for the geometric representation of the road network will remove the inaccuracies caused by the errors during the creation process as well as the necessity of upgrading both existing and non-existing roads (which may concern large amount of data to be exchanged or updated).
- The distance driven will cover the areas that are not registered as roads (i.e. off-road driving) or not captured on the map (or other sources of the information about the road network).
- Reduction in the communication costs of upgrading the charge objects database
- Possibility to deploy the algorithm of the distance calculation regardless of whether a thin or thick architecture is adopted.

Despite all the advantages listed above, the approach without the knowledge of the road network may run into the following difficulties (although these issues may be addressed during the future developed of the approach):

- Weak or no GPS signal coverage leads to the incomplete location data received
- There may be other issues identified during subsequent development

PRINCIPLE

The main principle of this solution is to use only the data provided by the generic GPS device to calculate the distance and detect the road type. Due to the fact that this is the only source of the data related to the location and the speed of the vehicle, the maximum available information items are logged and used in this approach. The data from the GPS device are captured in the form of text-based NMEA (ref. to

NMEA standards) sentences, and are of the following types (4):

- Global Positioning System Fixed Data (GGA) contains time, position and fix type data.
- GNSS DOP and Active Satellites (GSA) contains GPS receiver operating mode, satellites used in the position solution and Dilution of Precision (DOP) data.
- GNSS Satellites in View (GSV) contains the number of GPS satellites in view, satellite ID numbers, elevation, azimuth and Signal-to-Noise Ratio (SNR) values. One GSV sentence can provide data for only four satellites; therefore 3 sentences may be needed for the full information. It may contain more satellites than stated in GSV, as it may list those not used for the position determination.
- Recommended Minimum Specific GNSS Data (RMC) contains time, date, and position and speed data.
- Garmin Estimated Error (RME) contains estimated vertical, horizontal and overall spherical errors. For this sentence is provided by the vendor (i.e. it is not standard NMEA sentence), it is subject of further development, whether any data provided will be used.

Other data may be used during the time of the further development. In order to minimize the impact of the inaccuracy in the data received, the latter has to go through the filtering process. The filtering process will perform an analysis on the series of the data and mark those assessed as inaccurate (this very likely concerns the latitude/longitude/altitude data). It is an area for further development whether those marked as inaccurate will be removed or amended. The connection to the GPS device is done via GPSd software (reference to the website), which provides both text and binary output of the sentences listed above.

ASSUMPTIONS

The following general assumptions have been defined as a basis for the data collection, analysis of the series and distance calculation:

- The only source of data related to location, speed, or any other geographic matter will be obtained via generic GPS device (GlobalSat BU-353). No connection to the odometer or any other vehicle's data sources will be used.
- The data will be collected and stored only during the time of the journey. The analysis and distance calculation process will take place once the journey is finished (i.e. simulation of the thin client architecture). It may be a subject of further development whether to perform optimization of analysis and calculation process in order to analyze the data and calculate interim distance on the fly (i.e. simulation of the thick client architecture).
- It is assumed that the following data (see Table 1,2,3 and 4) will be received from the GPS device on defined intervals (some of the data items may be of 'NULL' value due to the weak or loss of the signal):

NMEA sentence type	Item	Format	Description (4)
GGA (1 sentence per 1 second)	UTC time	hhmmss.sss	UTC time the fix has been taken
	Lat S/W indicator	ddmm.mmmm 'S', 'W'	Latitude and South/West indicator
	Lon N/E indicator	dddmm.mmmm 'N', 'E'	Longitude and North/East indicator
	Position Fix Indicator	0 - 6	Indicator of whether the fix is valid (1 SPS mode, 2 DGPS mode) or invalid (0)
	Satellites used	04	Satellites used to determine the latitude and longitude
	HDOP	1.0	Horizontal Dilution of Precision (the smaller the more accurate the fix should be)
	MSL Altitude Units	12 m	Altitude Units
	Height of Geoid Units	46.9 m	Mean sea level above the reference ellipsoid (WGS84)

Table 1 – GGA sentence items

NMEA sentence type	Item	Format	Description (4)
GSA (1 sentence per 1 second)	Mode 1	'M', 'A'	Manual (forced) mode or 2D automatic mode
	Mode 2	1, 2, 3	Indicator of how many satellites used (2D for less than 3 satellites, 3D for more than 3 satellites)
	Satellite Used	02	ID of the satellite used (lists all visible satellites)
	PDOP	1.8	Position Dilution of Precision
	HDOP	1.2	Horizontal Dilution of Precision
	VDOP	1.5	Vertical Dilution of Precision

Table 2 – GSA sentence items

NMEA sentence type	Item	Format	Description (4)
RMC (1 sentence per 1 second)	Status	'A', 'V'	Status of the fix (active/void)
	Speed over ground	25.6	Speed in knots units
	Course over ground	309.62	Track angle in degrees
	Date	ddmmyy	Date the data has been received

Table 3 – RMC sentence items

NMEA sentence type	Item	Format	Description (4)
GSV (1 sentence per 5 seconds)	Number of Messages	2	Number of messages the sentence is split into
	Message ID	1	ID of the message (i.e. part of the sentence)
	Satellites in View	05	Number of satellites used to determine the position
	Satellite ID	1 – 32	ID of the satellite
	Elevation	78 (max 90 degrees)	Elevation
	Azimuth	45 (max 359 degrees)	satellite
	SNR (C/No)	46 (min. 0, max. 99)	Often referred to as a signal strength, however the interpretation of the number varies from manufacturer to manufacturer

Table 4 – GSV sentence items

- Three different algorithms for the distance calculation will be used:
 - Great circle distance between two points on the perfect sphere (1)
 - Haversine formulae algorithm (1)
 - Vincenty algorithm (1)

The distance calculation algorithms specification or definition is out of the scope of this work.

- The minimum distance resolution with regard to the distance calculation algorithms is as follows (due to better understanding, the latitude and longitude figures are in the format as specified in Table 1):

Point 1 Latitude	Point1 Longitude	Point2 Latitude	Point2 Longitude	Perfect sphere (m)	Haversine (m)	Vincenty (m)
5126.7313N	00007.6733W	5126.7314N	00007.6733W	0.2527	0.2186	0.2186
5126.7314N	00007.6733W	5126.7314N	00007.6734W	0.1715	0.1156	0.1158
5126.7314N	00007.6734W	5126.7315N	00007.6735W	0.1855	0.1855	0.1854

Table 5 – Minimum distance resolution

SCENARIOS

The data has been collected via a GPS device attached to the roof of the vehicle, during a continuous journey between two locations within the UK. The route leads through both urban (UB part of the journey, see Table 6 and Figure 1) and inter-urban environments (IB part of the journey, see Table 7 and Figure 2) and has been driven in the morning between 7:00 am and 9:00 am.

The following table presents the information related to the urban part of the journey with the start point at the Tierney Road within London Zone 2 and the end point at the section of the M4 motorway in the outer area of London.

Point	Location	Latitude (deg)	Longitude (deg)	Altitude
START - A1 UB-part start	Tierney road SW2 4QH London	Google (7)	Google	N/A
		51.444659	-0.324298	
		OS Map (2)	OS Map	
END - B1 UB-part end	M4 section in the outer area of London	Google	Google	N/A
		51.492437	-0.324538	
		OS Map	OS Map	
		51.445094	-0.128606	
		51.492481	-0.349947	

Table 6 – The start/end co-ordinates of the urban part of the journey

The following picture displays the urban part of the journey.

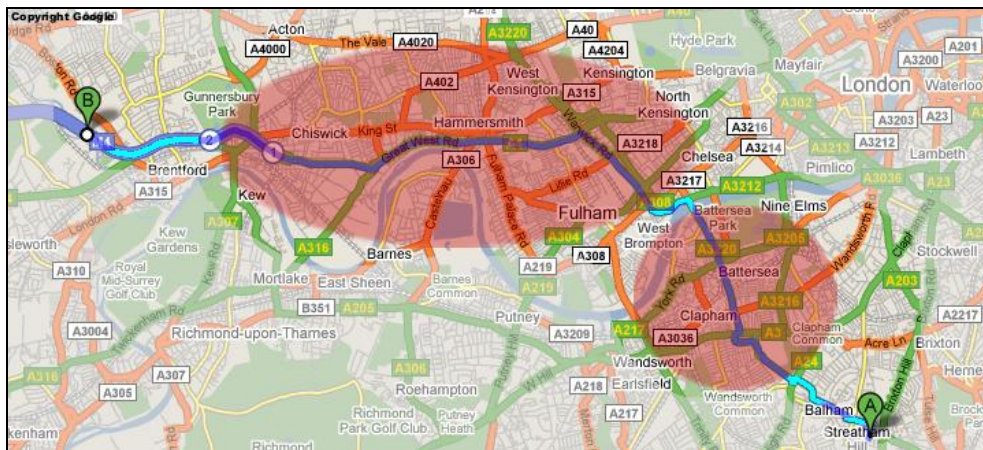


Figure 1 – Urban part of the journey (with congestion areas marked)

The approximate distance of the urban part of the journey is 11.1/17.76 miles/km (7). The following table presents the information related to the inter-urban part of the journey.

Point	Location	Latitude (deg)	Longitude (deg)	Altitude
START - A2 IB part start	M4 section in the outer area of London	Google (7)	Google	N/A
		51.492437	-0.324538	
		OS Map (2)	OS Map	
END - B2 IB part end	Madejski stadium, Outer area of Reading	Google	Google	N/A
		51.423386	-0.984452	
		OS Map	OS Map	
		51.492481	-0.349947	
		51.424153	-0.983831	

Table 7 – The start/end co-ordinates of the inter-urban part of the journey

The following picture displays the inter-urban part of the journey, with start point at the section of motorway M4 in the outer area of London and ends at the Park & Ride at the Madejski Stadium in the outer area of Reading.

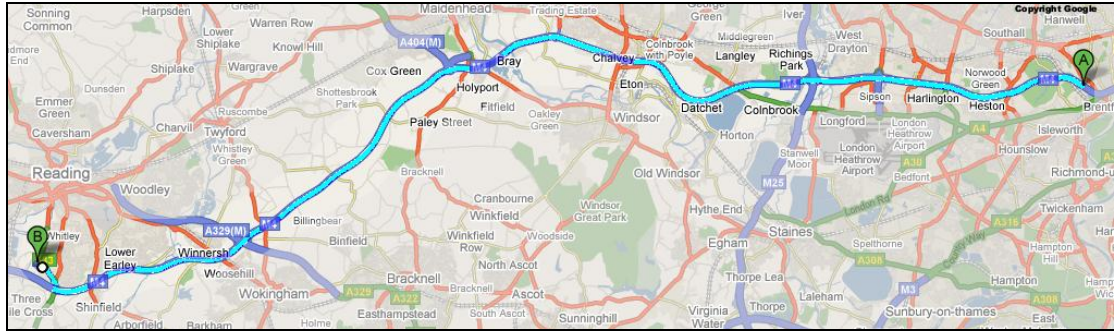


Figure 2 – The Inter-urban part of the journey

The approximate distance of the inter-urban part of the journey is 32.4/51.84 miles/km (7).

The urban area (see Figure 1) experiences rather congested traffic and the second half of the urban part of the journey leads through high built-up areas. By contrast, traffic in the inter-urban areas is freer flowing and without the major delays caused by the congestions.

The summary of the scenario is as follows:

- The environment tested includes both urban and inter-urban
- Congestion and high built-up areas in the urban part of the journey
- Flowing traffic in the inter-urban part of the journey
- The journey is continuous without any intentional brakes
- Fixed attachment of the GPS device to the roof of the vehicle

INITIAL FINDINGS

The journey from A1 (Tierney Road) to B2 (Madejski Stadium) took 103 minutes and 52 seconds and 44.877 sentences in total have been received via GPS device. The following table presents co-ordinates of start, end and border point between urban and inter-urban areas as received via GPS device.

Point	Latitude (deg)	Longitude (deg)	First GGA UTC time	Last GGA UTC time
Trip start (UB part start)	51.445522	-0.127955	05:58:28	N/A
M4 section (UB end /IB start)	51.495107	-0.349742	07:02:33	
Trip end (IB part end)	51.423973	-0.98465	N/A	07:42:16

Table 8 – Measured co-ordinates

The following chart illustrates the number of satellites visible during the time of the journey for every 30th location fix due to better presentation purposes.

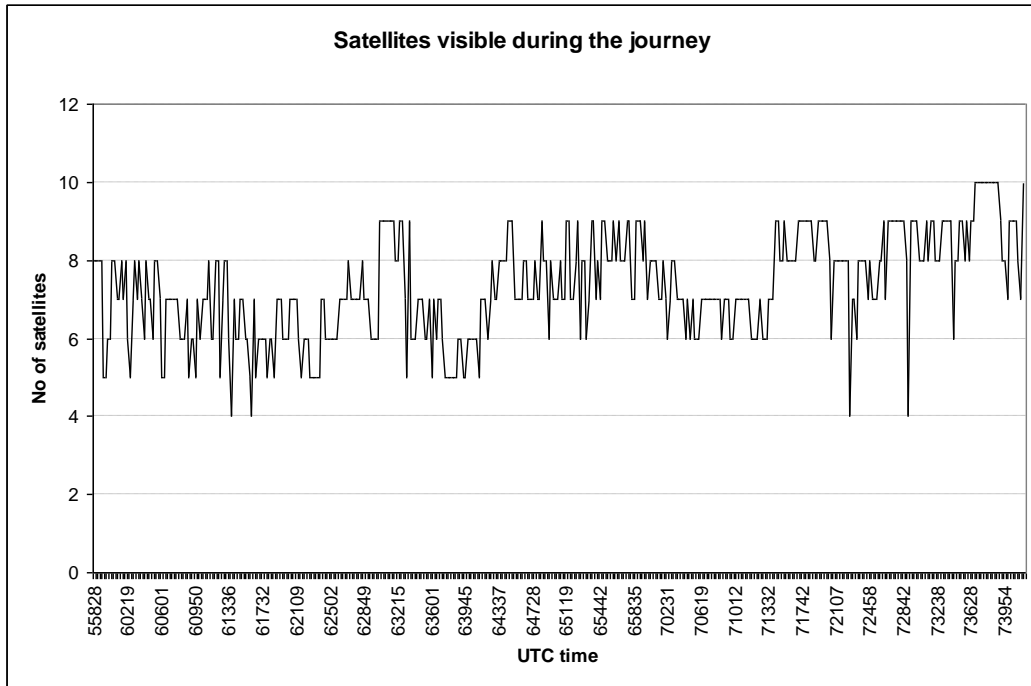


Figure 3 – Satellites visible during the journey

The number of satellites wavers between 5 and 8 initially and between 6 and 9 at the end of the UB part, and, as expected, there are rises in between 6 and 10 satellites with occasional deviation to only 4 satellites.

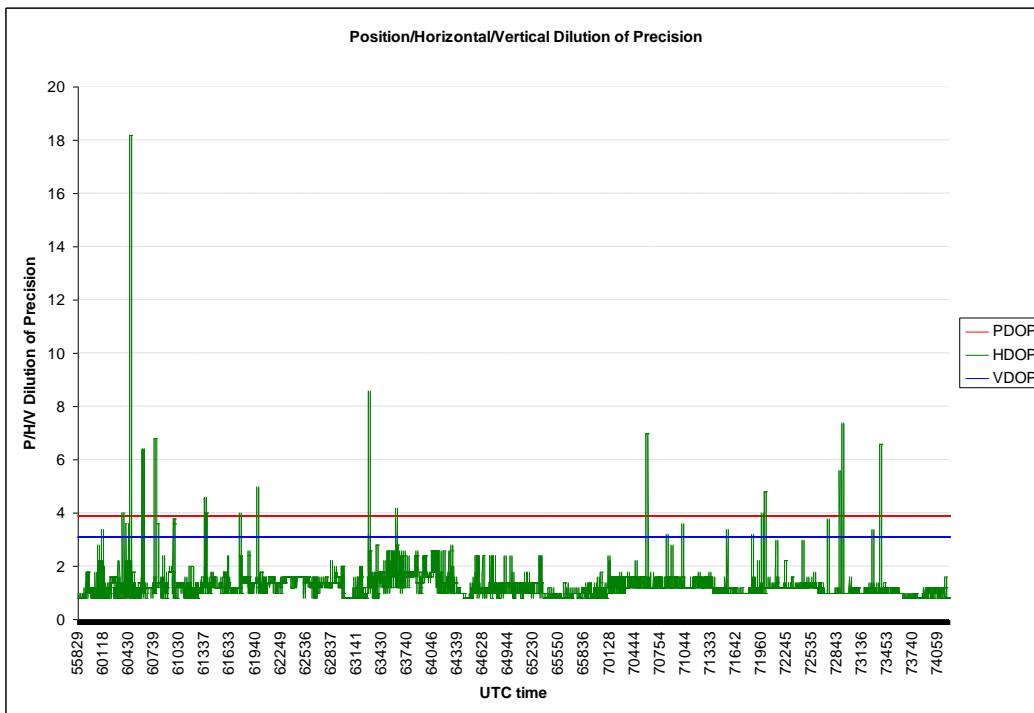


Figure 4 – Dilutions of Precision (DOP) during the journey

The DOP figure describes the geometric strength of satellite configuration on GPS accuracy (the principle is that the higher value of DOP the lower strength). The Horizontal DOP ranges from very good to moderate strength of configuration on GPS accuracy over the whole journey with denser waiver to good strength in the UB part

of the journey.

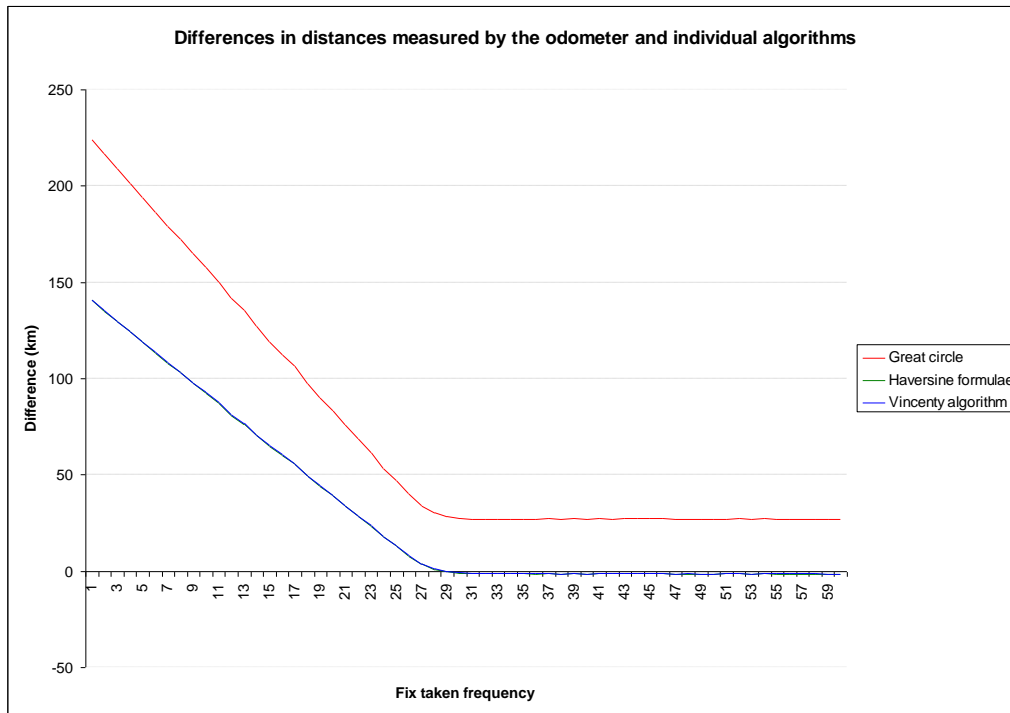


Figure 5 – Difference in distance measured by odometer and individual algorithms

In order to establish the ideal frequency for receiving the location data from the GPS device, the various intervals for the location fixes in series have been tested with three algorithms for the distance calculation. Following this, the results have been compared with the reading from the odometer. The chart above shows that the Vincenty algorithm and Haversine formulae provide very close results if the interval is set between every 28th to 33rd location fix.

CONCLUSION

This paper describes the principles and the initial findings of an alternate approach to the DBC based on GNSS. The major difference between the existing approaches and the suggested approach is the removal of both the matching process and the need for the representation of the road network subject to charge. The aim is to provide a solution for the distance calculation, applicable to any system which is using GPS, without using any additional measured data. The initial findings show that the parameters whose meaning indicates the level of the reliability of the GPS configuration are not sufficient enough to assess the accuracy of the location data. The variable interval between location fixes in series for the distance calculation potentially indicates one of the methods of tackling the inaccuracies that might be accumulated by using every location fix. However this has only been tested in one particular scenario and therefore would benefit from further research. Assessing the parameters indicating the reliability of the GPS configuration, the speed, and the variable intervals between location fixes will help with to establish the accuracy of the position received at the time and reaching the most accurate distance. It is anticipated that the results from further research will be presented at the time of the conference.

REFERENCES

- (1) Movable Type – Information Design & Management, Distance <http://www.movable-type.co.uk/>, the source of the distance calculation algorithms
- (2) Ordnance Survey, Britain's national mapping agency, Get-a-map TM, <http://www.ordnancesurvey.co.uk>
- (3) SiRF Technology, Inc., SiRF Binary Protocol Reference Manual, Revision 2.3, December 2007
- (4) SiRF Technology Inc., NMEA Reference Manual, Revision 1.7, August 2006
- (5) B. Hofmann-Wellenhof, H. Lichtenberger, J. Collins, GPS: Theory and Practice, Springer-Verlag, Austria, March 1997
- (6) ISO TC 204/SC N, Road transport and traffic telematics – Electronic fee collection – Application interface definition for electronic fee collection (EFC) based on Global Navigation Satellite Systems and Cellular Network (GNSS/CN), rev. 12, December 2004
- (7) Google Maps, <http://maps.google.com>