Mapping Linear Networks Based on Cellular Phone Tracking

Ronen RYBOWSKI, Aaron BELLER and Yerach DOYTSHER, Israel

Key words: Cellular Phones, Cellular Network, Linear Networks, Mapping.

ABSTRACT

The paper investigates the ability of accurately mapping linear road networks from cellular phone tracks.

Cellular phone tracking records the location of a cell phone for a string of time stamps. Such tracking does require an active call because of roaming. There are various methodologies of determining the location of a cell phone based on the signal strength to nearby antennae (base stations). Cellular networks which support the location functionality have a Location Manager (MLC) that translates the signal strengths to various base stations to the coordinates of the cell phone.

Standards like the US E911, which has been the engine behind the rapid technological development in the field, required the US cellular providers to measure the location in accuracy of 50-100 meters in 67% of the time and up to 150 meters in 95% of the time since the year of 2003.

We plan to use multiple tracks of cell phones location data provided by cellular providers to map linear road networks. Though cell phone locations are not accurate, we plan to use statistical methods to extract an accurate mapping of linear road networks. This possible since we are assuming that the multiple tracks will behave as random variables. This would allow rapid mapping where there is cellular coverage without any other mapping data. The technology would also enable updating linear road networks almost on the fly. Furthermore, since more and more cellular phones will have an embedded GPS, the same technology could be used along with such GPS data.

The paper reviews the current findings of this research, the development of a new geometric-statistical approach for filtering inaccurate locations and extracting approximately the original path from simulate data of cell phones locations. Moreover we will discuss future research which will include using true data to create a linear road network.

Mapping Linear Networks Based on Cellular Phone Tracking

Ronen RYBOWSKI, Aaron BELLER and Yerach DOYTSHER, Israel

INTRODUCTION

Historically mapping has been important for governmental, economic and military purposes and mostly dealt with roads, borders and properties. Today mapping is used as a basis planning, construction, government and military needs.

The common mapping methods today include global positioning systems (GPS), Airborne and Satellite Photogrammetry and surveying. Managing, displaying and analyzing the data are carried out by Geographic Information Systems GIS).

GPS is one of the most common and available systems that provide accurate location for navigating all over the world (Zogg J. M., 2002). Signals which carry the satellite location and the time of transmission are sent by the satellite and received by ground receivers. Calculation of location requires data from at least four satellites.

A less accurate but more prevalent technology for measuring location is based on cell phones. The cell phones are constantly tracked by the cellular providers at an accuracy of about 50 to 100 meters. Since cell phones are ubiquitous there is a potential for developing statistical methods that can quickly collect, create and update spatial data over a wide area based on cellular technologies. Such technology is being used to calculate real time traffic information (Drane et al., 1998) and other location based applications, but has not yet been used to actually acquire the coordinates and nodes of linear road networks.

The purpose of our research is to develop a mechanism for mapping linear road networks accurately with the use of multiple tracks of cell phones location data provided by cellular providers. In chapter 2 we shall describe the cellular technology with emphasis on the ability to measure cell phones location. On chapter 3 we shall present the simulator developed for building data base of cellular tracks for the research needs. Chapter 4 will follow the practical experiments that were done so far. Chapter 5 will present the future plans on the research issues.

1. CELLULAR TECHNOLOGY

Development in cellular technology has let to the mass use of cell phones, making this technology to an integral part of everyday life. Cellular companies have expanded their services to include location based services (LBS), which enables numerous possibilities such as rescue, real time road traffic analysis (Drane et al., 1998), billing based on location, LBS advertising, accumulated tracking data is allows cellular companies to plan more effective deployment of cellular networks, studying the social characteristics based on time-space and a true and valid planning for public systems (Ahas & Ular, 2005).

LBS is enabled by the ability to accurately measure the location of a cellular device. The location can be determined by the cellular device itself or by the cellular network. When the cell phone has an independent measuring device (e.g. GPS) the location determination is called “Device Based”. When the location of the phone is calculated by various measurements between the cell phone and the bases stations, the location determination is called “Network Based”. There are two types of network based location determination:

A. **Cell ID based -** calls are monitored by nearby base stations that provide the initial spatial boundary of the phone. This is further refined by technologies for measuring the transmission strength and the spatial distribution of the base station (using triangulation) (Kiran et al., 2007).

B. **Network based -** the measurements of the transmission strength (also defined as the distance from the basis station) and the signal direction are combined from the nearby base stations. The location of the cell phone is intersected by the distance and angles (Drane et al., 1998).

Accurate determination of cell phone location infringes on user's privacy (Fosca et al., 2005). Thus any use of cellular location to map linear road networks will have to overcome this problem. There are numerous possibilities: anonymity, user agreement and user ability to turn on and off location tracking and other means. This is why some of the research is based on simulation data. Currently there are many programs using cell location data such as:

- Emergency services
- Knowledge discovery Giannotti et al. (2005).
- Changxuan et al. (2005) offered a method of improving the public transportation planning using movement patterns taken from cellular tracking data.
- Forecasting heavy traffic requires the attribution of measured location to a mapped linear road network (Nanne, 2001).

In any case, we will not discuss privacy issue in this paper and only assume that it is possible to overcome this issue.

2. SIMULATOR FOR CELL PHONES TRACKING

Our initial research is based on simulating cell phone tracking to test feasibility and develop technologies. The simulator was built to create a collection of cell phones locations on a mapped linear road network. The mapped linear road network and other parameters defined by the simulator user are the input. The output is two sets of locations: “real” locations along the road, and “measured” locations that represent the location if it was calculated by the cellular providers existing technology. Using the simulator creates a controlled working environment for the mapping process that we try to develop on this research.

The following figure is a screenshot of the simulator:

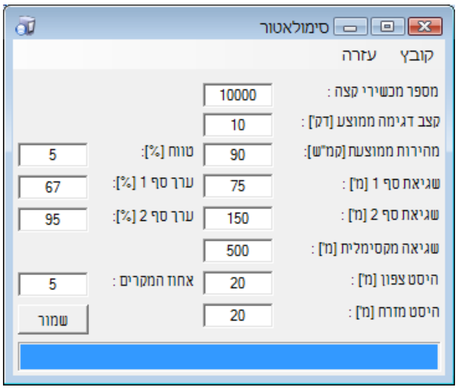

Figure 1 - The parameters: Right row (from top to bottom): a. Number of sampled cell phones; b. Average sampling rate for each phone (minutes); c. Average speed along the road (km/h); d. Threshold error 1 and 2 (meters); e. Maximal error (meters) for the range above threshold 2; f. Deviation to the north/south and deviation to the east/west (meters); Left row (from top to bottom) – a. Deviation from the average speed (percentage); b. Threshold value 1 and 2 (percentage); c. Occurrence of deviation (percentage);

After loading the chosen road, the **number of sampled cell phones** are defined. Each cell phone will be sampled in equal time intervals according to the **average sampling rate (minutes)**. The **average speed along the road (km/h)** is set to all the devices, but there is a permissible deviation according to the **deviation percentage from the average speed**. The "measured" location must match the maximal deviation range from the "real" location. This is set according to **threshold errors 1 and 2,** and the **maximal error (meters)**. The occurrence percentage in each of the deviation ranges is set by the **threshold values 1 and 2**. For some cases a deviation error is activated in the **north/south** and **east/west directions** according to the **occurrence percentage** of the defined deviation.

3. PRACTICAL EXPERIMENTS

Note that simulations create random tracks and the situation in the ”real world” may require further refinement of the algorithms used. The measured locations of cell phones are not accurate enough for mapping needs. The locations have 3 kinds of errors: random, deviation and severe.

Filtering the severe errors and smoothing the other errors is essential for the first step of mapping. A geometric-statistic method was developed to deal with the errors and an approximate release of the original linear road is also done, with no reference to connections between locations (e.g. a single cell phone holding a number of locations).

The method's steps are: first, an organized geometric network of cells is spread above the location data space. Then, each of the locations is attributed in space to one of the network's cells. A “representation point” is determined for each cell by calculating the center of gravity for all the locations attributed to the specific cell. The network cells are classified into four categories according to the number of locations attributed to each cell: in the upper category are the cells holding the maximal number of locations and in the lower category are the cells holding the minimal number of locations.

The approximate road extrication process is calculated by adding some of the representation points from the two top categories according to a defined algorithm. In such way the severe errors are filtered, and the other errors are smoothed.

Operating the method on a number of cellular tracking sets (produced by the simulator) has reached the following conclusions:

A. The method can extricate an approximate linear road networks.
B. The maximal improvement for the approximate road in accordance with the crude data is about 70%.
C. There is a minimal value for the number of cell phones needed to reach an optimal result.

A number of questions were raised:

A. What is the optimal size for network cells?
B. Does the road angle in relation to the network have an effect?

To answer these questions a vast number of tests were taken in front of different sets of locations. The following conclusions were reached:

A. The optimal size of a cell network is a function of threshold error 2 (as defined in the simulator) set by the user.

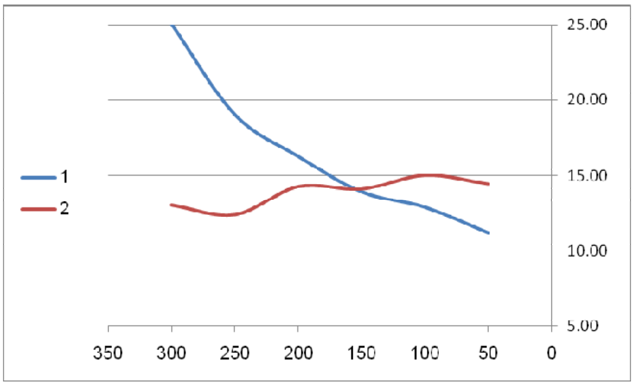

Figure 2: X axis – Size of network cell (meters); Y axis – Distance (meters); Graph 1 – The distance of the original road turning points to the approximate road; Graph 2 – The distance of the approximate road turning points to the original road. The graph's intersection point has a value of 150 meters, and that is the value of threshold error 2 defined to this set of data.

B. The road angle in relation to the network is insignificant.

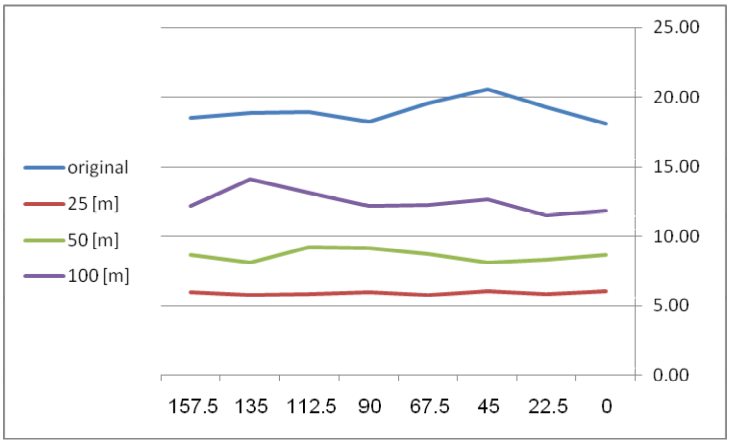

Figure 3: X axis – Network angle in relation to the road (angles); Y axis – Distance (meters); The four graphs present the average distance of the points after filtered from the original road in the different angles. The blue graph presents the preliminary situation. The red, green and purple graphs present distance from networks in 25, 50 and 100 meters size respectively. The network angle in relation to the original road is insignificant.

After using the geometric-statistical method, filtered locations of the cell phones and an approximate traffic line in ±10 meters accuracy are acquired.

A possible method to improve the level of accuracy is based on matching polynomials by segments. The original road is cut into segments (e.g. linear segments) and the polynomials are matched to each segment by its type. A set of measured locations was built along a linear line to find if this method works. The geometric-statistical method was operated on the data and a matching polynomials algorithm was operated on the representation points from the two upper categories. The result shows that using this method increase the mapping accuracy to ±3 meters.

4. CONCLUSION AND FUTURE WORK

Currently, most research tries to match locations to an existing digital road map, for determining factors such as traffic, but not to map the roads themselves from scratch.

The research operated a geometric-statistical method on different kinds of sets built by simulator and operated matching polynomials methods on segments. This work suggests a significant potential for extraction an original linear road networks out of measured data from cell phones. The extraction of the original traffic line will permit accurate mapping. The methods and tools developed in this research will allow an accurate and rapid mapping of road networks by using only the cell phones location data.

More and more cellular phones are being equipped with GPS devices. If the above research program works, including cellular devices with GPS technology will speed up the convergence of the linear road networks.

Future work will focus on the following issues:

1. Building an algorithm to match polynomials by segments which are identified on the basis of the approximate road. After finishing the mapping process from the first step of crude data to the last step of mapping the original road, the process will be tested on real cellular location data.
2. Collecting cellular location data from accurately mapped road networks.
3. Operating the algorithm on the real location data and testing the product's compatibility to existing accurate maps.
4. Since the positioning of base stations may cause biased results, we will need a method to eliminate this bias. It would help to use cellular tracks from multiple cellular companies to reduce bias.
5. Develop a methodology for detecting and defining the nodes of linear road networks.
6. Develop a methodology for detecting one way, two way roads (and their center lines) and turn restrictions.
7. Develop a methodology for detecting the number of lanes in each direction and their centerlines.
8. Develop a methodology for detecting bridges and clover leaf structure.
9. Testing the effect of the different parameters on the quality of the solution.

REFERENCES

1. Ahas R., Ular M., 2005. Location Based Services - New Challenges for Planning and Public Administration. Futures, 37, pp. 547-561.
2. C. Drane, M. Macnaughtan, C. Scott, 1998. Positioning GSM Telephones. IEEE Communications Magazine, vol. 36, no. 4, pp. 46-59.
3. Changxuan Pan, Jiangang Lu, Shan Di, Bin Ran, 2005. A Cellular - Based Data Extracting Method for Trip Distribution.
4. F. Giannotti, A. Mazzoni, S. Puntoni, C. Renso, 2005. Synthetic Generation of Cellular Network Positioning Data. In Proc. of ACM GIS 05, Bremen.
5. Fosca Giannotti, Mirco Nanni, Dino Pedreschi1, Chiara Renso, 2005. GeoPKDD Geographic Privacy-aware Knowledge Discovery.
6. Kiran G.S., Bhoolakshmi M., G. Varaprasad, 2007. Algorithm for Finding the Mobile Phone in a Cellular Network. IJCSNS International Journal of Computer Science and Network Security, VOL. 7 No. 10.
7. Nanne J. van der Zijpp, 2001. A MAXIMUM LIKELIHOOD MAP – MATCHING ALGORITHM. The 13th Mini - EURO Conference Mon Nov 26 2001.
8. Wise S., 2002. GIS Basics.
9. Wolf P.R., B.A. Dewitt, 2000. Elements of Photogrammetry: with applications in GIS.
10. Zogg J. M., 2002. GPS Basic.

BIOGRAPHICAL NOTES

Ronen Rybowski received an Engineering degree in Geo-information from the Technion – Israel Institute of Technology in 2005. He is currently a candidate on the M.Sc track in Mapping and Geo-Information Engineering at the Technion.

Dr. Aaron Beller received a Ph.D. in Mathematics from The Hebrew University of Jerusalem in 1977. Lectured in the Department of Decision Sciences, The Wharton School, University of Pennsylvania and the School of Business Administration, Hebrew University of Jerusalem from 1977 – 1986. Worked in numerous software compnaies in Israel and the US. Currently a GIS developer and consultant.

Prof. Yerach Doytsher graduated from the Technion – Israel Institute of Technology in Civil Engineering in 1967. He received a M.Sc. (1972) and D.Sc. (1979) in Geodetic Engineering also from Technion. Until 1995 he was involved in geodetic and mapping projects and consultations within the private and public sectors in Israel. Since 1996 he is a faculty staff member in Civil Engineering and Environmental at the Technion, and currently the Dean of the Faculty of Architecture and Town Planning. He also heads the Geodesy and Mapping Research Center at the Technion.

CONTACTS

Mr. Ronen Rybowski
Technion – Israel Institute of Technology
Faculty of Civil and Environmental Engineering
Geva Binyamin 63, Geva Binyamin
Israel
Tel. +972-544-866463
Email: RonenRy@Gmail.com

Dr. Aaron Beller
Technion – Israel Institute of Technology
Faculty of Civil and Environmental Engineering
Email: aaron.beller@gmail.com

Prof. Yerach Doytsher
Technion – Israel Institute of Technology
Faculty of Architecture and Town Planning
Technion City
Haifa 32000
Israel
Tel. + 972 4 8294001
Fax + 972 4 8295641
Email: doytsher@technion.ac.il