STATISTICAL INVESTIGATION OF GPS-BASED LOCALIZATION OF VEHICLES

Rajmund Drenyovszki[1], Lorant Kovacs[1*], Krisztian Barsony[1], Istvan Pinter[1] and Bence Csak[1]

[1] Kecskemet College, Faculty of Mechanical Engineering and Automation,
Izsaki ut 10, H-6000 Kecskemet, HUNGARY
Corresponding author e-mail: lorant.kovacs@gamf.kefo.hu

Abstract
As part of an intelligent traffic management system development project we are building an in-vehicle GPS module for algorithms requiring position information. The aim of this paper is to present our statistical approach which is used to investigate the estimation of GPS receivers accuracies and to elaborate signal processing methods to improve precision.

Keywords:
GPS differential error, GPS location, covariance matrix

1. Introduction

Users of intelligent traffic management systems are growing worldwide. Traffic information is used not only by vehicle drivers but they are also available for public transport users, pedestrians and cyclists [1]. From the viewpoint of traffic and transportation the role of European integration is even more valuable. The need to develop cooperative intelligent traffic systems and services is increasing, applications use cross border data more extensively. Broad set of technical solutions are available that enable us to build intelligent traffic management systems so it is time to develop algorithmic solutions to solve the most relevant problems posed by our society. As many universities and research centers, Kecskemet College also joined the research activities related to intelligent traffic system developments. The "Smarter Transport – ICT support for cooperative traffic systems" (TÁMOP-4.2.2.A-11/1/KONV) project is about driver assistance and sustainable traffic solutions for vehicles, where vehicles communicate with each other (vehicle-to-vehicle) and with infrastructure elements (e.g. roadside units). In our project we are working on research fields that need GPS data processing.

Proactive pollution reduction with ad hoc networks capable of vehicle-to-vehicle and vehicle-to-infrastructure communication

Reducing environmental impact is feasible with preventive algorithmic solutions. On the one part technological basis for algorithmic developments are present in cars in the form of embedded systems (more specifically Electronic Control Units), on the other part researchers need to elaborate vehicle-to-vehicle and vehicle-to-infrastructure communication solutions to be able to control traffic in a whole. It is possible to reduce environmental impact with reactive actions. Vehicles can send desired destination information prior to the trip, or in case it is not available, habitual route information can be communicated to a central traffic control system. Pre-trip knowledge about traffic can help the system to make optimal decisions so that preventing polluted gridlocks becomes possible. To reach these aims we need to collect GPS location information, learn habitual route behaviors and build adequate knowledge base.

Development of novel driver assistant systems with ad hoc networks capable of vehicle-to-vehicle and vehicle-to-infrastructure communication

One of the most effective ways to make traffic safety better is the improvement of the efficiency of driver assistant systems. Communication also opens up new ways towards the development of totally new kind of services in this field. One potential application can be the forecasting of risky situations based on database of drivers' profiles and motion and velocity vectors of nearby vehicles. Driving profiles are learned from vehicle dynamics and GPS location information. The driving assistant can send alert signals to own and nearby vehicles drivers according to the GPS data, forecasted vehicle path and dynamics.

Vehicles providing environmental data as mobile sensors platforms, data acquisition and communication tools

In case of vehicle-to-infrastructure communication capability, data coming from on-car sensors can be transmitted to central systems and much more information can be gathered from the environmental parameters as ever before in a distributed manner. If the system receives precise GPS positions, it can generate maps containing environmental parameter data and can make GSM/radio coverage measurements which can support several applications as background information.
We can see from the descriptions of our project aims that our main task is to collect and process the location data coming from GPS devices. First

of all we make measurements with multiple GPS receivers to be able to estimate GPS location errors with statistical investigations. Development of an intelligent traffic management system requires the elaboration of differential accuracy of GPS localization. If we can define a statistical method to improve the precision with multiple GPS receivers, we can build systems like collision avoidance which demand 1-2 m precision or better. This paper aims to rely on pure GPS data, and does not use sensor fusion techniques, which are available in the literature [2], most frequently using data coming from an accelerometer or gyroscope.
Sources of errors in GPS are from the differences of satellite positions, orbital fluctuation, multipath, relativistic and atmospheric effects, clock and rounding inaccuracies [3]. In consideration of the error sources and with attention on our application fields, following properties of measurement situations are the most important: accuracy-speed dependence, radius of trajectory, effect of multipath (high building, vs. rural environment) and weather conditions. In our measurements 3 different types of GPS receivers were installed in a car. It is important to note that these devices are not capable to use EGNOS system to make positioning more precise. The primal issue is that on a moving vehicle the exact GPS position is not available, so we do not have a fixed position as a reference point. The only opportunity to make implications is to use differential signal from multiple GPS receivers. In the followings the reader can find our model, evaluation of the measurement data and our final conclusions.

2. The model

In this paper an effective signal processing method will be introduced in order to improve the precision of GPS-based localization of vehicles. The base of the method comes from statistical analysis of the signals of multiple receivers. In the model the following assumptions were made:

- sensor fusing techniques such as combining the GPS signal with the signal of accelerometers are excluded;
- We can access only the global output of the GPS receiver devices (we cannot reach the raw data of time measurements);
- Possibly signal processing elements have been applied in the receiver devices – the algorithms are not known for us;

In the mathematical model we assume generally N receivers that measure a series of position in a synchronized way. For the sake of simplicity, without any loss of generality, we assume N=3. Let us denote the signal of the ith receivers by X_i as follows

$$\begin{aligned} X_1 &= X_0 + \mu_1 + \xi_1 \\ X_2 &= X_0 + \mu_2 + \xi_2 \\ X_3 &= X_0 + \mu_3 + \xi_3 \end{aligned} \tag{1}$$

where X_0 denotes the exact position, μ_i is the DC offset of the ith receiver and ξ_i represents Gaussian measurements noise with zero mean and σ_i standard deviation

$$\begin{aligned} \xi_1 &\sim N(0, \sigma_1) \\ \xi_2 &\sim N(0, \sigma_2) \\ \xi_3 &\sim N(0, \sigma_3) \end{aligned} \tag{2}$$

respectively. The measurement noise of the different receivers are assumed to be independent, and hence uncorrelated:

$$E(\xi_i \xi_j) = 0 \tag{2.1}$$

The basic problem is, that in the case of a moving vehicle, we have no information about the exact position X_0, therefore we can elaborate the statistical properties of difference sequences. In order to do this we define the following variables:

$$\begin{aligned} d_{12} &= x_1 - x_2 = \mu_1 - \mu_2 + \xi_1 - \xi_2 \\ d_{23} &= x_2 - x_3 = \mu_2 - \mu_3 + \xi_2 - \xi_3 \\ d_{31} &= x_3 - x_1 = \mu_3 - \mu_1 + \xi_3 - \xi_1 \end{aligned} \tag{3}$$

Using the properties of Gaussian random variables [4] the statistics of the difference signals are

$$\begin{aligned} d_{12} &\sim N(\mu_1 - \mu_2, \sigma_{12}) \\ d_{23} &\sim N(\mu_2 - \mu_3, \sigma_{23}) \\ d_{31} &\sim N(\mu_3 - \mu_1, \sigma_{31}) \end{aligned} \tag{4}$$

where

$$\begin{aligned} \sigma_{12}^2 &= \sigma_1^2 + \sigma_2^2 \\ \sigma_{23}^2 &= \sigma_2^2 + \sigma_3^2 \\ \sigma_{31}^2 &= \sigma_3^2 + \sigma_1^2 \end{aligned} \tag{5}$$

respectively. Furthermore we can calculate the covariance matrix **K** of the difference signals, where by definition

$$\boldsymbol{K} = \begin{bmatrix} E(\tilde{d}_{12}^2) & E(\tilde{d}_{12} \cdot \tilde{d}_{23}) & E(\tilde{d}_{12} \cdot \tilde{d}_{31}) \\ E(\tilde{d}_{23} \cdot \tilde{d}_{12}) & E(\tilde{d}_{23}^2) & E(\tilde{d}_{23} \cdot \tilde{d}_{31}) \\ E(\tilde{d}_{31} \cdot \tilde{d}_{12}) & E(\tilde{d}_{31} \cdot \tilde{d}_{23}) & E(\tilde{d}_{31}^2) \end{bmatrix} \tag{6}$$

and

$$\tilde{d}_{12} = d_{12} - E(d_{12}) = \xi_1 - \xi_2$$
$$\tilde{d}_{23} = d_{23} - E(d_{23}) = \xi_2 - \xi_3 \qquad (6.1)$$
$$\tilde{d}_{31} = d_{31} - E(d_{31}) = \xi_3 - \xi_1$$

are the centralized difference variables. The expectations in the covariance matrix can be easily calculated as

$$E(\tilde{d}_{12}^2) = E((\xi_1 - \xi_2)^2) = E(\xi_1^2 - 2\xi_1\xi_2 + \xi_2^2) = \sigma_1^2 - \sigma_2 \qquad (7)$$

$$E(\tilde{d}_{12} \cdot \tilde{d}_{23}) = E((\xi_1 - \xi_2)(\xi_2 - \xi_3)) = E(\xi_1\xi_2 - \xi_2^2 - \xi_1\xi_3 + \xi_2\xi_3) = -\sigma_2^2 \qquad (8)$$

where we used the assumption that the measurement noises are uncorrelated. As a result

$$\mathbf{K} = \begin{bmatrix} \sigma_1^2 + \sigma_2^2 & -\sigma_2^2 & -\sigma_1^2 \\ -\sigma_2^2 & \sigma_2^2 + \sigma_3^2 & -\sigma_3^2 \\ -\sigma_1^2 & -\sigma_3^2 & \sigma_1^2 + \sigma_3^2 \end{bmatrix} \qquad (9)$$

3. Measurement results

For validation of the model described in Section 2 the signal of three different GPS receivers has been registered in urban environment installed in a car. Since the exact trajectory of the vehicle is not known, only the differential data were used for calculations. The types of the receivers that were used are

- Sony Ericsson Xperia sk17i Mini Pro
- Garmin GPS18x-5Hz
- Bluetooth connected G-Sat BT-328

All the GPS receivers has been sampled by 1 Hz sample frequency. The measurement time was approximately 11 minutes resulting altogether 668 coordinate data per receiver.
We made difference calculations separately for the latitude and longitude coordinates, because the Gaussian assumption can be used only for coordinate data (the Euclidean distance is a positive valued random variable, with chi-square distribution [5] in the case of Gaussian coordinate-data.)

Results for latitude coordinates

In Figure 1 the difference signals as a function of time can be seen, while in Table 1 we summarized the statistical parameters of the difference signals. We used the following notations:

- $d_{12:}$ Xperia-Garmin
- $d_{23:}$ Garmin-Bluetooth
- $d_{31:}$ Bluetooth-Xperia

Note, that there is a nonzero mean, i.e. a constant offset between the receivers. In the case of the Xperia-Garmin difference, this offset is almost 5 meters. The standard deviations are approximately 3-5 meters, but the maximum of the differences is of approx. 40 m. For more information see Figure 2 which depict the histograms of the difference signals. We also performed chi-square hypothesis test to validate the normality of the difference signals, the H0 hypothesis of normality should be rejected with a negligible probability (see Table 1), i.e. the Gaussian hypothesis for the measurement uncertainty validated by our measurements.

	d_{12}	d_{23}	d_{31}
mean	4.8024	-1.4500	-3.3524
standard deviation	3.0778	4.4750	5.3314
prob. of H0 reject.	8.36e-143	0	2.38e-241

Table 1. Statistics for latitude difference values

The covariance matrix of the difference signals can be used to estimate the individual uncertainty of the receivers, as explained in Section 2, the covariance matrix for the latitude data is

$$\mathbf{K} = \begin{bmatrix} \mathbf{9.4727} & \mathbf{-0.5369} & \mathbf{-8.9358} \\ \mathbf{-0.5369} & \mathbf{20.0254} & \mathbf{-19.4885} \\ \mathbf{-8.9358} & \mathbf{-19.4885} & \mathbf{28.4243} \end{bmatrix}$$

from which the estimation of the individual standard deviations of the receivers are

$$\sigma_{Xperia} = \sqrt{8.9358} = 2.99m$$
$$\sigma_{Garmin} = \sqrt{0.5369} = 0.73m$$
$$\sigma_{Bluetooth} = \sqrt{19.488} = 4.41m$$

As one can see, the accuracy of the three receivers are very different (regarding the latitude coordinate), Garmin has the smallest standard deviation of 0.73 m.

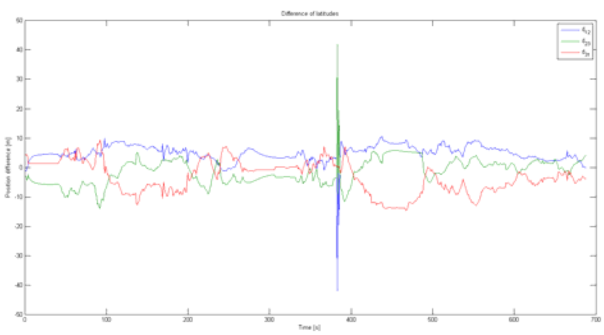

Figure 1. Differences of latitude coordinates

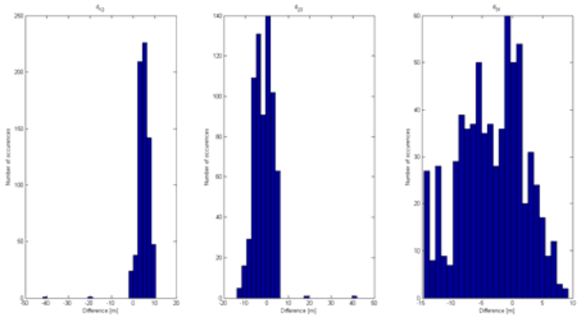

Figure 2. Histograms of differences of latitude coordinates

Results for longitude coordinates

In Table 2 there are the descriptive statistics of the longitude coordinate data. The offset is much lower for longitude (<0.5m). The hypothesis of normality can be accepted by a high probability for the longitudes as well. In Figure 3 and 4 the difference signal and the histograms of longitude data are depicted, respectively.

	d12	d23	d31
mean	0.4969	-0.4580	-0.0388
standard deviation	4.8955	9.2832	7.8148
prob. of H0 reject.	1.81e-005	4.64e-163	3.80e-029

Table 2. Statistics for longitude difference values

The covariance matrix:

$$\mathbf{K} = \begin{bmatrix} 23.965 & -24.536 & 0.5702 \\ -24.536 & 86.178 & -61.64 \\ 0.5702 & -61.64 & 61.0719 \end{bmatrix}$$

from which the standard deviations of the individual receivers:

$$\sigma_{Xperia} = \sqrt{0.5702} = 0.75m$$

$$\sigma_{Garmin} = \sqrt{24.536} = 4.95m$$

$$\sigma_{Bluetooth} = \sqrt{-61.64} = 7.85m$$

Note, that for the longitude coordinate the Xperia has the lowest standard deviation of 0.75 m, and (as well as for latitude) Bluetooth shows the poorest performance.

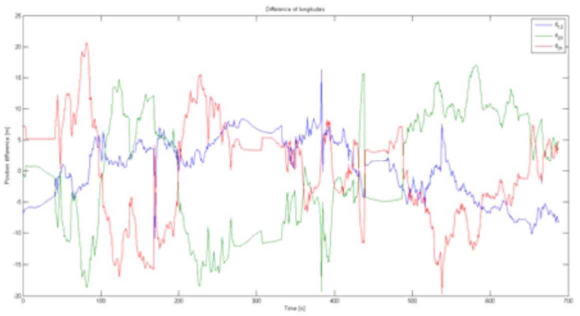

Figure 3. Differences of longitude coordinates

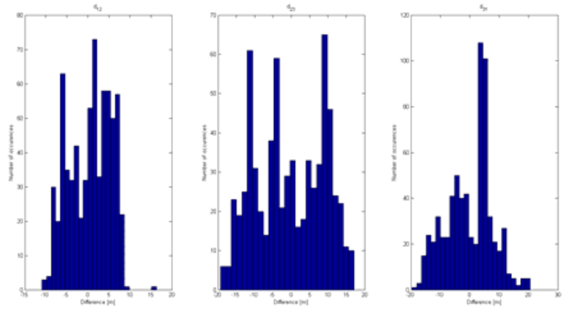

Figure 4. Histograms of differences of longitude coordinates

As a conclusion one can see, that in the case of this three type of receivers the longitude coordinate data should be read from the Xperia, while the latitude from the Garmin receivers, having smaller variance than can be reached by averaging.

4. Conclusion and future work

In this paper we introduced statistical investigation to recognize the statistical behavior of commercial GPS receivers. The Gaussian nature of the uncertainty of the measurement has been proven. Using the special form of the correlation matrix, a method has been introduced to estimate the individual accuracy of the receivers by differential measurements. In the future, for the sake of statistical reliability much more data point will be required. For control more receivers from the same type should be incorporated. The accuracy as a function of different parameters related to mobile vehicle applications (such as velocity, shape of trajectory, rural vs. urban environment) should be systematically detected.

5. Acknowledgement

This research is supported by TÁMOP-4.2.2.C-11/1/KONV-2012-0012: "Smarter Transport" - IT for co-operative transport system - The Project is supported by the Hungarian Government and co-financed by the European Social Fund.

6. References

[1] J. Pfister, B. Horváth, T. Kulpa, T. Tvrzsky, “GSP-technical feasibility and achieve results“

[2] C. Hide, T. Moore, M. Smith, “Adaptive Kalman Filtering for Low-cost INS/GPS“. The Journal of Navigation, 56, pp. 143-152, 2003

[3] “Sources of Errors in GPS,” [Online] Available: http://www.kowoma.de/en/gps/errors.htm [Accessed: 10-Sept-2013].

[4] J. Wu, “Some Properties of the Gaussian Distribution”, 2004

[5] E. Hung, L. Xiao, R. Y.S. Hung, “An Efficient Representation Model of Distance Distribution Between Uncertain Objects”, Comp. Int., Vol. 28, Issue 3, pages 373-397, Aug 2012.