# Method of Collecting, Processing and Storing Geolocational Data Received From Mobile Devices[1]

M. L. Voskoboinikov, R. K. Fedorov
V.M. Matrosov's institute of system dynamics and control theory SB RAS
Lermontov str., 134, Post Box 292
664033, Irkutsk, Russia
mikev1988@mail.com, fedorov@icc.ru

**Abstract.** The article considers the method of calculating the average speed of traffic along the roads of the Irkutsk region, Russia. The method implementing software is based on mobile technologies, distributed computing and the use of the principles of open access to spatial data. An Android application and a Web service have been designed for the collection, processing and storage of geolocation data.

**Keywords**: mobile technologies, distributed computing, GPS

## 1. Introduction

In the Irkutsk region of Russia, anthropogenic impact on the environment is being investigated. One of the main factors of anthropogenic impact is the transport availability of the environmental objects. To determine the transport availability of the environmental objects, the average speed of traffic along the roads of the Irkutsk region is required to be assessed. At the moment there is no information on the average speed of vehicles at specific time intervals. In the earlier study, the average speed of traffic was assessed based on information on the type of road surface.

The use of public mapping services provided by Yandex, Google and others to calculate the average traffic speed for research purposes is impossible due to the lack of open access to the initial geolocation data, such as the speed of individual vehicles, through their API. This circumstance requires the implementation of a system for collecting, processing and storing geolocation data received from mobile devices. In connection with the foregoing, it is necessary to organize the collection, storage and processing of traffic data on the road network of the Irkutsk region.

One of the handiest ways to get real speed of traffic is to use mobile devices equipped with GPS/GLONASS-sensors, permanent access to Internet and opportunity of install custom applications. This way of collecting data uses a concept reffered to as the phone as a sensor [1]. According to the authors in [2], it is quite correct to talk about the standard means of collecting sensory data using mobile devices. Based on the concept of the phone as a sensor in [3], the authors give examples of systems based on

[1] This work was supported by the Center for Collective Use of the IICN ISEC and the RFBR grant, No. 16-07-00411.

the collection of sensory data using mobile devices, as well as practical implementation of the system to monitor the vehicle's compliance with the prescribed route.

## 2. Problem definition

To obtain more accurate and complete data, several mobile devices are required to be used at once. In this case, there is a significant increase in the volume of initial data, which leads to the complexity of collecting, storing and processing large amounts of geolocation data.

To solve the above problem, the authors suggest the creation of new methods, software and technology solutions based on:

– mobile technologies;
– distributed computing;
– principles of open access to spatial data and Open Geospatial Consortium (OGC) standards.

## 3. Method and software

This method is based on the use of several mobile devices to collect geolocation data. The use of mobile devices has several obvious advantages:

1) there is no need to install special expensive equipment with GPS / GLONASS-sensors on vehicles;

2) the majority of modern mobile phones with GPS/GLONASS-sensors has an acceptable accuracy of geolocation for further analysis of geolocation data and the identification of regularities.

To solve the task, the following software systems were developed:

– Android application for collecting geolocation data together with the accuracy, time, speed and altitude information;
– a specialized Web service for processing, storing, analyzing and visualizing collected data.

The initial data for calculating the average speed of traffic on the road network of the Irkutsk region consist of:

1) geolocation data containing the movement speed of the mobile devices;

2) database of roads of the Irkutsk region.

## 4. Implementation of Android-based applications

To implement the Android-application, the integrated development environment Android Studio was chosen.

The Android-application process execution consists of several stages.

At the first stage, after the application is launched, the Internet availability of connection to Web service is checked. If there is the connection, a request for registration to the Web service is sent to obtain an unique identifier that will be used later to link the data to a specific user to track the movements on the map of the region, otherwise the application displays an error message and terminates.

At the second stage, a request is sent to the service to obtain an application configuration that specifies the interval for accessing the geolocation service. In accordance with the resulting configuration is performed periodically appeal to the geolocation service for access geo-location data such as latitude, longitude, altitude, speed and accuracy.

At the third stage, the received geolocation data is sent to the Web service in the format of the serialized JSON object via the HTTP protocol. If there is no access to the Web service at the time moment of sending data, the data is saved to the local SQLite database.

The next time of the application start, if the application has an access to the Web service, the data stored in the local database is sent to the Web service wit the same way as above.

## 5. Implementation of the geolocation data Web-service

In accordance with the main aim, the implemented Web service should correspond to the following requirements:

1) establishing and maintenance of HTTP connections with a variety of mobile devices;
2) distributed storage and processing of a large volume of constantly updated geolocation data;
3) online processing and analysis of the spatial data;
4) the data visualization on an interactive map of Irkutsk region.

When choosing the technologies for implementing the Web service, the following criteria were taken into account:

– openness;
– availability of a complete documentation.
– ensuring stable operation in high load mode.

The software platform Node.js was chosen as the basis for the implementation of the Web-service, as well as the REST architecture of distributed applications. As a database management systems, the object-oriented rational database server PostgreSQL was chosen, which allows efficient processing large amounts of data using spatial indexes R-Tree / GIST.

The table structure “location” is given in Table 1.

**Table 1.** Table structure "location"

| *Column name* | *Column type* |
|---|---|
| id | bigint |
| coord | point |
| datetime | timestamp with time zone |
| provider | character varying(32) |
| accuracy | double precision |
| Uid | integer |
| altitude | double precision |
| speed | double precision |
| Text | character varying(255) |
| Rid | integer |

To ensure the scalability of the system resources, it was decided to share the load between the three servers. A separate server has been used, which performs two tasks:

1) execution of Web-services;

2) load balancing between multiple servers.

Two servers were used for data storage and processing. The data is distributed randomly between them.

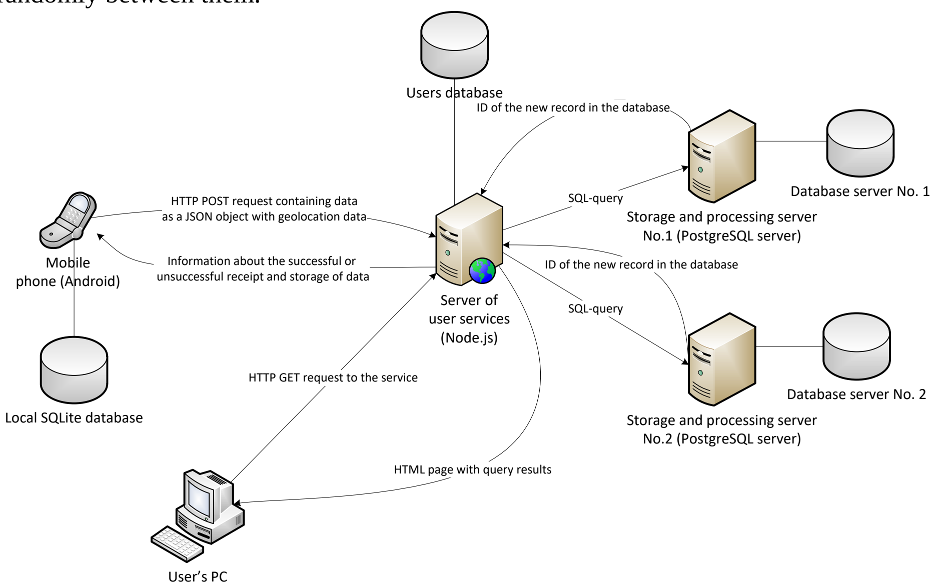


**Fig. 1.** The system architecture

A PostGIS extension is used to process and analyze geolocation data. For the visualization of geolocation data, an open freely distributed JavaScript library Leaflet was used.

To ensure confidentiality, the Web service does not store any user personal data, such as IMEI of the mobile device, telephone number and other data, which can be used by the attackers in illegal activities.

Fig. 2 is a block diagram of the algorithm for storing the mobile data. The algorithm consists of four steps:

*Step 1*. The JSON object or JSON object array received from the mobile devices via the HTTP protocol arrives at the system input. The object contains the following data:

- coordinates of the mobile device: latitude, longitude, altitude;
- speed of the mobile device;
- the accuracy of the data;
- data source type, i.e., GPS/GLONASS or mobile network;
- timestamp of the data acquisition taken from the sensors of the mobile device.

*Step 2*. A SQL query is executed, which uses the coordinates of the mobile device obtained at the previous step and searches for the nearest road in the Irkutsk region road table. The result of the query is the road identifier ID.

*Step 3*. The obtained ID is saved together with the JSON object data.

*Step 4*. After successful saving data, a corresponding notification is displayed on the screen of the mobile device.

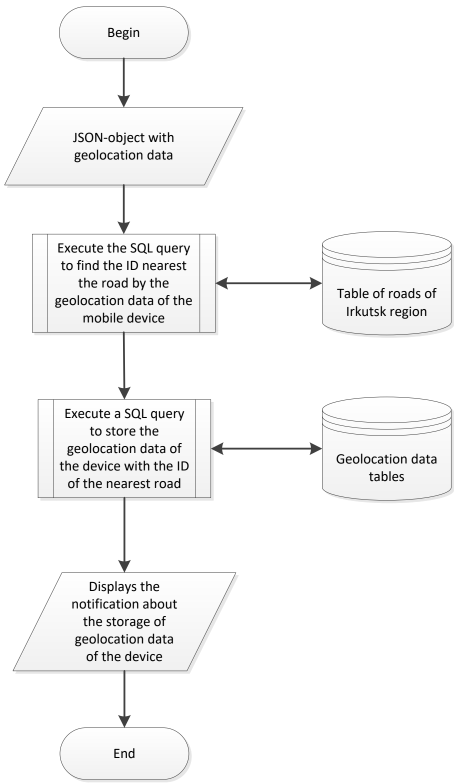


**Fig. 2.** Block diagram of the algorithm for storing the mobile data

## 6. Mathematical model

To calculate the average traffic speed in a distributed data storage and processing system, a mathematical model is developed. Below are the symbols of the mathematical model:

K – the number of servers processing data in a distributed system;

$L_k$ – the number of roads on which observations were made on the k-th server;

$m_{kl}$ – the number of stored observations on the speed of vehicles along the $L_k$-th road on the k-th server;

$v_{ikl}$ – the i-th value of the speed of vehicles on the k-th server on the l-th road;

The arithmetic mean value of the speed of the l-th road on the k-th server can be obtained by the formula:

$$V_{lk} = \frac{\sum_{i=1}^{m_{kl}} v_{ikl}}{m}, k = 1 \ldots K, l = 1 .. L_k \quad (1)$$

The arithmetic mean of the traffic speed for all servers can be obtained from the formula:

$$V_a = \frac{\sum_{k=1}^{K} V_{lk}}{K} \quad (2)$$

Weighted average value of the speed of the l-th road on all servers can be obtained by the formula:

$$V_W = \frac{V_{lk} * m_{kl}}{\sum_{i=1}^{K} m_{kl}}, k = 1 \ldots K, l = 1 .. L_k \quad (3)$$

The weight in this model is $m_{kl}$. The more stored observations on the speed of vehicles along the $L_k$-th road on the k-th server, the more weight the k-th server has in calculating the weighted average speed.

## 7. Implementation of analytical part of the Web service

To calculate the average speed of traffic on the road network of the Irkutsk region, the data presented in Shapefile format were used. It has previously been converted into a PostgreSQL table using the shp2pgsql utility.

The Fig. 3 is a block diagram of the algorithm for calculating and displaying the average speed of traffic on the outskirts of the Irkutsk region. The algorithm consists of five steps:

*Step 1.* The PNG image of the OSM map of the Irkutsk region is uploaded to the HTML page.

*Step 2.* oading the vector map of the road network of the Irkutsk region from the set of files format Shapefile.

*Step 3.* The Web service sends a request to two PostgreSQL servers to calculate the average traffic speed for each road, using the road table of the Irkutsk region.

*Step 4.* Data received from two PostgreSQL servers is combined to obtain the final average speed of traffic along the roads of the Irkutsk region.

*Step 5.* Each object of the road loaded on the HTML page is assigned the corresponding value of the average traffic speed, which can be obtained by clicking on it.

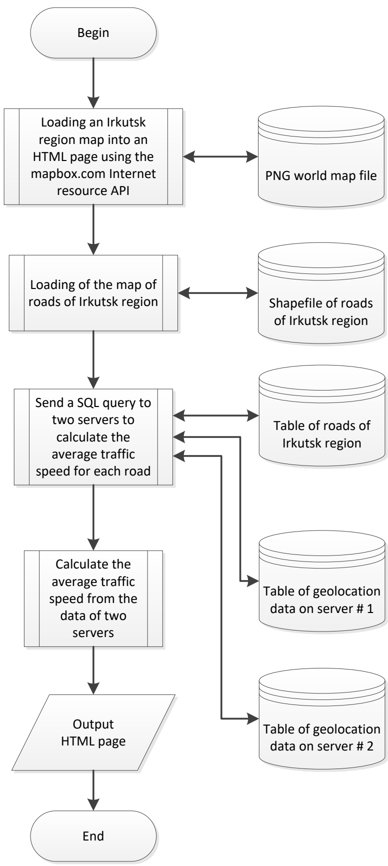


**Fig. 3.** Block diagram of the algorithm for calculating the average speed of traffic on the outskirts of the Irkutsk region

## Conclusion

The result of the work is a system for collecting, processing, storing and visualizing geolocation data received from mobile devices. Data on the average speed of traffic along the roads of the Irkutsk region were obtained. The next stage of the study is to determine the transport availability of the environmental objects of the Baikal region on the basis of the collected information.

## References

1. Lane, N.D., Miluzzo, E., Lu, H., et al. (2010). A survey of mobile phone sensing. Communications Magazine, IEEE, Vol. 48, No. 9. P. 140-150.
2. Namiot, D., Manfred, S.-S. *On Open Source Mobile Sensing*. Internet of Things, Smart Spaces, and Next Generation Networks and Systems. Springer International Publishing, 2014. P. 82-94.
3. Neznanov, I.V. Namiot, D.E. Control of transport routes via mobile phones. International Journal of Open Information Technologies. 2015. Vol. 8. P. 30-39 (in Russian)