



TRANSPORT AND TELECOMMUNICATION INSTITUTE

VICTOR KREBS

**RESEARCH ON LOCALIZATION METHODS OF
TRANSPORTATION NETWORKS OBJECTS USING SPATIAL
DATABASES**

Summary of promotion work

to obtain the scientific degree of Doctor of Science in Engineering (Dr.sc.ing.)
Scientific area “Transport and Communications”, scientific subarea “Telematics
and Logistics”

**Scientific supervisor:
Dr.sc.ing., B. Tsilker**

RIGA – 2015

UDK 656:004

K 82

Transport and Telecommunication Institute

Krebs V.

Research on localization methods of transportation networks objects using spatial databases: Summary of the promotion work. Riga: Transport and Telecommunication Institute, 2015.

In the submission kept the author's style and design.

ISBN 978-9984-818-72-6

© V. Krebs, 2015

© Transport and Telecommunication Institute, 2015

**THE PROMOTION WORK PRESENTED TO THE
TRANSPORT AND TELECOMMUNICATION INSTITUTE
TO OBTAIN THE SCIENTIFIC DEGREE –
DOCTOR OF SCIENCE IN ENGINEERING (Dr.sc.ing.)**

OFFICIAL REVIEWERS:

_____,
_____,
_____.

_____,
_____,
_____.

The defence of the thesis will be held at _____ o'clock at the special Promotion Council of Transport and Telecommunication Institute on award of a doctor's degree at 1, Lomonosova Street, Conference Hall 130, Riga, Latvia, tel. +371 67100594, fax: +371 67100535.

CONFIRMATION

I hereby confirm that I have developed the promotion work that is presented to the promotion Council of Transport and Telecommunication Institute to obtain the scientific degree of Doctor of Science in Engineering. The promotion work has not ever been presented to any other promotional council to obtain the scientific degree.

_____, 2015

V. Krebs

The promotion work is written in English, it contains an introduction, 5 chapters, conclusions, 52 figures, 15 tables, 120 pages, and 3 appendixes. Bibliography contains 151 sources.



This work has been partly supported by the European Social Fund. Project Nr. 2009/0159/1DP/1.1.2.1.2/09/IPIA/VIAA/006 (The support in Realisation of the Doctoral Programme “Telematics and Logistics” of the Transport and Telecommunication Institute).

CONTENTS

ANNOTATION.....	6
1. THE RELEVANCE OF THE PROBLEM AND MOTIVATION OF THE RESEARCH.....	7
2. THE METHODOLOGY AND METHODS OF INVESTIGATION.....	8
3. THE OBJECT AND SUBJECT OF RESEARCH.....	10
4. RESEARCH GOALS AND OBJECTIVES.....	10
5. HYPOTHESES OF THE STUDY.....	11
6. THE METHODOLOGY AND METHODS OF INVESTIGATION.....	11
7. SCIENTIFIC NOVELTY.....	11
8. THE PRACTICAL VALUE AND IMPLEMENTATION.....	12
9. THE THESES SUBMITTED FOR DEFENSE.....	12
10. APPROBATION OF THE RESEARCH.....	13
11. THESIS STRUCTURE.....	13
12. SUMMARY OF WORK CHAPTERS.....	14
Chapter 1. (Localization of objects in intelligent transportation systems and related tasks).....	14
1.1 Intelligent Transport Systems and Telematics.....	14
1.2 Transport Ad-hoc networks, VANET, V2V and I2V.....	14
1.3 Location based services and localization.....	15
1.4 Localizability.....	16
1.5 Coverage estimation in VANET networks.....	16
1.6 LBS categories.....	19
Chapter 2. (Localization methods).....	19
2.1 Localization taxonomy.....	19
2.2 Range-free localization.....	21
2.3 Range-based localization.....	22
Chapter 3. (Analysis of relative position and characteristics of reference points on the object localization error in VANET networks).....	25
3.1 Graph vertices coordinates error characteristics.....	25
3.2 Trilateration error modeling.....	26
3.3 Measurement errors comparison.....	26
3.4 Conclusions.....	28
Chapter 4. (Scene analysis algorithm).....	28
4.1 Graph embedding.....	28
4.2 Localization as graph embedding, problem formulation.....	29
4.3 Localization with additional information resources.....	30
4.4 Embedding algorithm.....	31
4.5 Simulation results.....	35
Chapter 5. (Spatial queries over sensor networks).....	35
5.1 Spatial queries.....	35
5.2 Proposed method.....	37

5.3 Simulation environment.....	38
5.4 Spatial query profiling	39
5.6 Experiment data	41
5.7 Results.....	44
CONCLUSION.....	45
PUBLICATIONS WITH AUTHOR'S PARTICIPATION	48

ANNOTATION

The thesis of Viktors Krebs (Victor Krebs) “Research on localization methods of transportation networks objects using spatial databases”. The scientific supervisor is Dr.sc.ing., Boris Tsilker.

The main goal of the research is to explore the possibility of using alternative information resources in localization tasks, providing decision-making support for cooperative transportation systems, in cases when classical localization methods are not applicable.

Relevance of the topic is explained by the increasing need to improve the accuracy and availability of the wireless networks nodes localization methods in transportation, as well as the need for further integration of information systems of vehicles and road infrastructure.

The research considers existing localization methods, their characteristics, applicability and limitations imposed by the use of wireless networks in real traffic. Particular attention is paid to the scene analysis methods based on measured distances to various landmarks and knowledge of the measurement. Provided experimental data showing the influence of the different transportation situation scenarios on the resulting distribution of errors in the coordinates of the localized object. Especially noted situations where localization becomes very difficult or impossible at all.

Author formulates requirements for transport systems cooperative localization methods and shows a possible way to implement them, taking in account scenarios when the classical methods are not applicable.

As a result, author proposes a new original approach to object localization, relying on additional information resources. To accomplish this, new proposed localization method based on the embedding of the graph distance between the nodes of a wireless network on the plane with additional, alternative information resources as a set of constraints that exclude alternative graph realizations.

Proposed method has been realized both as “pure” algorithmic implementation, and as a compact, optimized implementation in the form of spatial queries to the Geographic Information System. A simulation-spatial analytical model has been developed to allow investigate and validate the proposed method. Implementation of the method is studied at different, close to the actual embedding scenarios. Data, collected within experiment cycles, has been processed, analyzed and evaluated.

Main results of the thesis presented are universal in nature and shows that the proposed method can be widely used as a source of missing information for other localization methods, or as the source of information for the period of time when the classical methods of localization are unavailable.

1. THE RELEVANCE OF THE PROBLEM AND MOTIVATION OF THE RESEARCH

Ensuring the ever-increasing requirements for sustainability and competitiveness of European mobile transportation systems creates the need for certain innovations. The most promising type of such innovations are Intelligent Transport Systems (ITS) that combine the most advanced solutions in the field of information and communication technologies. The potential of ITS capable of accomplishing a wide range of tasks related to the transport sector in all its diversity. Suffice it to mention the electronic payment system for transport, dynamic traffic management, including adaptive speed limits, parking assistance, support for navigation applications in real-time, operational information transmission and processing for stability control and collision avoidance systems. This goes far beyond the range of tasks assigned to ITS. In addition to the marked, ITS designed to improve the interaction between different transport modes in multimodal transport systems.

It is obvious that to achieve these goals one of the main requirements for ITS is the high degree of communication development that allows to maintain intensive information flows specific operational decisions support.

Significant expansion of the range of applications covered by the ETS, and reducing development costs can be achieved integrating information systems of vehicles and road infrastructure. Open information systems, united in one standard platform can ensure compatibility and interoperability between the infrastructure systems and equipment. This should accelerate the development of cooperative systems, based on the exchange of information between vehicles and between vehicles and the road infrastructure.

In this regard, it should be noted separately two interrelated tasks that require new and more effective solutions. The first of these is the development of new cooperative systems and their evaluation from the standpoint of the possibility of establishing a unified approach to the implementation strategy in ITS. And the other, related to first one - definition of specifications and standards for communication in cooperative systems such as “infrastructure to infrastructure” (I2I), “vehicle to infrastructure” (V2I), “vehicle 2 vehicle” (V2V), taking into account that the objects in both cases are regarded as the wireless network.

Until now issues related to capacity and effectiveness of applications aimed at improving traffic safety (communication type V2V) and similar issues related to applications that focus on optimizing traffic flow (communication type V2I) remains open.

One of the central and indispensable elements in the majority of cooperative ITS applications is localization (position estimate) of vehicles. Any measurement data lose their value without knowledge of the exact coordinates of the place where they were received. Generally, localization methods are used in wireless networks to determine the coordinates of nodes using the a priori known coordinates of the individual components and the results of some measurements, such as distances, angles, and other connectivity. It should be noted that the methods of localization in wireless networks is not a trivial GPS methods extension or methods based on radars.

Localization of objects in the ITS-based wireless networks have a number of features, such as a variety of types of measurements, work in difficult conditions, where it is impossible to provide line of sight between nodes, limited communication and computing capabilities of nodes and dynamically changing external environment and network configuration. However, before the actual localization must determine whether the node is localizable, i.e., whether there can be found a unique coordinates showing the position of the node in the area, on the basis of available data.

Reliable information about the coordinates of the nodes of wireless networks in the ITS opens up new possibilities in the development of a rapidly growing segment of applications based on location (Location Based Services, LBS).

Despite the undeniable progress in the localization methods that are using together the data received from the GPS and built-in sensors of different types, still there are situations when the availability and accuracy of localization are not high enough. Accumulated errors in localization system based on motion sensors and the mirroring effect of a signal in an urban environment adds an error component unacceptable for many modern applications what requires further search for opportunities to improve the accuracy and reliability of the localization methods.

Significant improvement of the results of localization could be achieved employing digital maps of the area and opportunities for visual recognition of landmarks. The problem here is that the segmentation and recognition of landmarks is still extremely difficult and inadequately costly process. In this case seems promising multilevel data summarizing schemes, which uses the data obtained from different independent sources. However, the creation of such systems highlights the new challenges associated with synchronization, resource management, and the need for high computing power.

In practice, there may be situations where none of the wireless nodes has no information about its coordinates. For some applications, knowledge of absolute position is not critical, and can be limited by the relative coordinates of the nodes in the network. Moreover, improved methods of relative localization may provide a means to improve the methods of absolute localization. In this case, the assignment of all network nodes virtual coordinates in Euclidean space is a special case to solve the problem of graph realization in the plane, which is a daunting task requiring a new, more effective decisions.

In recent years a number of studies were devoted to the localization of wireless networks in order to find new, efficient and stable algorithms that take into account the mobility of the nodes, as well as a shortage and unreliability of information resources. This indicates that the problem of localization remains an open challenge requiring continuing search for solutions. This fact was the motivation for the present work aimed at studying the possibility of using new, additional sources of information for locating objects in the cooperative transport systems, in cases where classical methods are not applicable.

2. THE METHODOLOGY AND METHODS OF INVESTIGATION

In recent years, researches in the field of cooperative systems have focused on improving the accuracy of localization as a critical problem. Attention is paid to this problem in the European Action Plan to support the deployment of ITS, Roadmap to a Single European Transport Area (2011). It is particularly important to have an accurate real-time data in collision avoidance systems, navigation systems, and control systems. It is assumed that the required improvements can be achieved by multi-modal localization and by combining different types of measurements from different data sources. Despite the fact that the development of such methods has achieved a certain progress, there are still situations where localization accuracy decreases to unacceptable limit. Moreover, modern localization systems do not provide sufficient reliability and allow a situation where localization becomes impossible.

As an example can be mentioned differential GPS (DGPS) and the assisted GPS (A-GPS) - two of the most advanced technology GPS, which allow highly accurate position and high speed determining the object. Nevertheless, the use of the GPS receiver as the only source

of data for locating the vehicle may become unreliable, particularly in densely built urban areas where the satellite signal is distorted significantly. Literature Drawil & Basir (2010), Sliety (2007), Stockus (2000), Skog & Handel (2009) describes a number of solutions that offer information to supplement GPS measurements with dead reckoning, which allows improving the accuracy and reliability of localization.

Dead reckoning (DR) - a method of determining the next position of mobile object in a series of short time intervals, based on the known direction of travel, speed, and data on the previous position. The method is simple, but it leads to incremental accumulation of errors and requires frequent resetting at short intervals to perform correction. One method of correction is to use GPS measurements to eliminate the accumulated DR errors in the moments when measurements are available.

Another approach (scene analysis) involves the detection and recognition of targets on the ground. This allows a comparison of the DR and GPS data with data from spatial databases. In particular, Fuerstenberg & Weiss (2005), Jabbour, Bonnifait & Cherfaoui (2006), Jabbour, Cherfaoui & Bonnifait (2006), Rae & Basir (2007), Weiss et al. (2005) in their researches indicates that additional information can be obtained either by using a laser scanner for recognizing the pavement edge or a range finder which can measure the range to detected targets. Further, the data received from the DR, GPS and Geographic Information System (GIS) are joined and processed in order to provide more accurate information about the location of the object, even under conditions of significant GPS signal distortion. Given that the GIS data stores digitized roadmap, buildings plan and other landmarks, the stage of training for these systems is usually not required.

Since digital maps can provide a large amount of data about the environment, the visual data may also find use in the localization methods. However, as noticed Jabbour, Cherfaoui & Bonnifait (2006), due to the fact that the image processing requires extensive computational resources, GIS databases, as a rule, stores and recognizes only certain key images. Thus, at each moment, the uncertainty of individual data sources can be analyzed and compensated using the error covariance matrix that ensures the least mean square error values.

Another promising approach is the cooperative localization, assuming the presence of several vehicles connected by wireless means of communication. As an example, one can consider positioning system indoor "RADAR", Bahlandv & Padmanabhan (2000). In this case, the data necessary for localization may be supplemented not only by various types of measurements but also by measurements taken on different network nodes. This allows keeping the availability of nodes localization, either indoors or even underground that is, not having visibility of GPS satellites. In this situation, the network nodes that have their own location data are so-called "anchors" for the other nodes in the network.

Despite some success in developing methods of cooperative localization, a number of questions remains open. For example, the same set of measurements can allow multiple network configurations, thus existing localization methods can return incorrect data about the coordinates of the network nodes. In addition, most techniques of localization based on the optimization techniques, simply assigns to not localizable nodes coordinates local minimum. All this can lead to incorrect operation of the application, based on the coordinates of the nodes in the network

Using the scene analysis algorithms along with other methods of localization allows obtaining the location and eliminating erroneously detected topological network configuration. Especially effective is an iterative implementation of both these methods. In this case, the scene analysis algorithms are close to the pattern recognition problems, in which the observed physical characteristics are compared with the data and reference images available localization system.

Usually, the localization methods based on the scene analysis come down to two basic steps: collection of the area information and the position estimate based on the measured physical parameters. The two most common approaches in the evaluation of the coordinates are not parameterized statistical method of k-nearest neighbor (kNN) and probabilistic methods.

KNN method involves the received signal strength preliminary data collection and creation of the corresponding database, and then, in the process of localization, it involves the database search for closest data corresponding to observed data by the least-squares error. A probabilistic approach implies the possibility of the existence of n locations for the observed set of measurements, calculating the probability according to Bayes formula.

There is another approach successfully working in indoors environment developed for Bluetooth networks and presented by Forno, Malnati & Portelli (2007). The device sends queries to other network members with different power levels, and depending on the response of other mobile nodes, estimates a relative distance. This method requires at least one “anchor” node with known coordinates, and provides accurate localization up to 1.88 meters.

Important factors that affect the quality of the scene analysis localization methods are the density and location of the networks “anchors”. This especially concerns algorithms, which uses the measured distance between the mobile nodes. At the same time, there are additional sources of data that can significantly improve localization solutions. Since the advent of GIS, it has been presented many new methods of spatial analysis based on digital maps. GIS applications allow to store, modify and query both spatial and non-spatial attributes for a particular position.

Spatial queries to the GIS can be an important tool for managing wireless networks formed by vehicles. Such queries can be adapted for each specific task, thereby avoiding the need for changes in the network configuration. In essence, such a network should become a single agent distributed to the user, delivering data from the region of interests as needed.

Thus, localization algorithms can use these digital maps available through GIS, as a set of constraints and can be reduced to the placement of the graph given in the form of a matrix of known distances between members of the cooperative transport system, and that was the main idea of this study. It should be emphasized that the proposed localization methods in are not universal for any possible scenario, however, can successfully complement the classic methods of localization and replace them in an environment where, due to lack of information resources classical methods are not applicable.

3. THE OBJECT AND SUBJECT OF RESEARCH

The objects of research are methods of localization of mobile objects in transportation systems, organized based on wireless networks.

The subject of the study is to investigate the possibility of localization of mobile objects of transport systems in environments of the limited availability of spatial data.

4. RESEARCH GOALS AND OBJECTIVES

The goal of this work is to study the possibility of using alternative information resources for decision-making support in cases when classical localization methods of cooperative transport systems objects are not applicable.

To achieve this goal, the following objectives have been formulated:

1. Review and classification of localization methods used in modern cooperative transportation systems;
2. Identify the main problems associated with localization in cooperative transportation

- systems, and formulate the factors that contribute to the lack of information resources;
3. To analyze the applicability of the existing classical methods and their combinations in scenarios with limited access to information resources;
 4. To propose and develop an alternative method for localization based on spatial queries to geographic information systems, used as an additional source of data;
 5. Build a model based on spatial database, which provides supports digital maps fragments, placement of wireless nodes at random coordinates with the given parameters, the ability to set the conditions of the line of sight between nodes, specifying the coordinates of “anchors”, as well as support for spatial queries and indexing spatial data;
 6. Carry out numerical experiments for different localization scenarios and check the validity of propositional method. Visualize the results and analyze them;
 7. Present the evidence-based recommendations for the use of the method to decision-making support in cases when classical localization methods of cooperative transportation systems are not applicable.

5. HYPOTHESES OF THE STUDY

The study presents the following hypothesis:

1. Placement of the graph distance between nodes of the cooperative transportation system with additional geographic information resources can serve as a means of localization in conditions where other methods are inapplicable;
2. The results of the localization obtained by the proposed method can be used as the source of the missing information for the classical methods of localization or as a temporary source of information for the period of inoperability of the classical localization methods.

6. THE METHODOLOGY AND METHODS OF INVESTIGATION

To achieve the goals set in this research, were used methods of computational geometry, probability theory and graph theory, as well as the theory of simulation systems and combined analytical and simulation. The research is based on materials of scientific thematic publications, collections of articles of international conferences, the European Commission reports, the results of scientific research papers and monographs.

To test the hypotheses and the proposed methods numerical experiments were carried out. Appropriate software was created based on spatial database Oracle Spatial 11g and software platform Java. Calculations and visualization carried out in a Microsoft Office Excel 2010. The data obtained were used in several cycles of research, in particular evaluation of relevance, context research for the development of the following requirements to the application, the definition of eligibility criteria and evaluation of the results, according to the procedure described in A. Hevner, S. Chatterjee (2010).

7. SCIENTIFIC NOVELTY

The scientific results obtained in the thesis are as follows:

1. Proposed a GPS independent approach to the localization of the nodes of wireless networks in transportation systems using alternative information resources;

2. Developed and implemented the algorithm for GPS independent localization task using a matrix of measured distances between objects and vector digital map of the area as an additional set of constraints to exclude alternative object placements;
3. Proposed a compact and effective implementation of the method using spatial queries to the database of indexed geolocation data. Implemented analytical and simulation models;
4. Performed validation of the model, investigated the scope of the method and possible scenarios for its use.

8. THE PRACTICAL VALUE AND IMPLEMENTATION

The results showed that the proposed localization method could be used:

1. As an independent method, in cases when for various reasons the classical methods based on the measured angles and distances to points with known coordinates cannot be applied;
2. To compensate temporary dysfunction of other localization methods due to external interference and other obstacles to obtaining the necessary measurements;
3. For data correction in localization systems that utilizes integration of data obtained from several sources (data fusion).

Specified the opportunities of method expansion for three-dimensional localization support and compatibility with any type of digital maps, supporting vector representation of the data.

Developed models implemented as a set of programs on Oracle Spatial platform 11g, allows simulating different scenarios to investigate the localization and the effectiveness of the method in various areas of transport infrastructure configurations.

9. THE THESES SUBMITTED FOR DEFENSE

Within the research, following theses has been formulated:

1. Localization of objects in the wireless transportation network using GPS devices as a single data source can become not reliable enough, especially in the densely built-up city. Using geolocation context and integration with other data sources can substantially compensate this disadvantage;
2. There is a need to develop new and improve existing centralized localization algorithms that are resistant to lack of information resources. Algorithms should provide absolute and relative location estimation and remain functional with both static and mobile objects. As the basis of these algorithms is advisable to use scene analysis methods;
3. If you have a matrix of the measured distances between objects wireless vehicles network, real coordinates of at least one of the objects of the network or transportation infrastructure and digital maps, localization task can be reduced to the distances graph placement on the plane, given the constraints of the digital map. Moreover, the matrix range measurements may not be complete;
4. Placement of distance graphs in the plane can be performed algorithmically, based on the Euclidean coordinates propagation, limiting the alternative placements of non-rigid graph fragments by digital map vector data;
5. The proposed localization method can be implemented using spatial queries,

allowing to make it compact and to take advantage of the out of the box advanced technology of spatial indexing, and compatibility with existing digital geolocation data formats;

6. Employing proposed algorithm and additional data sources allows compensating the possibility of temporary GPS dysfunction, as well as, if necessary, correcting the errors encountered by other localization methods.

10. APPROBATION OF THE RESEARCH

The basic concepts and findings were reported and discussed at scientific conferences and seminars. The author has 9 publications on the subject of the thesis including indexed at peer reviewed scientific literature citation databases such as Index Copernicus, Google Scholar, EBSCO, SCOPUS, ERA, DOAJ and ProQuest Index. The research results have been used in the project COST IC0906 “Improving the accuracy of real-time positioning of moving objects in mines”.

11. THESIS STRUCTURE

The thesis consists of an introduction, five chapters, conclusion, bibliography, and three appendixes. The paper contains 120 pages, and includes figure 52 and 15 tables. Bibliography includes 151 references. The structure of the work is as follows.

Introduction devoted to examine the relevance of the thesis research; there are formulated goals and objectives of the study, described scientific novelty and practical value of the results.

First chapter discusses three important problems in the development of Location Based Services for intelligent transport systems: localization of wireless network objects, localizability and network coverage evaluation. Conclusions are made that, despite of significant progress in the localization methods development, based on the combined use of different sensors and GPS, still there are situations in which these tools and methods are ineffective or do not apply at all.

Second chapter describes the classification of localization methods and the types of measurements required for the implementation of the algorithms. Closely reviewed algorithms that are using as an input conditional distance determined by the number of transit transitions between nodes. In addition, this chapter pays attention to the algorithms that are more accurate, requiring measurement of Euclidean distances, angles, time, signal propagation delay and other physical parameters of the environment based on optimization techniques, numerical methods and the methods of computational geometry.

Third chapter gives the characteristic of network nodes localization errors resulting from inaccuracies in the measurement of distances to reference points. The influence of the relative placement of reference points on the resulting error localization. The created model and a number of experiments carried out to investigate the properties of the probability distribution of the resulting error localization depending on the probability distributions of the error distance to reference points. Tested behavior with not only a conventional normal distribution but also with the exponential distribution, and some combinations of different types of error distributions.

Fourth chapter summarizes in practice the general and theoretical problems of wireless sensor networks localization task. Pointed out possible directions in solving such problems in cases where the classical methods of localization are not applicable. Proposed the conceptual idea of novel localization method involving additional information resources. Developed a model, used for a preliminary verification of the algorithm and its performance evaluation, as

well as computational complexity and possible deficiencies that requires further development.

Fifth chapter shows a possible implementation of the localization algorithm for wireless networks with additional information resources. Closely reviewed and defined features of the implementation and customization of software solutions based on spatial queries. Chapter shows the model building principles, which enables different localization scenarios simulation. Specific scenarios and parameters have been selected for the study and analysis of the algorithm in this particular implementation. Conducted a series of experiments, collected and processed data, more accurately characterized the range of arising problems and opportunities for real life application of the proposed algorithm. Usage constrains of such application also considered.

Conclusion presents an overview of the most significant results of the thesis, discussion and recommendations on opportunities to further develop the proposed solutions.

Appendix contains a list of publications of the author, the key specification of spatial functions and data structures used in the implementation of the localization method using spatial queries and the main classes of localization applications based on spatial queries.

12. SUMMARY OF WORK CHAPTERS

Chapter 1. (Localization of objects in intelligent transportation systems and related tasks)

1.1 Intelligent Transport Systems and Telematics

Telematics, as an interdisciplinary domain of study Computer Science and Technology, includes telecommunications, transportation technology, safety issues, the provision of information services, and related engineering tools. Of particular note are applied problems of telematics in the transport sector, especially the problem of positioning and tracking of vehicles using global navigation systems and mobile networks, as well as wireless communication between vehicles to exchange information on the location, speed, dangerous situations on the road, etc. Such communication involves organizing temporary Ad-hoc local wireless networks and should be based on wireless devices and sensors, installed in vehicles.

Cooperative communication on the road includes the mutual communication between vehicles and between vehicles and infrastructure. Data received from vehicles by the infrastructure devices usually sent to the centralized server for joint processing.

1.2 Transport Ad-hoc networks, VANET, V2V and I2V

The nature of problems solved in transport systems involves the use of decentralized self-organizing wireless networks whose nodes can independently move in all directions, closing and establishing connections with neighboring nodes. Communication type Vehicle-to-Vehicle (V2V), provided by such networks, allow dynamic and anonymous data exchange between neighboring vehicles.

Cars can serve as sources of information, its consumers and, if necessary as data relaying entities. The necessary information can be obtained from the preset infrastructure devices that via communication of type Vehicle-to-Infrastructure (V2I) extends the capabilities of applications far beyond conventional embedded sensors.

However, despite the considerable amount of research work in the field of VANET and ITS applications, practical implementation faces a number of challenges that have only partial solutions. Among these problems is the lack of a unified technology platform capable of supporting all the necessary applications.

Solving these problems requires standardization with mandatory involvement of the automotive industry. It is necessary to adopt standards that are aligned with technical solutions that have been successfully implemented and used by manufacturers. The use of a single standard for transmitting and receiving equipment, the specification of the level of transport communication (Fig. 1) in the model of OSI (Open Systems Interconnection) allow solving such problems as geographical addressing, multiple message routing, security and management of communication channels.

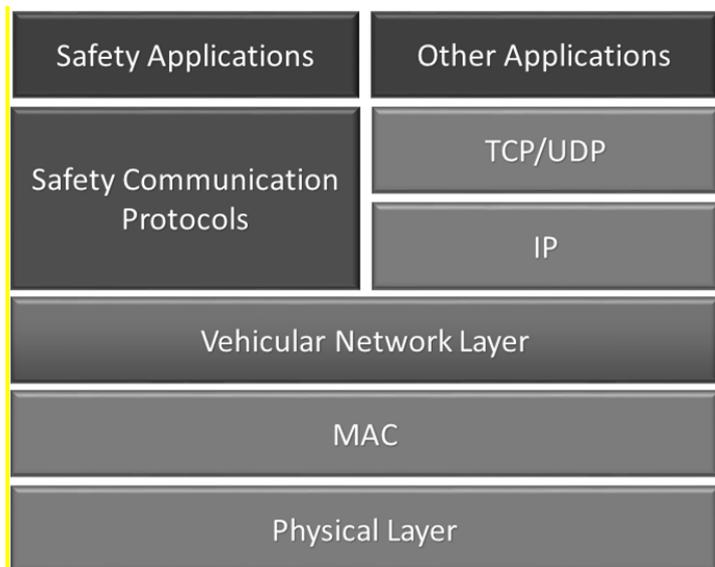


Fig. 1. Transport communication in the OSI model

1.3 Location based services and localization

Key and becoming increasingly common technology platform for the implementation of applications in vehicles Ad-hoc networks are Location-based services (LBS), services based on knowledge of the location. Supporting a wide range of communication types, from short-range Bluetooth, to global satellite technology (Fig. 2), LBS are designed to provide a basic understanding of the circumstances and the spatial context of the problem being solved by application.

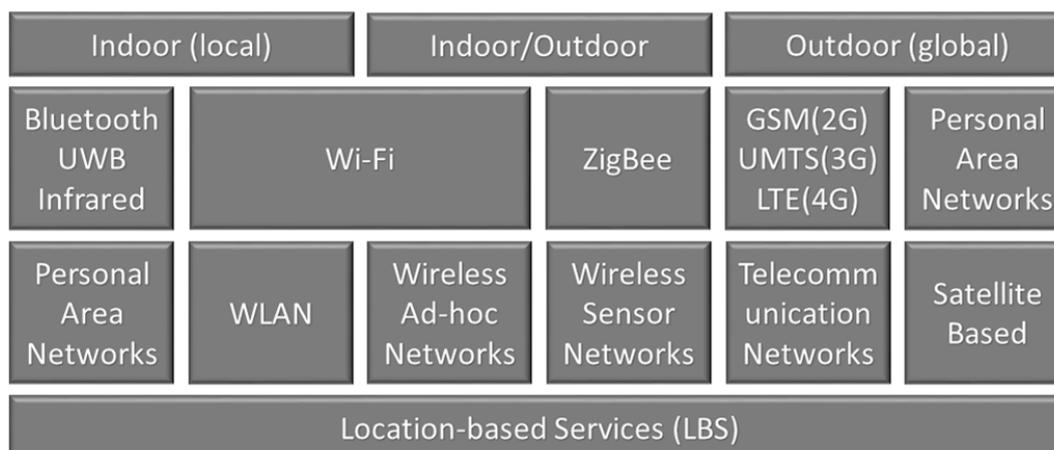


Fig. 2. LBS technologies

In simple terms, the fundamental problem of localization can be formulated as the problem of estimating the coordinates of a point in two-dimensional or three-dimensional space coordinate system formed by some other reference geometry data.

Modern localization systems still relies on the basic data such as distances and angles,

which may be obtained through indirect measurement by the related parameters, such as signal strength, signal arrival time, signal arrival time difference, phase difference and other data. New technologies allowed not only to obtain more accurate results, but also to create new architectural localization systems solutions, such as cooperative systems and data fusion systems.

Object localization methods in wireless networks are typically used for estimating the location of the sensor with an initially unknown coordinates with respect to some of the sensors, whose coordinates are known a priori and are available for the necessary inter sensor measurements. Sensors with a priori known coordinates are called anchors. The anchors coordinates can be obtained using global positioning systems or as a result of the installation of the anchor in geodetically certain point.

1.4 Localizability

Localizability of wireless network characterizes the possibility of unique localization in the chosen system of coordinates of all nodes in the network. Network localizability analysis is means that having a set of network nodes with known coordinates (anchors), a set of nodes whose coordinates are unknown, and the measured distances between a set of nodes, to determine whether it is possible to calculate the coordinates of all non-localized nodes. This fundamental algorithmic problem in the theory of wireless networks is the subject of a large number of recent studies. For example, Aspnes et al. (2007), Eren et al. (2004), Mao et al. (2007), So & Ye (2007).

Network with a given set of known coordinates of nodes and a set of known distances between nodes is considered uniquely localizable in the case when for the initial data there is a unique set of coordinates of all nodes in the network. Unique network localizability in two-dimensional space can be described by the graphs rigidity theory. Unique localizability depends only on the combinatorial properties of the network and is completely determined by the network graph distances and anchor nodes set. Edges of the graph in this case correspond to the distances, and vertices to the network nodes. Vertices of the graph are connected by an edge if the distance between the nodes of the network is explicitly known.

Based on the technique of measuring the distance, the actual location of the Ad-hoc wireless network can be modeled as a graph of distances $G = (V, E)$. Where a set of vertices V is a set of wireless communication devices in the network, and a set of edges E represents the set of distances between the communications devices (i, j) , in the case where the distance between points i and j may be measured, or both have known coordinates of the vertices and are anchors. For each edge, we can use the function $d(i, j): E \rightarrow R$ to get measured distance values between i and j .

Within this model, it is necessary to answer the question of whether the resulting graph distances representing the wireless network is localizable. What means that for each vertex v of the graph G , given the set of possible constraints H (known coordinates of the anchors vertices), there exists a unique location $p(v)$ such that $d(i, j) = \|p(i) - p(j)\|$ for all pairs (i, j) in the set of edges E , and respecting the set of constraints H , where $\| \cdot \|$ denotes the Euclidean distance on the two-dimensional plane.

1.5 Coverage estimation in VANET networks

In addition to the problems of tracking, location and location calculation, one of the fundamental problems in wireless VANET networks is the problem of network coverage

estimation. Despite the fact that the coverage estimation algorithms well studied, usually, they take into account only the uniform (isotropic) sensor model. In practice, the coverage range of the sensor in different directions often is substantially different. This means that the maximum allowable distance between the sensors is largely dependent on their orientation, and such a sensor network is called anisotropic.

Coverage estimation methods

1.5.1 Geometric method

In contrast to isotropic, directional sensor cannot cover the area over the entire circumference and has a certain sector of coverage. Therefore, coverage area can be considered as a sector on the two-dimensional plane. The following parameters completely characterize the sensor i , where: (x_i, y_i) coordinates of the sensor on the plane, θ - the maximum angle of coverage, R_s - the maximum sensitivity range of the sensor, beyond which the target detection is not possible and the vector $\vec{d}_{i,j}$, which divides the coverage vector.

If the object k , located in \vec{t}_k and is directed sensor i located in \vec{l}_i , then object presence in a coverage zone can be defined by an algorithm which is called presence in target sector test, Target In Sector (TIS).

For this it is necessary to find a distance vector $\vec{v}_{i,k}$, defining the distance from the sensor i to the object k .

$$\vec{v}_{i,k} = \vec{t}_k - \vec{l}_i \quad (1)$$

Next, it is necessary to determine whether the distance vector is covered by sensor i , using the following test:

$$\vec{d}_{i,j}^T \vec{v}_{i,k} \geq \|\vec{v}_{i,k}\|_2 \cos\left(\frac{\theta}{2}\right) \quad (2)$$

Then, check whether the sensor can detect the object:

$$\|\vec{v}_{i,k}\|_2 \leq R_s \quad (3)$$

In case when conditions 2 and 3 are met, the test result is true, and the sensor i covers the object k .

1.5.2 Voronoi diagram

The approach is based on the Voronoi tessellation, which is widely used to determine the coverage in isotropic sensor networks. The idea is that to introduce a model of the sensor in the form of an ellipse with a certain orientation, to simplify the construction and analysis of the chart, eventually reducing the problem to the isotropic model.

Anisotropic Voronoi region for sensor i in this case is defined as:

$$V_i^* = \left\{ q \in Q \mid \|q - p_i\|_{L_i} \leq \|q - p_j\|_{L_j}, \forall j \neq i \right\} \quad (4)$$

Where the point q belongs to the set of points in the region Q , and the distance expressed in non-Euclidean metric:

$$\|q - p_i\|_{L_i}^2 = (q - p_i)^T L_i (q - p_i) \quad (5)$$

Where the matrix L_i can be expanded as $L_i = F_i^T F_i$, where

$$F_i = \begin{bmatrix} \left(\frac{c}{a} & 0\right) \\ \left(0 & \frac{c}{b}\right) \end{bmatrix} \begin{pmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{pmatrix} \quad (6)$$

Where θ_i angle (orientation) i -sensor, and $(a, b, c) > 0$, the ellipse parameters.

In this case, the Voronoi regions of the polygons becomes figures described by the second order curves which leads to discontinuities in the regions and occurrence of regions without generators and with undetermined coverage. In order to bring the chart to normal view it is required to make another assumption - that all sensors are oriented in the same direction. This substantially reduces the legitimacy of the method

1.5.3 Probabilistic method

This approach does not determine the coverage area for each sensor, as is done in the Voronoi diagram. Instead, the calculated probability of object detection by the sensor. Moreover, several sensors total probability may be taken into account and, in case when the object is in the multiple sensors coverage area simultaneously.

The model depends not only on distance but also on the orientation of the sensor relative to the region of interest. Suppose each sensor has a limited coverage area Q_i with a maximum range R and the direction Θ . The possibility of covering each sensor is limited by radial distance and angle relative to the object that should fall into the coverage area.

Mathematically, the sensor i characteristics are dependent on the distance d_i , orientation α_i of sensor i with relatively to the target q . The sensor coverage area defines as follows:

$$Q_i = \{q \in Q \mid d_i \leq R \wedge |\alpha_i| \leq \Theta\} \quad (7)$$

Where:

$$d_i = \|q - s_i\|$$

$$\alpha_i = \cos^{-1} \left(\frac{(q - s_i)(\cos\theta_i \sin\theta_i)}{\|q - s_i\|} \right) \quad (8)$$

$$\Theta \in \left(0, \frac{\pi}{2}\right)$$

It is also necessary to make an assumption that the sensor i has coverage only in its region Q_i .

$$p_i(q) = 0, \frac{\partial p_i(q)}{\partial d_i(q)} = 0, \frac{\partial p_i(q)}{\partial \alpha_i(q)} = 0 \text{ if } q \notin Q_i \quad (9)$$

Thus, the probabilistic model of the sensor is:

$$p_i(q) = \begin{cases} \frac{(d_i - R)^2 (\alpha_i - \Theta)^2}{R^2 \Theta^2} & \text{if } q \in Q_i \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

As can be seen from the review, the geometric estimating method has serious shortcomings that limits its use. First, the need of detected object approximate coordinate knowledge in advance. Also, it is impossible to determine the probability of the object detection, since the method gives only a binary estimation - in the zone or outside the zone.

1.6 LBS categories

To formulate the requirements for the localization algorithms and problems caused by them it makes sense to classify them dividing localization data consuming applications into four types:

Traffic safety: applications that support traffic safety need the most accurate localization data, including association with digital maps. In this case, localization errors exceeding 5 meters are unacceptable. This is especially actual for lane identification problems, where the requirements for the error value are even higher and do not exceed 1 meter.

Mobility: accuracy requirements are less stringent, application allows localization error up to 8 meters. The problem is reduced to the identification of the road on which the vehicle resides, without specifying the lanes.

Efficiency: Allowable limit of error is within 10 meters and can be provided by a large number of vehicles involved in the information exchange. Unlike safety applications where the number of participants is limited both temporally and spatially, in efficiency and mobility applications large number of users can enable alternative localization results verification methods.

Telematics: telematics applications mainly operate by zones and the problem reduces to the identification of areas in which the vehicle resides. Allowable error may be of the order of 15 meters and more. On the other hand, these applications often may handle large amounts of data.

To sum up the above, the requirements for localization methods can be characterized as follows:

- Accuracy: defined as the closeness of the position estimate of the vehicle to its actual location;
- Availability: the number of vehicle location estimates, which can be obtained per time unit;
- Response time: the time required by the localization procedure to acquire position estimate;
- Integrity: the level of trust that can be applied to the vehicle position estimation results.

Chapter 2. (Localization methods)

2.1 Localization taxonomy

Localization algorithms for wireless networks can be classified by the structure of the network, the conditions of applicability and limitations in the use of various algorithms. In order to provide a clearer picture and better assess the possibilities when selecting localization algorithms, further we propose a classification that largely takes into account and generalizes approaches mentioned in J. Hightower, R. Want and G. Borriello (2000), M. Bylund and F. Espinoza (2001), Krebs V. and Tsilker B. (2013).

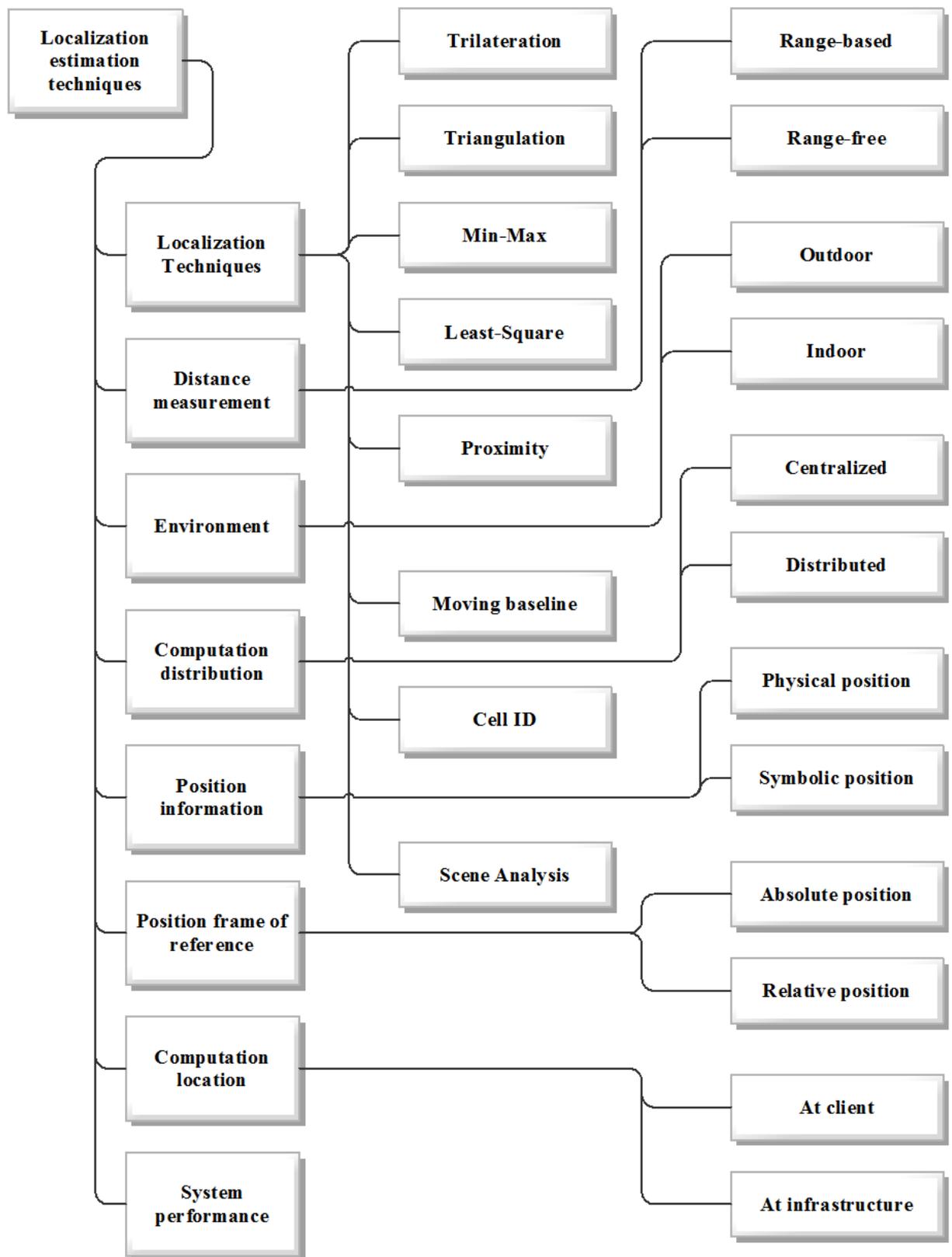


Fig. 3. Localization methods taxonomy

Scheme of the localization methods main taxonomic factors shown in Fig. 3.

Range free and range based methods: A significant part of the localization methods assumes a network node, with a precisely known location in a global coordinate system. Such a node called a beacon or an anchor, and such methods called techniques with a reference point. The essence of these methods can be formulated as the problem of finding a consistent set of nodes in the network locations given the available information about the relationship between nodes and between nodes and anchor. Methods without a reference point do not require specialized equipment or anchors to use data received from the outside of network. Instead of calculating the absolute nodes location, these methods allow to determine only the relative positions of the nodes in the local coordinate system with an arbitrary starting point.

Centralized and distributed localization methods: In the case of centralized algorithms, basic calculations are performed on a dedicated computer or a node located outside the network, while distributed algorithm computational load can be shared between all network nodes. This classification is directly connected with the formulation of the localization problem. If the problem is formulated as a global optimization, such as convex optimization, the node performing the calculations, should have all the data on the network distances and calculations must be performed centralized. Implementation of centralized algorithms, usually leads to collisions and concurrence in the calculations during the necessary communications network with a central hub.

2.2 Range-free localization

Depending on the object, carrying out measurement and location calculation, range free location methods can be divided into three main types:

- **Mobile device positioning:**
Device measures all the necessary data, approximates and calculates its own location relying only on this data;
- **Network positioning:**
Network measures all necessary parameters for localization, approximates and calculates the location of the device while the device remains passive throughout the process;
- **Hybrid positioning:**
Location calculation operation is distributed between the device and the network. As a rule, the device collects the necessary data for the algorithm, and the network, having a large computing power, makes the necessary calculations.

Next, let's consider some of the most common range free localization algorithms.

Sum-Dist: This method is described in A. Savvides, H. Park, M. Srivastava (2002) article “The bits and flops of the N-hop multilateration primitive for node localization problems” and is the most simple solution for estimating the distance from the selected node to anchor node. It consists of adding the distances of all the transitions between nodes during network traversal. Each node sends a message which includes a unique node ID, the local coordinates and the length of the path, which is initiated by zero in the first round. When a node receives a message, it adds the distance to the transmitting node to the length of the path and sends a message further. The advantage of this method is that it is extremely simple, and the algorithm is very fast. Moreover, the sheer number of computations is very small. On the other hand, there is an obvious drawback, which consists in the accumulation of errors in the distance estimation on each transition and distribution of errors over the network.

DV-Hop: Sum-dist algorithm drawback is that it collects and distributes the error as network

traversal. In large networks, as well as in networks with a large distance between the nodes, these errors can reach a considerable value. Instead of summing distances containing errors, more robust alternative provides method based on topological information for counting the number of transit passages. This approach is called DV-hop and described in Niculescu and Nath (2003), as well as the Hop-TERRAIN in Savarese et al. (2002).

Essentially, DV-hop consists of two waves of network traversal. After the first wave, and in Sum-dist algorithm, nodes receive data about its location and the minimum number of hops to the anchor nodes. The second wave gauge traversal uses known transitions number that allows calculating new adjusted node location. Refinement consists of multiplying the number of hops to the average length of the transition. Once the node specifies its location, it passes the calibration data to the following nodes.

Network positioning: The advantage of network positioning for the mobile wireless devices is that it requires no additional hardware for location measuring and calculating. Need for of such additional hardware significantly increases cost of portable wireless devices. The most common example of network positioning is Cell-ID method, which idea is to use the GSM cells with known geographic coordinates to calculate positions. Each GSM station has its own unique ID/Mac address. Thus, knowing the mobile station ID, it is possible to estimate the approximate location of the user.

In general, the accuracy Cell-ID method depends on the density of the mobile network in the area. In suburban areas, covering cell size can be up to 30 kilometers, while in the cities, the coverage may be considerably less, up to 10 meters.

2.3 Range-based localization

Virtually all existing today, numerous range-based techniques of localization are based on the below listed basic methods and algorithms.

Triangulation: Triangulation is the process of positioning, based on angular measurements relative to a known reference points. Historically, the triangulation is most often used in geodesy, as in this application area angles are easily and accurately measured by simple methods. Applications of this method in the context of wireless localization systems are considered in detail in D. Niculescu and B. Nath articles “Ad hoc positioning system (APS) using AOA” (2003) and “Error characteristics of ad hoc positioning systems (APS)” (2004).

Trilateration: Trilateration also is one of the most widely used positioning methods, which implementations in the ad-hoc systems discussed in detail in A. Savvides, C. Han, and M.B. Strivastava article “Dynamic fine-grained localization in ad-hoc networks of sensors” (2001). Trilateration in two-dimensional space requires the measurement of distances up to three non collinear reference points, or four non collinear reference points in three-dimensional space.

Multilateration: Multilateration is a navigation technique based on the measurement of the difference in distance to two or more reference points, which transmit signals at certain points in time. In contrast to the measurement of absolute distances and angles, the difference in the distance measurement can use arbitrarily large number of control points required for the desired accuracy. The difference of signal arrival time τ from n reference points to the reference point $n + 1$ makes it possible to calculate the coordinates in n dimensions (Fig. 4).

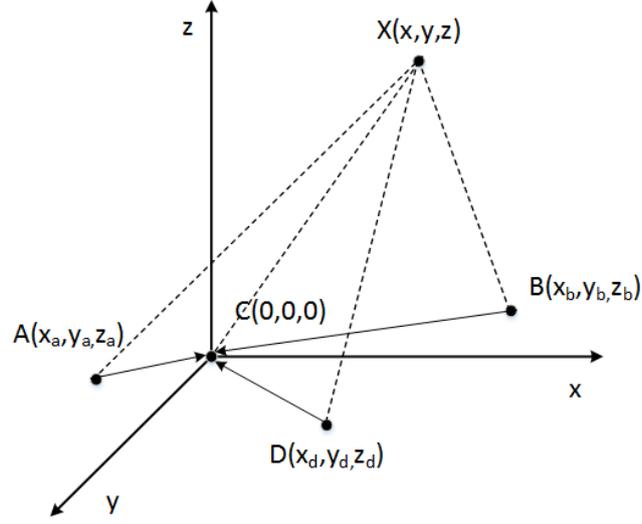


Fig. 4. Multilateration

$$\begin{aligned}
 \tau_a &= \frac{1}{c} \left(\sqrt{(x - x_a)^2 + (y - y_a)^2 + (z - z_a)^2} + D_{BC} - \sqrt{x^2 + y^2 + z^2} \right) \\
 \tau_b &= \frac{1}{c} \left(\sqrt{(x - x_b)^2 + (y - y_b)^2 + (z - z_b)^2} + D_{BC} - \sqrt{x^2 + y^2 + z^2} \right) \\
 \tau_c &= \frac{1}{c} \left(\sqrt{(x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2} + D_{DC} - \sqrt{x^2 + y^2 + z^2} \right)
 \end{aligned} \tag{11}$$

Where: τ time difference of signal arrival, c speed of light.

Min-Max: In some cases, a simplified method is applied where the position is taken at the intersection of the diagonals of the rectangle formed by the intersection of minimal bounding squares circles, centered at the reference points and a radius equal to the measured distance (Fig. 5).

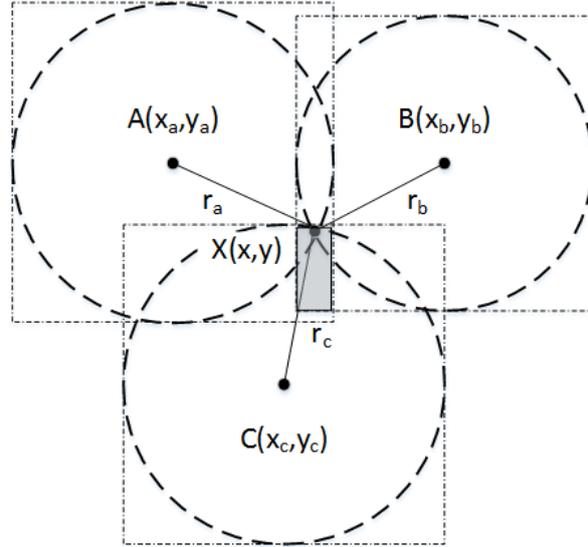


Fig. 5. Min-Max

$$[\max(x_i - r_i), \max(y_i - r_i)] \times [\min(x_i - r_i), \min(y_i - r_i)] \tag{12}$$

Wherein: (x, y) coordinates of localizable point, r_i radius of the circle. The possibilities of using this method in localization systems are considered by R. Parker and S. Valaee (2006) in the article ‘‘A Robust Min-Max Localization’’.

Clustering localization: A distinctive feature of the cluster localization method proposed by David Moore, John Leonard, Daniela Rus, and Seth Teller (2004), is that it remains functional even with a significant component of the error in the measurements.

Least squares method: Having information about the distances from the object X at unknown location to three reference points A , B and C with known coordinates it is possible to set up a system of equations defining the relationship between the coordinates X , reference points coordinates, and the distance to them. The resulting system of three or more equations with two unknowns: (x_x, x_y) . Since the number of unknowns exceeds the number of equations, the system is over defined and generally has no unique solution. Nevertheless, it has the least square solution.

$$\begin{aligned}(x_x - a_x)^2 + (x_y - a_y)^2 - a_r^2 &= a_\Delta^2 \\(x_x - b_x)^2 + (x_y - b_y)^2 - b_r^2 &= b_\Delta^2 \\(x_x - c_x)^2 + (x_y - c_y)^2 - c_r^2 &= c_\Delta^2\end{aligned}\tag{13}$$

Least squares solution is a unique solution (x_x, x_y) , which minimizes the sum of squared residuals $(a_\Delta^2 + b_\Delta^2 + c_\Delta^2 + \dots)$. The system of equations is nonlinear in the parameters of the solution (x_x, x_y) and to obtain the solution it requires the use of a non-linear least squares algorithm. More details of such algorithms are considered in the literature on numerical methods, e.g., Press et al., Numerical Recipes in C (1992).

Proximity factor: Localization method based on the proximity factor described by M.J. Feuerstein (2008) in “Urban and Indoor Location using Pattern Matching of Wireless Network Measurements”, can be formulated in terms of graph theory. A wireless network is represented as a graph $G(V, E)$. Subset of nodes $H \subset V$, know their own location $(p_1, p_2, p_3 \dots p_m)$. Proximity factor assessment usually is interpreted in two models:

- Adjacency matrix
- Distance Matrix

The task is to estimate locations $(s_1, s_2, s_3 \dots s_{n-m})$ of the V-H set of other network nodes.

Moving baseline: This localization method of wireless networks, as well as many others, is based on the fact that the network nodes are devices capable of measuring distances and to communicate with a subset of other network nodes. Method, based on these data, suggests to reconstruct globally consistent network topology, however, unlike the previous methods, implies that the nodes are in motion. Reconstructed nodes paths, using hyperbolic approximation, trilateration, subgraphs alignment and turning points recognition allow to solve these challenges and to evaluate the location of moving network objects.

Scene analysis: Localization method based on scene analysis suggests the possibility of monitoring some space around the object being localized. Typically, this is done using a set of sensors covering the required space and capable of perceiving visual, infrared, or other data. Mainly consist of two stages. The first step is to gather information about the environment, which localization will be performed, and development of an environment identification labels database. The most common tags used for data rates are received signal strength RSS (Received Signal Strength). The two most common methods of this type are the method of k-nearest neighbor (kNN), also known as the radio map, and the probabilistic method.

Chapter 3. (Analysis of relative position and characteristics of reference points on the object localization error in VANET networks)

An important part of the localization problem is to provide the most accurate origin estimation of the localized object in a difficult measurement conditions, noise, distortion or absence of the signal line of sight. Other factors affecting the accuracy of localization are reference point location scenarios used for coordinates calculations, as well as the parameters of the probability distribution and the nature of errors, both in the measured ranges and errors resulting localization with inaccurate data.

3.1 Graph vertices coordinates error characteristics

Sources of error in the localization system can be divided into three main classes, Savides A., Garber W., Adlakha S., Moses R., Srivastava B., (2003): placement errors, channel errors and algorithmic errors. Placement errors associated with both the initial configuration of the network, its density, size, and inaccuracies in the anchor nodes placements with a priori known coordinates. Channel errors determined by the properties of the measured physical parameters that allow calculating the distance between nodes of the network, and by properties of the communication channel, used for measured data exchange. The magnitude of this error component is largely dependent on used technology and the environment changes in which localization tasks are performed. The overall assessment of the channel component of the measurement error is extremely difficult and must be done individually for specific technological solutions.

In the solutions using the principle of sharing coordinates where each newly localized node can become a new reference point, measurement error component, usually one way or another is accumulated, what in turn makes estimation task even more difficult.

Considering trilateration algorithm as an example in, Zheng Yang, Yunhao Liu (2010) proposed to introduce a metric of trilateration quality: Quality of Trilateration (QoT) based on the observation of trilateration accuracy, depending on relative geometrical arrangement of reference points.

QoT metric describes the differences in the placement of control points numerically and allows selecting the optimum combination from the possible. This mechanism also allows marking those placing of anchor points, which will have a significant uncertainty of result and a higher probability of reflection.

Trilateration quality t defined as: $Q(t) = \int_p f_t(p)dp, p \in Disk(p, R)$, where $t = Tri(s, \{1,2,3\})$ means node s trilateration, based on three reference points s_i .

In this expression $f_t(p) = \prod_{i=1}^3 f_{s_i, s_j} (d(p, p(s_i)))$, the probability density function for any point p in the plane, where $p(s)$ means real node s location; $d(s_i, s_j)$ is the real distance between two adjacent nodes s_i and s_j ; $f_{s_i, s_j}(x)$ probability distribution ($x \in [0, +\infty]$ - the distance value); expression $Disk(p, R)$ defines the area of the disk with center p and radius R . The parameter R is specific to each application and accuracy requirements to localization.

Fig. 6, (a), (b) and (c) shows three trilateration scenarios. Assuming a normal probability distribution of measurement error, the first scenario (a) shows the most common type of scene configuration. The second scenario (b) shows an extreme case, when all three reference points are approximately on one line and there is a high probability of such errors as a mirror reflection of the localized node. The third scenario (c) shows the ideal localization conditions when the

anchor points are equidistant from a localized node and from each other.

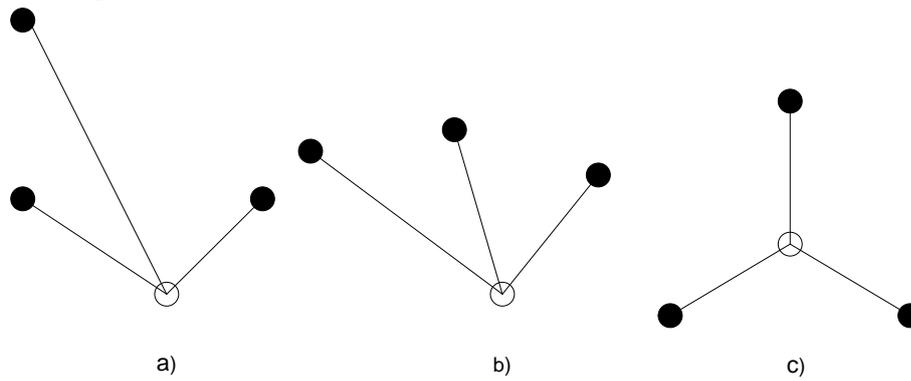


Fig. 3. Three sample scenarios of trilateration

3.2 Trilateration error modeling

To analyze and assess the impact of the reference point location scenarios against which the desired coordinates are calculated, as well as the parameters of the probability distribution and the nature of the error a series of experiments were planned and conducted.

Three previously considered localization scenarios were implemented. For comparative evaluation three algorithms were selected supporting localization on three reference points: cluster algorithm, min-max algorithm and Levenberg-Markarth implementation of least squares method. As a result of the first iteration of the simulation the best of the algorithms with the greatest precision was selected, also taking into consideration execution speed. Then, for each of the scenarios localization error was modeled using Gaussian probability distribution, and finally the combined distribution of errors and their impact on the resulting localization error evaluated.

Experiments showed that while the execution time of considered algorithms, regardless of the anchor location scenarios differs insignificantly. Clustering algorithm was for about five to ten percent slower than the other two. For this reason, execution time of the algorithm further will not be considered.

Least squares algorithm and clustering algorithm showed approximately the same location accuracy (within statistical error) for an ideal (A) and conventional (C) scenarios. However, for a bad scenario (B), a cluster algorithm was more stable and about 50% of the cases provided successful localization results, while the least squares method stable fell into a mirrored location situation.

Min-Max algorithm showed ten times less accuracy than the first two, and for this reason in the future will not be considered. Further, only the cluster algorithm will be consider as more accurate and, at the same time, in contrast to the method of least squares stable with respect to bad scenario.

3.3 Measurement errors comparison

Independent error estimates of each of the x and y coordinates have the same distribution law and the same parameters as the distance measurement error (Fig. 7).

Localization error (distance between the actual point and the resulting localization pint) correspond to the Rayleigh distribution (Fig. 8), which is consistent with the definition of “If X and Y are independent Gaussian random variables with zero mean and equal variances, the total distribution of the random variable is Rayleigh: $= \sqrt{x^2 + y^2}$ ”.

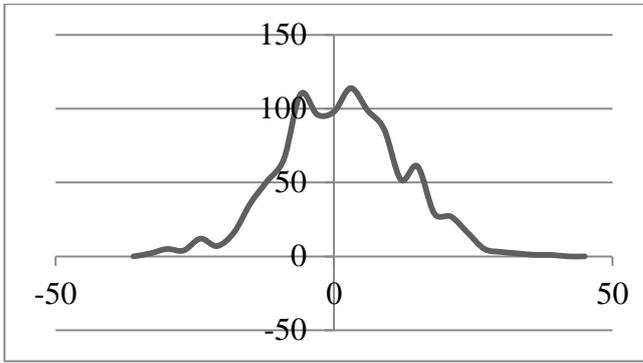


Fig. 7. Independent error distribution by x u y coordinates for scenario A

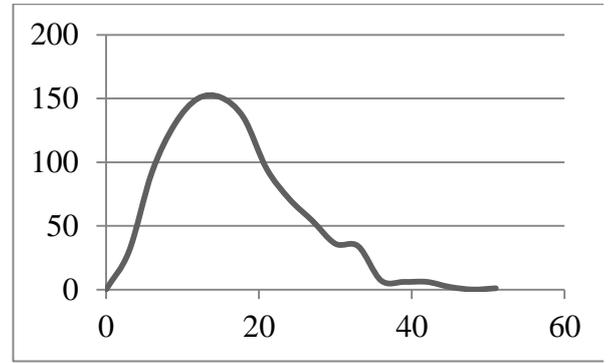


Fig. 8. Resulting error distribution for scenario A

In cases of non-ideal scenario localization error probability distribution of each of the coordinates x , y (Fig. 9 and 10) can vary significantly, while, as the resultant probability distribution of error localization (Fig. 11) may remain constant.

Since the distribution of localization error for each of the coordinates is “mixed” distribution of arithmetic sums of the errors distributions of measured distances, for better understanding the components of the overall localization error separate experiments were carried out to localize with distance measurement errors distributed both exponentially and with Gaussian distribution having the mean different from zero.

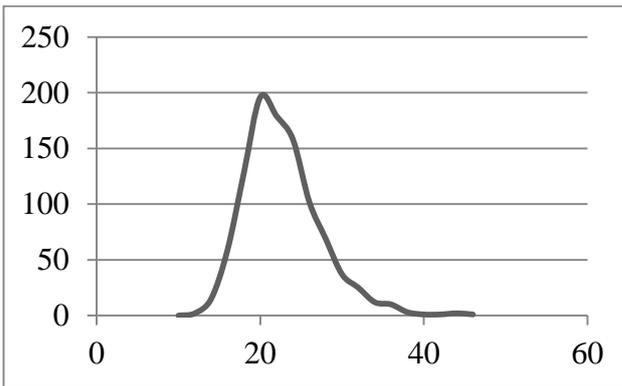


Fig. 9. Error distribution by x coordinate for scenario C

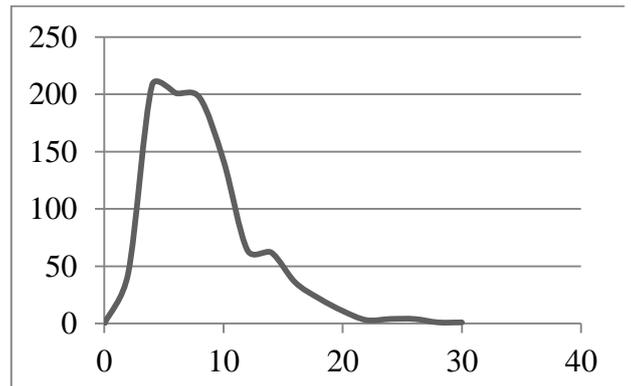


Fig. 10. Error distribution by y coordinate for scenario C

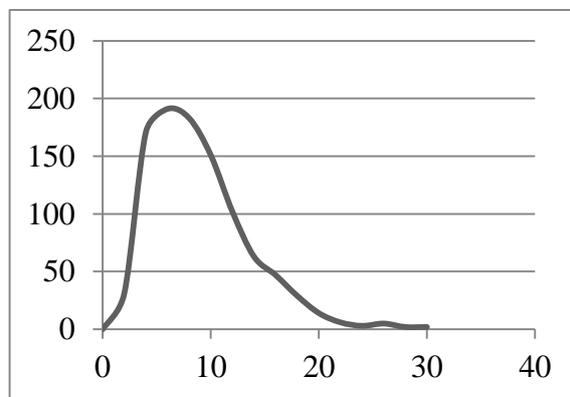


Fig. 11. Resulting error distribution for scenario C

Experiments showed that in this case the total localization error always distributed according to Reiss, while the distribution of errors for each coordinate separately, depending on the scenario can be either Gaussian distribution (in the ideal scenario) or Rayleigh distribution

(in bad scenarios). A similar pattern observed in case when distance measurement error distributed by Rayleigh.

3.4 Conclusions

Errors distribution in the case of localized coordinate with Gaussian distribution has the same parameters as the distribution of the distance measurement error. Other measurement error distributions, in localized coordinates, depending on the scenario, become either Gaussian distribution or a Rayleigh distribution and are summed. The total localization error (the distance between the calculated and the actual location of the object) has a lower value for the study because it does not indicate the direction (vector) localization of error, and has a Rayleigh distribution or Rice depending on the scenario in the case of the a non-zero expectation.

Simulation has shown that superposition of random errors in the localization process has the same nature as the superposition of harmonic oscillations with a random phase. It is well studied and has a developed mathematical description, what, if necessary, can be used in tasks of the distribution of localization errors analysis.

Chapter 4. (Scene analysis algorithm)

4.1 Graph embedding

Another approach to the problem of localization is to consider it as similar to the problem of graphs embedding. Obviously, for the nodes in the proximity of the reference point, the distance to this point can be estimated by direct measurement. Using certain methods of coordinate propagation distance to the reference point can be estimated for the second level of the neighbors, separated by two hops. Thus, by traversing entire network can be examined. If the graph is sufficiently connected and the lengths of its edges are known, it could be reconstructed on the plane.

Thus, the problem of localization can be viewed as the problem of reconstructing a set of locations of sensors based on a variety of known distances between sensors, being within a certain radius around each sensor. In this case, some of the sensors can be anchors, but for the placement in local coordinate system it is not so important. At its core, the problem can be formulated as the problem of determining whether a given graph can be physically implemented in two-dimensional space.

In contrast to the classical methods of localization, suggesting that the network must contain a significant number of nodes anchors, embedding a graph with the coordinate propagation does not need anchors for the local coordinate systems and therefore requires much less anchors for binding to the global system. All nodes can be initialized with random coordinates, which as a result of the algorithm execution and data exchange between nodes must match with coordinates that do not contradict the known distances.

The resultant coordinates will have certain degree of freedom in terms of orientation and translation, but the scale will be respected. Subsequent processing may include absolute positioning with reference to the three or four reference points, which makes it possible to eliminate the additional degrees of freedom. As possible set of constraints that reduces the number of required reference points, additional geolocation information can be used.

Formally, the problem of placement of the graph can be described as follows: placement of G on the surface Σ is a projection of G on Σ , where point on the surface Σ associated with the vertices of the graph G , but simple arcs of surface Σ associated with edges of the graph G , moreover:

- The end point of the arc associated with edge e , are the points associated with the vertices of e ;
- Neither arc does not include the points associated with other nodes;
- Two arc never intersect at a point that is not a final for both arches;

In other words, placing a graph is such its location on the surface, in which the edges of the graph intersect only at their endpoints.

Another problem is closely connected with the problem of graphs embedding and networks localization. This is the problem of determining the rigidity of the graph. In the context of the graph G realization task it is considered as a function mapping vertices of G at points in two-dimensional Euclidean space. In general, the implementation of the graph is considered to be admissible if constraints are met in form of distances between pairs of vertices i and j , where the edge $(i, j) \in E$. This, as mentioned above, means that the $d(i, j) = \|(p(i) - p(j))\|$ for all $(i, j) \in E$. Two realization of G are equivalent if they are identical for all symmetry operations: reflections, translations and rotation. Distances graph G should have at least one consistent realization that corresponds to the actual location of the topological network. This implies that G is connected and has at least four vertices.

Graph is generically rigid if it cannot be deformed continuously in any of its realization, without violating distances constraints. Graph is generically globally rigid if in translation, reflection, rotation has only one unique realization. Realization of graph considered common in the case where the coordinates of its vertices are algebraically independent. Here and in further on we will only talk about the common realization of graphs.

Summarizing conditions precluding uncertainty in the placement of graphs, Jackson and Jordan proposed the theory of necessary and sufficient conditions for the global rigidity of graphs.

Theorem 1. Graph with $n \geq 4$ vertices is globally rigid in two dimensions if and only if it is three connected and redundantly rigid.

Based on Theorem 1, the property of the global rigidity of the graph can be tested by polynomial time using Pebble algorithm, Nissanka B. Priyantha, Hari Balakrishnan, Erik Demaine, Seth Teller (2003) and network traversal algorithms.

4.2 Localization as graph embedding, problem formulation

Assume that N nodes labeled $1, \dots, N$, are located in unique locations in a physical region. Also, assume that there is some mechanism, which allows each node to discover neighboring nodes by using communications with these nodes. This mechanism in addition allows estimating the distance to the neighbor nodes. Each discovered neighboring node gives one undirected edge $e = (i, j)$ for a graph G , that describes known nodes.

Thus, given the graph consisting of n vertices $G(V = \{1, \dots, n\}, E)$, and its Euclidean length $l_{i,j}$ for each edge $(i, j) \in E$. Denote the two-placement arrangement of the graph as $x, y \in \mathbb{R}^n \wedge n$, where the coordinates of vertex i are denoted as $p_i = (x_i, y_i)$, and the distance between the vertices i and j as

$$d_{ij} = \|p_i - p_j\| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (14)$$

Not yet taking into account the measurement error, we can say that there is a physical placement of network nodes, which implements given length of edges of the distances graph $d_{i,j} = l_{i,j}$. The aim is to reproduce this placement, and, placement corresponding to these requirements may not be unique. Not unique placement is acceptable within such operations as rotation, translation and reflection, however, the measured distance must clearly fix placement scale.

For some graphs configurations placement can be non-unique, and have the possibility of transformation, other than rotation, translation and reflection.

Theorem 2. Assume that N network located in the coordinate system \mathbb{R}^d , where the number of dimensions $d = 2$ or $d = 3$, is composed of $m > 0$ anchors located at points $p_{m+1}, p_{m+2}, \dots, p_n$, and $n - m > 0$ nodes with unknown coordinates at the points $p_{m+1}, p_{m+2}, \dots, p_n$. For the case where $d = 2$, suppose that there are at least three anchor nodes. Accordingly, in case where $d = 3$ at least four anchors. We denote as \mathbb{F}_p points formation corresponding to the position of nodes p_1, p_2, \dots, p_n , and links between the points marked as pairs of neighbor nodes and anchors in N . For $d = 2$ and $d = 3$ cases, the problem of localization can be solved if and only if the formation \mathbb{F}_p globally rigid. Thus, all nodes in the globally rigid subgraph having at least three anchors are localizable.

In addition, please note that in the real world the measured distance $l_{i,j}$ between network nodes contain an error component $l_{i,j} = d_{i,j} + \varepsilon_{i,j}$. This means that the optimal placement solution might not exist. In such cases, there is an additional problem to minimize the difference between the actual deployment of network and topological placement, obtained by calculations.

4.3 Localization with additional information resources

Digital maps that have already become widely available location-based resource that carry a lot of potential additional source of data for the object localization methods in wireless transportation networks. Using raster data type in such problems is not a trivial task, but using vector data type in the task of distance graphs placement on the ground can be successful and without significant additional costs.

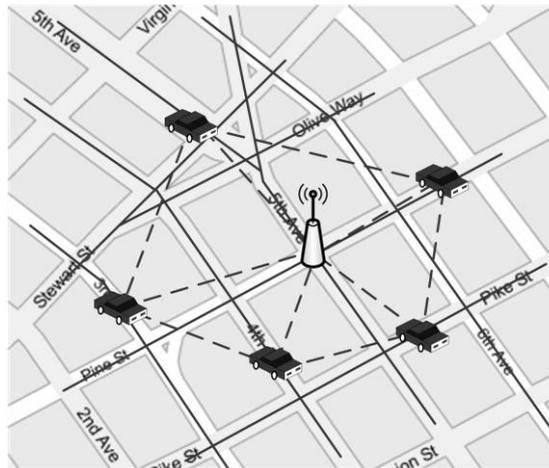


Fig. 12. Vertices of distance graph embedded into geo-location data graph

These sort of data viewed in the context of road infrastructure, can provide substantial assistance to improve localization processes in transport, excluding those locations at which the vehicle cannot be located physically as well as to refine the results of localization, comparing them to known geometric arrangement of the roadway. In a sense, the vector component of the

digital maps can be considered as the global graph, inside which the vertices of the subgraph formed by the wireless network nodes are placed, while vertices of the subgraph can or cannot match with vertices of the global graph, and can be placed on its edges.

Fig. 12 shows a graph of distances formed by distances measured between neighbor vehicles, as well as the distances to the preset infrastructure anchors with known coordinates. This graph should be embedded in the graph formed by the street digital map so that all vehicles are in line with its edges or vertices. Using these additional restrictions, in some cases, objects can be located, even if the graph of distances is not rigid by itself. Moreover, even acyclic graphs (trees), with such a method may be placed correctly, unlike with other previously discussed classical techniques.

Another significant advantage of this approach is the possibility of successful placement with less and formally insufficient number of anchors. Essentially, to determine the search window for embedding into global graph one anchor can be sufficient.

4.4 Embedding algorithm

The main idea of the approach based on localization with iterative placement of the graph is to use the known data on the coordinates of nearby landmarks together with the coordinates of known anchor nodes to localize the remaining may not sufficiently connected with each other anchor and nodes with unknown coordinates. Without any loss of applicability, this approach fits static network located in two-dimensional space. For simplicity, first let us consider the case where each sensor of the network has the ability to measure the distance to all other network sensors, regardless of the distance.

Given distances graph $G(V, E)$ and the graph of digital maps $I(W, F)$. To start the algorithm, at least one node n in the network $1, \dots, N$ must be an anchor $b_0 \in N$ with known coordinates x_0, y_0 in the chosen coordinate system.

We use the definition of “neighbor” to denote the set of nodes $v \in N$, whose distance from the starting anchors node is known and available for computational processes. Another definition that is required to introduce is a reflection nodes s , meaning the whole set of points of intersection of graph elements of digital map $\{s | s \in W \vee s \in F\}$ and the circle circumscribed around the anchor node b_0 with coordinates x_0, y_0 . The diameter of the circle R assumed to be equal to the distance d node v being localized.

Coordinates of s given by:

$$\begin{cases} x = x_0 + R \cos \varphi \\ y = y_0 + R \sin \varphi \end{cases} \quad 0 \leq \varphi < 2\pi \quad (28)$$

The physical meaning of this definition is to display all valid physical location of this node in the existing topological context (in this first example, the simple grid). General block diagram shown in Fig. 13.

Given the above, the algorithm is as follows:

1. For each anchor node b , select all neighboring nodes v and group them based on whether there exists measured distances between each v in the group (connected nodes). In the case of unconnected nodes group can consist of a single node, and when all nodes are connected, respectively, form only one group (Figure 29);

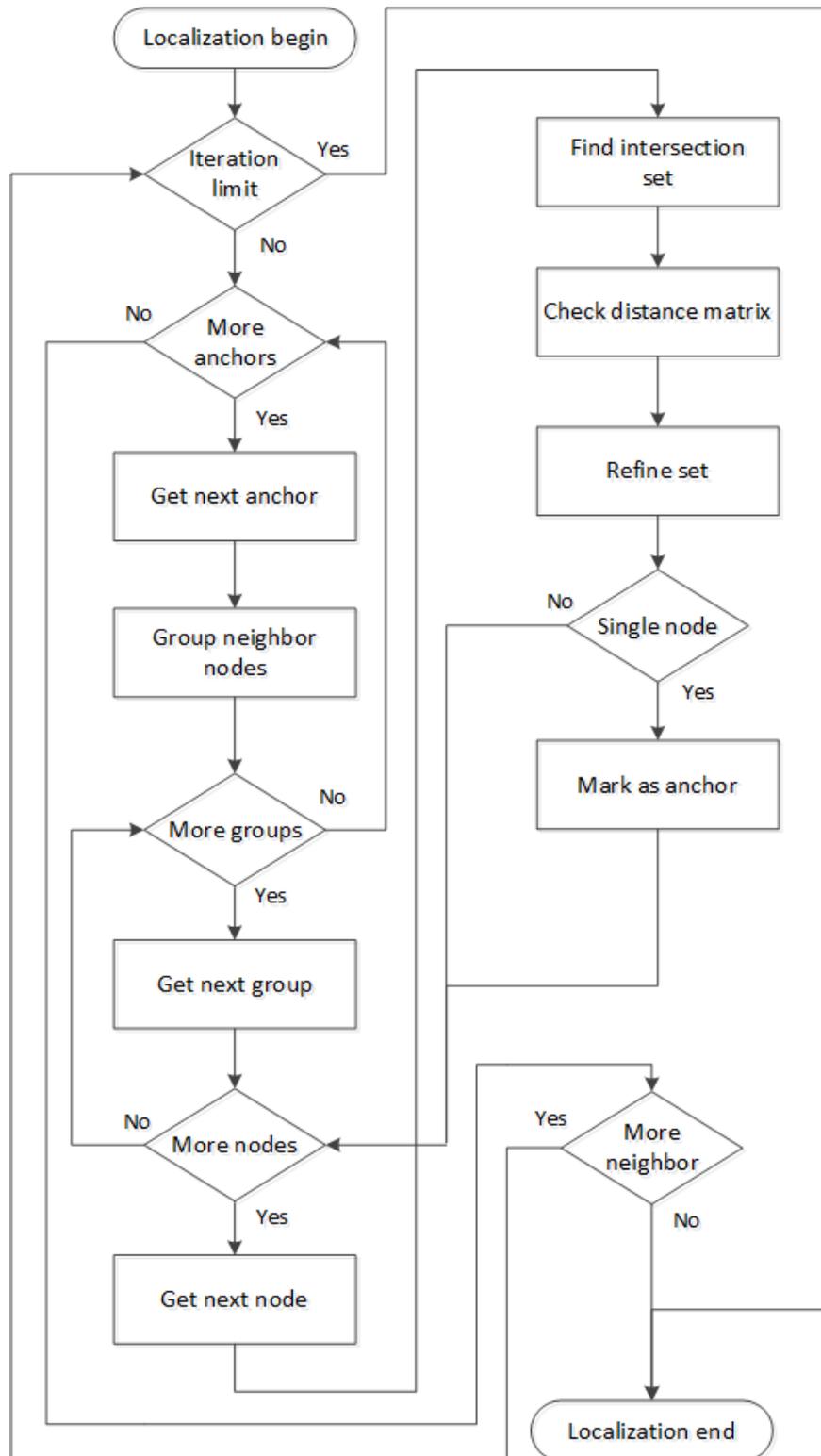


Fig. 13. Localization algorithm scheme

2. In each group for each node, calculate the coordinates of its reflections s_n , by finding the points of intersection of the graph elements I , and a circle with the center with coordinates of the node v_n and a diameter equal to the known distance from that node to anchor b_n . Step 2 of the algorithm illustrated by Fig. 30;
3. For each node in the group v_n calculate the coordinates of the set of possible reflections s_n ;
4. If the group contains more than one node for each pair $s_n s_{n+1}$ check whether the distance between the known distance between v_n and v_{n+1} ;

5. Exclude from node combinations nodes that do not match a known distance measurement, taking into account a given by a model certain tolerance to errors;
6. If the group remains the only one combination of reflections $s_n, s_{n+1}, \dots, s_{n+m}$, the nodes in the group are localized;
7. The algorithm repeated for anchors and nodes, localized at the previous iterations of the algorithm and, from this moment, also considered as additional anchors, until any anchor has at least one neighboring node that is not localized. Stopping criteria of the algorithm is the localization of all nodes in the network or the reaching a predetermined maximum number of repetitions.

Modeling environment settings: Table 1 demonstrates pseudo-code, which gives an overview of the localization algorithm.

Functions `find_neighbors` (find neighboring nodes), `split_to_groups` (split nodes into groups), `compute_all_combinations` (find all nodes combinations) and `evaluate_shadow_layout` (evaluate placement reflections) described above are trivial operations of computational geometry and combinatory.

For algorithm simulation and debugging the following settings were selected: a network of eleven nodes, ten of which are not localized, and each has a measured distance for up to six neighbor nodes, one anchor node and grid, representing in this case the elements of a digital map.

After performing the first iteration of the algorithm unique placement has been found for six not-located nodes that in turn became anchor nodes. The second iteration led to the localization of three more nodes (4, 7 and 10), and the third iteration of step 3, localized the last non-localized node 10.

Table 1. Localization algorithm pseudo code

```

while(stop conditions are not met){
Input: beacon_array[];
Input: node_array[];
for(n = 0; n < beacon_array.length; n++){
    neighbors[] = find_neighbors(beacon_array[n], node_array[]);
    neighbor_groups[] = split_to_groups(neighbors[]);
    for(m = 0; m < neighbor_groups.length; m++){
        for(i = 0; i < neighbor_groups[m].length; i++){
            shadows[] = find_shadow_locations(neighbor_groups[m][i]);
            neighbor_groups[m][i].add(shadows[]);
        }
        shadow_combinations[] = compute_all_combinations(neighbor_groups[m]);
        possible_combinations[] = evaluate_shadow_layout();
        for(k = 0; k < possible_combinations[].length; k++){
            if(possible_combinations[k] is only one possible){
                localized_node = node_array.pop(neighbor_groups[m][node]);
                localized_node.isBeacon = true;
                beacon_array.push(localized_node);
            }
        }
    }
}
}
}

```

Another thing that requires mentioning is the fact that in this case the goal was only in

the evaluation of performance of the algorithm itself, and considers only the simplified localization scenario using the grid as a digital map.

Computational complexity: In general, the complexity of the algorithm can be estimated by a function depending on the volume of work performed by algorithm and the size of the input data. The algorithm has a complexity of $O(f(n))$, if the when input size incises to N dimension , the running time is increased at the same rate as the function $f(N)$.

Estimating the order of complexity of the algorithm, it is necessary to use only the part that grows the fastest. Assume that the working cycle described by expression N^3+N . In this case, the complexity is equal to $O(N^3)$. Consideration of the fastest-growing part of the function allows evaluating the behavior of the algorithm with increasing N . Calculating O constant factors in the expressions can be ignored. Algorithm with a working step $3N^3$ treated as $O(N^3)$. This makes the dependence of the ratio $O(N)$ from the task size changes more obvious.

If one procedure calls another one, it is necessary more thoroughly estimate the overall complexity. If it executes a certain number of instructions (eg, printing), is virtually has no effect on complexity estimation. However, if the called procedure executes $O(N)$ steps, this function can significantly increase algorithm complexity. If the procedure is called inside the loop, the impact can be much greater.

If the inner loop of one procedure there is a call to another procedure, the complexities of both procedures are multiplied. In this case, the complexity of the algorithm is: $O(N2) * O(N3) = O(N5)$. If the main program calls a procedure at a time, theirs complexities are added: $O(N2) + O(N3) = O(N3)$.

Based on the above, we estimate the computational complexity of the localization algorithm, which was built by the examined model first. For this purpose, we simplify the pseudo code given in Table 1, to cycles taking as example code given in Table 6. Since, in contrast to the example considered, the algorithm is not a function with one variable, the computational complexity is the sum of the number of repetitions I , the number of anchors B , the number of localized nodes N , and also depends on the number obtained in the process of computing node groups G and reflections for each node S .

Table 2. Simplified localization pseudo code

```

for 1 to I // stop conditions are not met
begin
for 1 to B // beacon_array length
begin
for 1 to G // neighbor_groups
begin
end;
for 1 to N // possible_combinations
begin
end;
end;
end;

```

Strictly speaking, referring to the combination of the location of nodes in terms of combinatory we talk about combinations without repetition, the number of n for which over k is the binomial coefficient:

$$C_n^k = \frac{n!}{k!(n-k)!} \quad (16)$$

Consequently, the total number of combinations will be the sum of combinations of the reflections over maximum number of reflections for one node in all groups of connected nodes. In addition, given the fact that the presence of three or more anchors, localization becomes possible by classical methods, the number of anchors B shall not be greater than two, and justified number of algorithm repetitions cannot be greater than the number of non-localized nodes. Consequently, the function of maximum computational complexity is as follows:

$$f(N) = N \cdot 2 \cdot \left(\frac{N}{G} + \frac{n!}{k!(n-k)!} \right) \quad (17)$$

Thus, the computational complexity depending on the connectivity of the graph (number of groups) can vary from $O(N^2)$ to $O(N!)$.

It should be noted, an optimistic scenario of the proposed algorithm computational complexity $O(N^2)$ is acceptable, and pessimistic, the maximum variant computational complexity of $O(N!)$ is too expensive, and requires optimization algorithm. A possible way of this optimization will be discussed in the next chapter.

4.5 Simulation results

It is obvious that the problem of localization, solved the above algorithm as a classical problem of placement of graphs is NP-complete in the class NP, to which can be reduced any of the problems solved in polynomial time. As can be seen, the algorithm on its steps 4 and 5 actually makes an exhaustive search to avoid conflicting combinations of nodes. Unfortunately, over certain network size time of the exhaustive search can become prohibitively large.

On the other hand, along with a continuous increase in productivity of modern computers, large enough NP - complete problems have already become quite rapidly solved. For example, the traveling salesman problem with the complexity of 2000 cities, now is solved by brute force is an acceptable period of time. And when the data is well structured, the complexity of the problem being solved can be increased up to 14000 cities. In other words, there is a certain distance between the modern empirical results of various implementations of algorithms and theoretical studies known to us.

At the same time, we should pay attention to the fact that meanwhile there was considered only a simplified model with the digital map in the form of grid. Further, more realistic scenarios will be considered, with fragments of maps, consisting of more complex geometries. Therefore, the functional tests should include and adequately handle such scenarios as well.

Chapter 5. (Spatial queries over sensor networks)

5.1 Spatial queries

Based on location spatial requests are related to the type of requests, the result of which depends, also on a location of the requesting client. Effectiveness of management and processing of spatial queries becomes critical with the development and growth of wireless and mobile technologies. In this context, the implementation of localization algorithm based on spatial queries assumes that the network user has the ability to establish a direct connection with the information server and receive answers to their queries on demand.

As one of the possible implementation of approach to this problem Oracle Spatial 11g package was chosen. This well-known and well-proven software package contains a database and all the necessary functions and procedures that allow you to store a wide range of spatial data, as well as to organize a quick and efficient access to data and their analysis. Spatial data are in this case, the main characteristics of the location of the real or conceptual objects and their relationships in the real or conceptual space of their existence.

Oracle Spatial 11g provides the user with SQL schema and functions that provide storage, retrieving and updating sets of spatial elements, as well as requests to database platform Oracle 11g. Spatial database platform consists of the following main components:

- Schemes (MDSYS), which determines the order of storage, the syntax and semantics of supported geometric data types;
- Mechanism of spatial indexing;
- Operators, functions, and procedures for performing queries in a designated location, spatial connection requests and other analytical operations;
- Functions and procedures for debugging and tuning;
- Topological data model;
- Network data model;
- GeoRaster, functionality allowing storing, index, query, and analyze raster data and associated metadata.

The spatial component of the considered data structures, hereinafter referred to by the term “geometry” is a geometrical representation of the object shape in a certain coordinate space.

The method of obtaining the compound or spatial data uses a model consisting of two queries. Such model design means two independent successive operations performed to obtain the results of the spatial query. The combination of the results of each of the two operations is the final result of the query. The first operation acts as a filter for a large amount of data. It significantly reduces the number of candidate solutions for their further processing. The second, more accurate filtering operation works with a small amount of data and returns the final result of the query

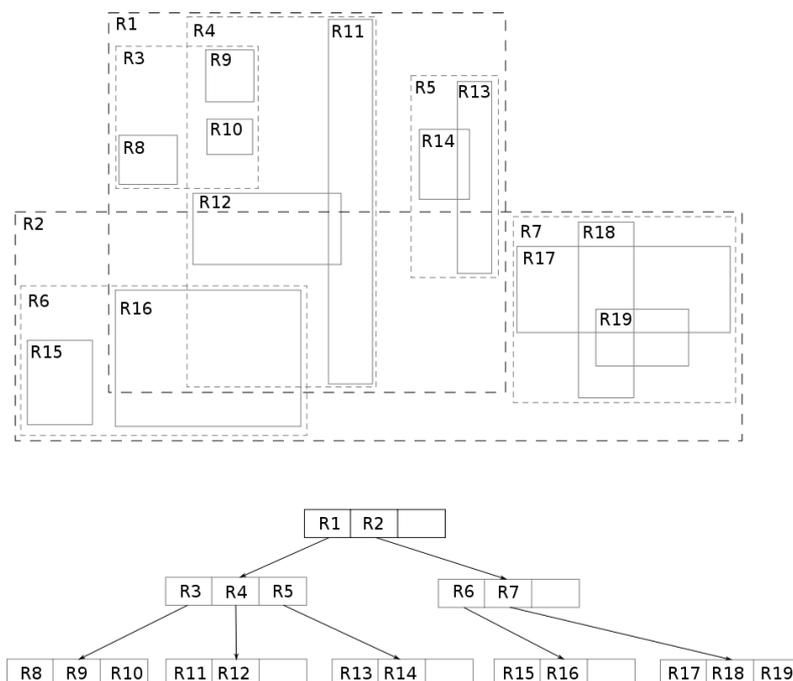


Fig. 14. R-tree for two-dimensional rectangles

The possibility of spatial database indexing is one of the key new features in Oracle Spatial product. A spatial index, as well as any other index, provides a mechanism to restrict the search window. In this case, the limitation is based on the spatial criteria such as geometries intersection, overlapping, etc.

This functionality provides the use of R-tree indexing, a tree-like data structure that is used for to access spatial data, i.e., for multi-dimensional information indexing, such as geographic data with two-dimensional coordinates. This data structure divides the space into a set of hierarchically nested and possibly overlapping rectangles (two-dimensional space). In the case of three-dimensional or multidimensional space, it will be a set of rectangular parallelepipeds (cuboids), (Fig. 14).

In 2004 was published a new, improved algorithm, Antonin Guttman (2004), which determines the priority R-trees. An improved algorithm not only showed the results at the level of the most effective modern methods, but its optimality for worst-case scenarios.

5.2 Proposed method

Computational complexity of the localization algorithm, proposed in the fourth chapter of this research, can reach unacceptably high values in various scenarios, in some cases up to $O(N!)$. In this connection, it seems appropriate to test its implementation as a method of spatial queries, which should allow largely optimizing the performance of the algorithm and reducing the influence of the localization scenario on requirements to computing resources.

Using spatial query functionality described above, consider the implementation of the algorithms mentioned in the form of spatial SQL operations. For the implementation of the algorithm two tables will be required, each of which has a column of type `SDO_GEOMETRY`. The tables use a primary key and record type identifiers. Other columns are system and are not directly related to the performed algorithm logic.

The first table `AREA_MAP` designed to store digital maps, including an anchor point b with known coordinates as required by the considered network localization scenario. Furthermore, during scenario processing in the same table are recorded not localized nodes $1 \dots n$, then for those whose distance from the anchor node is taken as known, constructed set of circles C , with a center having a coordinate (b_x, b_y) and radius r_i equal to the distance to the node i , the distance d_i to which was assumed as known.

The second table `STAGING_MAP`, is the staging table, needed to store temporarily the intermediate results of the algorithm. Spatial indexing procedure not described here specifically, since it is not part of the localization algorithm logic. However, for spatial queries high performance, spatial indexes update must be performed after any changes in the digital map or the location of the anchors.

The whole localization process is performed by two SQL operations. The first operation, calculates all the points of intersection of the circles C and digital map geometries elements, and stores these points in the staging table with a mark indicating to which of the circles c_i each given point belongs.

The operation uses a primary filter `ANYINTERACT` (any geometric interaction), to narrow the search window, and then function `SDO_INTERSECTION`, which have primary responsibility for the calculation of the total number of points of intersection. Last, the third parameter of the function, map tolerance, in this case set to value of 0.005. Description and examples of the filter functions `ANYINTERACT` and `SDO_INTERSECTION` usage are given in Appendix N2.

The second stage of the localization algorithm, SQL operation 2, is a selection of the set of unique points of intersection from the staging table, satisfying conditions of the corresponding matrix of known distances, both between nodes and between nodes and anchor.

In the case where the graph formed by nodes, is rigid enough or additional information obtained by means of a digital map, makes it rigid enough, the result of the query will be a unique set of points corresponding to the real topological arrangement of network nodes. However, when information is insufficient, multiple localization results are possible, one of which is real and the other will be products of rotation in the case of one anchor, or products of reflection in the case of the two anchors.

5.3 Simulation environment

As previously, the modeling is performed using a discrete simulation model to perform testing of the possible ways of events development, and the behavior of the object by variation of some or all of the model parameters. Furthermore, the model can be described as geometrically graphic, as it is presented largely by geometric objects and images appropriately coded in spatial database.

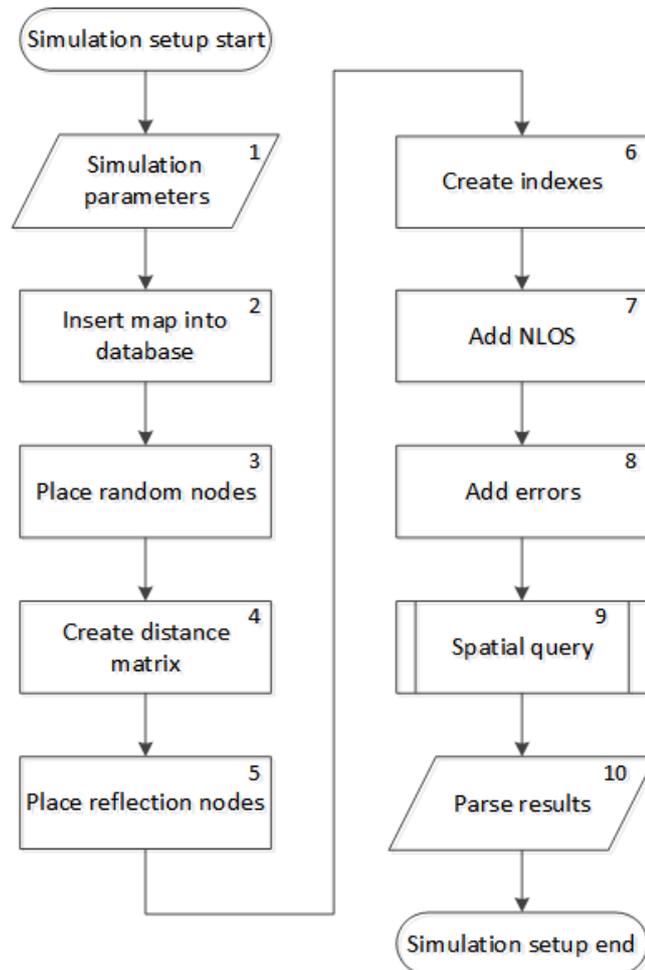


Fig. 15. Simulation algorithm

In this model, an algorithm includes a simulation environment preparation and setting up the scenario. Algorithm block diagram shown in Fig. 15.

The task of blocks 1 to 8 is to prepare digital maps, placement of network nodes according to the selected scenario and setting of simulation configuration parameters and

indices. Localization is performed by spatial queries in block 9. The objective of the latter, the 10th block is processing the results and preparation of data for visualization. The main classes of localization applications are given in Appendix N3.

5.4 Spatial query profiling

An important stage in the implementation of both spatial and conventional SQL queries is their profiling, collection characteristics of the program in order of further optimization that includes the search for an optimal query execution plan of all possible for a given query and process of query or the database structure changes to reduce usage of computing resources while executing the query. In the implementation of the model on the Oracle Spatial platform, the built-in Tuning Optimizer was used for profiling and debugging.

Table 3. Original query plan

Id	Operation	Name	Rows	Bytes	Cost(%CPU)	Time
0	SELECT STATEMENT		1	330	55 (0)	00:00:01
1	NESTED LOOPS		1	330	55 (0)	00:00:01
2	NESTED LOOPS		1	300	50 (0)	00:00:01
3	NESTED LOOPS		1	270	45 (0)	00:00:01
4	NESTED LOOPS		1	240	40 (0)	00:00:01
5	NESTED LOOPS		1	210	35 (0)	00:00:01
6	NESTED LOOPS		1	180	30 (0)	00:00:01
7	NESTED LOOPS		1	150	25 (0)	00:00:01
8	NESTED LOOPS		1	120	20 (0)	00:00:01
9	NESTED LOOPS		1	90	15 (0)	00:00:01
10	NESTED LOOPS		1	60	10 (0)	00:00:01
* 11	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01
* 12	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01
* 13	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01
* 14	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01
* 15	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01
* 16	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01
* 17	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01
* 18	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01
* 19	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01
* 20	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01
* 21	TABLE ACCESS FULL	AREA_MAP	1	30	5 (0)	00:00:01

The resulting report of this tool is performed detailed analysis of requests and recommendations for its improvement, followed by rational justification for the proposed steps. Recommendations of performance analysis of the original spatial queries are presented in Table 3. The plan of the query contains 10 nested program cycles and 11 accesses to a full reading of the table containing a digital map.

Each step of the query plan execution time is insignificant and showed as a minimum value for the data format of the results time table. Data size is small as well, which is typical for vector geometry data type, and is only 330 bytes. At the same time, CPU usage when running a query reaches 55%. In this regard, it is proposed to optimize the structure of the database that is to introduce additional composite index by type of geometry of a digital map.

Now in the query plan there was only one table full access to digital maps and 10 index-based accesses to the type of map element geometry. As it can be seen, result of CPU utilization decreased down to 25%. Therefore, to support localization implemented by proposed method, twice as less computing power will be needed, what is a good profiling result.

5.5 Scenario evaluation

The purpose of the simulation is to evaluate the performance of the proposed algorithm under different network sizes on different configurations of digital maps, different visibility conditions, as between the nodes, and the visibility of the anchor, as well as to evaluate the stability of the algorithm to the distance measurement errors. For the simulation of the localization algorithm were chosen two random fragment of typical urban road infrastructure. Both fragments were transferred to database as composite spatial geometries and hereinafter referred are as digital map F and G. It is worth mentioning that as a map F area with a high degree of symmetry was deliberately selected and which in localization process should provide a high probability of errors like rotation, reflection, and flex, J. Zhang, M. Zhu, D. Papadias, Y. Tao and DL Lee (2003). Map is G more asymmetric and belongs to the optimistic scenario simulation.

Evaluation starts with generating a predetermined number of nodes with a uniform distribution of random coordinate within geometric elements of the map. One random node is chosen as an anchor with a priori known coordinates. It is assumed that each node has a line of sight, and thus the measured distances to a given number of neighbor nodes. The remaining nodes are not in line of sight (NLOS). Also, specified the number of nodes with NLOS to the anchor node. By default, the distance to the anchor node is known. To emulate the distance measurement error, to the tolerance parameter added normally distributed random variable with expectation 0 and variance given, which if necessary can distort the results of the request within the predefined limits.

For each of the map different simulation scenarios have chosen, which consider the possibility of successful localization for networks of different sizes and different degrees of the graph connectivity, which is reflected in the proportion of nodes with NLOS. Scenarios parameters discussed in more detail below. At the end of each stage of the simulation, program returns detailed data log. For clarity and ease of understanding, the results are processed and presented in the form of three-dimensional surface diagrams.

Each chart shows the results of localization depending on the scenario parameters. This can be uniquely localized network, a number of possible placement solutions or absence of results in the case where no solution is found or the number of possible placements is too high and exceeds a predetermined threshold value. The number of solutions shown on the axis Y, in

the range from 0 to 50. For practical reasons, the number of solutions more than 50, has been considered as not determined and taken as a failure. NLOS percentage ratio between network nodes shown on the axis X, and NLOS percentage ratio between nodes and anchor on the axis Z.

5.6 Experiment data

Experimental verification of the proposed algorithm included wireless network scenarios localization simulation using spatial queries, and consisted of three main stages. The objective of the first step was to explore in general form of network algorithm performance in a wide range of parameter values as well as the size of the network and nodes visibility conditions NLOS for various localization scenarios. The result should become the initial data on the impact of the scenarios and the critical region of the parameters of the environment when the quality of the algorithm begins to deteriorate and further localization becomes impossible. The second phase of the study is aimed at a more detailed study of a narrow range of the environment parameters values to obtain more accurate data on the properties and the most critical parameters of the algorithm. The purpose of the third stage was to simulate different measurement errors values for the algorithm provided in the critical range that gives an opportunity to evaluate the order of the measuring accuracy requirements to network hardware.

For the first localization scenario map G was selected. Percent of nodes that are unable to measure the distance, and hence having NLOS with the anchor or NLOS with neighboring nodes, ranging from 70% to 90% of the total number of nodes, incrementing of 5% each iteration of the simulation. Furthermore, the localization scenario was repeated for networks of 5, 25 and 45 nodes. Thus, a simulation of the selected placement on the map consisted of 75 localization scenarios runs. Totally 3 random placement of nodes and anchors on the map was made, so the total number of runs in one simulation algorithm was equal 225.

The results of the simulation for networks consisting of 25 and 45 nodes showed 100% successful localization for all selected NLOS values. Non-unique results emerged only in the network with the fewest nodes. In Fig. 16 is given a three-dimensional graph of the first placement on the map G network consisting of five nodes.

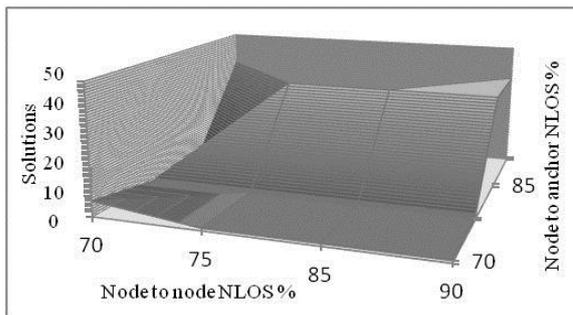


Fig. 46. Map G, 5 nodes, placement 1

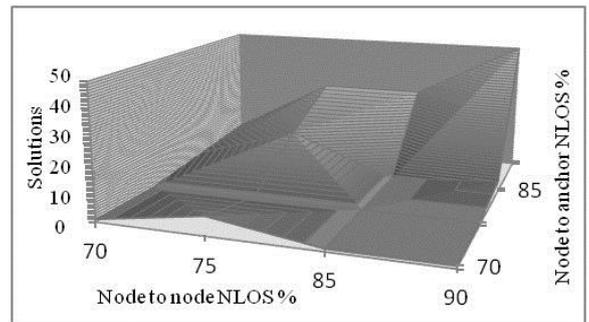


Fig. 17. Map G, 5 nodes, placement 2

Similar results were obtained in two other random placements. Networks of 25 and 45 nodes showed only the unique successful localization results. Localization of network consisting of five nodes, in some cases, showed some non-unique results as shown in Fig. 17 and Fig. 18.

As can be seen from the above diagrams greatest, stable localization, up to 90% NLOS, is observed in poor visibility between nodes. Although intuitively it is assumed that a larger amount of data must always improve the results, one can see here that on the contrary, the excess data in some cases adds from two to five alternative placements. In turn, a large number of NLOS from node to the anchor generates up to 50 alternative placements, because of the high degree of map G symmetry and thus the possibility of transformation of rotation.

Next, simulation was repeated for the same scenarios but the map F, having more asymmetric properties. Similarly, stable results with 100% successful localization were obtained for networks of 25 and 45 nodes. A network of 5 nodes at large NLOS degree shows alternative non-unique placements. All three placements are shown in Figures 19, 20 and 21.

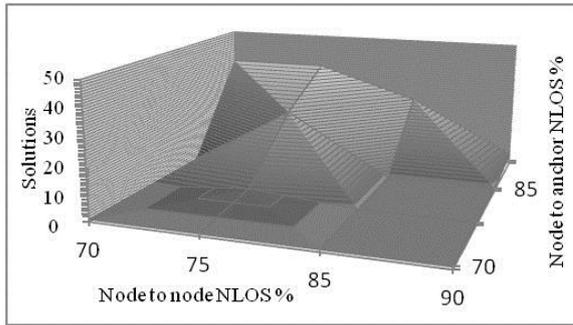


Fig. 18. Map G, 5 nodes, placement 3

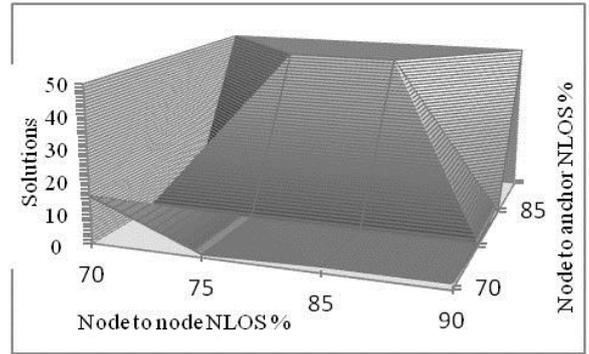


Fig. 19. Map F, 5 nodes, placement 1

It should be noted that the range and number of non-unique placements for large values of NLOS in this scenario have increased. This illustrates the greater network graph susceptibility to transformations such as reflection and flex what is characteristic for the map with a lower degree of symmetry. However, none of the scenarios has not gone beyond the predefined maximum of non-unique solutions, when the result of localization is treated as unsuccessful.

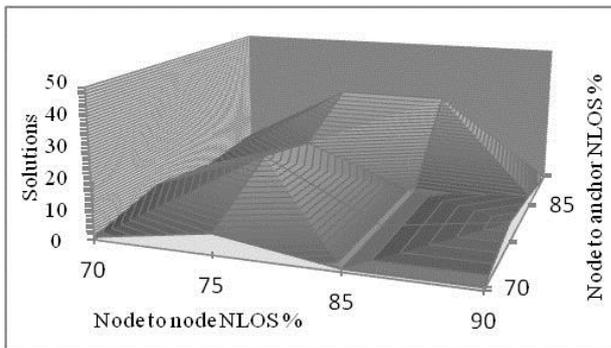


Fig. 20. Map F, 5 nodes, placement 2

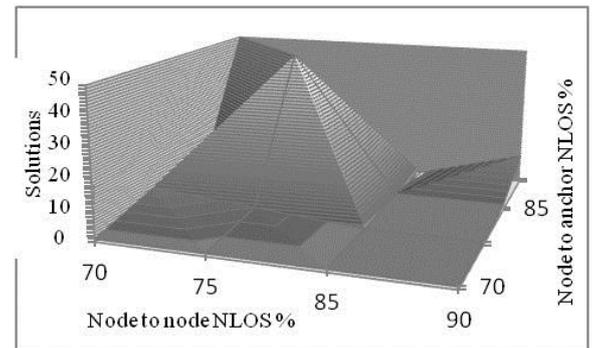


Fig. 21. Map F, 5 nodes, placement 3

Previous simulation showed that when the size of the network exceeds 20 nodes and the number of nodes without a line of sight NLOS is up to 80% of the total number of nodes, algorithm shows stable successful results. Therefore, in further simulations chosen NLOS range will be from 85% to 97% and the size of the network will be decreased from 30 nodes to 5 in decrements of 5 nodes. This allows getting a more qualitative view to analyze and plot parameters of the algorithm on the interval when it stops working.

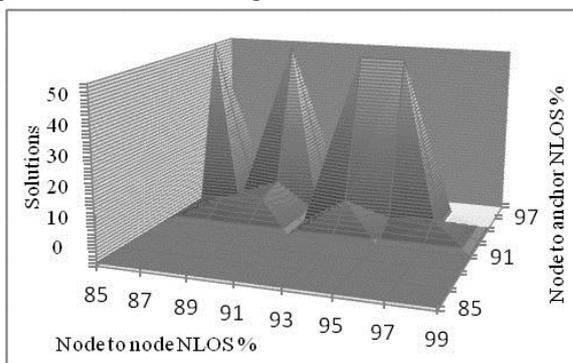


Fig. 22. Map G, 30 nodes

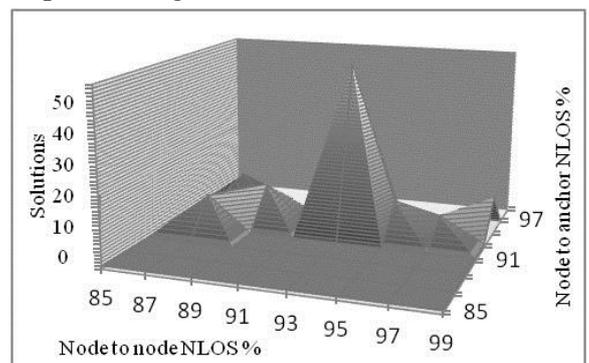


Fig. 23. Map G, 25 nodes

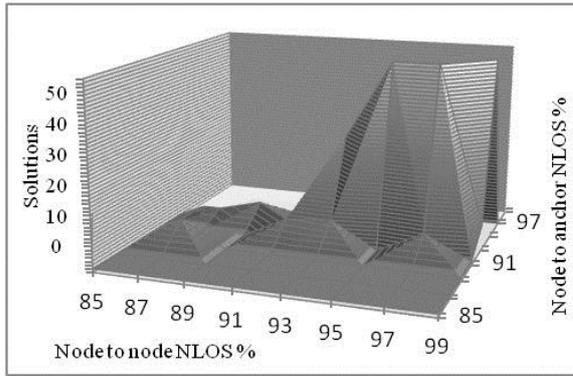


Fig. 24. Map G, 20 nodes

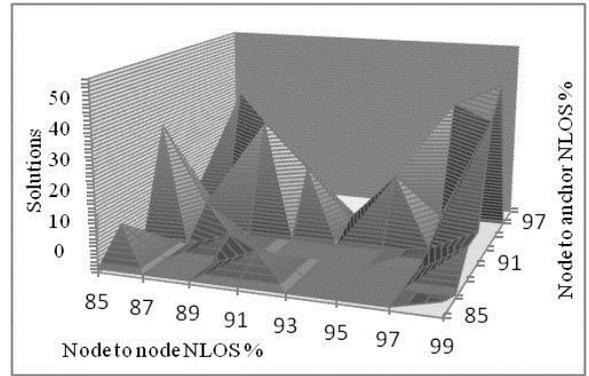


Fig. 25. Map G, 15 nodes

Moreover, since the previous simulation results showed insignificant dependence on the type of map and the number of random placements, further simulations were carried out on the map G with one placement.

Diagrams presented in Fig. 22 to 27 shows the results of simulations carried out with six parameters mentioned above.

Localization scenario of network consisting of 30 nodes, as shown in Fig. 22, shows a stable and successful localization and unique results over the entire range of selected NLOS size from node to node, provided that the NLOS from nodes to the anchor does not exceed 90%.

Moreover, at the maximum NLOS from node to node, when NLOS from the nodes to anchor, exceeding 95%, the first undefined results occurred (Fig. 23). This means that the limit is reached when the number of possible alternative placements becomes computationally difficult and localization algorithm stops operating. While the size of the network decreases, as shown in Figures 24, 25 and 26, deterioration of localization results progresses.

Last, the most pessimistic scenario with a network size of 5 nodes and extremely low degree of connectivity of the graph, the results of the localization of which are shown in Fig. 27, does not result in a single unique location. The whole area beyond 90% of NLOS from nodes to the anchor is not defined. However, with such a small network size, 90% NLOS would mean that only one node has information about the anchor and its range what it is insufficient when the network configuration graph is poorly connected.

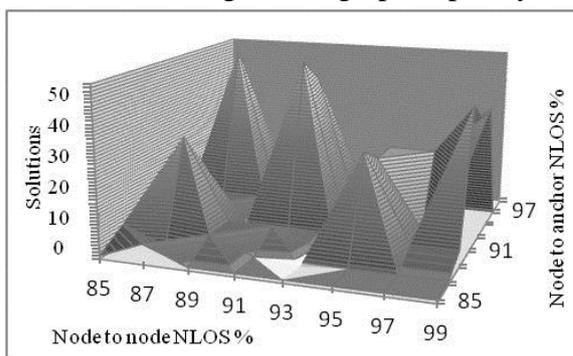


Fig. 26. Map G, 10 nodes

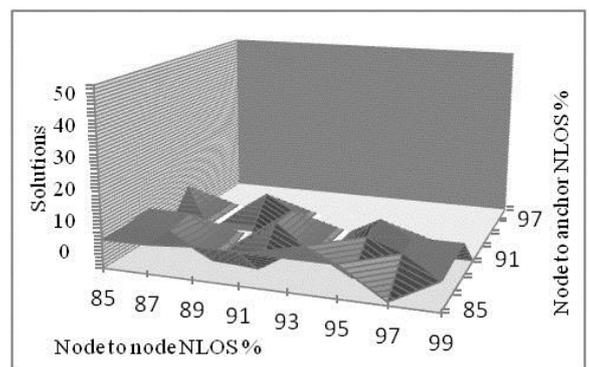


Fig. 27. Map G, 5 nodes

The last three scenarios were intended to simulate and evaluate overall impact of the measurement error magnitude. Previously, the default measurement error was set at 0.05m which corresponds to the boundary of statistical error, and does not significantly affect the results. In further simulations error size was set as normally distributed random variable with zero mean $\mu = 0$ and standard deviation $\sigma = 1$ for the first scenario, $\mu = 1$, $\sigma = 5$ for the second

and $\mu = 5$, $\sigma = 25$ for the third scenario respectively. This means that the magnitude of the measurement error can reach 25 meters.

It should be noted that at the largest allowable error of 25 meters, the localization algorithm functioning was obviously impossible. Since the error limit of 25 meters is comparable to the placed network size, the size of the geometric component of a digital map and the distance between them in the context of urban development, this method of localization becomes disoriented and gives no results. Therefore, in further we consider only the first two scenarios parameters $\mu = 0$, $\sigma = 1$ and $\mu = 1$, $\sigma = 5$.

Simulation results of the last two scenarios are reflected in the Figures 28 and 29. Measurement errors are implemented with tolerance parameter of spatial query, what in turn erodes the value of the measured distances. In both cases, settings used are the most approximate to the real. Studied network has size of 20 nodes and NLOS values ranged from 10% to 75%. The results showed that the measurement error of 1 meter on the map G caused the emergence of alternative graph placements from 1 to 10. Accuracy up to 5 meters on given map causes stable appearance of more than 40 alternative graph placements over the entire range of the simulation parameters.

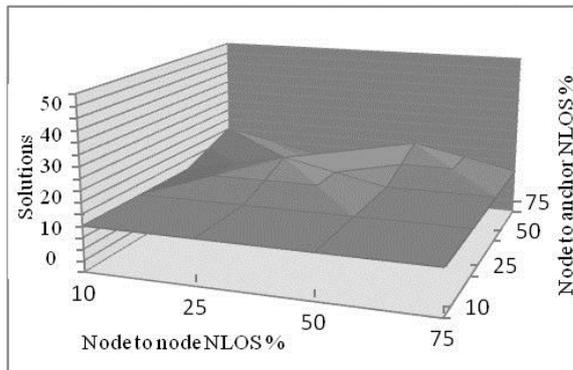


Fig. 28. Map G, 20 nodes, error 0m – 1m

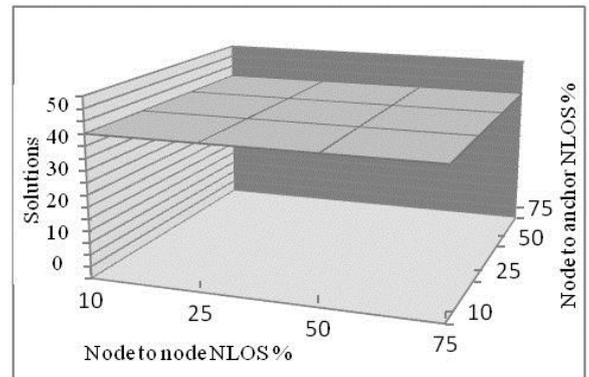


Fig. 29. Map G, 20 nodes, error 1m – 5m

5.7 Results

The results of simulations showed performance of the proposed algorithm, but also the fact that the measurement errors and the lack of line of sight between the nodes in the network may lead to non-unique alternative placements or uncertain results of localization by embedding the distances graph on a digital map. In addition, you will notice that there is a certain critical density of the network, above which the quality of the localization increases significantly and remains stable.

Intuitively it is expected and is clear that the localization accuracy should increase with the density of the network increasing. Increasing the number of neighbor nodes for each node means more constraints to embed the distance graph of by spatial query. However, after a certain critical density further improvement does not occurs any more. Moreover, given the considerable measurement error, the network with high density will be more susceptible to alternative placements.

It should be noted that almost all of the various localization algorithms proposed for wireless networks have much in common in their structure:

- Measure the distance between the non-localized nodes and anchors;
- Identify the location of the node relative to the anchor;
- Optionally, accurate coordinates of the nodes using the location data of neighboring nodes.

However, all of these algorithms for the successful operation require at least three anchors and rigid graph of distances between nodes. That is the reason why the proposed method of scene analysis using spatial queries deserves special attention as a method which allows localizing objects in cases where other methods are not applicable for fundamental reasons. Of particular interest in this regard are scene analysis techniques, combined with the placement of the graph distances and terrain maps. This feature allows the use of standard digital maps, which are usually equipped with transport navigation systems and existing communication systems.

However, the proposed realization of the described localization method on the basis of spatial queries can solve the problem of high computational complexity, what is typical for scene analysis methods. All this makes it possible to translate the problem of the transportation systems mobile objects localization in terms of limited spatial data availability from theoretical considerations to practical implementation.

CONCLUSION

This thesis is devoted to problems of objects localization in corporate transportation networks. Particular attention is paid to the methods of localization of ad-hoc networks with additional alternative sources of information in terms of inability to use global positioning systems.

Relevance of the work is determined by the need of mobile objects reliable positioning in a variety of environmental conditions. Unfortunately, most of the “classical” methods to solve this problem do not achieve the desired result in all possible embodiments and traffic environment. Some new perspectives in solving this problem reveals the rapid development and spread of autonomous intelligent devices and communications equipment, thanks to which there are additional opportunities for localization of objects equipped with such devices, by cooperative interaction of such objects.

At present, it is known a significant amount of research devoted to the new heuristic localization methods beyond the scope of classical computational geometry. Such methods are known as methods of the scene analysis. Nevertheless, many fundamental localization problems based on the analysis of the topology of the scene remains unsolved.

During the research following results were achieved:

1. The review of the classical methods of localization, the most frequently used in transport systems. Review and classification provided for the most promising, from the point of view of the author, new approaches to solving the problems of localization. It includes also a description of advantages, disadvantages and features of these approaches and their implementation methods;
2. The detailed analysis of localization system requirements in ad-hoc corporate transportation networks. The features associated with the various conditions of member access to the network information resources and the dynamic changes in the network configuration. Defined types of relevant software applications and their sensitivity to coordinates accuracy and actuality;
3. Evidences were provided that in the corporate transportation systems, depending on the type of application, both the relative and the absolute global coordinates of nodes may be relevant. Also, it is noted that the localization method should consider that the network within a long period time, may stay both static and dynamically changing. Furthermore, due to network reconfiguration serious constraints can be imposed on the possibilities of distributed computing. However, the most serious

limiting factor in the application of classical localization methods is generally a high probability of insufficient number of reference points or lack of precise topological data required for localization. At the same time, devices with online access to the information network opens up new alternatives for obtaining missing data;

4. Justified the choice of the localization method, based on the scene analysis, by enabling the collection and use of data on the possible location and network location context to reconstruct temporarily unavailable or create new reference points, as well as receive other additional geometric data;
5. Formulated and developed a model based on spatial database, what provides support for digital maps prototyping. Support includes the ability to store topological, network and other specific objects metadata, the ability for debugging and profiling. This model provides a good opportunity to analyze different localization scenarios of transportation networks;
6. Developed the method of localization by placing a graph given by the matrix of distances between localizable objects on a graph, describing the environment to which localizable objects belongs, when all the vertices of the distance graph are located on the edges or match with the vertices of the environment;
7. The method is able to use digital maps and other additional data available in GIS systems, in order to create constraints that increase rigidity of graph being embedded and thereby eliminate uncertainty and excluding alternative placements. In some scenarios, the method may preserve its operability even when there is only one reference point whose coordinates are known;
8. Conducted validation and analysis of the results of applying the developed method for different localization scenarios. The above scenarios included networks of different size, the most common configuration fragments of digital maps, emulation of lack of line of sight between network nodes and various distance measurement errors.

The experimental results suggest that the topological data provided by modern GIS-systems can be effective alternative information resource for decision-making support in cases when of the classical methods of objects localization in the cooperative transport systems are not applicable. Positioning based on a combination of classical methods and spatial queries, helps to compensate the possibility of temporary dysfunction of GPS, as well as, if necessary, correct the errors encountered in the other localization methods.

The results of the thesis can serve as the basis for further improvement of scene analysis methods. The basis for this improvement is use of not only topological but other additional information resources as embedding constraints. The most promising in this respect are the following resources:

- Exogenous resources, where each localization scene has its patterns of signal propagation. Information on such patterns stored in the dedicated spatial database layer can be transparently included in the topological queries;
- Patterns of objects that take into account the type of vehicle and its inherent size and other parameters. Data transmission on the type of vehicle within the localization application communication protocol may be used for further measurement calibration and error compensation;
- Information about the objects proximity. Proximity factor, without any significant improvements can be estimated by spatial query and can serve as an additional indicator that eliminates some alternative placements if they arise as a result of the

proposed localization method;

- Information about the line of sight between the network nodes. The experiments showed that the proposed method, in some cases, might demonstrate good results with a very high proportion (up to 90%) of non-line-of-sight between network nodes. At the same time, the lack of direct line of sight between the network nodes, with high probability occurs due to the presence of obstacles. If the digital map data is not only about the road infrastructure, but also about objects that can serve as obstacles, the data that are associated with data visibility, could be another, an additional set of constraints with lower priority. With this improvement, it becomes possible to locate the network with weakly connected distance graphs where other localization methods usage is very difficult or impossible in principle;
- Infrastructure resources. These include extensive computational capabilities of modern transportation infrastructure IT hardware, allowing to identify store and share data on the vehicles involved in the localization and, further, to take into account their movement between scenes.

Mentioned above ways of improving the localization software solutions will bring the quality of applications solving reliability, safety and efficiency problems of cooperative systems in transportations telematics to higher level and will simplify its implementation.

PUBLICATIONS WITH AUTHOR'S PARTICIPATION

1. Krebs V., Tsilker B. *Coverage Evaluation Approaches for Intelligent Transportation Systems Based On Anisotropic Sensor Networks*. Proceedings of the 10th International Conference “Reliability and Statistics in Transportation and Communication”, 2010, Riga, Latvia, pp.199-206.
2. Krebs V., Tsilker B. *Coverage Estimation in Synergic networked Intelligent Transportation Systems for Non-Isotropic Nodes*. Transport and Telecommunication, volume 11.No.4, 2010, Riga, Latvia, pp. 36-45
3. Krebs V., Tsilker B. *Vehicle Dynamic localization in intelligent transportation systems based on sensor networks*. Transport and Telecommunication, volume 12.No.3, 2011, Riga, Latvia, pp. 63-72.
4. Krebs V. *Analysis of the influence of the characteristics and relative position of the reference points on the error localization of objects in VANET networks*. Research and Technology – Step into the Future, volume 7. No 1, 2012, Riga, Latvia. p. 125.
5. Krebs V., Tsilker B. *VANET nodes localization with limited reference points availability*. In: Proceedings of the 12th International Conference “Reliability and Statistics in Transportation and Communication” (RelStat'12), 17–20 October 2012, Riga, Latvia, pp. 285-295.
6. Krebs V. *Spatial Query Localization Method in Limited Reference Point Environment*. International Journal of Mathematical Sciences. World Academy of Science, Engineering and Technology 75, 2013, pp 1242 – 1249.
7. Krebs V., Tsilker B. *VANET object localization specifics*. In: Proceedings of the 13th International Conference “Reliability and Statistics in Transportation and Communication” (RelStat'13), 16–19 October 2013, Riga, Latvia, p. 249-256.
8. Krebs V. *STSM Scientific Report improving the accuracy of real-time positioning of moving objects in mines*. COST-STSM-ECOST-STSM-IC0906-041113-036757, Luleå University of Technology, Sweden November 15, 2013.
9. Krebs V. *Spatial query localization method in limited reference point environment: an empirical study*. European Scientific Journal August 2014 edition vol.10, No.24. pp. 48 – 64.