UDC 519.8 Вестник СПбГУ. Прикладная математика... 2017. Т. 13. Вып. 1

R. V. Voronov

THE PROBLEM OF OPTIMAL PLACEMENT OF ACCESS POINTS FOR THE INDOOR POSITIONING SYSTEM

Petrozavodsk State University, 33, Lenina pr., Petrozavodsk,
185910, Russian Federation

This paper deals with an indoor positioning system. The system is based on the use of wireless local area network access points. A location calculation engine is based on Bayesian algorithms. Location accuracy depends on the number and placement of access points. This paper considers the mathematical model and the method of solving the problem of optimal access point placement for indoor positioning system. The criteria for evaluating the quality of the access points placement is the mathematical expectation of the localization error. We consider two strategies for localization of a mobile object. It is demonstrated that, for some strategies, the addition of access points can possibly increases the expectation errors, for example, the strategy selecting the most probable zone. A strategy, guaranteeing that the addition of access points does not lead to an increase in the expectation errors is proposed. An algorithm for solving the optimization problem is developed. We present the result of testing the algorithm on real data. Refs 16. Fig. 1.

Keywords: indoor location, placement of access points.

Р. В. Воронов

ЗАДАЧА ОПТИМАЛЬНОГО РАЗМЕЩЕНИЯ ТОЧЕК ДОСТУПА СИСТЕМЫ ПОЗИЦИОНИРОВАНИЯ ОБЪЕКТОВ В ПОМЕЩЕНИИ

Петрозаводский государственный университет, Российская Федерация,
185910, г. Петрозаводск, пр. Ленина, 33

Статья посвящена системам определения местоположения мобильных объектов внутри помещений, основанным на использовании беспроводных сетей точек доступа. Метод расчета местоположения базируется на байесовском алгоритме. Точность определения местоположения зависит от числа и мест расположения точек доступа. Предлагаются математическая модель и метод решения задачи оптимального размещения точек доступа системы позиционирования мобильных объектов в помещении. Критерием оценивания качества размещения точек доступа служит математическое ожидание ошибки определения местоположения объекта. Рассмотрены две стратегии выбора алгоритмом позиционирования искомой позиции объекта. Показано, что для некоторых стратегий возможна ситуация, когда добавление точек доступа приводит к увеличению математического ожидания ошибки определения местоположения объекта. Такой стратегией является, например, выбор наиболее вероятной зоны. Разработана стратегия, для которой гарантируется, что добавление точек доступа не приводит к росту математического ожидания ошибки определения местоположения объекта. Построен алгоритм решения поставленной оптимизационной задачи. Приводится результат тестирования алгоритма на реальных данных. Библиогр. 16 назв. Ил. 1.

Ключевые слова: локация внутри помещений, расстановка точек доступа.

Introduction. The article considers the problem of mobile objects location inside a building. As examples of the application of the indoor positioning systems, we can specify monitoring of children in child care centers, goods search on the store stock,

Voronov Roman Vladimirovich — PhD of technical sciences, associate professor; rvoronov76@gmail.com

Воронов Роман Владимирович — кандидат технических наук, доцент; rvoronov76@gmail.com

DOI: 10.21638/11701/spbu10.2017.106

locating medical equipment, personnel and elderly patients in the hospital, locating of the miners in the mines in the emergency situation, controlling position the group of law enforcement officers during the execution special operations, tracking firefighters inside a burning building, locating sniffer dogs searching for explosives, the definition frequented places in public buildings, in robotics [1, 2], etc.

Special positioning system is used where global navigation satellite systems is inadequate, for example, in buildings and indoor areas. Indoor positioning systems uses different technologies, including wireless sensor networks [3, 4]. For the construction of such systems various wireless technologies are used: Wi-Fi, ZigBee, nanoLOC, UWB, Bluetooth, etc. The access points of the network have been installed in the building. Each access point has its own coverage area. The location of the object is determined by processing the received signal characteristics (time of flight or signal strength). Sometimes the plan of building is initially divided into set of zones. Physically zones correspond to the small room of the rooms or corridors. As a rule, zones do not contain walls and partitions. During the installation of the location system signal strength map of access points is built [5, 6]. Typically, the signal strength of the access point map is formed from the average values of its signal strength in each zone ("fingerprinting"). When mobile object is broadcasting, it registers signals from all access points. This signal strength vector is compared to a map and zone, for which signal values the most similar to the registered values, is selected [3, 4, 6]. When this positioning quality significantly depends on the number and placement of stationary access points.

During the installation of positioning systems the always to reduce the number of used access points to minimize the cost of the equipment. Therefore, the following iterative scenario is often used. First, access points in some areas of the building is set. The next step is to hand-learn systems, calculate signal strength maps, and then if needed in different parts of the building add a few more access points, improving location accuracy. The disadvantage of this approach is the time required to conduct such an experiment. An alternative is to use mathematical modeling to estimate the quality of the access points placement and to make recommendations for its improvement.

It is assumed that we know the set of possible positions of installation of stationary wireless access points, and for each such position is known probability distribution of signal strength in all zones from the access point. Note that the parameters of the probability distribution of signal strength can be estimated on the basis of mathematical models of signal distribution in the buildings of [7–9].

It is proposed to estimate the quality of the positioning system using the mean error location, which is averaged distance between the actual object location and the location defined by positioning system. The main feature of the models is the use of the signal strength distribution in zones.

Related works. The paper [10] provides a method of combining maximum coverage space requirements and reducing errors in determining the location of mobile objects. A mathematical model for determining the error locations of mobile object based on the variability of the signal strength measurements. This model is based only on general assumptions about the work of localization algorithm used.

The overall expected error is determined by the formula

$$E = \frac{\int\int_{A\times A} d(x,\hat{x})P(x \mid \hat{x})w(\hat{x})dxd\hat{x}}{\frac{1}{|A|}\int_A w(x)dx}, \tag{1}$$

where x — valid coordinates of region A; $\hat{x}$ — mobile object position; $d(x,\hat{x})$ — the distance

between points x and $\hat{x}$; $P(x \mid \hat{x})$ — conditional probability distribution; $w(\cdot)$ — weight coefficient.

Let τ be the signal strength threshold, below which the signal is considered unacceptable. It defines the set A_C of all zones in the networked area that receive sufficient coverage by at least one access point:

$$A_C = \{x \in A \mid \bar{\mu}(x) \geqslant \tau\},$$

here $\bar{\mu}(x)$ — the strongest signal from access points in zone x.

Size of A_C is considered as a measure of coverage:

$$C = |A_C|. \tag{2}$$

Access points placements is determined by various criteria of optimization: minimizing localization error (1), maximizing coverage area C (2), combination of these two criteria. To find the placement of access points, which minimizes the average location error and maximizes coverage, it is suggested to use a combined objective function:

$$C' = E + \gamma \frac{1}{C}.$$

The authors experimentally determined constant value $\gamma = 2500$. For each option, the placement of access points defined by the conditional probability distribution $P(x \mid \hat{x})$ and average strength of the signal.

The authors have proposed different methods of solving the optimization problem. The most efficient method is local search algorithm with restrictions. In this algorithm is allowed the transition to the state with the worst value of the objective function, but it is prohibited to move the "recently" displaced access point.

A disadvantage of the proposed model is that for each zone only one, the most probable combination of signal strengths from the access points in this area is considered. It may lead to solutions, in which the access points are arranged too closely to each other. Note that, in practice, the access points are placed at a certain distance from each other. Thus, in placement models of access points for each zone the distribution of signal strength of combinations of access points must be considered, not only the most likely values.

The paper [11] considers the problem of optimal placement of access points. It is assumed that the distance from the mobile unit to the access point is determined by the time of arrival, or by the received signal strength. To solve this problem we apply the method of least squares. However, deterministic signal propagation model assumes.

In the works [12, 13] problem of optimal placement of access points is reduced to the problem of covering. The model is deterministic, excluding accidents. The paper [14] refers to the NP-complexity of the problem of choosing locations for access points. In this article deterministic signal propagation model assumes also.

In general it can be noted that the access points placement problem in locating system is a little studied. The majority of the works, such as [15] are dedicated to the problem of providing a given signal strength in areas.

Mathematical model. We introduce the notation for a mathematical model. Given a positioning system $L = (V, d, I, S, u_D^*)$, where

- V — a set of zones, $V = \{v_1, \ldots, v_n\}$;
- d — metric on V;
- I — a index set of access points placements (we shall call, for brevity, set of access points);

- S — a set of signal strength values, $S \subset \mathbb{Z}$;
- u_D^* — positioning function, $D \subset I$.

To simplify the notation, we assume that $d(j_1, j_2) = d(v_{j_1}, v_{j_2})$, here $v_{j_1}, v_{j_2} \in V$.

We introduce the random variables ξ, η_i, $i \in I$, where

1) ξ takes values from the set $J = \{1, \ldots, n\}$ and its value indicates the the zone number;

2) η_i takes values from the set S that we will be interpreted the signal strength of the object sensor measurements from the access point with the number i.

Random variables η_i $(i \in I)$ is assumed are mutually conditionally independent given the random variable ξ, i. e.

$$P(\eta_{i_1} = s_{i_1}, \ldots, \eta_{i_2} = s_{i_2} \mid \xi = j) =$$

$$= P(\eta_{i_1} = s_{i_1} \mid \xi = j) \ldots P(\eta_{i_2} = s_{i_2} \mid \xi = j).$$

Suppose that the probability

$$P(\xi = j) = p_j, \qquad j \in J,$$

$$P(\eta_i = s \mid \xi = j) = q_{ijs}, \qquad i \in I, \quad s \in S, \quad j \in J,$$

$$\sum_{s \in S} q_{ijs} = 1, \qquad i \in I, \quad j \in J.$$

We introduce some additional notation:

- $\overline{s}$ — signal strength vector recorded by the mobile object from the set I access points, or any subset;
- $S(D)$ — the set of all possible signal strength vectors from the subset of access points $D \subset I$.

Formally positioning is described by a family of functions $\{u_D^*(\cdot)\}_{D \subset I}$, assigning to each signal strength vector from access points some zone:

$$u_D^* : S(D) \to V.$$

Next it will be considered two ways of zone selection when signal strength vector is received:

a) positioning function returns most likely zone;

b) positioning function returns the zone whith the minimum conditional expectation of errors positioning.

The values of the random variable $d(v_\xi, u_D^*(\overline{\eta}))$ are called location error. The expectation of location error is

$$F(D) = M(d(v_\xi, u_D^*(\overline{\eta}))) = \sum_{\overline{s} \in S(D)} \sum_{j=1}^{n} P(\xi = j, \overline{\eta} = \overline{s}) d(v_j, u_D^*(\overline{s})).$$

We will use $F(D)$ as a criterion for comparing subsets $D \subset I$ of access points. We will study the change of criteria in the case of adding a new access point.

Now we formulate the optimization problem.

Problem P. Given a number of m. It is required to find a subset of the access points $D \subset I$, such that $|D| = m$ and function $F(D)$ takes a minimum value.

Theorem 1. *Problem P is NP-hard.*

The proof is analogous to that NP-hard of problem presented in the article [16].

Selecting the most likely zone. Let $D \subset I$ be a subset of the access points. Consider the case where the positioning function returns the zone with the maximum a posteriori probability:

$$u_D^*(\overline{s}) = \arg \max_{j=1,\dots,n} \{P(\xi = j \mid \overline{\eta} = \overline{s})\}.$$

Let $\Psi(\overline{s})$ be the set of the most likely zones for signal strength vector $\overline{s} \in S(D)$:

$$\Psi(\overline{s}) = \left\{ j \in J \;\middle|\; P(\xi = j \mid \overline{\eta} = \overline{s}) = \max_{j' \in J} \{P(\xi = j' \mid \overline{\eta} = \overline{s})\} \right\}.$$

Let $\Delta(j, \overline{s})$ be the distance between the zone $j \in J$ (where in fact the object is located) and the farthest of the most likely zones for $\overline{s} \in S(D)$:

$$\Delta(j, \overline{s}) = \max_{j' \in \Psi(\overline{s})} \{d(v_j, v_{j'})\}.$$

Let $\nu(j, \overline{s})$ be the zone farthest from the zone of $j \in J$ of the most likely zones for $\overline{s} \in S(D)$:

$$\nu(j, \overline{s}) = \arg \max_{j' \in \Psi(\overline{s})} \{d(v_j, v_{j'})\}.$$

So $\Delta(j, \overline{s})$ — the distance between the zones v_j and $\nu(j, \overline{s})$.

It's obvious that

$$d(v_j, u_D^*(\overline{s})) \leqslant \Delta(j, \overline{s}).$$

Hence the expectation of determining object location error is bounded from above:

$$F(D) \leqslant \sum_{\overline{s} \in S(D)} \sum_{j \in J} P(\overline{\eta} = \overline{s},\ \xi = j) \Delta(j, \overline{s}).$$

Let us find the conditions under which the addition of access points results in a lack of increase of the upper bound of the objective function $F(D)$.

We will start with the simplest case. Let $D = \emptyset$.

Let $\Psi(\emptyset)$ be the set of numbers most likely zones, i. e. zones with the maximum a priori probability of the mobile object location:

$$\Psi(\emptyset) = \left\{ j \in J \;\middle|\; P(\xi = j) = \max_{j' \in J} \{P(\xi = j')\} \right\}.$$

Let $\Delta(j)$ be the maximum distance between the zone v_j and the most likely zone:

$$\Delta(j) = \max_{j' \in \Psi(\emptyset)} \{d(v_j, v_{j'})\}.$$

The expectation of the distance between a randomly selected zone (in accordance with their a priori probabilities) and the farthest from it the most likely zone is

$$F' = \sum_{j \in J} P(\xi = j) \Delta(j).$$

Suppose that there is only one access point (to be specific with number 1).

Lemma 1. *If for all $s \in S$ exists $j \in J$ wherein*

$$P(\xi = j \mid \eta_1 = s) > \frac{1}{2},$$

then

$$\sum_{s\in S}\sum_{j\in J} P(\eta_1 = s,\ \xi = j)\Delta(j,s) \leqslant \sum_{j\in J} P(\xi = j)\Delta(j).$$

P r o o f. It is obvious that the conditions of lemma, it follows that for any $s \in S$ done $|\Psi(s)| = 1$, that is, for any signal strength, there is exactly one most likely zone. But then the value $\nu(j,s)$ are independent of j, denote them $\nu(s) = \nu(j,s)$.

It's obvious that

$$\sum_{s\in S}\sum_{j\in J} P(\eta_1 = s,\ \xi = j)\Delta(j,s) = \sum_{s\in S}\sum_{j\in J} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)\Delta(j,s) =$$

$$= \sum_{s\in S}\sum_{j\in J\setminus\{\nu(s)\}} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)\Delta(j,s) +$$

$$+ \sum_{s\in S} P(\eta_1 = s)P(\xi = \nu(s) \mid \eta_1 = s)\Delta(\nu(s),s). \tag{3}$$

As $\nu(\nu(s),s) = \nu(s)$ then $\Delta(\nu(s),s) = 0$. Consequently,

$$\sum_{s\in S} P(\eta_1 = s)P(\xi = \nu(s) \mid \eta_1 = s)\Delta(\nu(s),s) = 0. \tag{4}$$

Let $\nu^* \in \Psi(\emptyset)$ be is selected zone. Given the triangle inequality

$$d(j,\ \nu(s)) \leqslant d(j,\ \nu^*) + d(\nu^*,\nu(s)),$$

we get

$$\sum_{s\in S}\sum_{j\in J\setminus\{\nu(s)\}} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)\Delta(j,s) =$$

$$= \sum_{s\in S}\sum_{j\in J\setminus\{\nu(s)\}} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)d(j,\ \nu(s)) \leqslant$$

$$\leqslant \sum_{s\in S}\sum_{j\in J\setminus\{\nu(s)\}} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)(d(j,\ \nu^*) + d(\nu^*,\nu(s))) =$$

$$= \sum_{s\in S}\sum_{j\in J} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)(d(j,\ \nu^*) + d(\nu^*,\nu(s))) -$$

$$- 2\sum_{s\in S} P(\eta_1 = s)P(\xi = \nu(s) \mid \eta_1 = s)d(\nu^*,\nu(s)). \tag{5}$$

Considering (3)–(5), and

$$d(j,\ \nu^*) \leqslant \Delta(j),$$

we get

$$\sum_{s\in S}\sum_{j\in J} P(\eta_1 = s,\ \xi = j)\Delta(j,s) \leqslant \sum_{s\in S}\sum_{j\in J} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)(d(j,\ \nu^*) +$$

$$+ d(\nu^*, \nu(s))) - 2\sum_{s\in S} P(\eta_1 = s)P(\xi = \nu(s) \mid \eta_1 = s)d(\nu^*, \nu(s)) \leqslant$$

$$\leqslant \sum_{s\in S}\sum_{j\in J} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)(\Delta(j) + d(\nu^*, \nu(s))) -$$

$$- 2\sum_{s\in S} P(\eta_1 = s)P(\xi = \nu(s) \mid \eta_1 = s)d(\nu^*, \nu(s)) =$$

$$= \sum_{s\in S}\sum_{j\in J} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)\Delta(j) + \sum_{s\in S}\sum_{j\in J} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)d(\nu^*, \nu(s)) -$$

$$- 2\sum_{s\in S} P(\eta_1 = s)P(\xi = \nu(s) \mid \eta_1 = s)d(\nu^*, \nu(s)) = \sum_{s\in S}\sum_{j\in J} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s) \times$$

$$\times\ \Delta(j) + \sum_{s\in S} P(\eta_1 = s)d(\nu^*, \nu(s)) \left(\sum_{j\in J} P(\xi = j \mid \eta_1 = s) - 2P(\xi = \nu(s) \mid \eta_1 = s)\right).$$

The first term in the resulting sum is simplified as follows:

$$\sum_{s\in S}\sum_{j\in J} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)\Delta(j) = \sum_{j\in J}\sum_{s\in S} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s)\Delta(j) =$$

$$= \sum_{j\in J}\Delta(j)\sum_{s\in S} P(\eta_1 = s)P(\xi = j \mid \eta_1 = s) = \sum_{j\in J} P(\xi = j)\Delta(j) = F'.$$

Considering

$$\sum_{j\in J} P(\xi = j \mid \eta_1 = s) = 1,$$

we get

$$\sum_{s\in S}\sum_{j\in J} P(\eta_1 = s,\ \xi = j)\Delta(j, s) \leqslant F' + \sum_{s\in S} P(\eta_1 = s)d(\nu^*, \nu(s))(1 - 2P(\xi = \nu(s) \mid \eta_1 = s)).$$

Since all $s \in S$

$$P(\xi = \nu(s) \mid \eta_1 = s) > \frac{1}{2},$$

then

$$1 - 2P(\xi = \nu(s) \mid \eta_1 = s) < 0,$$

and

$$\sum_{s\in S} P(\eta_1 = s)d(\nu^*, \nu(s))(1 - 2P(\xi = \nu(s) \mid \eta_1 = s)) \leqslant 0.$$

Consequently

$$\sum_{s\in S}\sum_{j\in J} P(\eta_1 = s,\ \xi = j)\Delta(j, s) \leqslant F'.$$

The lemma is proved. □

Here is an example in which the conditions of Lemma 1 are not met. Suppose there are three zones (v_1, v_2), one access point and two signal strength values (s_1, s_2). Let the distance between the zones be

$$d(v_1, v_2) = 10, \quad d(v_2, v_3) = 10, \quad d(v_1, v_3) = 20.$$

Let the a priori probability zones are as follows:

$$P(\xi = 1) = 0.3, \quad P(\xi = 2) = 0.4, \quad P(\xi = 3) = 0.3.$$

Assume that the conditional probability of the signal strength in the zones are

$$P(\eta_1 = s_1 \mid \xi = 1) = 0.75, \quad P(\eta_1 = s_1 \mid \xi = 2) = 0.5, \quad P(\eta_1 = s_1 \mid \xi = 3) = 0.25,$$

$$P(\eta_1 = s_2 \mid \xi = 1) = 0.25, \quad P(\eta_1 = s_2 \mid \xi = 2) = 0.5, \quad P(\eta_1 = s_2 \mid \xi = 3) = 0.75.$$

It is easy to verify that in this case, $F' = 6$, $F(D) = 7$ and $F' < F(D)$, that is, the use of access points reduces the accuracy of the location.

Let us turn to the case when $D \neq \emptyset$.

Theorem 2. *Let $D \subset I$, $i \in I$, $i \notin D$ and for all $\overline{s} \in S(D \cup \{i\})$ exist $j \in \{1, \ldots, n\}$, that $P(\xi = j \mid \overline{\eta} = \overline{s}) \geqslant \frac{1}{2}$ then*

$$\sum_{\overline{s} \in S(D \cup \{i\})} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \xi = j) \Delta(j, \overline{s}) \leqslant \sum_{\overline{s} \in S(D)} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \xi = j) \Delta(j, \overline{s}).$$

P r o o f. Let the signal strength vector $\overline{s} \in S(D)$ from access points of set $D \subset I$ is fixed.

We denote $P'(A) = P(A \mid \overline{\eta} = \overline{s})$ the conditional probability of some event A given signal strength vector $\overline{s}$ from access points of set D.

Let i be index added access point.

Suppose that for all $s \in S$ there is a zone $v = \nu'(s)$, which satisfies $P'(\xi = j \mid \eta_i = s) > \frac{1}{2}$.

Let $\Delta'(j)$ be the distance between the zone v_j and the outermost zone $v_{j'}$ with the maximum probability value $P'(\xi = j')$, $\Delta'(j, s)$ — the distance between the zone v_j and zone $\nu'(s)$.

Then, according to the Lemma 1,

$$\sum_{s \in S} \sum_{j \in J} P'(\eta_i = s, \xi = j) \Delta'(j, s) \leqslant \sum_{j \in J} P'(\xi = j) \Delta'(j).$$

Multiply the inequality by $P(\overline{\eta} = \overline{s})$, we get

$$\sum_{s \in S} \sum_{j \in J} P'(\eta_i = s, \xi = j) P(\overline{\eta} = \overline{s}) \Delta'(j, s) \leqslant \sum_{j \in J} P'(\xi = j) P(\overline{\eta} = \overline{s}) \Delta'(j).$$

Given that

$$P'(\eta_i = s, \xi = j) P(\overline{\eta} = \overline{s}) = P(\eta_i = s, \xi = j \mid \overline{\eta} = \overline{s}) P(\overline{\eta} = \overline{s}) = P(\overline{\eta} = \overline{s}, \eta_i = s, \xi = j)$$

and

$$P'(\xi = j) P(\overline{\eta} = \overline{s}) = P(\xi = j \mid \overline{\eta} = \overline{s}) P(\overline{\eta} = \overline{s}) = P(\overline{\eta} = \overline{s}, \xi = j),$$

we have the inequality

$$\sum_{s \in S} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \eta_i = s, \xi = j) \Delta'(j, \eta_i = s) \leqslant \sum_{j \in J} P(\overline{\eta} = \overline{s}, \xi = j) \Delta'(j).$$

Since $P'(\xi = j) = P(\xi = j \mid \overline{\eta} = \overline{s})$, that $\Delta'(j) = \Delta(j, \overline{s})$. Seeing

$$P'(\xi = j \mid \eta_i = s) = P(\xi = j \mid \overline{\eta} = \overline{s}, \eta_i = s)$$

we get $\Delta'(j,s) = \Delta(j,(\overline{\eta} = \overline{s}, \eta_i = s))$ and

$$\sum_{s\in S}\sum_{j\in J} P(\overline{\eta} = \overline{s}, \eta_i = s, \xi = j)\Delta(j,(\overline{s},s)) \leqslant \sum_{j\in J} P(\overline{\eta} = \overline{s}, \xi = j)\Delta(j,\overline{s}). \qquad (6)$$

Inequality (6) holds for any $\overline{s} \in S(D)$, hence

$$\sum_{\overline{s}\in S(D)}\sum_{s\in S}\sum_{j\in J} P(\overline{\eta} = \overline{s}, \eta_i = s, \xi = j)\Delta(j,(\overline{s},s)) \leqslant \sum_{\overline{s}\in S(D)}\sum_{j\in J} P(\overline{\eta} = \overline{s}, \xi = j)\Delta(j,\overline{s}),$$

or the same

$$\sum_{\overline{s}\in S(D\cup\{i\})}\sum_{j\in J} P(\overline{\eta} = \overline{s}, \xi = j)\Delta(j,\overline{s}) \leqslant \sum_{\overline{s}\in S(D)}\sum_{j\in J} P(\overline{\eta} = \overline{s}, \xi = j)\Delta(j,\overline{s}).$$

This proves the theorem. □

Lemma 1 is a special case of Theorem 2. To show this is sufficient to put in the theorem $D = \emptyset$.

It can be concluded that the way, when the positioning function returns most likely zone, has the following disadvantage. Expectation of location error may increase, when the new access point is added.

Selecting the zone with a minimum expectation errors. Let positioning function returns the zone for which will be the minimum conditional expectation of errors positioning:

$$u_D^*(\overline{s}) = \arg\min_{j=1,\ldots,n}\left\{\sum_{k=1}^{n} P(\xi = j \mid \overline{\eta} = \overline{s})d(v_j, v_k)\right\}.$$

In this case, the task of finding the point $u_D^*(\overline{s})$ is called Fermat—Torricelli—Steiner problem.

Introduce the zone designation, for which will be the minimum expectation to randomly selected zones:

$$u^* = \arg\min_{u\in V}\left\{\sum_{j\in J} P(\xi = j)d(u, v_j)\right\}.$$

Lemma 2. *For all $D \subset I$*

$$\sum_{\overline{s}\in S(D)}\sum_{j\in J} P(\overline{\eta} = \overline{s},\ \xi = j)d(v_j, u_D^*(\overline{s})) \leqslant \sum_{j\in J} P(\xi = j)d(v_j, u^*).$$

P r o o f. Using elementary calculations, we get

$$\sum_{\overline{s}\in S(D)}\sum_{j\in J} P(\overline{\eta} = \overline{s},\ \xi = j)d(v_j, u_D^*(\overline{s})) = \sum_{\overline{s}\in S(D)}\sum_{j\in J} P(\overline{\eta} = \overline{s})P(\xi = j \mid \overline{\eta} = \overline{s})d(v_j, u_D^*(\overline{s})) =$$

$$= \sum_{\overline{s}\in S(D)}\left(P(\overline{\eta} = \overline{s})\sum_{j\in J} P(\xi = j \mid \overline{\eta} = \overline{s})d(v_j, u_D^*(\overline{s}))\right) \leqslant$$

$$\leqslant \sum_{\overline{s}\in S(D)}\left(P(\overline{\eta} = \overline{s})\sum_{j\in J} P(\xi = j \mid \overline{\eta} = \overline{s})d(v_j, u^*)\right) =$$

$$= \sum_{\overline{s} \in S(D)} \sum_{j \in J} P(\overline{\eta} = \overline{s}) P(\xi = j \mid \overline{\eta} = \overline{s}) d(v_j, u^*) =$$

$$= \sum_{j \in J} \sum_{\overline{s} \in S(D)} P(\xi = j) P(\overline{\eta} = \overline{s} \mid \xi = j) d(v_j, u^*) =$$

$$= \sum_{j \in J} \left(P(\xi = j) d(v_j, u^*) \sum_{\overline{s} \in S(D)} P(\overline{\eta} = \overline{s} \mid \xi = j) \right) = \sum_{j \in J} P(\xi = j) d(v_j, u^*).$$

The lemma is proved. □

Substantially Lemma 2 means, that the average value of the error location determination of the object will not be greater if the system uses the location data received from the access points.

We now show that when adding new access points expectation determining object location error does not increase.

Theorem 3. *Let $D \subset I$, $i \in I$, $i \notin D$ and*

$$u_D^*(\overline{s}) = \arg \min_{u \in V} \left\{ \sum_{j \in J} P(\xi = j \mid \overline{\eta} = \overline{s}) d(v_j, u) \right\},$$

then

$$\sum_{\overline{s} \in S(D \cup \{i\})} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \xi = j) d(v_j, u_{D \cup \{i\}}^*(\overline{s})) \leqslant \sum_{\overline{s} \in S(D)} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \xi = j) d(v_j, u_D^*(\overline{s})).$$

P r o o f. Let the signal strength vector $\overline{s} \in S(D)$ from access points of set $D \subset I$ is fixed.

Let $P'(A) = P(A \mid \overline{\eta} = \overline{s})$ be the conditional probability of any event A given signal strength vector $\overline{s}$ from access points of set D.

Let i be index added access point.

We introduce the notation

$$u'(s) = \arg \min_{u \in V} \left\{ \sum_{j \in J} P'(\xi = j \mid \eta_i = s) d(u, v_j) \right\}$$

and

$$u' = \arg \min_{u \in V} \left\{ \sum_{j \in J} P'(\xi = j) d(u, v_j) \right\}.$$

Then, according to the Lemma 2,

$$\sum_{s \in S} \sum_{j \in J} P'(\eta_i = s,\ \xi = j) d(v_j, u'(s)) \leqslant \sum_{j \in J} P'(\xi = j) d(v_j, u'). \tag{7}$$

Multiply the inequality (7) by $P(\overline{\eta} = \overline{s})$, we get

$$\sum_{s \in S} \sum_{j \in J} P'(\eta_i = s, \xi = j) P(\overline{\eta} = \overline{s}) d(v_j, u'(s)) \leqslant \sum_{j \in J} P'(\xi = j) P(\overline{\eta} = \overline{s}) d(v_j, u').$$

Given that

$$P'(\eta_i = s, \xi = j)P(\overline{\eta} = \overline{s}) = P(\eta_i = s, \xi = j \mid \overline{\eta} = \overline{s})P(\overline{\eta} = \overline{s}) = P(\overline{\eta} = \overline{s}, \eta_i = s, \xi = j)$$

and

$$P'(\xi = j)P(\overline{\eta} = \overline{s}) = P(\xi = j \mid \overline{\eta} = \overline{s})P(\overline{\eta} = \overline{s}) = P(\overline{\eta} = \overline{s}, \xi = j),$$

we have

$$\sum_{s \in S} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \eta_i = s, \xi = j) d(v_j, u'(s)) \leqslant \sum_{j \in J} P(\overline{\eta} = \overline{s}, \xi = j) d(v_j, u').$$

Since $P'(\xi = j) = P(\xi = j \mid \overline{\eta} = \overline{s})$, then $u' = u^*_D(\overline{s})$. It's obvious that $u'(s) = u^*_{D \cup \{i\}}(\overline{s}, s)$ and

$$\sum_{s \in S} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \eta_i = s, \xi = j) d(v_j, u^*_{D \cup \{i\}}(\overline{s}, s)) \leqslant \sum_{j \in J} P(\overline{\eta} = \overline{s}, \xi = j) d(v_j, u^*_D(\overline{s})). \quad (8)$$

The inequality (8) holds for any $\overline{s} \in S(D)$, so

$$\sum_{\overline{s} \in S(D)} \sum_{s \in S} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \eta_i = s, \xi = j) d(v_j, u^*_{D \cup \{i\}}(\overline{s}, s)) \leqslant$$

$$\leqslant \sum_{\overline{s} \in S(D)} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \xi = j) d(v_j, u^*_D(\overline{s})),$$

or the same

$$\sum_{\overline{s} \in S(D \cup \{i\})} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \xi = j) d(v_j, u^*_{D \cup \{i\}}(\overline{s})) \leqslant \sum_{\overline{s} \in S(D)} \sum_{j \in J} P(\overline{\eta} = \overline{s}, \xi = j) d(v_j, u^*_D(\overline{s})).$$

The theorem is proved. □

Solution algorithm. NP-difficult problem of choosing locations for access points means that it is necessary to apply practical solutions heuristics. For many combinatorial optimization problems solution method is based on the sequential execution of the greedy algorithm and local search. The following algorithm is an adaptation of this method for solving the problem of locating positioning system access points.

Greedy algorithm. At the beginning we search an optimal location of a single access point. This position is fixed and is not changed. Then, given the location of the first access point, we search an optimal position of the second access point. This position is also fixed, and so it goes on, until we have found the location of all access points. Thus, in the greedy algorithm locations of all access points are consistently determined.

Steps of greedy algorithm are:

STEP 0. Let $D_0 = \emptyset$. Go to STEP 1.

STEP k. We find an access point $i \in I$, for which the minimum value of the objective function $F(D_{k-1} \cup \{i\})$. Let

$$D_k = D_{k-1} \cup \{i\}.$$

If $k = m$, then let $D = D_k$ and STOP; else otherwise go to STEP $k + 1$.

Local search. With a greedy algorithm we search a combination of locations for m access points. Then we change placements each access point trying to improve the value

of the objective function. This continues for as long as the change in the position of each access point will not lead to an improvement in the objective function.

Local search steps are:

STEP 1. Initial placement of m access points D_0 we form using a greedy algorithm. Let $k = 0$. Go to STEP 2.

STEP 2. Let $k' = k$, $D' = D_k$. Loop through all access points $i' \in D'$. For all i' loop over all $i \in I$. Thus, if

$$F(D_k \backslash \{i'\} \cup \{i\}) < F(D_k),$$

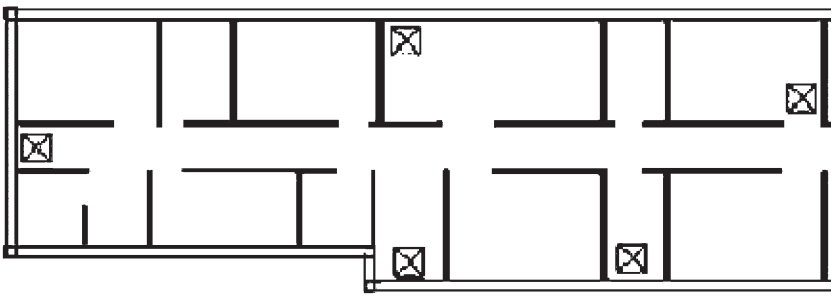
then assign

$$D_{k+1} = D_k \backslash \{i'\} \cup \{i\},$$

and let $k = k + 1$.

STEP 3. If $k' = k$, then let $D = D_k$ and STOP; else go to STEP 2.

This approach lies at the basis of the developed software system. We tested the developed algorithm and software system on real data. Part of second floors of IT-park of Petrozavodsk State University with an area of 550 m^2 was divided into 80 zones. Figure shows five access points placements that have been determined by developed algorithm.



Placement of five access points on the floor of the IT-park
of Petrozavodsk State University

Testing location system showed that the resulting placement of access points provided in the 85% of the error in the detection of mobile objects not exceeding two meters. Such precision location was recognized quite valid and confirmed the practical applicability of the developed algorithm.

Conclusion. The proposed approach of determining the locations of the access points may be used in indoor positioning systems based on the signal strength map of the access points, for example, Wi-Fi, ZigBee, Bluetooth, nanoLOC, and generalized by using other types of sensors.

In addition, the proposed methods can be used to identify zones of the building in which the mobile object positioning error is greatest. From a practical point of view, it generates a recommendations on how to add the access point to improve the accuracy of mobile objects positioning. Described approaches and algorithms have been applied in the development of the indoor positioning technology RealTrac.

References

1. Morozova N. S. Virtual'nye formacii i virtual'nye lidery v zadache o dvizhenii stroem gruppy robotov [Virtual formations and virtual leaders in formation control problem for group of robots]. *Vestnik*

of Saint Petersburg University. Series 10. Applied mathematics. Computer science. Control processes, 2015, issue 1, pp. 135–148. (In Russian)

2. Sotnikova M. V. Algoritm avtomaticheskogo uderzhaniya kolesnogo robota na vizual'no zadannoj linii [Algorithm for visual path following by wheeled fully actuated mobile robot]. *Vestnik of Saint Petersburg University. Series 10. Applied mathematics. Computer science. Control processes*, 2016, issue 1, pp. 99–108. (In Russian)

3. Chandrasekaran G., Ergin M., Yang J., Lui S., Chen Y., Gruteser M., Martin R. Empirical evaluation of the limits of localization using signal strength: Beyond Cramer—Rao bounds. *Processing of IEEE SECON*, 2004, pp. 406–414.

4. Elnahrawy E., Li X., Martin R. The limits of localization using signal strength: A comparative study. *Processing of IEEE SECON*, 2004, pp. 406–414.

5. Galov A., Moschevikin A. Bayesian filters for ToF and RSS measurements for indoor positioning of a mobile object. *Processing of the Intern. Conference on Indoor Positioning and Indoor Navigation* (IPIN-2013). Montbeliard, France, October 28–31, 2013, pp. 310–317.

6. Galov A. S., Moschevikin A. P., Voronov R. V. Combination of RSS localization and ToF ranging for increasing positioning accuracy indoors. *Processing of the 11th Intern. Conference on ITS Telecommunications* (ITST), 2011, pp. 299–304.

7. Ata O. W., Ala'Eddin M. S., Jawadeh M. I., Amro A. I. An indoor propagation model based on a novel multi wall attenuation loss formula at frequencies 900 MHz and 2.4 GHz. *Wireless Personal Communications*, 2013, vol. 69, no. 1, pp. 23–36.

8. Chrysikos T., Georgopoulos G., Kotsopoulos S. Site-specific validation of ITU indoor path loss model at 2.4 GHz. *IEEE Intern. Symposium on a World of Wireless, Mobile and Multimedia Networks & Workshops*, 2009, pp. 1–6.

9. Molisch A. F., Balakrishnan K., Chong C. C., Emami S., Fort A., Karedal J., Schantz H., Schuster U. IEEE 802.15.4a *channel model-final report*. Technical Report, Document IEEE 802.1504-0062-02-004a, 2005.

10. Battiti R., Brunato M., Delai A. Optimal wireless access point placement for location-dependent services. Technical Report, 2003. Available at: http://eprints.biblio.unitn.it/489/1/DIT-03-052-withCover.pdf (accessed: 30.11.2016).

11. Chen Y., Francisco J. A., Trappe W., Martin R. P. A practical approach to landmark deployment for indoor localization. *Sensor and Ad Hoc Communications and Networks*, 2006. *SECON'06. 3rd Annual IEEE Communications Society on IEEE*, 2006, vol. 1, pp. 365–373.

12. Gondran A., Caminada A., Fondrevelle J., Baala O. Wireless LAN planning: a didactical model to optimise the cost and effective payback. *Intern. Journal of Mobile Network Design and Innovation*, 2007, vol. 2, no. 1, pp. 13–25.

13. Kim T., Shin J., Tak S. Cell planning for indoor object tracking based on RFID. *Mobile Data Management: Systems, Services and Middleware, MDM'09. Tenth Intern. Conference on IEEE*, 2009, pp. 709–713.

14. Farkas K., Huszák Á., Gódor G. Optimization of Wi-Fi access point placement for indoor localization. *Journal IIT* (*Informatics & IT Today*), 2013, vol. 1, no. 1, pp. 28–33.

15. Vilović I., Burum N. Location optimization of WLAN access points based on a neural network model and evolutionary algorithms. *Automatika*: *časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije* [*Automatika*: *Journal for control, measurement, electronics, computing and communications*], 2015, vol. 55, no. 3, pp. 317–329.

16. Voronov R. V., Moschevikin A. P. Primenenie uslovnoj jentropii pri formirovanii rekomendacij po razmeshheniju bazovyh stancij v lokal'nyh sistemah pozicionirovanija [Use of conditional entropy for optimal disposition of base stations in local positioning systems]. *Information Technology*, 2014, no. 10, pp. 11–16. (In Russian)

For citation: Voronov R. V. The problem of optimal placement of access points for the indoor positioning system. *Vestnik of Saint Petersburg University. Applied mathematics. Computer science. Control processes*, 2017, volume 13, issue 1, pp. 61–73. DOI: 10.21638/11701/spbu10.2017.106

Статья рекомендована к печати проф. А. П. Жабко.

Статья поступила в редакцию 11 декабря 2016 г.

Статья принята к печати 19 января 2017 г.