Type-2 Fuzzy Rule-Based Model of Urban Metro Positioning Service

A.R. Gimaletdinova <argimaletdinova@edu.hse.ru>
K.Y. Degtiarev <kdegtiarev@hse.ru>
National Research University Higher School of Economics (HSE),
3, Kochnovsky Proezd, Moscow, 125319, Russian Federation

Abstract. In the last few years there has been a growing interest in route building oriented mobile applications with the following features of navigation and sending timely notifications about arrival. Despite the large body of existing knowledge on navigational services, there has been an important issue relative to positioning accuracy. The paper discusses a possible solution to comparison problem, which is linked to the determination of the closeness to destination metro station through finding a difference between user's current coordinates and fixed-point coordinates. With this end in view, fuzzy logic approach is used to develop Routes Recommender System (RRS) that utilizes linguistic variables to express the vague and uncertain term 'closeness to…'. The paper provides detailed explanation of each variable considered in the fuzzy inference system (FIS), set of fuzzy rules in line with graphical representation of system's output. Based on Mamdani model, we propose a set of test cases to check maintainability of the model and provide a description about received results. At a later time, an Android-based mobile application aimed at public transport route building will be developed whose notification system will be based on our model`s implementation presented. It should be emphasized that the paper examines potentials of the modeling approach based on interval type-2 fuzzy sets (IT2FS) that attract much attention these days in various research studies and conventional Mamdani fuzzy inference system (MFIS) as applied to real and rather topical problem. The significance of developing such models may be of a high demand for appropriate representation of factors that are inherently vague and uncertain. Hence, this study may also contribute to future research on similar topics.

Keywords: positioning service; mobile applications; fuzzy modeling; GPS; WiFi; cellular networks; public transport; interval type-2 fuzzy sets; fuzzy inference system; fuzzy matching of coordinates; uncertainty; fuzziness

DOI: 10.15514/ISPRAS-2017-29(4)-6

For citation: Gimaletdinova A.R., Degtiarev K.Y. Type-2 Fuzzy Rule-Based Model of Urban Metro Positioning Service. *Trudy ISP RAN/Proc. ISP RAS*, vol. 29, issue 4, 2017, pp. 87-106. DOI: 10.15514/ISPRAS-2017-29(4)-6

1. Introduction

Over the past decade positioning techniques have become common in almost all branches of industry. In particular, nowadays vast majority of phone models are provided with GPS-module that can be enabled in different cases. Positioning feature is rather common to mobile applications supporting navigational services, and the latter can be used by people in urban transport. The purpose of the paper is to exploit the potentialities of fuzzy logic regarding recommender system with the navigational service. Such service may solve the problem of frequently encountered disorientation of passengers in unfamiliar terrain and allow to pave routes between stations of interest (case of urban transportation system). The potential application may notify a passenger about forthcoming arrival, when he/she is situated closely to the end station. The main purpose in the present context is to determine a deviation between current and end points (stations). Consequently, it leads to the serious problem, since we cannot precisely assert whether a user is close to the end station or not. It occurs because there is a need to estimate the smallest difference (delta) between current and end-point coordinates and then set rule(-s) to classify user`s location.

The issue of applying fuzzy logic to positioning, tracking and transportation attracts attention of researchers. Selected publications have focused on indoor positioning. For example, Chen C.-Y., Yung J., et al. [1] studied indoor positioning technique based on received signal strength and fuzzy approach; they showed experimentally that such method has better performance as compared to geometric triangulation method [2] – actually, the same objective was pursued in the research by Teuber A. and Eisfelller B. [2]. The fuzzy system to control train automatic stop, with the emphasis on stop accuracy, was developed by Yasunobu S., Miyamoto S. and Ihara H. in [3]. It is evident that the practical application of fuzzy logic to positioning or transportation subject matter cannot be considered as exclusive one, however, the issue of positioning in metro should be studied in detail.

As it was mentioned above, the study is devoted to indoor positioning within the metro transportation system. We make an attempt to develop a fuzzy model of metro stops allowing to send timely destination notifications to passenger. It is clear that we do not know exact minimum and maximum distances between stations or the moment when the application should send a reminder. Uncertainty has many faces and forms of manifestation. As stated by George J. Klir and Mark Wierman, "uncertainty involved in any problem-solving situation is a result of some information deficiency; … information may be incomplete, fragmentary, not fully reliable, vague, contradictory, or deficient in some other way" [4]. Hence, when we do not know or cannot obtain exact values/parameters of some phenomena (e.g. distances between points, the location of some moment on a time scale), we need to deviate from type-1 fuzzy sets as a general framework to handle *vagueness* (for more information see seminal papers "Fuzzy Sets" (1965) and "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning – I" (1975) by L. Zadeh) to more general type-2 fuzzy sets that allow to reflect the uncertainty in adequate, more thorough manner, or, put it precisely, to model it. In the work interval type-2 fuzzy sets (IT2FS) are used; to ensure computational efficiency, the preference is given mainly to piecewise linear functions (trapezoidal shape) as upper and lower membership functions of IT2FS.

The rest of the paper is organized as follows: the second section explains the main problem that the paper is devoted to. Section 3 provides definition of linguistic variable (LV) and describes those variables and their linguistic values represented in the form of type-2 membership functions that are used in the inference process (Mamdani's fuzzy model); explanations on domains (universal sets) for each variable are also adduced in this section. The following section 4 makes emphasis on fuzzy rules that serve as a basis for fuzzy system (model developed), covers short comments on type-reduction defuzzification methods used in the study; results of experiments with the system under different values of input variables are presented in both tabular and graphical forms. The last, 5th section of the paper concludes explicitly mentioning the ways of further elaborating upon the subject.

2. Problem definition and general comments

One of the main issues we have to deal with is to find a user's position. Current position obtained should be compared with fixed station's coordinates (e.g. end-point of the route or interchange point to other line) – it will allow to say where is a user now. If he/she is close to one of the points, the application should signal to him about it, thus the understanding and definition of the word "close" becomes essential. The factor of closeness is treated unequally by different people, and a nearby object for one person can be far away for another one. It means that estimation of closeness relates to certain difficulties and, as a consequence, we cannot associate crisp numbers as a basis for possible values of the variable "close". Therefore, the only way to describe closeness at a first approximation is to set a numerical interval of its possible values and to use it at further processing steps.

We may assume that in the beginning the application gets start and end points of the route (input data), then it ensures passenger tracking using one of the positioning technologies (GPS, WiFi, Cellular Networks) and compares his/her current location with the one of key points. According to [5] and practical everyday experience, GPS

Table 1. Approximate accuracy for different positional techniques [6, 7]

№	Technique	Min accuracy (m)	Medium accuracy (m)	Maximum accuracy (m)
1	GPS	2	11	20
2	WiFi	10	80	150
3	Cellular	100	800	1500
	Average	37.3	297	557
	Average for №2 and №3*	55	440	825

* The last row is calculated without GPS characteristics (signal in metro is bad)

has poor accuracy indoors, including metro, therefore, we do not consider GPS positioning accuracy to calculations shown in Table 1. As already mentioned before, we will use numeric interval to represent difference between fixed and current coordinates.

Received data concerning current position can be inaccurate, because positioning techniques used in the phone do not guarantee 'ideal' precision of geographical coordinates supplied because of various objective reasons (e.g. tracking indoors or underground, inferior quality of signal from provider, etc.). Thus, it makes sense to emphasize another overt source of *fuzziness*, which relates to fuzzy (vague) matching of coordinates – latitude and longitude indicators will be analyzed separately.

3. Fuzzy logic model: definition of linguistic variables and their values

Firstly, we should select input-output variables for fuzzy system and provide necessary explanations. All significant internal and external factors, in which uncertainty shows itself, must be analyzed; this is an important stage in development of the model. Internal factors signify that certain issues depend solely on application itself (its realization), and some tuning steps can lead to better results. On the contrary, external factors indicate that there is obvious reality that is not dependent on realization per se – these are the factors that most people are familiar with, viz. bad quality of signal from provider, poor WiFi coverage, etc.

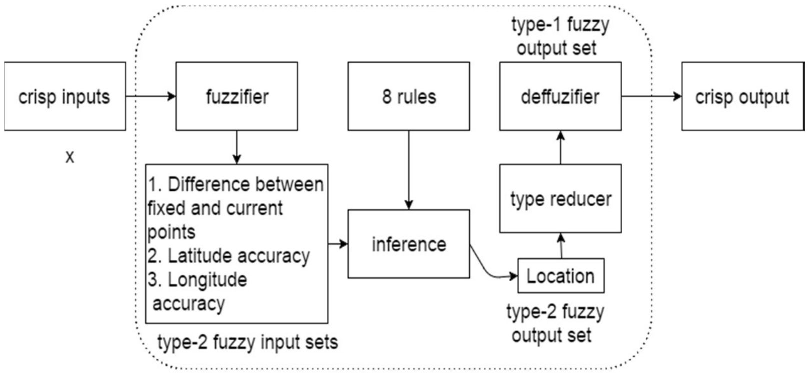

Fig. 1. Fuzzy model with names of linguistic variables in use

In order to explain internal factors, it is necessary to detect variables at the level of passenger's tracking; the central operation here is obtaining a current position. Once it is done the difference between fixed point coordinates and current point must be determined – we call this difference (i.e. variable) "*Difference between fixed and current points (delta)*".

If we use WiFi and Cellular Networks, it means that the application is going to get coordinates with different accuracy, even at the same place without any movements.

We have to admit that the factor of fuzziness definitely becomes apparent in the problem of coordinate matching, and the accuracy will depend on chosen positioning method (see Table 1). We will combine two techniques mentioned above, and because of that Table 1 contains cells with calculated average accuracy. In the paper, we take into account possible accuracy of latitude and longitude – let's name these variables as "*Latitude accuracy*" and "*Longitude accuracy*". The output of fuzzy system will represent position respective to metro station. All these variables as inputs and output of rule-based system to be used are shown in Fig. 1.

Variables mentioned above in the text are *linguistic variables*, i.e. their values are words (or, phrases) of natural language; formally, these values are fuzzy sets, and they are represented by membership functions. In general, a linguistic variable is defined as a tuple $\langle L_v, T(L_v), U, G, M \rangle$, where L_v is the name of the variable (e.g. $L_v \equiv$ "*Latitude accuracy*"), $T(L_v)$ is the set of labels of variable's L_v linguistic values $l_1,..,l_n$ (term-set of L_v; e.g. $l_i \equiv$ 'insignificant', etc.). The names (labels) are generated using syntactic rule G, the meaning $M(l_i)$ is associated with each value l_i, $i=\overline{1,n}$, from $T(L_v)$; $M(l_i)$ is a fuzzy set (respective membership function) defined on a universe of discourse (domain) U. The latter must be defined for all input and output variables introduced earlier. Thus, every variable is characterized by its own set of acceptable values and membership functions for each such value l_i, $i=\overline{1,n}$.

3.1 Linguistic variable "delta" and its values

Earlier we were talking about the difference between fixed and current points (so-called *delta*). What does it really mean? The value that expresses the difference falls into the interval $[0,a]$, where real-valued $a>0$ (deviation is analyzed in absolute magnitude); its left bound (0) means that passenger's coordinates are similar (better to say, close) to some fixed point. We assume that the application should notify a passenger outright before a given destination, when he/she is at the station that precedes terminal station of the route, or at some later moment. Consequently, we consider the average distance between two stations, and a passenger should have enough time to alight from the railway (metro) carriage without effort.

To calculate the biggest difference between coordinates, we should estimate the average distance between any two stations in the metro and double it, because at this moment it will be not an urgent question to notify a passenger about the arrival as he/she still has to go two or more stations more. Following [8, 9], the mean distance between stations in Moscow metro is equal approximately to 1,780 meters. Hence, *a* value signifies the biggest possible difference, i.e. 40,075,000 meters (the length of Earth's equator) = 360° (circle grade measure).

$$1,780 \times 2 = \frac{1,780 \cdot 2 \cdot 360^\circ}{40,075,000} \approx 0.032^\circ \qquad (1)$$

Therefore, values of *delta* are limited to the interval $[0, 0.032]$ (in degrees) that relates to domain (universal set) U, over which linguistic variable $L_v^{(1)} \equiv$ "*delta*" is defined. Yet, why do we talk about linguistic variable in that case? In the everyday life people prefer to use words or phrases of the natural language as a habitual terms (values) for description of phenomena they are dealing with in their diverse activities. In case of *delta* variable such attached to it terms as 'big', 'small', etc., on one hand, form a solid ground for communication within the professional medium allowing almost uniform apprehension of the meaning of these values. On the other hand, their inherent uncertainty has to be adequately modeled when used in computational methods. In particular, we may introduce 2 linguistic terms 'small' and 'bigger' (difference between coordinates) as applied to the variable *delta*. Since type-1 membership functions (T1MF) are precise, i.e. the degree of belongingness $\mu(x)$ of each generic element x to corresponding fuzzy set is a crisp number, T1MF cannot represent the typical uncertainty intrinsic to estimates $\tilde{\mu}(x)$ (tilde sign emphasizes the fact that these degrees are not reducible to ordinary numbers).

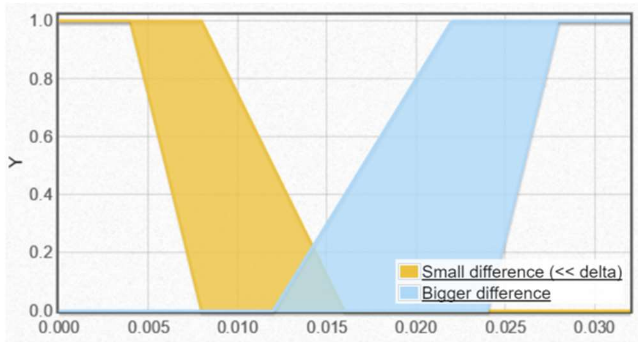

Fig. 2. Difference between fixed and current points (values of "delta")

Linguistic values can be represented in the form of interval type-2 fuzzy sets (IT2FS); the latter are characterized by Lower (L) and Upper (U) membership functions that bound the area called footprint of uncertainty (FOU). The shape of this region allows to express the uncertainty in $\tilde{\mu}(x)$ estimates obtained, providing "additional degrees of freedom … to handle MF uncertainties" [10]. For each $x \in U$, where U is a universe of discourse under consideration, all points in the range $[\mu^{(L)}(x), \mu^{(U)}(x)]$ may have equal unitary weights, i.e. secondary membership function defined on this interval is constant one. For practical reasons, such IT2FS seem to be convenient enough, accurate from the standpoint of giving proper weigh to uncertainty represented and most easily understood by stakeholders. Henceforth, just this kind of T2FS is used in the model with the direction of attention toward piecewise-linear type (trapezoidal case) of L and U membership functions.

Firstly, it is needed to define trapezoidal MF in terms of L and U functions' parameters for each linguistic value (term) – all calculations are done in accordance with (1). We assume that $L_v^{(1)} \equiv "delta"$ is associated with the term-set $T(L_v^{(1)}) = \{l_1, l_2\} = \{\text{'small difference','bigger difference'}\}$ with 2 elements (Fig.2). The upper function (U) for the term 'small difference' of the variable *delta* can be characterized by parameter's set A(0,0), B(0,1), C(0.008,1) and D(0.016,0); the x-coordinate of the point C is the average of x-coordinates of parameters B (B_x) and D (D_x), the latter is the distance between any 2 stations. In much the same way, for the lower function (L) corresponding parameters are A(0,0), B(0,1), C(0.004,1) and D(0.008,0). The 4-tuple of the upper function (U) that represents the linguistic term 'bigger difference' of *delta* is A(0.012,0), B(0.022,1), C(0.032,1) and D(0.032,0), where $A_x = \left(D_x^{(L'small')} + D_x^{(U'small')}\right)/2 = 0.012^\circ$, both x-coordinate C_x and D_x are set to maximum difference 0.032 (1), the value of B_x is calculated as a mean of two neighboring points $\left(A_x + C_x\right)/2 = 0.022^\circ$. It's worth noting that not-yet-application will receive latitude and longitude coordinates as input data, so values are bound to degrees, but not meters. For the lower function (L) set of its parameters takes the form A(0.024,0), B(0.028,1), C(0.032,1), D(0.032,0); again, the value that relates to maximum difference appears here, the B_x value is obtained much as shown above, and A_x equals to the sum of $A_x^{(U'bigger')}$ and the width of the left tail constituting an approximate half of the distance between stations (890 m) converted to degrees.

3.2 Linguistic variable "latitude/longitude accuracy" and its values (terms)

$L_v^{(2)} \equiv "latitude/longitude\ accuracy"$ is the next variable to consider in the paper. As the telephone receives positioning information due to a correction to be made for the accuracy, it must be taken into account in calculation of difference between fixed and current points. The variable $L_v^{(2)}$ is defined on the interval $[0, b]$, where $b > 0$ is the maximum of average accuracy as shown in the last row of Table 1. The not-yet-application doesn't allow to use GPS in metro, so we consider combined usage of WiFi and Cellular Networks. All calculations shown below are based on values summarized in Table 1, and they are performed in line with (1), i.e.

$$440\ \text{m} \approx 0.00395^\circ;\ 825\ \text{m} \approx 0.007^\circ$$
$$55\ \text{m} \approx 0.00049^\circ;\ \text{mean of min and medium}$$
$$(55 + 440\ \text{m}) \approx 0.00444^\circ \tag{2}$$

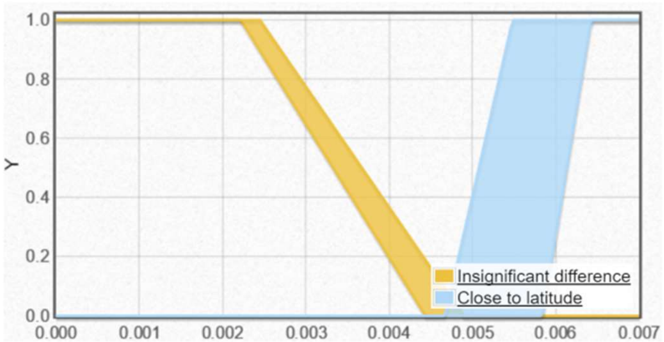

Fig. 3. Values of variables "latitude/longitude accuracy" (same graph)

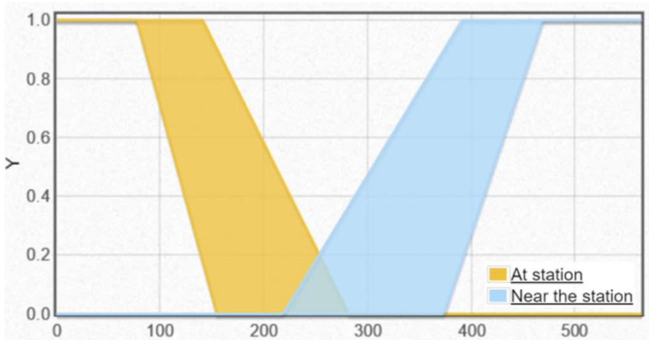

Fig. 4. Two values of the linguistic variable "location"

Thus, the universe, on which variable $L_v^{(2)}$ is defined, results in $U = [0, 0.007]$ (2). The upper function (U) for the term 'insignificant difference' of the variable $L_v^{(2)}$ can be characterized by parameter's set A(0,0), B(0,1), C(0.00245,1) and D(0.0049,0); C_x is calculated as the arithmetic mean of B_x and D_x, which is the minimal average accuracy shown in Table 1. As for the lower function (same linguistic term is considered), the values of its parameters are A(0,0), B(0,0), C(0.00222,1) and D(0.00444,0). First two parameters reflect perfect accuracy at the position; C_x is obtained as before, while D_x value corresponds to (2). Linguistic values 'close to latitude/longitude' should be viewed separately, because for each component of coordinate's pair factor of inaccuracy (its measurement) sounds alike, but still differently. Their presence leads to more stable model (Fig. 1) and helps to improve the results attained. The upper function (U) is determined by parameters A(0.00467,0), B(0.00548,1), C(0.007,1) and D(0.007,0), i.e. $A_x = \left(D_x^{(L'\text{insig.diff}')} + D_x^{(U'\text{insig.diff}')}\right)/2 = 0.00467^\circ$, B_x is an arithmetic mean of A_x and C_x. For the lower function (L) parameters are specified as follows: A(0.00584,0), B(0.00642,1), C(0.007,1) and D(0.007,0), where $A_x = \left(A_x^{(U'\text{close to latitude}')} + 0.007^\circ\right)/2 = 0.00584^\circ$, B_x is calculated much as it is done in the case of upper function (U), both C_x and D_x are equated with the value of 0.007° that stands for minimum accuracy (or, maximum inaccuracy) – corresponding values are shown in Fig. 3. In the case concerned, only non-negative values of accuracy are considered; if calculations lead to negative result, we use its modulus.

3.3 Linguistic variable "location" and its values (terms)

The variable $L_v^{(3)} \equiv \text{"}location\text{"}$ is the next matter under discussion – actually, it expresses the location, as it arises from variable's name, of a passenger due to indications related to previously mentioned variables. The variable is represented graphically in Fig. 4. We introduce two values (fuzzy sets) of $L_v^{(3)}$, namely, they are 'at station', i.e. main region that must be reached to notify a user about the arrival, and 'near the station'. The standard length of Moscow metro' platform is appr. 155 meters (8 train carriages), the longest station is "Vorobyovy Gory" – its length is about 282 meters [11]. The universe of discourse U the variable $L_v^{(3)}$ is defined on can be denoted as $[0, c]$, where the right bound *c* equals to the double length of the longest platform in the metro. For the upper function (U) as a constituent of IT2MF representing value 'at station', we set the following parameters: A(0,0), B(0,0), C(141,1) and D(282,0), where C_x is a half of the longest station (282 m) in the Moscow's metro. The parameters of the lower function (L) of IT2MF are A(0,0), B(0,0), C(77.5,1) and D(155,0) with C_x calculated as the arithmetic mean of B_x and D_x coordinates. We suggest to model the linguistic value 'near the station' with the IT2MF, whose upper function (U) is characterized by A(218.5,0), B(391.25,1), C(564,1) and D(564,0); the value of A_x is obtained as $\left(D_x^{(L'\text{at station}')} + D_x^{(U'\text{at station}')}\right)/2 = 218.5$ (in meters), B_x is the mean of A_x and C_x x-coordinates, both C_x and D_x are equal to 564 meters (double length of the longest platform). Similarly, parameters of the lower function (L) are A(373.5,0), B(468.75,1), C(564,1) and D(564,0), where A_x (x-coordinate of the first

parameter) equals to $A_x^{(U'\text{at station}')} + 155 = 373.5$ that takes into account the length of the standard metro train, i.e. the latter will have direct influence on the spread of the left tail of the membership function. As before, coordinate B_x is the average of A_x and C_x (468.75 meters), and non-negative values are considered.

Rather detailed description of linguistic variables and their values is important for deeper understanding of fuzzy logic system (its model), the use of interval type-2 membership functions to represent uncertainty inherent in verbal values introduced and with the regard for specific character of possible implementation of the system in the code. To a large extent, the definition of a very small number of linguistic variables' values pursues two plain objects – namely, (1) to obtain the initial "non-overloaded" (in terms of number of values and fuzzy rules) variant of the system to perform experiments with and to lay a ground for further analysis, tuning parameters and rule base, revealing drawbacks, etc., and (2) to examine the general idea of using type-2 fuzzy sets in recommendation services that are actively advancing as it applies to enormous market of mobile devices.

4. Rules of the fuzzy model (Inference System) and experiments conducted

The core of the fuzzy inference system (FIS) as shown in Fig. 1 is a set of linguistic values represented in the form of fuzzy sets, If-Then rules having a generic form "If {*antecedent*} Then {*consequent*}" and fuzzy reasoning scheme; the latter just operate on a given rules along with specified inputs to derive system's outputs or conclusions. The experts' understanding of the phenomenon under study and their knowledge of the domain field provide a basis for formation of the primary version of rule-base, in which linguistic variables $L_v^{(1)} \equiv \text{"}latitude/longitude\ accuracy\text{"}$ and $L_v^{(2)} \equiv \text{"}delta\text{"}$ are used in antecedent part of fuzzy rules (input of the system), whereas $L_v^{(3)} \equiv \text{"}location\text{"}$ operates as system's output (its terms form consequent part of rules). The evident transparency of the rule-base in general is substantiated here by a specific fact of simplicity and lucidity of both linguistic values submitted for consideration and existing relations between them. To the opinion of authors, such situation can be viewed as an advantage in terms of efforts needed to design the rule-base. However, it does not mean that the subsequent fine-tuning of rules as well as values of variables will not be needed – most likely, this stage is unavoidable in practice regardless of the system at hand. At the moment, the rules can be represented in the following form:

Rule 1	If *delta* is 'small difference' and *latitude accuracy* is 'insignificant difference' and *longitude accuracy* is 'insignificant difference' Then *location* is 'at station'

Rule 2	If *delta* is ′small difference′ and *latitude accuracy* is ′close to latitude′ and *longitude accuracy* is ′insignificant difference′ Then *location* is ′near the station′
Rule 3	If *delta* is ′small difference′ and *latitude accuracy* is ′insignificant difference′ and *longitude accuracy* is ′close to longitude′ Then *location* is ′near the station′
Rule 4	If *delta* is ′small difference′ and *latitude accuracy* is ′close to latitude′ and *longitude accuracy* is ′close to longitude′ Then *location* is ′near the station′
Rule 5	If *delta* is ′bigger difference′ and *latitude accuracy* is ′insignificant difference′ and *longitude accuracy* is ′insignificant difference′ Then *location* is ′near the station′
Rule 6	If *delta* is ′bigger difference′ and *latitude accuracy* is ′close to latitude′ and *longitude accuracy* is ′insignificant difference′ Then *location* is ′near the station′
Rule 7	If *delta* is ′bigger difference′ and *latitude accuracy* is ′insignificant difference′ and *longitude accuracy* is ′close to longitude′ Then *location* is ′near the station′
Rule 8	If *delta* is ′bigger difference′ and *latitude accuracy* is ′close to latitude′ and *longitude accuracy* is ′close to longitude′ Then *location* is ′near the station′

4.1 Test 1 (difference between fixed and current points (delta))

The first carried out experiment is related to checking the difference between fixed and current points (i.e. linguistic variable ″delta″) under the constant latitude/longitude accuracies equal to 0.00074 (step of delta's change is taken as 0.0032, number of steps equals to 10). IT2MF is an assortment of type-1 membership functions embedded between upper (U) and lower (L) functions. Each of these embedded functions (type-1) can be defuzzified, viz. converted to crisp number that represents generically corresponding fuzzy set (its membership function). The most commonly used method of defuzzification is called centroid [10]. The processing of type-2 fuzzy systems provides for the use of type reduction procedure (TRp) that can be seen as an expanded form of type-1 defuzzification resorting to Extension principle [12]. Each of rules **Rule i**, $i=\overline{1,8}$, "fires" and leads to obtaining output type-2 fuzzy set under a given input data. The union of these output sets and calculation of the centroid of resultant set is the essence of the centroid type reduction. Both theoretical framework and development of type reduction's use in type-2 fuzzy systems were

presented in publications by Karnik N.N. and Mendel J.M. [13, 14]. As applied to IT2FS (secondary membership function in that case is constant), TRp becomes simpler in comparison with generalized type-2 sets – the results of experiment (see the data above) using centroid type reduction (**CTR**) defuzzification as summarized in Table 2.

Table 2. CTR defuzzification

№	Difference between fixed and current points (delta)	Location
1	0.0032	84.398
2	0.0064	94.312
3	0.0096	139.576
4	0.0128	282.000
5	0.016	392.995
6	0.0192	392.995
7	0.0224	392.995
8	0.0256	448.347
9	0.0288	460.487
10	0.032	460.487

Table 3. CoSTR defuzzification

№	Difference between fixed and current points (delta)	Location
1	0.0032	84.398
2	0.0064	84.398
3	0.0096	84.398
4	0.0128	275.180
5	0.016	460.487
6	0.0192	460.487
7	0.0224	460.487
8	0.0256	460.487
9	0.0288	460.487
10	0.032	460.487

On the other hand, another TRp called center-of-sets type reducing approach (**CoSTR** – it is a family of defuzzification methods proposed up to now) can be used to substitute the consequent parts of rule-base by singletons at the centroid of corresponding fuzzy sets (Then-part of rules). Subsequent step is connected with obtaining the centroid of type-1 fuzzy set constituted by aforementioned singletons [10]. Calculated values that refer to test data (section IV, item's *A* preamble) are accumulated in Table 3.

It can be noticed that for a particular set of test data centroid TRp demonstrates better (i.e. smoother) approximation of the moderately growing exponential trend. Relative angularity (in Fig.5 it is not so strongly pronounced in comparison with Fig.6 case) relates to the use of piecewise linear (trapezoidal) functions representing fuzzy sets, certain (potential) drawbacks ascribed to rule-base design issues and small number of linguistic terms defined for each variable under consideration. However, even under these circumstances, results of centroid TRp indicate that it is more sensitive to accuracy changes (fine-tuning) than the second TRp. The second graph (Fig.6) visualizes marked broken line consisting of 2 constant levels, and one of those is rather lengthy. To a variable degree, both lines are increasing, and centroid TRp is preferable, since it considers specificity of all functions' values.

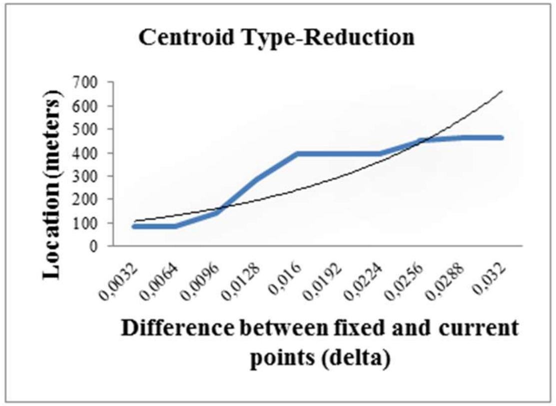

Fig. 5. Centroid type reduction method for "delta" variable

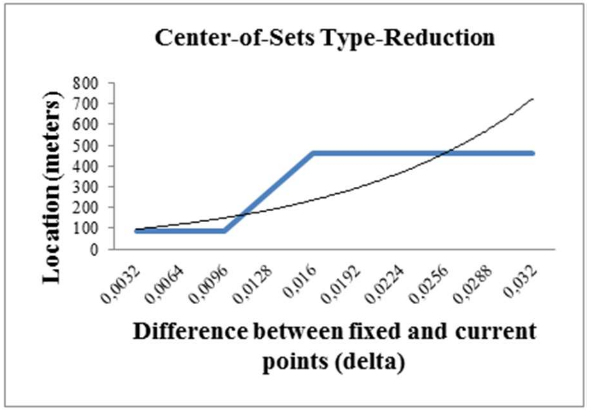

Fig. 6. Center-of-sets type reduction method for "delta" variable

4.2 Test 2 (latitude/longitude accuracy)

The second test relates to checking the latitude/longitude accuracy under constant difference between fixed and current points (delta) equals to 0.0032 (longitude accuracy is 0.00074 OR latitude accuracy is 0.00074, the number of steps is set to 10). Results are shown by Tables 4 and 5.

Here, situation retains characteristic features observed in Fig.5 and 6, i.e. centroid TRp also demonstrates better "behavior". The line (Fig.7) grows monotonously being smooth enough, except for x-coordinates falling into the real range [0.00518, 0.00666] (approx.). Lines shown in both graphs (Fig.7,8) follow the exponential trend (the less latitude/longitude accuracy, the less location accuracy observed).

Table 4. CTR defuzzification

№	Latitude / Longitude accuracy	Location
1	0.00074	84.398
2	0.00148	84.398
3	0.00222	84.398
4	0.00296	90.831
5	0.0037	99.770
6	0.00444	139.576
7	0.00518	392.995
8	0.00592	438.650
9	0.00666	460.487
10	0.0074	460.487

Table 5. CoSTR defuzzification

№	Latitude / Longitude accuracy	Location
1	0.00074	84.398
2	0.00148	84.398
3	0.00222	84.398
4	0.00296	84.398
5	0.0037	84.398
6	0.00444	84.398
7	0.00518	460.487
8	0.00592	460.487
9	0.00666	460.487
10	0.0074	460.487

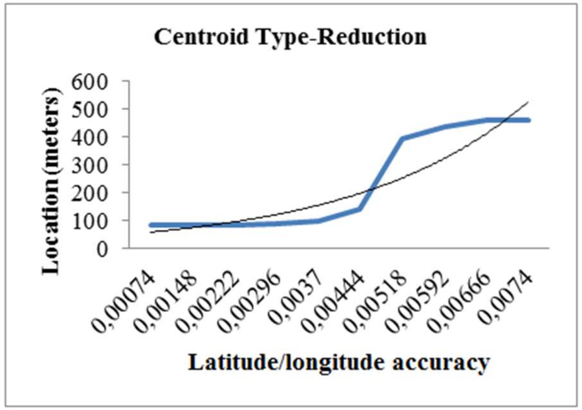

Fig. 7. CTR defuzzification for "latitude/longitude accuracy" variables

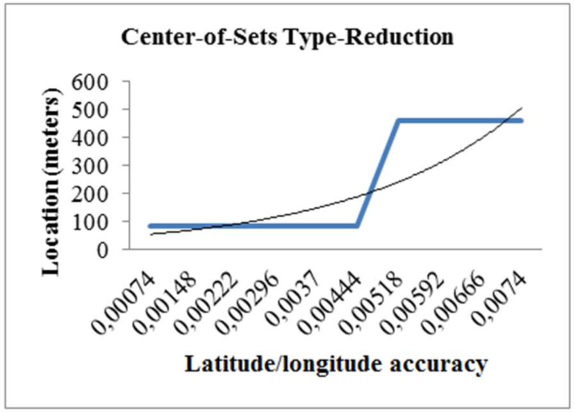

Fig. 8. CoSTR defuzzification for "latitude/longitude accuracy" variables

It should be explicitly mentioned that we have additionally tested rules using two defuzzification methods already mentioned before, namely, (1) centroid TRp and (2) center-of-sets TRp (approaches).

4.3 Test 3 (checking rules used in the model)

Table 6. CTR defuzzification results

№	Rule	Difference between fixed and current points (delta)	Latitude accuracy	Longitude accuracy	Location
1	1*a*	0.0032	0.00074	0.00444	139.576
2	1*b*	0.0064	0.00222	0.0037	99.770
3	1*c*	0.0096	0.00296	0.00296	139.576
4	1*d*	0.0128	0.0037	0.00222	282.0
5	1*e*	0.016	0.00444	0.00074	392.995
6	2*a*	0.0032	0.00518	0.00074	392.995
7	2*b*	0.0064	0.00444	0.00222	139.576
8-22	2*c*-5*b*	0.0096	0.00592	0.00296	392.995
23	5*c*	0.0256	0.00296	0.00296	446.153
24	5*d*	0.0288	0.0037	0.00222	441.641
25-29	5*e*-6*d*	0.032	0.00444	0.00074	392.995
30	6*e*	0.032	0.00666	0.00074	460.487
31-40	7*a*-8*e*	0.0192	0.00074	0.00666	392.995

Each rule was "fed" with 5 (five) test cases, thus each of Tables 6 and 7 covers $40 = 8 \times 5$ cases in total. Tests per rule are numbered in ascending order starting with [n]*a* and ending with [n]*d*, where [n] is the rule's number (index). For example, Rule

3 corresponds to sequence of labelings $3a, 3b, 3c, 3d$ and $3e$ used in Tables 6 and 7. Test data were generated according to intervals of each variable's domain (the same approach as in tests 1 and 2). Step of "*delta*" changes is 0.0032, while step for the latitude and longitude variables equals to 0.00074. Input values are mixed to ensure wider coverage and variety. The last column presents location according to test values and calculation method (TRp) selected. Last column's cells with light-grey shading determine 'At station' (≤ 282 meters) value (set), while other values show location near some station (linguistic value 'Near the station'). Both tables are wittingly shortened, because of recurrent location results.

Table 7. CoSTR defuzzification results

№	Rule	Difference between fixed and current points (delta)	Latitude accuracy	Longitude accuracy	Location
1	1a	0.0032	0.00074	0.00444	84.398
2	1b	0.0064	0.00222	0.0037	84.398
3	1c	0.0096	0.00296	0.00296	84.398
4	1d	0.0128	0.0037	0.00222	275.180
5	1e	0.016	0.00444	0.00074	460.487
6	2a	0.0032	0.00518	0.00074	460.487
7	2b	0.0064	0.00444	0.00222	84.398
8-40	2c-8d	0.0096	0.00592	0.00296	460.487

A defuzzification method computes the range of possible location values according to input data provided, and the last column of tables shows a mean value of interval bounds, e.g. 139.576 is a mean of $[0, 279.152]$ real-valued range obtained through defuzzification procedure.

5. Conclusion

The paper examined potentials of the modeling approach based on interval type-2 fuzzy sets (IT2FS) and conventional Mamdani fuzzy inference system (MFIS) as applied to real and topical problem related to passengers tracking in urban metro (positioning service by the example of Moscow city). Appeal and significance of developing and further analysis of such models may be of a high demand for appropriate representation of those factors that are inherently vague and uncertain. The aspects that provide for eventuality to discuss models with broad sections of stakeholders owing to model's transparency, abilities to tune their parameters and to carry out experiments (test runs) play a sound role in theory and from practical standpoint. Empirical studies had shown that design issues concerned with linguistic variables and their labelled values (or, terms) influence significantly fuzzy model's output. Test cases presented in the paper corroborate both the applicability and relevance of fuzzy logic-based approach to various problems emerging in the field of navigational services, passenger tracking based on positional technologies. As it was mentioned in section IV, the model that makes use of IT2FS and MFIS leads at the end to resultant intervals that can be calculated in genuine mobile applications without appreciable extra costs with the object of determining the distance to notify users about their arrival (approach) to station. Hence, the developed fuzzy (prototype) model helps to estimate exemplary limits for values of each variable examined. Due to promising test results (for the time being we can talk about model prototype only) and its potential practical applicability, the model (Fig.1) will be implemented in the Android-based mobile program aimed at building routes and notifying users about their destination.

From the standpoint of further theoretical research and topic evolvement, diverse types of membership functions together with fine tuning of their parameters as well as alternative type reduction defuzzification (TRDf) methods should be considered more thoroughly. Besides, by way of illustration GPS technique may beat its own path in IT2FS-based models as applied to ground transportation.

References

[1]. Chen C-Y., Yang J.-P., Tseng G.-J., Wu Y.-H. and Hwang R.-C. An Indoor Positioning Technique Based on Fuzzy Logic, in *Proc. International Multi Conference of Engineers and Computer Scientists (IMECS)*, 2010, pp. 854-857.

[2]. Teuber A. and Eissfeller B. WLAN Indoor Positioning Based on Euclidean Distances and Fuzzy Logic, in *Proc. Workshop on Positioning, Navigation and Communication (WPNC)*, 2006, pp. 159-168.

[3]. Yasunobu S., Miyamoto S. and Ihara H. A Fuzzy Control for Train Automatic Stop Control. *Transactions of the Society of Instrument and Control Engineers*, 2002, vol. E-2(1), pp. 1-9.

[4]. Klir G.J. and Wierman M. Uncertainty Formalizations. In: Uncertainty-Based Information. Elements of Generalized Information Theory, ser. Studies in Fuzziness and Soft Computing (#15), 2nd ed., Physica Verlag, Germany, 1999, 168 p.

[5]. Arigela L., Veerendra P., Anvesh S. and Hanuman K. Mobile Phone Tracking & Positioning Techniques. *Int. Journal of Innovative Research in Science, Engineering and Technology,* 2012, vol.2, pp. 906-913.

[6]. Gps.gov, “Official U.S. Government Information About the GPS and Related Topics, GPS Accuracy”, 2017. [Online resource]. Available: http://www.gps.gov/systems/gps/performance/accuracy/ [Accessed 27-Feb-2017].

[7]. Zeimpekis V., Kourouthanassis P. E. and Giaglis G. M., Mobile and Wireless Positioning Technologies, in *UNESCO Encyclopedia of Life Support Systems (EOLSS)*, vol. 6.108, EOLSS Publishers, France, 2007, [http://www.eolss.net].

[8]. Mosmetro.ru, “Metropoliten v tsifrakh”, 2017. [Online resource]. Available: http://mosmetro.ru/press/metropoliten-v-tsifrakh/ [Accessed 28-Jan-2017] (in Russian).

[9]. Nashemetro.ru, “Metro v tsifrakh.”, 2017. [Online]. Available: http://nashemetro.ru/facts.shtml [Accessed 13-Jan-2017] (in Russian).

[10]. Mendel J.M., Hagras H., Tan W.-W., et al. Introduction to Type-2 Fuzzy Logic Control. Theory and Applications (IEEE Press Series on Computational Intelligence), Wiley-IEEE Press, Piscataway, 2014, 376 p.

[11]. En.wikipedia.org, “Moscow Metro”, 2017. [Online resource]. Available: https://en.wikipedia.org/wiki/Moscow_Metro [Accessed 25-Jan-2017].

[12]. Mendel J.M. Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions, Prentice Hall PTR, Englewood Cliffs, 2001, 576 p.

[13]. Karnik N.N., Mendel J.M. Type-2 Fuzzy Logic Systems: Type-Reduction, in *Proc. IEEE Int. Conference on Systems, Man, and Cybernetics*, 1998, pp. 2046-2051.

[14]. Mendel J.M. Interval Type-2 Fuzzy Logic Systems and Perceptual Computers: Their Similarities and Differences. In: Sadeghian A., Mendel J., Tahayori H. (eds) Advances in Type-2 Fuzzy Sets and Systems. Studies in Fuzziness and Soft Computing, vol. 301. Springer, New York, 2013, pp. 3-18, doi: 10.1007/978-1-4614-6666-6_1.

Модель сервиса позиционирования в метро, основанная на правилах и нечетких множествах второго типа

А.Р. Гималетдинова <argimaletdinova@edu.hse.ru>
К.Ю. Дегтярев <kdegtiarev@hse.ru>
Национальный исследовательский университет «Высшая школа экономики»,
125319, Россия, Москва, Кочновский проезд, д. 3

Аннотация. За последние несколько лет возник значительный интерес к мобильным приложениям, ориентированным на построение маршрутов пользователей гаджетов; в таких приложениях наряду с важной функцией навигации также возможно отправление своевременных оповещений о прибытии к заданному месту назначения. Несмотря на большой объем имеющейся информации о специфике навигационных сервисов, актуальным остается вопрос относительно точности позиционирования. В данной статье рассматривается возможный подход к решению проблемы сравнения, связанного с определением близости пользователя к конечной станции его маршрута в метро. Такая близость определяется путем подсчета разницы в координатах между текущей позицией пассажира и фиксированной точкой. С целью создания Системы Рекомендаций Маршрутов (СРМ) была применен аппарат нечеткой логики, который использует лингвистические переменные для выражения имеющейся нечеткости (неопределенности) в понимании/восприятии вербального понятия «близость к …». В работе подробно объясняется каждая переменная, используемая в системе нечеткого вывода (англ. FIS), а также представляется набор нечетких правил ЕСЛИ-ТО модели. Для проверки стабильности модели (пока имеет смысл говорить о прототипе модели как первом шаге на пути дальнейшей ее проработки и изменения), основанной на схеме логического вывода Мамдани, рассматриваются несколько тестовых экспериментов с моделью, описываются получаемые результаты. В дальнейшем, планируется разработка мобильного Android-приложения, нацеленного на построение маршрутов городского общественного транспорта с возможностью использования представленной модели при реализации функции по отправлению своевременных оповещений о приближении к пункту назначения. Следует отметить, что акцент делается на использовании в модели интервальных нечетких множеств второго типа (англ. IT2FS), которые привлекают значительное внимание исследователей в настоящее время. Значимость задачи разработки подобных моделей определяется, в первую очередь, необходимостью адекватного учета тех факторов, которые по своей сути являются нечеткими (неопределенными). Данная работа, по мнению авторов, может помочь в продолжении и развитии исследований, связанных с этой же или подобными темами.

Ключевые слова: сервис для позиционирования; мобильные приложения; нечеткое моделирование; GPS; WiFi; мобильные сети; общественный транспорт; интервальные нечеткие множества второго типа; система нечеткого вывода; нечеткость; неточность

DOI: 10.15514/ISPRAS-2017-29(4)-6

Для цитирования: Гималетдинова А.Р., Дегтярев К.Ю. Модель позиционирования в метро, основанная на нечетких правилах второго типа. *Труды ИСП РАН*, том 29, вып. 4, 2017 г., стр. 87-106 (на английском языке). DOI: 10.15514/ISPRAS-2017-29(4)-6

Список литературы

[1]. Chen C-Y., Yang J.-P., Tseng G.-J., Wu Y.-H. and Hwang R.-C. An Indoor Positioning Technique Based on Fuzzy Logic, in *Proc. International Multi Conference of Engineers and Computer Scientists (IMECS)*, 2010, pp. 854-857.
[2]. Teuber A. and Eissfeller B. WLAN Indoor Positioning Based on Euclidean Distances and Fuzzy Logic, in *Proc. Workshop on Positioning, Navigation and Communication (WPNC)*, 2006, pp. 159-168.
[3]. Yasunobu S., Miyamoto S. and Ihara H. A Fuzzy Control for Train Automatic Stop Control. *Transactions of the Society of Instrument and Control Engineers*, 2002, vol. E-2(1), pp. 1-9.
[4]. Klir G.J. and Wierman M. Uncertainty Formalizations. In: Uncertainty-Based Information. Elements of Generalized Information Theory, ser. Studies in Fuzziness and Soft Computing (#15), 2nd ed., Physica Verlag, Germany, 1999, 168 p.
[5]. Arigela L., Veerendra P., Anvesh S. and Hanuman K. Mobile Phone Tracking & Positioning Techniques. *Int. Journal of Innovative Research in Science, Engineering and Technology,* 2012, vol.2, pp. 906-913.
[6]. Gps.gov, "Official U.S. Government Information About the GPS and Related Topics, GPS Accuracy", 2017. http://www.gps.gov/systems/gps/performance/accuracy/ [Дата обращения 27.02.2017].
[7]. Zeimpekis V., Kourouthanassis P. E. and Giaglis G. M., Mobile and Wireless Positioning Technologies, in *UNESCO Encyclopedia of Life Support Systems (EOLSS)*, vol. 6.108, EOLSS Publishers, France, 2007, [http://www.eolss.net].
[8]. Mosmetro.ru, "Метрополитен в цифрах", 2017. http://mosmetro.ru/press/metropoliten-v-tsifrakh/ [Дата обращения 28.01.2017] (in Russian).
[9]. Nashemetro.ru, "Метро в цифрах", 2017. http://nashemetro.ru/facts.shtml [Дата обращения 13.01.2017] (in Russian).
[10]. Mendel J.M., Hagras H., Tan W.-W., et al. Introduction to Type-2 Fuzzy Logic Control. Theory and Applications (IEEE Press Series on Computational Intelligence), Wiley-IEEE Press, Piscataway, 2014, 376 p.
[11]. En.wikipedia.org, "Moscow Metro", 2017. https://en.wikipedia.org/wiki/Moscow_Metro [Дата обращения 25.01.2017].
[12]. Mendel J.M. Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions, Prentice Hall PTR, Englewood Cliffs, 2001, 576 p.
[13]. Karnik N.N., Mendel J.M. Type-2 Fuzzy Logic Systems: Type-Reduction, in *Proc. IEEE Int. Conference on Systems, Man, and Cybernetics*, 1998, pp. 2046-2051.
[14]. Mendel J.M. Interval Type-2 Fuzzy Logic Systems and Perceptual Computers: Their Similarities and Differences. In: Sadeghian A., Mendel J., Tahayori H. (eds) Advances in Type-2 Fuzzy Sets and Systems. Studies in Fuzziness and Soft Computing, vol. 301. Springer, New York, 2013, pp. 3-18, doi: 10.1007/978-1-4614-6666-6_1.