EVALUATION OF DISTANCE BETWEEN PEDESTRIAN CROSSINGS BY STUDENTS IN ONE OF THE POLISH CITIES

M. KRUSZYNA[1]

In this paper, the distances between pedestrian crossings in twenty one places in the city of Wrocław, together with their evaluation by the researched groups of students, were analyzed. The database created from the collected questionnaires contains a set of two-dimensional variables: the distance between crossings and the rating of the students. The database set was analyzed using a fuzzy data mining approach to create particular clusters. Various numbers of clusters were analyzed, and the division of data into three clusters made it possible to relate the analysis to the LOS methodology. Each variable was enriched with a third dimension representing a membership value. The obtained evaluated distances are similar to values recommended in literature, although the distances highly evaluated by the students do not often occur in reality. This might suggest that there is the need to create new crossings, especially in the city centre, where pedestrian traffic is or should be important.

Key words: Crossings; Pedestrians; Evaluation; Level of Service; Fuzzy Grouping

1. INTRODUCTION

Recently, researchers have studied pedestrians' crossing behaviour with reference to the determinants of making decisions about crossing the street. De LAVALETTE at all [1] concentrated on topographical features (e.g. the number of lanes of traffic to cross), infrastructure (e.g. the presence of a central traffic island), control systems (e.g. traffic signals for vehicles and pedestrians), as well as on of the pedestrian's primary task (getting to school, going shopping, etc.). PAPADIMITRIOU [2] showed that the number of lanes, high traffic flows, and traffic lights increase the probability of crossing at junctions, and that one-way roads, roadside parking, and the presence of an attractor across the road (e.g. large shop) increase the probability of crossing at mid-block. SISIOPIKU and AKIN [3] claimed that the most influential factor in making a decision to cross at a designated crossing location is the distance of a crosswalk to pedestrians' destinations. The distance between neighbouring crossings could be an important factor to determine an individual pedestrian's decision to cross or not to cross

[1] E-mail: maciej.kruszyna@pwr.wroc.pl, mailing addresses: Wroclaw University of Technology, Department of Civil Engineering, Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland (job), tel.: 071 320 45 39, (Corresponding author).

a street. Evaluation of this distance by the users can be helpful to build crossings in the appropriate places.

Younger pedestrians are more eager to cross an unsafe street (between the crossings). Holland and Hill [4] wrote that younger people (17–24 years old), men and especially younger men who were not drivers had reportedly more positive attitude towards crossing in risky situations. Cavallo and Lobjois [5] distinguished three age groups of pedestrians (20–30, 60–70, and 70–80 years old) for which they observed similar safety margins and unsafe-decision rates. The results showed that elderly pedestrians preferred larger time gaps than younger ones, enabling them to compensate for their longer crossing times resulting from their slower walking speed.

The aims of this paper are: investigating how young pedestrians (students) evaluate the distance between crossings, defining (founding) specific classes of distances in relation to the Level of Service methodology, and fuzzy grouping (clustering) of values of distances into these classes.

2. Methods

2.1. Survey localisation and procedure

Wroclaw is one of the biggest cities in Poland (about 630.000 residents), with many Universities, and will become the European Capital of Culture in 2016. Between 2006 and 2011, six groups of university students, each about 40 - 50 persons, were asked to fill in a specially designed questionnaire in which they had to evaluate the distance between the crossings in specific places. On the basis of these questionnaires the database was formed.

The distance between pedestrian crossings in the chosen spots in the city centre was analyzed. The rate of traffic flow in those areas was high, which made illegal (outside the signalled or zebra crossing or during the red light signal) road crossing practically impossible. Two crossings and the space between them are called "a place" in this paper. Twenty eight places were selected at the beginning of the research, but two of them obtained a special status. The place with the shortest distance, marked with "S" and the place with the longest distance, marked with "L", were used as a "reference level" to the evaluation in other places. Therefore, they were not taken into consideration in the following analyses. The locations of all places are shown in figure 1. Exemplary photo of a chosen place (No.22) is shown in figure 2. Between 2006 and 2011, some of the streets and intersections were rebuilt. The distances between the crossings in places 4, 5, 7, 10 and 11 were changed and these places were not included in the following analyses either. For these reasons, twenty one places were analyzed.

The database formed on the data collected from questionnaires contains a set of two-dimensional variables: the distance between the crossings and the rating (grade) of the students.
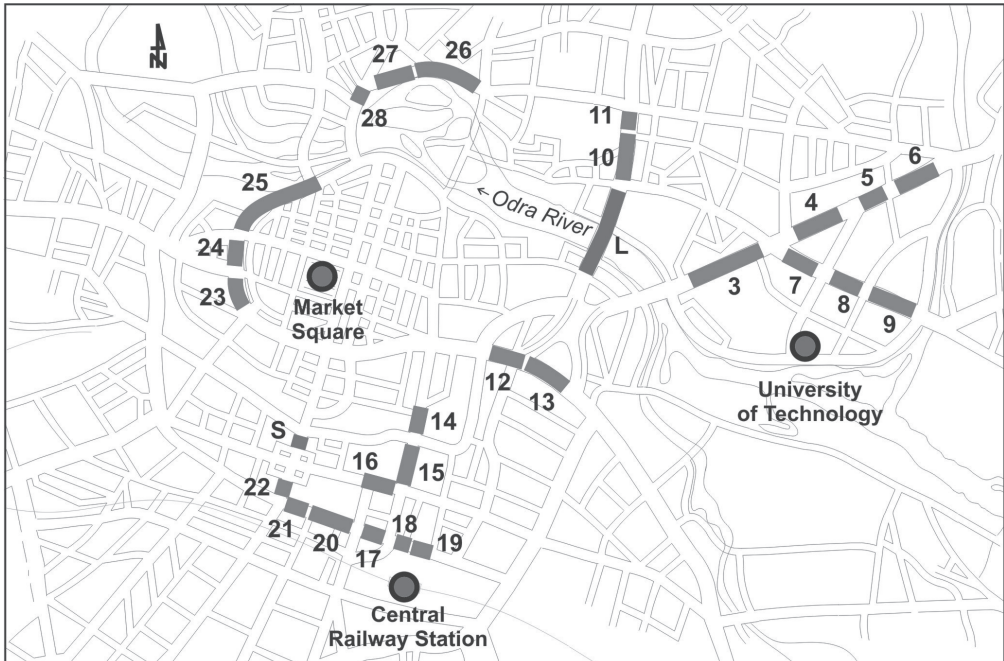
Fig. 1. Location of places where the distance between the crossings was evaluated

2.2. LEVEL OF SERVICE

Highway Capacity Manual [6] develops the conception of LOS so that it may be applied to describe pedestrian facilities (walkways and sidewalks, queuing areas, shared off-street paths, crosswalks), as well as other facilities along urban streets. LOS means "level of service" and it is a way to measure the quality of traffic facilities by describing operational conditions within a traffic stream. There exist six levels, from LOS A (representing the best operating conditions), to B, C, D, E, to LOS F (representing the worst operating conditions). Various performance measures were used to estimate LOS for particular places. For example, the average pedestrian space was used in the case of queuing areas, and the control delay was used to estimate LOS for signalized intersections.

It is possible to use the LOS methodology to evaluate the quality of pedestrians infrastructure (ways), particularly the convenience of crossing a road. The distance between crossings could be a measure to estimate LOS in such cases. The distance between pedestrian crossings is one of the parameters describing the quality of the transportation system. Until now, the problem of the distance between pedestrian crossings has been viewed from the drivers' perspective. The question was: how often can pedestrians cross the road and thus stop the vehicular traffic? Most frequently, the

distances between pedestrian crossings depend on the road category, as well as on the flow rate of the vehicular and pedestrian traffic in a given area. However, it should be remembered that the possibility to conveniently cross the road determines whether the road is an impediment to pedestrian mobility or not. Small distances between pedestrian crossings create a more pedestrian-friendly urban environment and reduce the rate of illegal road crossing instances.

As a rule, the boundaries between different LOS grades are defined in a very precise manner. For example, the average pedestrian delay in the range of 20-30 seconds at the signalized intersection denotes LOS C. A small difference in the value of a given parameter (such as a pedestrian delay) can mean a different LOS grade. 31 second delay already denotes LOS D. Not much smaller delay of 29 seconds (the difference of only 2 seconds) denotes LOS C. At the same time, there are objects that have been assigned the same grade, while the difference between the values is much greater. Both a 21 second delay and a 29 second delay qualify for LOS C, even though the difference is as big as 8 seconds. Precise values of LOS grades do not reflect the way human perception works. In the case of the latter, the expressions such as 'big' or 'small' are of greater importance than the precise values expressed in standardized units. Some researchers propose the use of fuzzy clustering to create boundaries between the LOS factors [7], [8].

2.3 Fuzzy grouping

Fuzzy grouping is an effective tool that has been used in technical fields for many years. Fuzzy rules and variables are also used in transportation engineering. This approach allows taking into consideration the human aspect of perceiving and evaluating certain parameters. Typical of this approach is rather descriptive than numerical grouping of delay (high, moderate, low as opposed to more than 30, between 10 and 30, and less than 10 seconds). The terms 'high' or 'low' are not precise - they are fuzzy. Fuzzy logic has therefore been applied to describe the values observed by users of transportation system. Apart from the example with speed [9], other parameters where fuzzy logic can be applied, include delays and journey time [10]. Apart from that, the fuzzy description can be used for movement patterns [11] and freedom of movement levels [12]. Groups with fuzzy borders are called clusters.

Because of the complexity of the problem, there are many fuzzy clustering methods. The most popular method is the classic Fuzzy C-means Method (FCM) and its modifications [13], [14]. The key problem with FCM method is the requirement that the number of clusters be determined a priori, right at the start. This number does not necessarily fit the type of the variable analyzed. Many models for optimizing the number of clusters have thus been developed. However, the FCM method is still widely used, especially when the samples are numerous and with little noise. Methods for developing fuzzy clustering models have been described in detail by Ali [15]. Moreover, the "classic" Fuzzy C-Means algorithm, developed by Bezdek [16] is still in use. This algorithm was used in this study too.

The cluster membership functions have traditionally been presented in triangular shapes. In fact, membership functions can be expressed as curves of various shapes. However, the exact shape of these curves is not indispensable for determining cluster boundaries.

The database set was analyzed using a fuzzy data mining approach to create particular clusters. Various numbers of clusters (3, 4, and 6) were analyzed, and the division of data into three clusters made it possible to relate the analysis to the LOS methodology. Each variable was enriched with a third dimension representing a membership value.

3. Results

3.1. Evaluation of distances between crossings

The highest possible grade which students could give was six, which denoted the shortest distance. The lowest possible grade, denoting the longest distance, was one. The shortest distance (in place "S") between crossings was 50m. The longest distance (in place "L") was 400 m. The students didn't know the precise value of distance; they only saw each place on photos (like in figure 2) and films. They had to assign a grade

Fig. 2. Example of place (No.22)

according to their own opinion and experience. At the beginning of the experiment the students were shown the places with the shortest and the longest distances between the crossings. They had to assign the highest and the lowest grade respectively. These two places were not taken into consideration in the next stage of the research. The characteristics of the 21 places are presented in table 1. The distances are given in column (2). The student's grades are given in columns (3) – (9), separately for subsequent years and with the calculation of weighted mean value. Figure 3 shows the dependence between the distance and mean student's grade.

Table 1

Summary of the results of surveys

Number of Place	Distance between the Crossings [m]	Grade of Students						mean
		in Year:						
		2006	2007	2008	2009	2010	2011	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
3	350	1,98	2,00	2,52	2,80	2,60	2,11	2,34
6	250	3,47	3,23	3,24	3,02	3,38	3,09	3,24
8	200	3,75	3,95	3,70	3,49	3,85	3,57	3,72
9	220	3,53	3,19	3,36	2,99	3,21	3,31	3,27
12	140	2,88	3,17	3,62	2,93	3,15	3,13	3,15
13	250	3,46	3,23	3,34	3,41	3,66	3,07	3,36
14	150	3,40	3,81	4,08	3,78	3,67	3,86	3,77
15	180	3,81	3,99	4,00	3,72	3,77	3,97	3,88
16	120	3,58	3,49	3,83	3,63	3,90	3,35	3,63
17	100	4,53	4,31	4,03	3,48	4,25	3,89	4,08
18	90	4,11	3,52	3,72	3,80	3,30	4,29	3,79
19	130	4,22	3,81	3,84	3,63	3,64	4,10	3,87
20	150	3,99	3,82	3,87	3,69	4,11	3,57	3,84
21	170	4,51	3,74	4,16	3,69	3,78	4,25	4,02
22	80	5,00	4,46	4,82	3,94	3,80	5,27	4,55
23	190	2,93	3,29	3,13	3,32	2,29	4,06	3,17
24	170	3,82	3,50	3,66	3,67	3,56	3,77	3,66
25	370	1,92	2,18	2,46	2,26	2,17	2,22	2,20
26	320	2,05	2,09	2,28	2,40	1,83	2,60	2,21
27	170	2,86	2,85	3,05	3,11	2,52	3,41	2,97
28	80	4,30	4,43	4,25	4,45	3,83	4,89	4,36

3.2. Fuzzy groups of evaluated distances

The three-dimensional variables were collected for each place in the research. These dimensions describe the distance between crossings, the rating (grade) of the students, and the membership value for each place. Every researched "place" is now represented (e.g. in figure 3) as a "point". Particular groups of points are called clusters. An example of clusters' borders for four clusters is shown in figure 3. Each point can be a part of one or more clusters, which is determined by the membership value. The membership value can be expressed in the values from zero (no membership) to one (full membership). Membership values for all points, calculated with the use of FCM-algorithm [16], are presented in table 2. The presented analysis concerns the cases of 3 and 4 clusters. The areas highlighted in this figure embrace the points with significant membership values (generally more than 0,50). Two points (representing places 8 and 19) belong to two clusters (their membership is less than 0,50).

Users' grades and values of distances create a set of data for further analyses. FCM-algorithm was used to divide the data into a certain number of clusters. The data set has J variables (here $J = 21$). Each variable is characterised by K dimensions x_k (here $K = 3$). General formulas for calculating the centre of a cluster, the membership value, the distance of each variable from the centre of a given cluster, and for checking the condition of stopping the calculation process have the following forms:

$$(3.1) \qquad v_{k,i,n} = \frac{\sum\limits_{j=1}^{J} \mu_{ij,n}^{m} \cdot x_k}{\sum\limits_{j=1}^{J} \mu_{ij,n}^{m}} \ , \quad k = 1,2, \dots K \ ,$$

$$(3.2) \qquad \mu_{ij,n+1} = \frac{1}{\sum\limits_{l=1}^{C} \left[d_{ij,n} \Big/ d_{lj,n} \right]^{\frac{2}{m-1}}} \ ,$$

$$(3.3) \qquad d_{ij,n} = \sqrt{\sum\limits_{k=1}^{K} \left(x_k - v_{k,i,n} \right)^2} \ ,$$

(3.4)
$$\frac{1}{J \cdot C} \sum_{i=1}^{C} \sum_{j=1}^{J} \left| \mu_{ij,n+1} - \mu_{ij,n} \right| \le \varepsilon ,$$

where: $v_{k,i,n}$ is the centre of cluster i for dimension k in step n, $\mu_{ij,n}$ is the membership value j to cluster i in step n, m is the weighted exponent for the possible membership (here $m = 2$), x_k is the analysed variable, k is the next number of dimension, K is the total number of dimensions that characterise each variable, i is the number of the next cluster, n is the number of the step in calculations of the FCM-algorithm, j is the number of next variable, J is the total number of variables, C is the total number of clusters (here $C = 3$ or 4), $d_{ij,n}$ is the distance (Euclidian) of variable j from the centre of cluster i in step n, e is the parameter whose value is the condition of stopping the calculation process (here $\varepsilon = 0,1$).
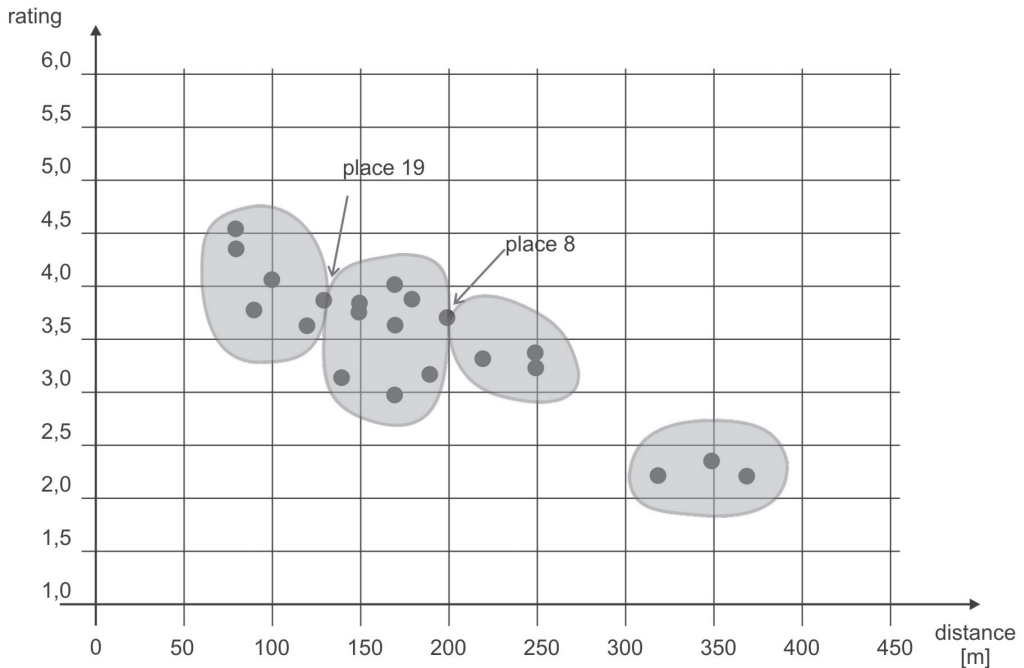

Fig. 3. Four fuzzy-clusters of examined database

Membership values for all places of research

Number of Place	Three-fuzzy-cluster case: membership value			Four-fuzzy-cluster case: membership value			
	cluster one	cluster two	cluster three	cluster one	cluster two	cluster three	cluster four
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3	0,00	0,00	1,00	0,00	0,00	0,00	1,00
6	0,09	0,72	0,19	0,02	0,03	0,93	0,02
8	0,00	1,00	0,00	0,04	0,47	0,46	0,03
9	0,03	0,95	0,02	0,01	0,07	0,91	0,01
12	0,87	0,12	0,01	0,21	0,73	0,05	0,01
13	0,09	0,72	0,19	0,01	0,03	0,93	0,03
14	0,70	0,28	0,02	0,07	0,90	0,03	0,00
15	0,10	0,88	0,02	0,02	0,91	0,06	0,01
16	1,00	0,00	0,00	0,68	0,27	0,04	0,01
17	0,97	0,03	0,00	0,98	0,02	0,00	0,00
18	0,93	0,06	0,01	1,00	0,00	0,00	0,00
19	0,97	0,03	0,00	0,43	0,50	0,06	0,01
20	0,70	0,28	0,02	0,07	0,90	0,03	0,00
21	0,26	0,72	0,02	0,00	0,99	0,01	0,00
22	0,90	0,08	0,02	0,98	0,01	0,01	0,00
23	0,02	0,97	0,01	0,04	0,72	0,22	0,02
24	0,26	0,72	0,02	0,00	0,99	0,01	0,00
25	0,01	0,02	0,97	0,01	0,01	0,03	0,95
26	0,01	0,04	0,95	0,01	0,03	0,08	0,88
27	0,26	0,72	0,02	0,00	0,99	0,01	0,00
28	0,90	0,08	0,02	0,97	0,02	0,01	0,00

Figure 4 shows the dependence between the distance and the membership value in the form of membership functions (as curves of various shapes). Places where the curves (bounded with following clusters) intersect, determine the borders between the classes of distances. Membership functions and their intersect-places determine the four groups of student's rate: "very good", "good", "bad", and "very bad".

A three-cluster case is presented in Figure 5 with the use of a simpler (triangular) form of membership functions. Three clusters make it possible to divide the distances into six groups. The borders between the classes of distances are determined by the

places where the membership functions intersect and where they reach the extreme values (zero or one). In the analysis of this case, one can relate to the LOS-methodology, as the distances were divided into natural classes, that is based on human perception. Each class of distances receives a name which denotes the grade given to the distances by the students. This name can mean the level of service (LOS), especially when the number of classes amounts to six. The names of three and six classes and also values of distances which are the borders of classes are presented in table 3.

Table 3

Names and values of distances for individual groups

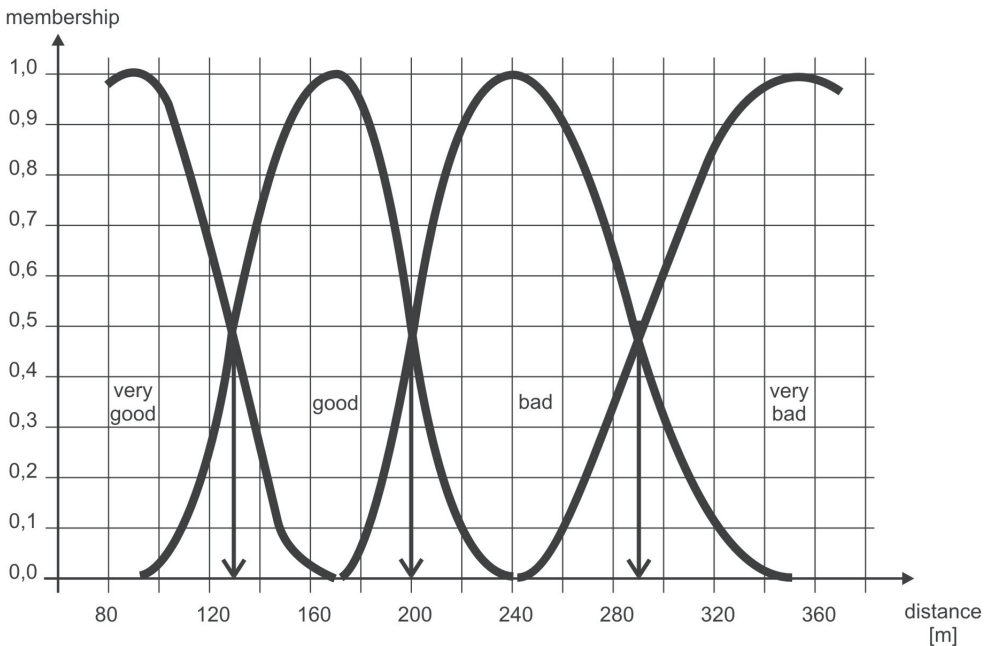| Distance [m]	Names for 3 classes	Names for 6 classes		LOS
less than 120	good	6	very good	A
120 ÷ 160		5	more good than average (good)	B
160 ÷ 200	average	4	more average than good	C
200 ÷ 275		3	more average than bad	D
275 ÷ 350	bad	2	more bad than average (bad)	E
350 and more		1	very bad	F

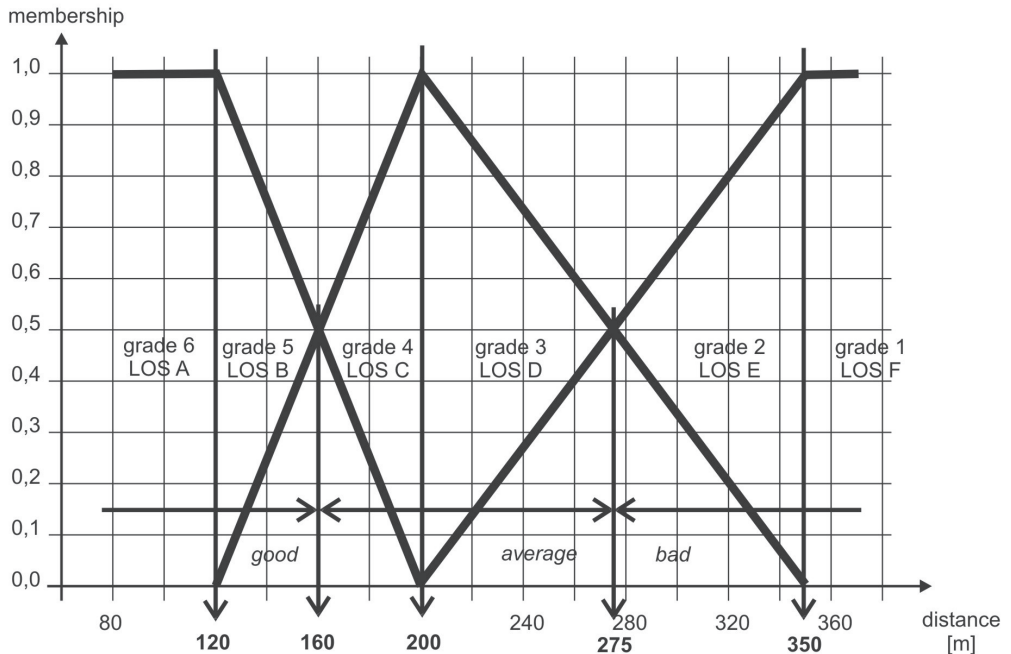Fig. 4. Membership functions for four-cluster case

Fig. 5. Three clusters and six groups of grade

4. DISCUSSION

4.1. COMMENTARY TO THE OBTAINED RESULTS

The research has shown how students evaluate the distance between pedestrian crossings. First, the analysis concerns only three classes of estimates. The distances smaller than 160m were considered 'good' (acceptable) while the distances greater than 275m were considered unacceptable. Also, 275m is more or less the maximum distance allowed by the road construction laws. In Poland, it is recommended by law that pedestrian crossings should be situated at least every 150-400m in city centres (depending on the level of urban development in that area). However, the results from this study suggest that distances between crossings greater than 200m are already evaluated as unacceptable (the membership of this sample is greater than 0). Only distances smaller than 120m are definitely acceptable (the membership in the class 'acceptable' equals 1). Such distances are also recommended by organizations promoting walking and friendly urban environments. For instance, the American society 'Walkable Communities' recommends that pedestrian crossings should be situated at least every 90m (300 ft). STRONG and YE [17] defined the "influence area" as 50 ft (about 30,5m) in either direction of a crossing. Thus, if the distance from one crossing to

another is greater than 61m, there is a zone between them that does not belong to their influence areas.

More detailed classification (into 6 clusters) allows applying the LOS method (by distinguishing between six classes) to evaluate the distances between the crossings. Levels A to F can be used to describe the quality of the road or a given part of a city in terms of pedestrian convenience while crossing the road. Also, by setting certain expected quality standards for particular road categories it is possible to choose the best type of distance between the crossings for a given road. The elements of transport infrastructure are usually designed to fit to LOS C. This study suggests that the optimal distance between pedestrian crossings should be between 160 and 200m. In the areas of greater importance to pedestrian traffic, the adherence to LOS B could be imposed. Then the distance between pedestrian crossings should measure between 120 and 160m.

4.2. Limitation of the study and future works

The research was conducted in one city only, on roads of similar categories, and on a unique pedestrian population. Further research is therefore necessary to generalize the obtained results.

The suggested boundaries between particular LOS-es could then be modified depending on the size of the city and vehicular, as well as pedestrian traffic flow rate. Interesting would be a comparison between the evaluation of distances and the tendency to cross the street illegally and in unsafe places (without crossings), as well as between the real distances and accident rates in particular places.

5. Conclusions

Despite the limitations mentioned above, this study could be useful to plan and design localisations of crossings. The obtained results, namely the evaluated distances, are similar to the values recommended in literature. Yet, the distances highly evaluated by students do not often occur in reality which means that there is a need to build new crossings, especially in the city centre where pedestrian traffic is or should be important.

Fuzzy grouping could be helpful to define the borders between classes of distances with reference to human perception. It was shown that in the case of the use of the FCM-algorithm the division of distances into four classes is the most distinct, but the use of the LOS methodology prompts that when it comes to human perception, the division of distances into three classes is more appropriate – the names of such classes ("good", "average", and "bad") seem to be clear and well understood.

References

1. de Lavalette, B.C., Tijus, Ch., Poitrenaud, T., Leproux, Ch., Bergeron, J., Thouez, J.-P., 2009. Pedestrian crossing decision-making: A situational and behavioral approach. Safety Science 47 (10), 1248-1253.

2. Papadimitriou, E., 2012. Theory and models of pedestrian crossing behaviour along urban trips. Transportation Research Part F 15 (1) 75–94.

3. Sisiopiku, V.P., Akin, D., 2003. Pedestrian behaviors at and perceptions towards various pedestrian facilities: an examination based on observation and survey data. Transportation Research Part F 6 (4), 249-274.

4. Holland, C., Hill, R., 2007. The effect of age, gender and driver status on pedestrians' intentions to cross the road in risky situations. Accident Analysis and Prevention 39 (2), 224-237.

5. Cavallo, V., Lobjois, R., 2007. Age-related differences in street-crossing decision: the effects of vehicle speed and time constraints on gap detection in an estimation task. Accident Analysis and Prevention 39 (6), 934-943.

6. Transportation Research Board, 2000. Highway capacity manual. Special Rep. No. 209, National Research Council, Washington, D.C.

7. Fang, C.F., Pecheux, K.K., 2009. Fuzzy data mining approach for quantifying signalized intersection level of services based on user perceptions. Journal of Transportation Engineering 135 (6), 349-358.

8. Kikuchi, S., Chakroborty, P., 2006. Frameworks to represent the uncertainty when determining the level of service. Transportation Research Record. 1968, Transportation Research Board, Washington, D.C., 53-62.

9. Adeli H., Karim, A., 2000. Fuzzy-wavelet RBFNN model for freeway incident detection. Journal of Transportation Engineering 126 (6), 464-471.

10. Hawas, Y.E., 2004. Development and calibration of route choice utility models: neuro-fuzzy approach. Journal of Transportation Engineering 130 (2), 171-182.

11. Yin, H., Wong, S.C., Xu, J., Wong, C.K., 2002. Urban traffic flow prediction using a fuzzy-neural approach. Transportation Research Part C 10 (2), 85-98.

12. Fang, C., Elefteriadou, L., Pecheux, K.K., Pietrucha, M.T., 2003. Using fuzzy clustering of user perception to define levels of service at signalized intersections. Journal of Transportation Engineering 129 (6), 657-663.

13. Zhang, J.S., Leung, Y.W., 2004. Improved Possibilistic C-Means Clustering Algorithms. IEEE Transactions On Fuzzy Systems, 12 (2), 209-217.

14. Leski, J., 2003. Towards a robust fuzzy clustering. Fuzzy Sets and Systems 137 (2), 215-233.

15. Ali, Y.M., Zhang, L., 2001. A methodology for fuzzy modeling of engineering systems. Fuzzy Sets and Systems 118 (2), 181-197.

16. Bezdek, J.C., 1981. Pattern recognition with fuzzy objective function algorithm. Plenum Press, New York.

17. Strong, Ch., Ye, Z., 2010. Spillover effects of yield-to-pedestrian channelizing devices. Safety Science 48 (3), 342-347.