Humanities and Social Sciences **2024**
Research Journal 31, No. 3 (2024), pp 153-172 **July-September**

Received: July 2023
Accepted: September 2024
DOI: 10.7862/rz.2024.hss.38

Marek ZBOROWSKI[1]
Witold CHMIELARZ[2]

ON THE APPLICATION OF THE CONVERSION METHOD FOR WEBSITE EVALUATION OF POLISH BANKS IN 2022

This article aims to examine the application of the new conversion method for evaluating electronic banking services. The conversion method consists of aggregating the average ratings of the respondents. Its intuitiveness results from the data collection process; the results collected from the respondents are converted using the algorithm presented in this article into the final result. The data was collected in 2022 using the CAWI method. The results of the calculations are compared with the results obtained by the scoring method, and the differences are interpreted. The conversion method reduces the subjectivity of the test results. The results indicate that economic factors have a large impact on users' opinions about the quality of banking services. This is additionally confirmed by research conducted among this group of customers. The main implication for researchers is to demonstrate the effectiveness of the conversion method in the analysis of electronic banking services.

Keywords: e-banking, evaluation methods of websites, scoring method, MCDA, Conversion method.

1. INTRODUCTION

The primary objective of this article is to assess the suitability and applicability of the Conversion method for evaluating websites and banking applications. To date, the long-term research the authors have been conducting since 2008 has shown that such an analysis can be both justified and highly beneficial in the following situations (Chmielarz, Zborowski, 2013):

- ranking websites from the perspective of the average e-banking user,
- indicating the bank with the highest website usability that can serve as a model for other banks,
- demonstrating which website attributes the average user considers most important and is most likely to rate highly,
- creating guidelines and a potential model for bank website designers,

[1] Marek Zborowski, University of Warsaw, Poland; e-mail: mzborowski@wz.uw.edu.pl (corresponding author). ORCID: 0000-0003-4762-127X.
[2] Witold Chmielarz, University of Warsaw, Poland; e-mail: witek@wz.uw.edu.pl. ORCID: 0000--0002-9189-1675.

- exploring existing methods for evaluating bank websites and determining which one is the best from the user's perspective (e.g., considering ease of evaluation and interpretation of results).

Decision-making, especially financial decision-making when facilitated through a website, is a complex task that involves numerous evaluation criteria, often conflicting or closely interconnected. The relevance of these criteria depends mostly on the user's own judgement. Therefore, the analysis and evaluation of web services and applications can aid in facilitating such decision-making processes. There are many multi-criteria methods, sometimes very sophisticated ones, which support such decisions. Multi-criteria methods generally do not yield an optimal value for a single indicator but rather provide a 'compromise' Pareto-optimal value (Roy, 1996). Multi-criteria decision-making methods emerged in the second half of the 1970s (Keenney, Raiffa, 1976; Nijkamp, Delft, 1977). Recent studies (Wątróbski et al., 2019) show that in most cases multi-criteria analysis methods are used to solve these problems (MCDA – Multi-Criteria Decision Analysis Method).

The simplest of these are scoring methods (Krzanowski, 2000), involving the attribution of a specific rating scale to distinguished criteria. Scoring methods, relying on the direct opinions of users, are classified as subjective. With large numbers of users, the subjectivity of this method generally decreases. The subjectivity associated with scoring methods can also be mitigated by employing a preference scale that is assigned to the respondents. This scale can be calculated as an average of the sample or assigned randomly. Nevertheless, even though its results are similar to other methods, e.g. TOPSIS (Chmielarz, Zborowski, 2018), it is not regarded as very sophisticated. However, it offers undeniable advantages in terms of its ease of application during surveys and subsequent interpretation of results as well as comprehensibility from the users' perspective. In this context, it can seem perplexing that the group of AHP/ANP methods (Saaty, 1990; 2008a) are seen as objective, especially since users are asked to compare relationships between very distant variables (in the case of financial portals, financial and strictly technical attributes are assessed simultaneously), and with the same number of attributes being compared, the user has to fill in an incomparably larger number of tables, in which the relationships between attributes are often determined intuitively. If respondents do not grasp the underlying assumptions of the method, the researchers should explain those assumptions to them, sometimes repeatedly. The question arises: Might this not occasionally impact the outcome of the data collection?

A similar scenario arises when employing methods from the group of PROMETHEE (Brans, Mareschal, 2005; Brans et al., 1986) and ELECTRE (Karagiannidis, Moussiopoulos, 1997). The concept of relevance weights is still understandable to the user, but how does it differ from preference weights? Is the concept of indifference weights or the veto threshold involved? As a result, users of portals often provide assessments in a haphazard manner. Examples of subjectivism present in various methods considered objective can be numerous and diverse (Saaty, 2008b). Another issue concerns the interpretation of the obtained results. In simple methods, the ranking is determined by the position achieved based on the average ratings received, with potential relations between them being established by preference indicators. However, there arises a challenge of how to interpret the results of relational methods in a manner that is fully understandable to the user. Certainly, it is possible to determine the ranking of the positions occupied by individual websites, using relational methods. However, calculating the differences

between these positions presents challenges that are more logical than technical in nature. Also, statistical methods may not offer contribute to a better understanding of the results.

It is important to acknowledge that banking websites and applications possess one specific characteristic. Unlike e-commerce sites, their evaluation is influenced not only by technical attributes but also by crucial financial attributes.

In the existing research, a research gap exists regarding the development of an evaluation method that enables non-professional users to collect data easily and accurately, while simultaneously addressing the challenging issue of substitutability.

In order to at least partially address this problem, a proprietary, Conversion method, was devised to evaluate web portals, specifically tailored for evaluating e-banking sites. This method takes into account the relationship between the results obtained and the calculated averages for individual portals and their attributes. Since its first version in 2008, this method has also been revised and verified while examining data related to other industries. In this article, we revisit the verification of this concept by comparing the results received with the Conversion method with those obtained by simple scoring methods.

In order to achieve the above objective, the following structure of the article was adopted. After an introduction to the topic and presentation of the purpose of the article, the second section presents a literature review on banking services and applications and the multi-criteria methods used to evaluate them. The third section describes the research procedure, evaluation methods and the characteristics of the research sample. The next section presents the study findings and their discussion. The last section contains the conclusions, limitations of the study and further conditions of the research.

2. LITERATURE REVIEW

Methods for evaluating IT systems, including assessing customer interactions, have long played a pivotal role in IT project management. The emergence of the Internet and the consequent shift of customer interactions to the virtual realm further confirmed the need for their application. The evaluation process takes place on two levels:

- internal – the system fulfils as many functionalities as possible that are necessary for the operation of the organisation (cost minimisation),
- external – the system facilitates contacts with customers, which is translated into the number of orders (maximising revenue).

The problem of evaluating information systems, or their relevant communication and distribution channels, revolves around assessing their usefulness from both the customer and bank perspectives, as well as their effectiveness in terms of applications (related to return on investment) (Boudreau et al., 2001; Myers et al., 1997). Researchers sought to develop universal metrics based on economic calculations (Melone, 1990) as well as individual metrics derived from measures such as labour intensity (Czarnacka-Chrobot, 2009). However, a universal measure has not been found, and individual assessment would have required the creation of as many measures as there are banks, which also seemed unreasonable. Therefore, it is now considered necessary to adapt the method of solving this problem to the decision-making situation (Wątróbski, 2016).

The evaluation of banking websites and applications was initially treated in the same way as the evaluation of e-commerce websites (Selz, Schubert, 1997; Whiteley, 2000; Evans, King, 1999). These were scoring methods and scoring methods with preferences, where a set of evaluation criteria is first defined and then assigned a specific rating scale. The evaluation criteria focus primarily on functional and technical factors. Their evaluation

by users is subjective, as the set of criteria always contains certain attributes preferred to varying degrees by individual users (e.g., readability of text or colour scheme). It is also possible to use a set of indicators tailored to the industry (Web Assessment Index method), where the desired results are obtained on the basis of speed, screen navigation, accessibility and content analysis. (Miranda et al., 2006). To achieve a larger and more reliable survey sample, researchers often employ simple assessment methods that can be easily understood by the average user. Instead of relying on complex indicators that are challenging to interpret, these methods focus on collecting data using simple criteria. (Dinitz et al., 2005; Guru et al., 2003; Saraswat, Katta, 2008; Mateos et al., 2001; Chiemeke et al., 2006; Miranda et al., 2006; Hadhémi, 2005; (Migdadi, 2008).

In such methods, the criteria are mostly limited to: functionality (ease of navigation, search functionality, readability); usability (catalogue of services, site map); visualisation (colours, background, graphics, fonts, etc.); reliability and accessibility; and quality and care with regard to design and performance.

From the point of view of users, multi-criteria methods can be divided as follows:

- simple methods, which are unambiguous and most commonly used in research, especially mass surveys. These methods are designed to be easy to apply in practice, accepted readily by users, and transparent in their interpretation. They can be developed based on a preference scale estimated at random or by the user, and they strongly depend on the attributes adopted for evaluation,
- relative, pairwise comparisons methods, whose basis are: AHP/ANP, Fuzzy AHP (T. L. Saaty, 1990), Fuzzy AHP (Liu et al., 2020) and Fuzzy ANP (Senturk et al., 2016), which allow to assess the importance of a given attribute in an absolute way and in comparison with others. It is difficult to apply, especially in a data collection procedure, due to the need to fill in many tables with just a few evaluation criteria. Although it is susceptible to rank reversal, this method remains highly popular and frequently utilized despite its drawbacks. Formally, this method has been occasionally criticized for its ambiguity, as it can lead to comparisons between characteristics that differ significantly. This was partly offset by the assumptions of the ANP method, making it even more difficult to use on a massive scale for surveys,
- parametric methods, in which respondents should assign values to certain additional parameters. Practitioners generally avoid using such methods due to their lack of specificity in defining their actual significance. Such methods include e.g. PROMETHEE II, the main issue with the use of such a method is - as in the AHP method - the necessity of preliminary education of respondents (taking place before the evaluation process) (Piwowarski, Ziemba, 2009),
- multi-stage methods, in the first stage, criteria and rating scales are defined (and data are collected, without scoring analysis, and in the second stage, various multi-criteria methods are applied. The study (Wątróbski, Jankowski, 2015) listed nearly 300 similar methods, based on various theoretical assumptions, considering or excluding the user's assumed preferences, distances from assumed optimum levels, etc. While the first stage of this method is typically straightforward and easily comprehensible for users, the second stage can present challenges. Users may encounter difficulties in accurately selecting the appropriate evaluation option and making decisions based on their interpretation,
- hybrid methods, e.g., those in which a combination of methods, their parallel use, or possibly the optimisation of previous simulation variants is used to eliminate

possible shortcomings of a single method or to bring them closer to the interpretative standards of the results.

From the user's point of view, the choice of an appropriate solution to a research problem based on one or more of the above-mentioned methods depends mainly on:

- the prior relevance of the selection of attributes, pertinent to the issue or industry under consideration, even if they appear to conflict with one another,
- the ease and intuitiveness of applying the evaluation scale and/or the scale of proposed preferences (when some criteria are more important to the user than others) during the input data collection,
- interpretation of the results, i.e. the ease and comprehensiveness of evaluating the data and making informed decisions based on them.

However, users, especially from small and medium-sized companies, are often less receptive to computationally complicated methodologies that involve complex calculation procedures and parameters that are challenging to interpret, which forces them to interpret the results themselves. In practice, users do not always request access to or details concerning the method of calculation, typically only seeking the final results. They are often not only interested in the results but also in having the opportunity to express their perspectives on how the results were obtained. Decision-makers of large companies trusting the mechanisms of business analytics are less interested in the method of obtaining the results, more in their interpretation and the resulting recommendations. Managers from small and medium-sized companies (SMES) prefer simple methods, with a relatively simple and intuitive process of obtaining results, which they then have to present to a less sophisticated audience.

In practical terms, striking a balance between these two tendencies would be highly beneficial. This can involve adopting a relatively simple and user-friendly approach to collecting data on the evaluation of banking services and applications. Simultaneously, it is essential to employ a data processing method that generates results that support decision-making processes. Users frequently opt for either the simplest method available, driven by the need to rapidly gather a research sample, or the method they are most familiar with. They may choose a method for which they have software that applies the algorithm for obtaining results. Also, they may already know how to present the decisions made on the basis of the calculations in a convincing manner, and select a method based on this criterion. Thus, it is a question of finding a method that balances, from the user's point of view, the effort required to obtain a result and, at the same time, a result that satisfies the user.

To address this dilemma, a potential solution emerged in the form of developing a new proprietary method for evaluating websites, i.e. the Conversion method (Chmielarz, Zborowski, 2013). This approach, used by many authors, has resulted in dozens of competing methods in recent years, which have produced different results for solving the same problem, described by the same set of data.

Another method is to compare the results of selected methods for evaluating websites with each other. The conversion method the authors developed, based on minimising the distance from the averages, seemed, like the AHP method, to eliminate the subjectivity of the respondents' approach in the research sample. The results of such a comparison for the evaluation of banking websites and applications in Poland, at the end of 2022, are presented in this article.

3. METHOD

3.1. Research procedure

The authors' extensive research into websites and banking applications over the years [e.g. (Chmielarz, Zborowski, 2015; 2019; 2020a; 2021)] has led to adopting the following research procedure:

- a bibliographic analysis of recent developments in the field of website analysis and evaluation, with particular emphasis on banking websites and applications,
- construction of a pilot version of the questionnaire to verify the correctness and comprehensibility of the questions,
- development and refinement of the prototype questionnaire on the basis of the above evaluation, preparing the final version of the questionnaire and adopting an unambiguous scale for the evaluation of attributes during data collection,
- random selection of groups of respondents and inviting them to complete the questionnaire using the CAWI (Computer-Assisted Web Interview) method,
- collection of completed questionnaires, data acquisition and preliminary verification of the correctness of filling in the questionnaires,
- selection of a method to evaluate the banking services in order to compare their quality with the author's conversion method,
- comparative analysis and discussion of the results of the obtained calculations and their statistics,
- drawing conclusions and making recommendations for further quality assessment of bank websites and banking applications.

3.2. Presentation of the applied methods

3.2.1. Scoring scale

The first assessment was based on a simple scoring method. Its principles are simple and do not require any specialist knowledge from the respondents. It can be simplified to include the following stages:

1. Each attribute adopted can score, at most, one point on a standardised Likert scale. A five-point, simplified, standardised Likert scale (Likert, 1932) was adopted to assess the individual criteria in the banks most frequently used by customers:
   a. 1.00 – fully meets the criterion,
   b. 0.75 – almost fully meets the criterion,
   c. 0.50 – moderately fulfils the criterion,
   d. 0.25 – minimally meets the criterion,
   e. 0.00 – does not meet the criterion.
2. Ratings from all respondents were aggregated and an average was calculated.
3. The averages were then totalled for the two cross-sections: banks and individual attributes, and their percentages of the total were counted, which proved necessary for comparison with the results of the Conversion method.
4. The next step was to relate the totals obtained to the maximum possible score (16 points in the banks' cross-section; 18 points in the attributes' cross-section).
5. The scores obtained provided an average assessment of the fulfilment of the usability and functionality of the attribute in relation to the best score and allowed a ranking to be presented in both crosssections.

Scoring method is undoubtedly burdened by a high degree of subjectivity in the case of individual respondents' assessments, but during mass surveys this subjectivity is minimised. Subjectivity can be reduced by introducing an individual (subjective) or objective (entropy-determined) preference scale.

Thus, the results obtained are not inferior (Chmielarz, Zborowski, 2018) to those received in other, more formalised and complicated methods. This method also has the advantage that there is no need to estimate additional indicators, which may be either incomprehensible or difficult to estimate for the respondents. As a result, minimal effort is dedicated to preparing participants to complete the survey, and the results cannot be directly affected. Consequently, there is also a higher return than with other methods, especially AHP/ANP, where dozens of tables have to be filled in for the same number of evaluation criteria. Their subsequent interpretation is also easy and comprehensible, both for analysts and users, and the results are easily presented graphically.

In order to reduce the presumed subjectivity of the evaluation of the banking services, the authors created the especially designed Conversion method (see below), which made it possible to relate the ratings provided by the respondents to the calculated average across the individual banks and across the individual criteria. Furthermore, the conversion method is based on the same source data, so that the respondent does not need to be familiar with the algorithm for calculating the results.

3.2.2. The Conversion Method

Below some assumptions for the Conversion method (Chmielarz & Zborowski, 2013) were presented: after constructing the experts' table of evaluations of average particular criteria for each website, the researchers need to perform the conversion with the established preference vector of the superior level criteria (Beaudrie et al., 2020), (Loach et al., 2016). Next, the authors perform the transformation of the combined scoring table into the preference vector (first converter):

The next steps are:

- constructing a matrix of distances from the maximum value for each criterion in every website:
  - establishing the maximum value

$$P_{i,max} = Max\{f_i(a_j), \dots, f_n(a_m)\} \text{ for } i = 1, \dots, n \text{ and } j = 1, \dots, m \quad (1)$$

  - establishing the matrix of the distances from the maximum value

$$\delta\,(f_i(a_j)) = P_{i,max} - f_i(a_j) \text{ for } i = 1, \dots, n \text{ and } j = 1, \dots, m \quad (2)$$

  - calculating the average distance from the maximum value for each criterion

$$\overline{F_{i,j}} = \frac{\sum_{j=1}^{m} \delta\left(f_i(a_j)\right)}{m} \quad (3)$$

- as a result of the above operation, constructing a matrix of differences in the distance from the maximum value and the average distance according to criteria,
- for each bank website: constructing conversion matrices – modules of relative distances of particular criteria to remaining criteria (the distance from the same criterion is 0), the obtained distances below the diagonal are the converse of the values over the diagonal:

- averaging criteria conversion matrices $\overline{n}\,m$ creating one matrix of average modules of values for all criteria

$$\overline{A}_{i,j} = \frac{\sum_{i=1,j=1}^{'}(\alpha_{i,j}-\alpha_{i+2,j})}{n} \tag{4}$$

- transforming the conversion matrix of criteria into a superior preference matrix (calculating squared matrix, adding up rows, standardization of the obtained preference vector; repeated squaring, adding up rows, standardization of preference vector – repeating this iteration until there are minimal differences in subsequent preference vectors).

As a result of the above operations, we establish a criteria conversion matrix Ta_{mx1}:

- Next, the authors performed a transformation of the scores presented by experts on the level of a matrix specifying expert websites' evaluations for particular criteria (second converter) (Beaudrie et al., 2020). The results have been obtained in an analogical way:
  - constructing a matrix of distances from the maximum value for each criterion and each website:
    – establishing the maximum value

$$P_{i,max} = Max\{f_i(a_j), \dots, f_n(a_m)\} \text{ for } i = 1, \dots, n \text{ and } j = 1, \dots, m \tag{5}$$

    – establishing the matrix of distances from the maximum value

$$\delta\,(f_i(a_j)) = P_{i,max} - f_i(a_j) \text{ for } i = 1, \dots, n \text{ and } j = 1, \dots, m; \tag{6}$$

    – calculating the average distance from the maximum value for each website

$$\overline{F_i} = \frac{\sum_{j=1}^{m}\delta\left(f_i(a_j)\right)}{m} \tag{7}$$

- constructing a matrix of the differences of deviations from the maximum value and the average distance of the features from the maximum,
- for each criterion: constructing a matrix of transformations (conversions) of the differences of the average distance from the maximum value between the websites, analogically as presented above (the distance for a particular feature in the same website from the same website is 0), values below the diagonal are the converse of the values over the diagonal,
- constructing a module matrix of transformations of the differences of the average distance from the maximum value between the websites, for each criterion

$$\overline{A}_{i,j} = \frac{\sum_{i=1,j=1}^{n,m}(\alpha_{i,j}-\alpha_{i+2,j})}{n} \tag{8}$$

- for each module matrix of transformation of the differences of the average distance from the maximum value between the websites, squaring it, adding up rows, standardization of the obtained ranking vector and repeating this operation until the obtained differences between two ranking vectors for each criterion will be minimal.

As a result of the above-presented operations we obtain a conversion matrix of websites' evaluations: Tf_{mx1}:

- using the obtained vectors to construct a combined ranking matrix – returning to the matrix where in its side-heading there are criteria, in the heading names of bank websites by appropriate transfer of the obtained preference vectors for each criterion,
- multiplying the matrix obtained in such a way by the previously calculated preference vector

$$T' = T\ f \otimes Ta \tag{9}$$

- results and conclusions (note: the lowest distances, in this case, are the most favourable, comparability adjustments to other methods can be obtained by subtracting these values from 1 and their repeated standardization).

The basis for the creation of the presented method was the assumption that it should be easy to apply. The objective has been reached, which is visible in the number of advantages presented below. The only disadvantage of the method is the fact that the transformation of the results of the survey is connected with carrying out many complex operations. The advantages of this method are:

- in the case of considering a large number of evaluation criteria or alternatives, there is no significant increase in the number of questions in the survey,
- the ease of application (similar to the realization of a scoring method) which results from the fact that in the survey form there are questions concerning the subjective evaluation of the element,
- there are no measures, as in the case of e.g., ELECTRE method – veto threshold, which may not be fully understandable for the respondent (Kizielewicz et al., 2020),
- the possibility of the application of the method with the participation of people who are not experts in a particular field,
- the result of the calculations which takes the form of the importance of the evaluations of the examined objects.

3.3. Sample characteristics

The website rankings used in this study were based on data collected in autumn 2022. A total of 738 people were surveyed, of which 356 people completed the survey questionnaire fully and correctly, resulting in a survey return rate of over 48%.

The first version of the survey was validated with the participation of a pilot sample of 50 people, conducted in an academic setting. Attributes were consulted with people in academia working on e-banking issues before being included in the survey.

As there were suggestions in relation to earlier studies (Chmielarz, Zborowski, 2018) questioning the feasibility and desirability of assessing financial and technical attributes at the same time, a survey was carried out to examine this issue.

It addressed the following questions:

- substitutability or equivalence of assessing technical and financial parameters,
- selection based mainly on financial attributes,
- competitiveness of technical and financial attributes,
- the decision to reject a bank whose attributes involved poor technical and good financial conditions.

In response to the first question, more than 60% of respondents expressed that the features such as aesthetically pleasing website design with attractive colours, easy readability, user-friendly navigation, and sufficient functionality *definitely cannot* or *rather cannot* be rated equivalently to the lower cost of the most important products or services in a banking website. This demonstrates a negative attitude towards evaluating banking websites by technical attributes alone. More than 16% had no opinion on the subject. The approach to financial services necessitates that providers of e-banking services incorporate financial considerations into their strategy for 'attracting' users.

This is further supported by the response to the question regarding the potential selection of a banking service with better price parameters, as 69% of the respondents answered with *rather yes* and *definitely yes*.

However, the question concerning the competitiveness of very good technical and financial conditions did not give such a unambiguous answer. Nearly 42% of respondents answered *rather yes*, 35% *rather no*, the *yes* and *no* answers were spread more or less equally. Respondents thus judged that there is a need to assess technical and financial attributes simultaneously.

The final question addressed the decision to reject a bank that possesses solely good financial attributes but its design is not appealing and performance is poor. The responses *rather yes* and *definitely yes* were selected by 56% of respondents, and *rather no* and *definitely no* by nearly 30%. This confirms the previous observation that financial factors, according to respondents, are not sufficient for evaluating banking websites.

In addition, the criteria finally established – the attributes of the banking services – were examined in terms of their comprehensibility and importance for the average website user. The list in the pilot sample included 30 items. After verification, corrections and removal of the least important criteria, 18 attributes (criteria) divided into three groups were included for the evaluation of each website. The sets included economic, technological and anti-crisis criteria. A detailed list of attributes is provided in Table 1.

To participate in the evaluation, respondents were required to assess the websites of a familiar electronic bank in comparison to the site of another banking service. This condition aimed to gather responses from experienced individuals who have dealt with various e-banking sites. Thus, a total of 712 complete evaluations of banking sites were obtained.

In addition to the relevance assessment, respondents stated their preferences in relation to the contribution of individual attributes to the quality rating of the banking service. In relation to the individual attributes, it turned out that they did not deviate particularly from the average of 5.55% and, when analysed in groups, the differences were also minimal (Table 1).

The *Significance Weight* appearing in Table 1 indicates how important the individual criteria are to the respondents in the evaluation procedure of the e-banking services. The higher the value of the weight, the more important the listed criteria are for study participants. On the other hand, the *Preference Weight* means that there is a large difference in ratings between the evaluated services. The greater the perception of a difference in ratings, the higher the weight value is. Table 1 shows the average of the values of all responses.

Respondents rated the sixteen most frequently used banking websites (A1, A2, …, A16) of the following banks: Alior Bank SA, Bank Handlowy w Warszawie SA, Bank Millenium SA, Bank Pocztowy SA, Bank Polska Kasa Opieki (PKO SA), Bank Polskiej Spółdzielczości, BNP Paribas SA, Credit Agricole Bank Polska SA, Getin Noble Bank

(currently: VeloBank), ING - Bank Śląski SA, mBank SA, Nest Bank SA, PKO Bank Polski SA (PKO BP: Inteligo i IKE), Santander Bank Polska SA, Santander Consumer Bank SA and Toyota Bank Polska SA. All banking services with less than five ratings are not included in this list - 16 banks' ratings were rejected.

Table 1. Averaged relevance and preference indicators for individual attributes

No.	Attributes	Significance weights	Preference weights
Economical factors			
C1	Monthly fee for card PLN/month	80.49	41.58
C2	Fee for a transfer to the parent bank	77.22	41.24
C3	Fee for transfer to another bank	81.22	41.44
C4	Fee for issuing a debit card	58.57	38.43
C5	Interest rate on savings accounts	68.99	40.06
C6	Interest rate on loans of PLN 10,000	58.31	38.45
C7	Interest rate on deposits of PLN 10,000	62.65	38.97
C8	Annual nominal interest rates on personal accounts	63.42	39.14
C9	Direct debit	51.57	38.52
C10	Account maintenance PLN/month	84.38	41.02
Technical factors			
C11	Additional services	53.79	41.75
C12	Account access channels	75.63	42.94
C13	Security	85.81	42.10
C14	Visualisation	58.00	44.05
C15	Navigation	65.73	42.60
C16	Scope of functionality	70.71	41.89
C17	Readability and ease of use	76.80	44.48
Anti-crisis factors			
C18	Anti-crisis measures	61.15	41.28

Source: own work.

The research sample was selected in a diverse manner using a combination of purposive and random sampling methods. The study was conducted within an academic environment, with randomly selected student groups, and a survey link was also shared on the Internet to reach a broader range of participants (*Respondenci Do Ankiet Online*, n.d.). The age range was thus between 19 and 35 years. Admittedly, this choice may have influenced the results of the survey (41 million people in Poland are potential customers of internet and mobile banking, more than 54% among registered customers are active users of internet banking and 44% active users of mobile banking. The surveyed age group represents more than 65% of users). The respondents surveyed included more than 70% women and nearly 30% men. 19% of study participants reported having a bachelor's degree or undergraduate level education. The remaining 80% of the respondents indicated having completed secondary education. The largest group of people came from large cities (more than 200,000 inhabitants) and 19% came from rural areas. A quarter came from small, medium-sized and large cities – up to 200,000 inhabitants. Among those surveyed, there were 52%

students, 31% working based on a contract for specific work, service contract or being self-employed and 17% of respondents were working on the basis of a contract of employment. The most common occupations were office workers (63%), service workers (16%), professionals (8%) and workers employed for simple technical work (7%). Most describe their financial situation as good (61%), very good (22%), average (16%) and sufficient (2%).

The data on the evaluation of banking services are generally relatively homogeneous and consistent. Once these were obtained, a reliability test in the form of Cronbach's alpha coefficient was applied. For all criteria (attributes), the Cronbach's alpha coefficient indicates internal consistency and the reliability of the sample was greater than 0.80. The internal consistency measure of the 16 dependent variables, based on Cronbach's alpha coefficient, was 0.85 (0.90 for Cronbach's alpha calculated from standardised items), for the 18 items in total.

4. ANALYSIS AND DISCUSSIONS OF THE RESULTS

The first part of the survey was conducted using a scoring method (without preferences), with equivalent attribute weights. Based on the data collected from the questionnaires distributed via the Internet, a summary table of averages from the respondents' ratings was created. This table was summarised by rows (individual attributes) and columns (individual banks). The totals obtained were then related to the maximum possible values (using a standardised Likert scale). In this way, the shares of the ratings of the individual banks and separately - of the attributes - in the highest possible rating of the banking services and/or applications were obtained. The results are presented in the following charts (Figure 1 and Figure 2).

The presented ranking shows that the Velo Bank service (the state-acquired Getin Noble Bank) enjoyed the highest reputation, with a score of more than 92% in terms of meeting the clients' evaluation conditions. With a slightly lower score (91%), the PKO Bank website came in second place, followed closely by the BNP Paribas service (90%). The lowest rated bank websites are: Bank Polskiej Spółdzielczości (46%), Bank Pocztowy (46%) and Bank Handlowy w Warszawie (48%). On the whole, however, the average ratings are high, with only 19% of the banking websites scoring below 50% in the maximum usability score. However, the distribution of ratings for individual services was significant, with 46% for the last three services in the ranking, 14% among the first eight highest rated services, 39% among those rated above 50%, and 2% among the three lowest rated services. As many as eleven banking websites out of the sixteen analysed were ranked above the average rating of 76%.

The calculations demonstrate that customers attributed the highest scores to the Getin Noble Bank website. Regrettably, following its acquisition by the state, the bank witnessed numerous account cancellations, which were not related to the usability of the services offered. At that time, the website has also undergone a major overhaul. Its evaluation will probably be included in next year's survey. Therefore, based on the ranking, PKO S.A. emerges as a potential benchmark for both customers and bank website designers, closely followed by BNP Paribas.

The results of the scoring methods of selected banks are shown in Figure 1.

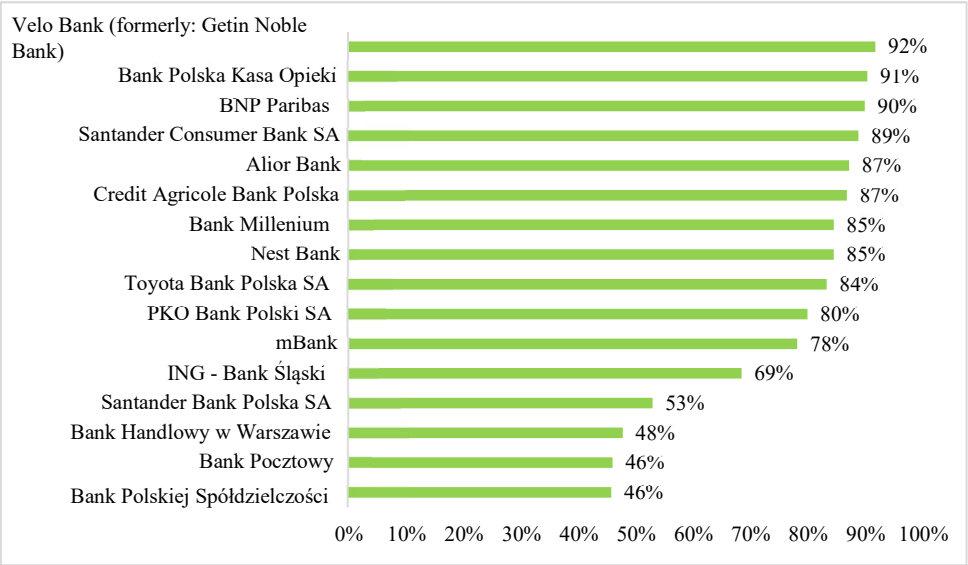

Figure 1. Scores obtained using the scoring method for selected banks

Source: own work.

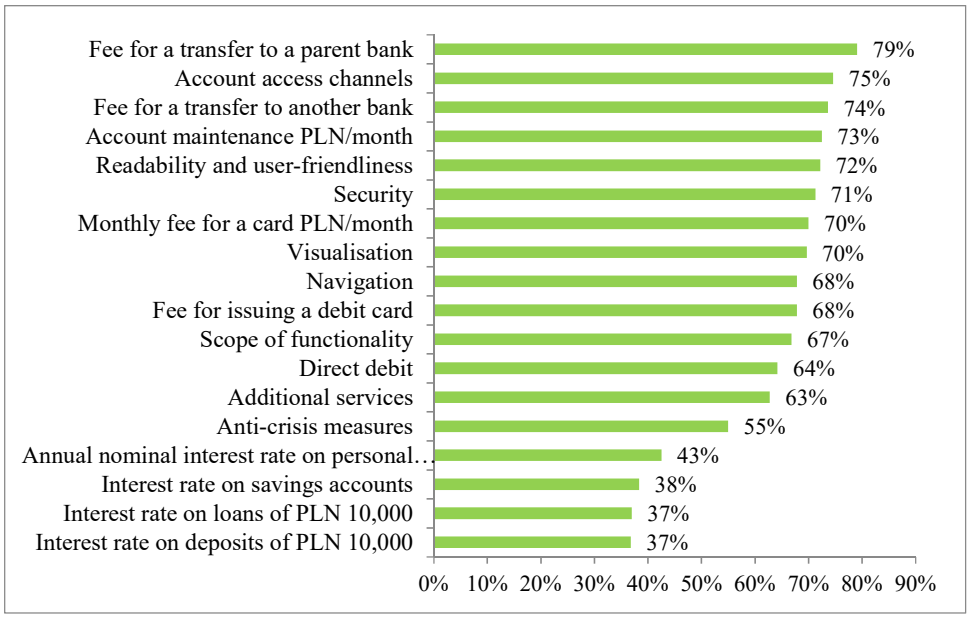

Figure 2. Results obtained with a scoring method according to attributes (evaluation criteria)

Source: own work.

The individual attributes of the banking services were rated slightly less positively on average (average attribute rating - 62%), although, as before, as many as 68% of the

attributes had a rating higher than average. The best scores (for financial data, the lower the attribute value, the higher the score) were for fees for transfers to the parent bank (79%) and the number and quality of account access channels (75%). Low fees for transferring funds to another bank and for account maintenance come next. It is worth noting that technical attributes: readability and ease of use and security features appear in the ranking from the fifth position onwards. Still, acceptance is above 70% of the possible highest rating. The worst rated financial factors were low interest rates on deposits and loans (37% each) and low interest rates on savings accounts (38%). The rating of the attributes of banking services is thus also linked to high inflation and indirectly to the whole policy of the state and the banking sphere. The spread of results here is slightly greater than in the previous statement (42%) overall, 24% in the above average group, 18% below average. The results are shown in Figure 2.

To compare the results of the scoring method and the conversion method, a form of standardization was employed. This involved transforming the results obtained from the first method to match the format used in the second method. This was done by relating the results for individual banks and attributes to the sum of the averages in each bank and for each criterion.

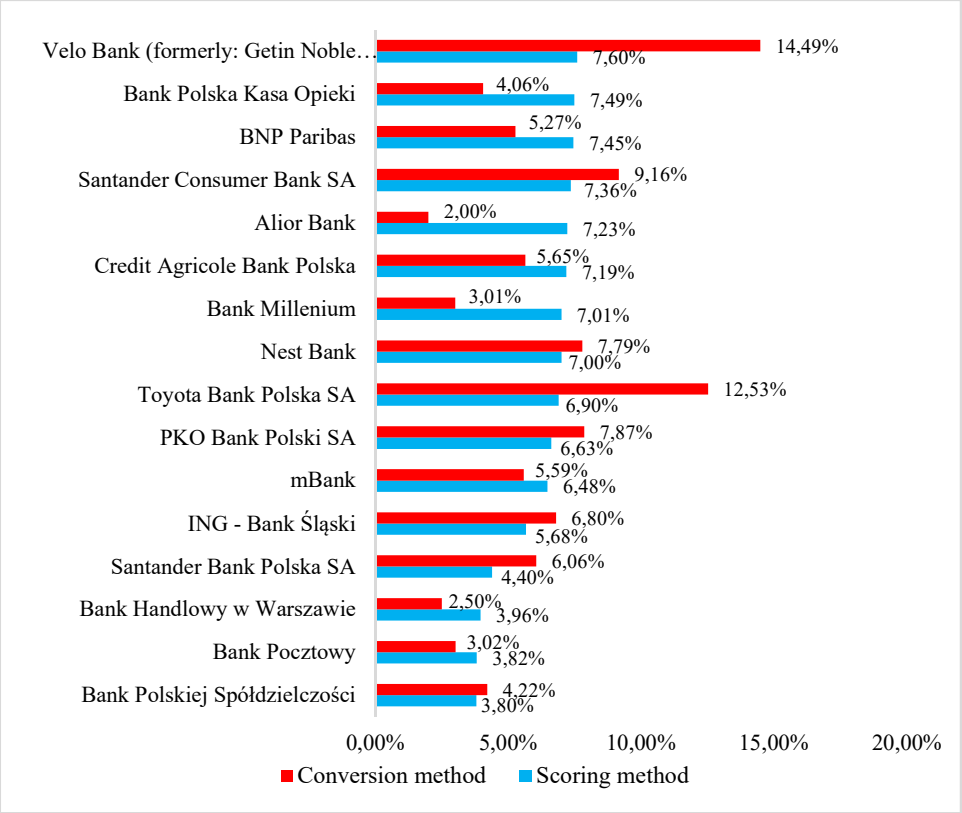

Figure 3. Comparison of results obtained by selected banks using the scoring method and the conversion method

Source: own work.

This allowed the ranking to be compared, on the one hand, according to the banks analysed and, on the other, according to the attributes the respondents evaluated. The reference to the average in the conversion method gave a completely different ranking of the banks compared to the results obtained by the scoring method. In Figure 3, it can be seen that while the results in the first position coincide, the online Toyota Bank, already in second position in the conversion method, was in third position in the scoring method. On the other hand, Santander Consumer Bank which previously came third, in the scoring method found itself in fourth place. Similarly, at the bottom end of the ranking, there are discrepancies between the two methods, although the list of the worst-rated bank websites is the same.

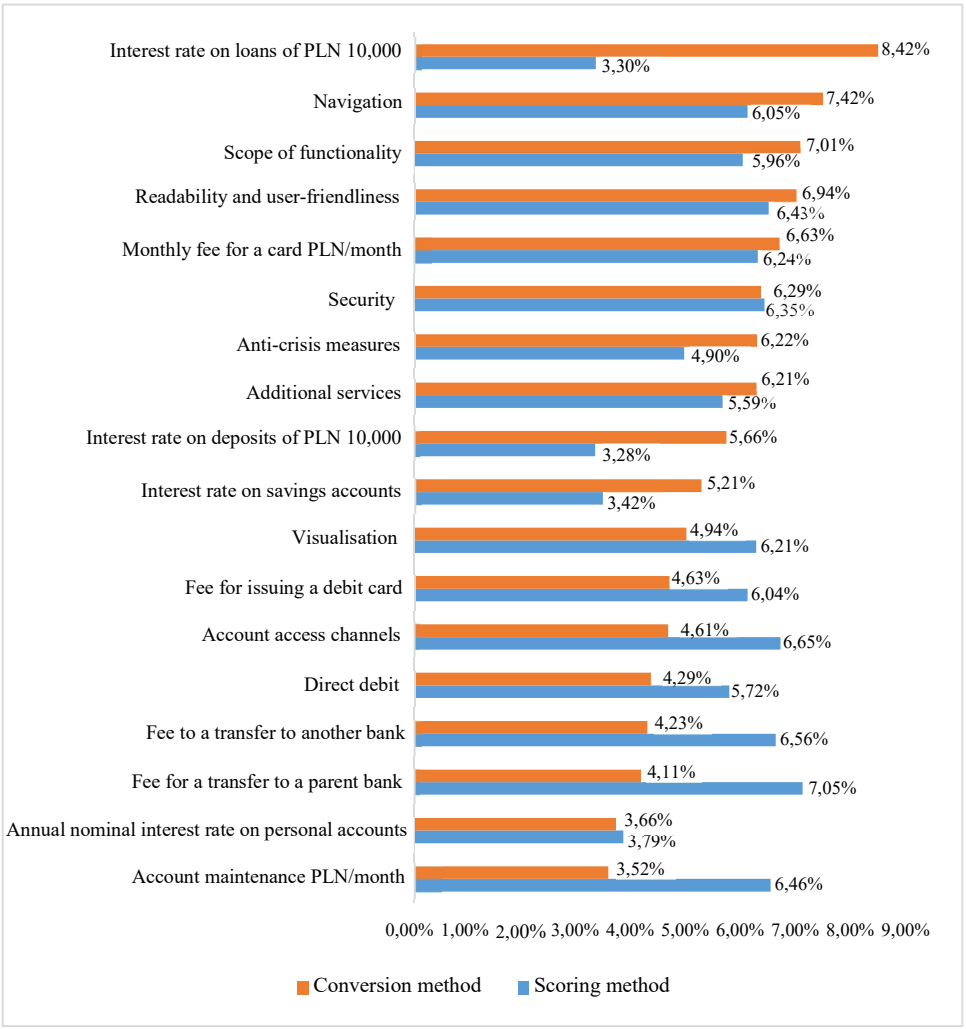

Figure 4. Comparison of the results obtained by attributes of selected banks using the scoring method and the conversion method

Source: own work.

A similar situation occurred in the assessment of attributes. In this case, the differences were even more pronounced. Unfavourable loan rates as significant criteria ranked first in the conversion method but significantly lower in the scoring method, and account management which ranked among the top attributes in the scoring method, ranked last in the conversion method.

A summary of the ranking places in the scoring method and the conversion method for the banks can be found in Table 2.

Table 2. Summary of bank rankings in the scoring method and conversion method

Bank	Scoring method	Position in the scoring method	Conversion method	Position in the conversion method	Euclidean distance
Velo Bank (formerly: Getin Noble Bank)	7.60%	1	14.49%	1	0
Bank Polska Kasa Opieki	7.49%	2	4.06%	12	100
BNP Paribas	7.45%	3	5.27%	10	49
Santander Consumer Bank SA	7.36%	4	9.16%	3	1
Alior Bank	7.23%	5	2.00%	16	121
Credit Agricole Bank Polska	7.19%	6	5.65%	8	4
Bank Millenium	7.01%	7	3.01%	14	49
Nest Bank	7.00%	8	7.79%	5	9
Toyota Bank Polska SA	6.90%	9	12.53%	2	49
PKO Bank Polski SA	6.63%	10	7.87%	4	36
mBank	6.48%	11	5.59%	9	4
ING - Bank Śląski	5.68%	12	6.80%	6	36
Santander Bank Polska SA	4.40%	13	6.06%	7	36
Bank Handlowy w Warszawie	3.96%	14	2.50%	15	1
Bank Pocztowy	3.82%	15	3.02%	13	4
Bank Polskiej Spółdzielczości	3.80%	16	4.22%	11	25

Source: own work.

The largest differences measured by Euclidean distance occurred in both rankings in case of the position of Alior Bank and Bank Polska Kasa Opieki, the smallest in the position of Velo Bank, Santander Consumer Bank and Bank Handlowy w Warszawie. These are mainly due to the evaluation of individual attributes and their relation to the average. Nonetheless, they appear to be, especially in relation to the shares in the assessment of individual banks. Efforts have been made to minimize these differences through research, utilizing penalty function coefficients to address the discrepancies in the averages after rows and columns. This approach follows the principles initially described in the study (Chmielarz, Zborowski, 2020b).

The rating of individual attributes in the analysed banks revealed more than twice as many differences. The variation in opinions regarding attribute evaluation is significantly stronger, resulting in greater disparity in the ranking of financial criteria.

Table 3. Summary of attribute rankings in the scoring method and the conversion method

Service	Scoring method	Position in the scoring method	Conversion method	Position in the conversion method	Euclidean distance
Fee for a transfer to the home bank	7.05%	1	4.11%	16	225
Account access channels	6.65%	2	4.61%	13	121
Fee for a transfer to another bank	6.56%	3	4.23%	15	144
Account maintenance PLN/month	6.46%	4	3.52%	18	196
Readability and ease of use	6.43%	5	6.94%	4	1
Security	6.35%	6	6.29%	6	0
Card monthly fee PLN/month	6.24%	7	6.63%	5	4
Visualisation	6.21%	8	4.94%	11	9
Navigation	6.05%	9	7.42%	2	49
Fee for issuing a debit card	6.04%	10	4.63%	12	4
Scope of functionality	5.96%	11	7.01%	3	64
Direct debit	5.72%	12	4.29%	14	4
Additional services	5.59%	13	6.21%	8	25
Anti-crisis measures	4.90%	14	6.22%	7	49
Annual nominal interest rates on personal accounts	3.79%	15	3.66%	17	4
Interest rates on savings accounts	3.42%	16	5.21%	10	36
Interest rates on loans of EUR 10 000	3.30%	17	8.42%	1	256
Interest rates on loans of EUR 10 000	3.28%	18	5.66%	9	81

Source: own work.

A comparison of the results obtained with the simple scoring method and the conversion method show that, as opposed to comparing it with other methods (Chmielarz, Zborowski, 2018), there is no convergence in this case. Making decisions based on such divergent results would prove challenging. This is not an isolated occurrence since comparisons (to the extent which was possible) with the results of the AHP method yielded similar discrepancies (Chmielarz, Zborowski, 2014).

5. CONCLUSIONS

The conversion method presented in this article can be used to evaluate websites and web applications. This is prompted by the simple way of collecting data identical to the scoring method and the automated conversion of results, which does not require the user to engage in these processes. The interpretation of the results can present a certain challenge. On this basis, it is possible to establish a ranking of both the banks analysed and the attributes assessing the usability of banking services.

The calculations made clearly indicate the results, but their comparison with the results of the simple scoring method shows high differences in the rankings, which in the article are assessed on the basis of Euclidean distance. The results can be considered relevant as

they help minimize the subjectivity of evaluations and provide a different perspective on the assessment of modern information technologies, similar to the AHP.

The article was subject to certain limitations. Firstly, the research sample was limited in scope and would need to be extended in order to generalise the results obtained. Secondly, the conversion method should also be compared with other multi-criteria evaluation methods.

The abovesaid limitations point to the directions for future research on the applicability of the Conversion method for the evaluation of websites. Further research should consist of comparing the results of this method with a selected group of other multi-criteria methods (e.g. Promethee, Electra, TOPSIS, etc.).

The issue examined in the study should be considered mainly from the point of view of future users. The focus ought to be on minimizing their difficulties in data collection for evaluation purposes and ensuring the simplicity and unambiguous interpretation of the obtained evaluations.

REFERENCES

Beaudrie, C., Corbett, J.C., Lewandowski, T.A., Malloy, T., Zhou, X. (2020). *Evaluating the Application of Decision Analysis Methods in Simulated Alternatives Assessment Case Studies: Potential Benefits and Challenges of Using MCDA. "Integrated Environmental Assessment and Management"*, 17(1). DOI: 10.1002/ieam.4316.

Boudreau, M.-C., Gefen, D., Straub, D.W. (2001). *Validation in Information Systems Research: A State-of-the-Art Assessment. "MIS Quarterly"*, 25(1), Article 1. JSTOR. DOI: 10.2307/3250956.

Brans, J.P., Vincke, Ph., Mareschal, B. (1986). *How to select and how to rank projects: The Promethee method. "European Journal of Operational Research"*, 24(2). DOI: 10.1016/03772217(86)90044-5.

Brans, J.-P., Mareschal, B. (2005). *Promethee Methods* [In:] Figueira, J., Greco, S., Ehrogott, M., eds., *Multiple Criteria Decision Analysis: State of the Art Surveys* (p. 163–186). Springer. DOI: 10.1007/0387-23081-5_5.

Chiemeke, S.C., Evwiekpaefe, A.E., Chete, F.O. (2006). *The Adoption of Internet Banking in Nigeria: An Empirical Investigation*. "*Journal of Internet Banking and Commerce"*, 11(3), Article 3.

Chmielarz, W., Zborowski, M. (2013). *Conversion Method in Comparative Analysis of e-Banking Services in Poland* [In:] Kobyliński, A., Sobczak, A., eds., *Perspectives in Business Informatics Research* (p. 227–240). Springer. DOI: 10.1007/978-3-642-40823-6_18.

—— (2014). *Wykorzystanie metody AHP/ANP w konfrontacyjnej metodzie projektowania wzorcowego systemów informatycznych. „Informatyka Ekonomiczna"*, 1(31).

—— (2015). *Comparative Analysis of Electronic Banking Websites in Selected Banks in Poland in 2014*. DOI: 10.15439/2015F43.

—— (2018). *Analysis of e-Banking Websites' Quality with the Application of the TOPSIS Method – A Practical Study. "Procedia Computer Science"*, 126, 1964–1976. DOI: 10.1016/j.procs.2018.07.256.

—— (2019). On a comparative analysis of internet banking services by the conversion method and the scoring method. "*Journal of Economics and Management"*, 38. DOI: 10.22367/jem.2019.38.02.

—— (2020a). *A Hybrid Method of Assessing Individual Electronic Banking Services in 2019. The Case of Poland. "Procedia Computer Science"*, 176. DOI: 10.1016/j.procs.2020.10.09.

—— (2020b). *Towards Sustainability in E-Banking Website Assessment Methods. "Sustainability"*, 12(17), Article 17. DOI: 10.3390/su12177000.

—— (2021). *Towards VES Function for Creating a Sustainable Method for Evaluating e-Banking Websites Quality. "Procedia Computer Science"*, 192. DOI: 10.1016/j.procs.2021.09.292.

Czarnacka-Chrobot, B. (2009). *Wymiarowanie funkcjonalne przedsięwzięć rozwoju Systemów Oprogramowania Wspomagających Zarządzanie na bazie uogólnionych danych historycznych.* „*Studia i Materiały Polskiego Stowarzyszenia Zarządzania Wiedzą"*, 23.

Dinitz, J.H., Dukes, P., Stinson, D.R. (2005). *Sequentially Perfect and Uniform One-Factorizations of the Complete Graph. "The Electronic Journal of Combinatorics"*, 1–12. DOI: 10.37236/1898.

Evans, J.R., King, V.E. (1999). *Business-to-Business Marketing and the World Wide Web: Planning, Managing, and Assessing Web Sites. "Industrial Marketing Management"*, 28(4), Article 4.

Guru, B., Shanmugam, B., Alam, N., Perera, J. (2003). *An Evaluation of Internet Banking Sites in Islamic Countries. "Journal of Internet Banking and Commerce"*, 8(2), Article 2.

Hadhémi, A. (2005). *An evaluation of internet banking and online brokerage in Tunisia.* "*Proceedings of the 1st International Conference on E-Business and E-Learning (EBEL)"*.

Karagiannidis, A., Moussiopoulos, N. (1997). *Application of ELECTRE III for the integrated management of municipal solid wastes in the Greater Athens Area. "European Journal of Operational Research"*, 97(3). DOI: 10.1016/S0377-2217(96)00252-4.

Keenney, R.I., Raiffa, H. (1976). *Decision Analysis with multiple conflicting objectives*. Wiley&Sons.

Kizielewicz, B., Wątróbski, J., Sałabun, W. (2020). *Identification of Relevant Criteria Set in the MCDA Process-Wind Farm Location Case Study. "Energies"*, 13. DOI: 10.3390/en13246548.

Krzanowski, W.J. (2000). *Principles of Multivariate Analysis: A User's Perspective*. Oxford University Press.

Likert, R. (1932). *A Technique for the Measurement of Attitudes* (No. 140). New York University.

Liu, Y., Eckert, C.M., Earl, C. (2020). *A review of fuzzy AHP methods for decision-making with subjective judgements. "Expert Systems with Applications"*, 161. DOI: 10.1016/j.eswa.2020.113738.

Loach, K., Rowley, J., Griffiths, J. (2016). *Cultural sustainability as a strategy for the survival of museums and libraries. "International Journal of Cultural Policy"*, 23. DOI: 10.1080/10286632.2016.1184657.

Mateos, M. B., Chamorro, A., Gonzalez, F.J.M., Gonzalez Lopez, O.R. (2001*). A New Web Assessment Index: Spanish Universities Analysis. "Internet Research: Electronic Application and Policy"*, 11(3), Article 3. DOI: 10.1108/10662240110396469.

Melone, N.P. (1990). *A Theoretical Assessment of the User-Satisfaction Construct in Information Systems Research. "Management Science*, 36(1), Article 1. DOI: 10.1287/mnsc.36.1.76.

Migdadi, Y.K. (2008). *Quantitative Evaluation of the Internet Banking Service Encounter's Quality: Comparative Study between Jordan and the UK Retail Banks. "Journal of Internet Banking and Commerce"*, 13(2), Article 2.

Miranda, F.J., Cortés, R., Barriuso, C. (2006). *Quantitative evaluation of e-banking web sites: An empirical study of Spanish banks. "The Electronic Journal of Information System Evaluation"*, *9*(2), Article 2.

Myers, B., Kappelman, L., Prybutok, V.R. (1997). *A Comprehensive Model for Assessing the Quality and Productivity of the Information Systems Function. "Information Resources Management Journal"*, 10. DOI: 10.4018/irmj.1997010101.

Nijkamp, P., Delft, A. van. (1977). *Multi-Criteria Analysis and Regional Decision-Making*. Springer US. Access on the internet: https://www.springer.com/gp/book/9789020706895.

Piwowarski, M., Ziemba, P. (2009). *Metoda PROMETHEE II w wielokryterialnej ocenie produktów. „Polskie Stowarzyszenie Zarządzania Wiedzą"*, 18.

Respondenci do ankiet online (n.d.). Retrieved 10 May 2023. Access on the internet: https://ankieteo.pl/program-doankiet/respondenci.

Roy, B. (1996). *Multicriteria Methodology for Decision Aiding*. Springer US. DOI: 10.1007/978-1-4757-2500-1.

Saaty, T.L. (1990). How to Make a decision: The Analytic Hierarchy Process. "*European Journal of Operational Research"*, 48(1), Article 1. DOI: 10.1016/0377-2217(90) 90057-I.

—— (2008a). *The analytic hierarchy and analytic network measurement processes: Applications to decisions under Risk.* "*European Journal of Pure and Applied Mathematics"*, *1*(1), Article 1. DOI: 10.29020/nybg.ejpam.v1i1.6.

—— (2008b). *Decision making with the Analytic Hierarchy Process. "Int. J. Services Sciences"*, 1. DOI: 10.1504/IJSSCI.2008.017590.

Saraswat, A., Katta, A. (2008). *Quantitative Evaluation of E-Banking Websites: A Study of Indian Banks. "Icfai University Journal of Information Technology"*, 3(4), Article 4.

Selz, D., Schubert, P. (1997). *Web Assessment – A Model for the Evaluation and the Assessment of Successful Electronic Commerce Applications. "Electronic Markets"*, *7*. DOI: 10.1109/HICSS.1998.655278.

Senturk, S., Binici, Y., Erginel, N. (2016). *The theoretical structure of Fuzzy Analytic Network Process (FANP) with Interval Type-2 Fuzzy Sets. "IFAC-PapersOnLine"*, 49(12). DOI: 10.1016/j.ifacol.2016.07.706.

Wątróbski, J. (2016). An Ontology Supporting Multiple-Criteria Decision Analysis Method Selection [In:] *Intelligent Decision Technologies 2016: Proceedings of the 8th KES International Conference on Intelligent Decision Technologies (KES-IDT 2016)* (p. 89––99). DOI: 10.1007/978-3-319-39630-9_8.

Wątróbski, J., Jankowski, J. (2015). Guideline for MCDA Method Selection in Production Management Area [In:] Różewski, P., Novikov, D., Bakhtadze, N., Zaikin, O., eds., *New Frontiers in Information and Production Systems Modelling and Analysis: Incentive Mechanisms, Competence Management, Knowledge-based Production* (Vol. 98, p. 119––138). Springer. DOI: 10.1007/978-3-319-23338-3_6.

Wątróbski, J., Jankowski, J., Ziemba, P., Karczmarczyk, A., Zioło, M. (2019). Generalised framework for multi-criteria method selection. *"Omega"*, 86. DOI: 10.1016/j.omega. 2018.07.004.

Whiteley, D. (2000). *E-commerce: Strategy, technologies and applications*. McGraw-Hill.