

PAPER • OPEN ACCESS

Comprehensive evaluation of usability at the mobile end interface

To cite this article: Sailan Wang *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **573** 012037

View the [article online](#) for updates and enhancements.

Comprehensive evaluation of usability at the mobile end interface

Wang Sailan¹, Li Bei², Zhu Yiran³

^{1 2}Architecture and Urban Planning College, Southwest Minzu University, Cheng Du, Si Chuan, 610041, China

*Corresponding author's e-mail: 14795748@qq.com

Abstract. Research on interface usability has always been qualitative. This thesis conducted a quantitative research on the evaluation of usability of mobile terminal interface adopting Analytic Hierarchy Process in Management and Fuzzy Evaluation Theory in Fuzzy Mathematics. Firstly, according to the factors influencing users' evaluation of usability of mobile terminal, an evaluation index framework of mobile application was established. Then, according to the fuzzy comprehensive decision based on the data of questionnaire survey, a comprehensive evaluation of usability of mobile application was made. Finally, a detailed elaboration on the practice and application of this evaluation method was offered with the example of a tourism mobile application.

1. Introduction

The netizen number in China reached 0.829 billion and the popularization rate was 59.6% by the end of December 2018, with a 3.8 percentage rise from the end of 2017, and the number of new netizens in the whole 2018 was 56.53 million. The scale of mobile phone users in China is about 0.817 billion, of whom 98.6% connect to the Internet, and the number of new mobile phone users is up to 64.33 million. [1] The dependence on the mobile Internet in entertainment, social interactions and life services has been gradually intensified. The Internet big data company QuestMobile issued China Mobile Internet User Behavior Insight Report 2019, which showed the refined operation of users in 2019 was increasingly treasured. It is reasonable that in the near future, the Internet user refined operation will be targeted at the maximization of user values, and the user activeness, retention or payment index will be improved through the operation means.[2]

2. Research review

Donald Norman proposed in 1980s that usability and apprehensiveness were two parallel attributes of products and any product to possess these two attributes should meet two conditions. First, predictability: users can predict the results of any behavior; Second, visibility: users can see the operational effect and directly judge whether the operation is efficient. This viewpoint was based on his research conclusions made on the use of daily commodities. Donald Norman elaborated usability in his book. Davis proposed the technology acceptance model (TAM) in 1986 and considered usability as one key influence factor on the acceptance of information technology by users in addition to other influence factors of usefulness and satisfaction to techniques. [3]Hartson defined the usability of information systems in 1998 and defined usability as the interfacial interactive efficiency, learnability, and user satisfaction. Microsoft proposed the famous Microsoft usability guide (MUG) in 2006, which became a key index of web usability evaluation.[4]



The research on usability has always been synchronized with technical development. In recent years, experts and researchers concerning software and web design have diverted to the mobile end. Jakob Nielsen, an expert of web usability, targeted at mobile phones and panel computers for the first time in his *Mobile Usability* published in 2012, which offered foundation and guidance for small-screen design. [5] Lyndon Cerejo in his *Mobile User Experience Elements* in 2012 systematically studied the 12 elements of mobile end user experience, including usability. [6] So far, the research on mobile end usability gradually becomes deeper and systematic.

3. Mobile end usability evaluation index framework

In 2006, China released the national standard GB/T16260, which clarified the five aspects of usability, including understandability, learnability, operability, attractiveness and compliability, and this standard is suitable for diverse computer software products and relevant information systems. The MUG from Microsoft includes five factors of contents, usability, promotion, customized services, and emotions. Usability is thought as the endeavors made by users during the web cognition process and is decided by structures, goals and feedback. Structures refer to the overall organizational structure of the web; goals mean whether or not the theme of the web is definite and understandable; feedback is the operation and progression information given by the web to users.

Thus, the mobile App usability is affected by the vision usability, structure usability and interaction usability and is a psychological process of judging and balancing the three aspects and also a comprehensive result from the joint action of the three aspects. According to hierarchical analysis, the usability of mobile Apps is considered as the overall goal, and then vision usability, structure usability and interaction usability are regarded as the standard indices to be used to build a mobile App evaluation index framework (Table 1).

Table 1. The evaluation index framework of mobile Apps

Overall goal G	Criterion index C	Plan index P	
Usability of mobile Apps	Vision usability (u_1)	Text structure u_{11} Picture quality u_{12} Metaphor of controls u_{13} Interface style u_{14} Text size u_{15}	P_1
	Structure usability (u_2)	Task flow u_{21} Guidance function u_{22} Information content u_{23} Newer guidance u_{24} Customization u_{25}	P_2
	Interaction usability (u_3)	Way of input u_{31} Way of output u_{32} Gesture operation u_{33} Feedback of errors u_{34} Help tips u_{35}	P_3

4. Evaluation methods of mobile App usability

4.1 Selection of evaluation methods

Since usability is one aspect of user experience and a psychological feeling, no quantitative method of usability can be absolutely precise. Due to the ambiguous relationship between the research targets, here the concept of fuzzy mathematics was introduced, which can comprehensively evaluate the mobile end usability.

4.2 Analytic hierarchy process (AHP) of fuzzy decision

The analytical hierarchy process (AHP) proposed by the American operational researcher T. L. Saaty [9] is a brief and effective method to make decision for complex problems. The tested problem from the overall goal G to the plan index P involved three layers: the target layer G, the criterion layer C and the plan layer P (Table 1). Then the importance of the decision-making plan, or namely the weights of decision plans P1, P2, ..., Pn relative to the target layer G, was determined through pairwise comparison. Thereby, a very satisfactory decision was made.

AHP involves four steps:

1) Clarify the problem and build a hierarchical structure.

2) Create a pairwise comparison judgment matrix $A = (a_{ij})_{n \times n}$, which should be a positive reciprocal matrix and elements of which should meet:

$$a_{ii} = 1, a_{ij} = 1/a_{ji} > 0, i, j = 1, 2, \dots, n.$$

3) hierarchical ranking.

After the judgment matrix A is built, find its largest eigenvalue λ_{\max} and then from the corresponding feature equation, find the eigenvector W. The normalized W is just the weights of importance of the elements in one layer relative those in the upper layer. This process is called hierarchical ranking.

4) consistency test.

In order to overcome the subjectivity and one-sidedness during the creation of the judgment matrix A, the consistency of A should be tested. When the random consistency test coefficient is $0 \leq CR < 0.1$, the consistency of A and the corresponding eigenvector are acceptable; otherwise, A should be adjusted. The consistency test is divided into simple test and comprehensive test.

4.3 Steps of fuzzy evaluation

According to the rationale of fuzzy comprehensive decision-making,⁹ the fuzzy evaluation of mobile App usability was finished as follows:

1) According to the mobile App evaluation index frame and the evaluation rationale, determine the factor set and evaluation set of the evaluation target.

The factor set is $G = \{u_1, u_2, u_3\}$; $u_i = \{u_{i1}, u_{i2}, u_{i3}, u_{i4}, u_{i5}\}$ ($i = 1, 2, 3$), where G is the overall goal or namely the usability of mobile Apps; the primary indices u_1 to u_3 indicate the vision usability, structure usability and interactive usability respectively; the corresponding secondary indices are u_{ij} ($i = 1, 2, 3; j = 1, 2, 3, 4, 5$).

The evaluation set is $V = \{v_1, v_2, v_3, v_4, v_5\}$, where v_1 to v_5 correspond to the "very high, high, neutral, low, very low" degree of usability, respectively.

2) Build a fuzzy judgment matrix and use AHP to determine the weights of the indices.

3) Fuzzily evaluate each factor u_k of G according to the grade indices of the evaluation set V, forming a single-factor fuzzy evaluation matrix.

$$R^{(k)} = \begin{pmatrix} r_{11}^{(k)} & r_{12}^{(k)} & \cdots & r_{1m}^{(k)} \\ r_{21}^{(k)} & r_{22}^{(k)} & \cdots & r_{2m}^{(k)} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1}^{(k)} & r_{n2}^{(k)} & \cdots & r_{nm}^{(k)} \end{pmatrix} = (r_{ij}^{(k)})_{n \times m}$$

Suppose N experts participate in voting evaluation; $r_{ij}^{(k)}$ means the secondary index u_{ij} in factor u_k in the above matrix gets the frequency of grade v_j in V. Admittedly, $r_{ij}^{(k)}$ means the degree of u_k belonging to v_j (here $k = 1, 2, 3; i, j = 1, 2, 3, 4, 5; n = m = 5$).

4) Mark the weight matrix of u_k as w_k , and then the comprehensive evaluation result of u_k is:

$$b_k = w_k R^{(k)} = (w_{k1}, w_{k2}, \dots, w_{k5}) \begin{pmatrix} r_{11}^{(k)} & r_{12}^{(k)} & r_{13}^{(k)} & r_{14}^{(k)} & r_{15}^{(k)} \\ r_{21}^{(k)} & r_{22}^{(k)} & r_{23}^{(k)} & r_{24}^{(k)} & r_{25}^{(k)} \\ r_{31}^{(k)} & r_{32}^{(k)} & r_{33}^{(k)} & r_{34}^{(k)} & r_{35}^{(k)} \\ r_{41}^{(k)} & r_{42}^{(k)} & r_{43}^{(k)} & r_{44}^{(k)} & r_{45}^{(k)} \\ r_{51}^{(k)} & r_{52}^{(k)} & r_{53}^{(k)} & r_{54}^{(k)} & r_{55}^{(k)} \end{pmatrix}$$

$$= (b_{k1}, b_{k2}, b_{k3}, b_{k4}, b_{k5})$$

Here b_{kj} means the corresponding weight of the sub-index in u_k in the evaluation set V ($k = 1, 2, 3$).

5) Comprehensively evaluate the 3 judgement elements u_1 to u_3 in G and its fuzzy evaluation matrix is:

$$R = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} \end{pmatrix}$$

If the weight of the index set in G is W , then the comprehensive evaluation result is:

$$B = WR = (b_1, b_2, b_3, b_4, b_5)$$

where b_1, b_2, b_3, b_4, b_5 are the weights of corresponding grades in the evaluation set V of G , or namely the degrees of usability. According to the maximum membership principle, if b_2 is the largest, then this mobile system is the most usable; if b_5 is the largest, then this mobile system is the least usable.

5. Case study

One tourism APP -- "bread trip" was used here and its usability was evaluated via questionnaire and analysis. In brief, 132 in-school students were asked to install and use this APP, and then were questioned. Questions of this questionnaire were set using a Likert5 scale (1 means "very much disagree" and 5 means "very much agree"). Then the indices of the plan layer P were measured. Finally 130 copies of questionnaire were returned, of which 125 copies were valid. The usability of this system was evaluated through comprehensive analysis of the fuzzy data set.

A judgment matrix concerning the importance of the criterion layer relative to the target layer was built:

$$G = \begin{pmatrix} 1 & 5 & 3 \\ 1/5 & 1 & 1/3 \\ 1/3 & 3 & 1 \end{pmatrix}$$

The eigenvector corresponding to the maximum eigenvalue $\lambda_{\max} = 3.0387$ was calculated by the summation method:

$$W = (0.6333, 0.1062, 0.2605)$$

This eigenvector was the weight vector of the primary factors u_1 to u_3 relative to the goal layer G . The random consistency test coefficient of $CR = 0.0334 < 0.1$ indicates the judgment matrix G meets the requirement of consistency test and this set of weights is acceptable.

Similarly, the weights concerning the importance of the secondary factors relative to the primary factors can be determined in the same way (Table 2).

Table 2 .Weight coefficients of factor u1

Secondary index	Judgment matrix	Weight vector w1	Consistency test coefficient CR
u11	1 5 3 4 5	0.4815	0.0369 < 0.1
u12	1/5 1 1/3 1/4 1/3	0.0580	
u13	1/3 3 1 1 2	0.1737	
u14	1/4 4 1 1 2	0.1778	
u15	1/5 3 1/2 1/2 1	0.1090	

Table 3. Weight coefficients of factor u2

Secondary index	Judgment matrix	Weight vector w2	Consistency test coefficient CR
u21	1 3 1/3 1/4 3	0.1551	0.0765 < 0.1
u22	1/3 1 1/3 1/4 1	0.0756	
u23	3 3 1 1/5 3	0.2150	
u24	4 4 5 1 5	0.4840	
u25	1/3 1 1/3 1/5 1	0.0703	

Table 4. Weight coefficients of factor u3

Secondary index	Judgment matrix	Weight vector w3	Consistency test coefficient CR
u31	1 1 1 5 3	0.2719	0.0189 < 0.1
u32	1 1 1 5 3	0.2720	
u33	1 1 1 5 7	0.3252	
u34	1/5 1/5 1/5 1 1	0.0597	
u35	1/3 1/3 1/7 1 1	0.0714	

For the vision usability factor u1, the fuzzy evaluation matrix of the secondary factors (namely the membership grade of the evaluation set V) can be determined from the questionnaire results and through statistical analysis.

$$R^{(1)} = \begin{pmatrix} 0.2760 & 0.4984 & 0.7513 & 0.9593 & 0.8407 \\ 0.6797 & 0.9597 & 0.2551 & 0.5472 & 0.2543 \\ 0.6551 & 0.3404 & 0.5060 & 0.1386 & 0.8143 \\ 0.1626 & 0.5853 & 0.6991 & 0.1493 & 0.2435 \\ 0.1190 & 0.2238 & 0.8909 & 0.2575 & 0.9293 \end{pmatrix}$$

Then from Table 2, the evaluation result of factor u1 is:

$$b_1 = w_1 R^{(1)} = (0.4815, 0.0580, 0.1737, 0.1778, 0.1090) \begin{pmatrix} 0.2760 & 0.4984 & 0.7513 & 0.9593 & 0.8407 \\ 0.6797 & 0.9597 & 0.2551 & 0.5472 & 0.2543 \\ 0.6551 & 0.3404 & 0.5060 & 0.1386 & 0.8143 \\ 0.1626 & 0.5853 & 0.6991 & 0.1493 & 0.2435 \\ 0.1190 & 0.2238 & 0.8909 & 0.2575 & 0.9293 \end{pmatrix} = (0.3280, 0.4832, 0.6859, 0.5723, 0.7056)$$

As for the structure usability factor u2 and the interaction usability factor u3, the fuzzy evaluation matrices of their secondary factors are listed in Table 5.

Table 5 .Evaluation set membership grade matrices of the secondary factors

Secondary factor	Weight index	Very high	High	Neutral	Low	Very low	Comprehensive consistency test coefficient C_R
u11	0.4815	0.2760	0.4984	0.7513	0.9593	0.8407	0.0841 < 0.1
u12	0.0580	0.6797	0.9597	0.2551	0.5472	0.2543	

u_{13}	0.1737	0.6551	0.3404	0.5060	0.1386	0.8143
u_{14}	0.1778	0.1626	0.5853	0.6991	0.1493	0.2435
u_{15}	0.1090	0.1190	0.2238	0.8909	0.2575	0.9293
u_{21}	0.1551	0.4173	0.4893	0.7803	0.1320	0.2348
u_{22}	0.0756	0.0497	0.3377	0.3897	0.9421	0.3532
u_{23}	0.2150	0.9027	0.9001	0.2417	0.9561	0.8212
u_{24}	0.4840	0.9448	0.3692	0.4039	0.5752	0.0154
u_{25}	0.0730	0.4909	0.1112	0.0965	0.0598	0.0430
u_{31}	0.2719	0.8687	0.4314	0.1361	0.8530	0.0760
u_{32}	0.2720	0.0844	0.9106	0.8693	0.6221	0.2399
u_{33}	0.3252	0.3998	0.1818	0.5797	0.3510	0.1233
u_{34}	0.0597	0.2599	0.2638	0.5499	0.5132	0.1839
u_{35}	0.0714	0.8001	0.1455	0.1450	0.4018	0.2400

The degree of usability was divided into five levels of "very high, high, neutral, low, very low" and thus was evaluated using a questionnaire. Users were asked to evaluate each secondary index according to five levels of "very high, high, neutral, low, very low". Finally, a decision was made via the fuzzy method.

The evaluation results of factors u_2 and u_3 are (Table 5):

$$b_2 = w_2R(2) = (0.7544, 0.4814, 0.4047, 0.5799, 0.2501)$$

$$b_3 = w_3R(3) = (0.4618, 0.4501, 0.5050, 0.5745, 0.1541)$$

Thus, the overall fuzzy evaluation matrix is:

$$R = \begin{pmatrix} 0.3280 & 0.4832 & 0.6859 & 0.5723 & 0.7056 \\ 0.7544 & 0.4814 & 0.4047 & 0.5799 & 0.2501 \\ 0.4618 & 0.4501 & 0.5050 & 0.5745 & 0.1541 \end{pmatrix}$$

The usability evaluation result of this system is:

$$B = WR =$$

$$(0.6333, 0.1062, 0.2605) \begin{pmatrix} 0.3280 & 0.4832 & 0.6859 & 0.5723 & 0.7056 \\ 0.7544 & 0.4814 & 0.4047 & 0.5799 & 0.2501 \\ 0.4618 & 0.4501 & 0.5050 & 0.5745 & 0.1541 \end{pmatrix}$$

$$= (0.4081, 0.4744, 0.6089, 0.5737, 0.5136)$$

The above results suggest the membership grade concerning the usability of this system to the evaluation set $V = \{v_1, v_2, v_3, v_4, v_5\}$ maximizes to 0.6089. According to the maximum membership principle, its usability is graded as v_3 , or namely the usability of this system is neutral. The comprehensive consistency test coefficient $CR < 0.1$ indicates this result is acceptable. The above computation was conducted on Matlab.

6. Conclusions

The mobile commerce is upsurging and shows a prospective future. Research on mobile commerce has never stopped, ranging from the commercial success and huge market to the popular discussion in the academia. The usability of interactive design becomes a key factor deciding the success or failure of a mobile end product, and usability evaluation is a very difficult process due to the complexity and subjectivity. This study was aimed to quantitatively evaluate usability. This study has two contributions from the theoretical and practical perspectives.

Theoretically, this study modestly enriches the theoretical research on usability of interactive design, especially at the level of quantitative research. The target was the popular mobile end interface, and the methods were analytical hierarchy process and fuzzy evaluation in the field of management. The usability of the mobile end was evaluated using a quantitative method, which offers a new method for quantitative research in the future.

From the practical perspective, research on usability is the key factor deciding the survival or failure of mobile products. Our findings can guide the management practice of mobile product

development and provide practical meaning for research & development institutions. Enterprises can evaluate the usability of products at the research & development stage or before release into the market, and thereby can modify and improve the usability problems in advance, which will reduce the loss of users due to usability problems.

As for limitations, the tourism APP was evaluated only from the user perspective, and the users were generally in-school students, which may have caused some bias. Nevertheless, since this study was aimed to investigate the evaluating method of mobile end usability, this case study only suggests the practical application of this evaluation method. In the future, more comprehensive and richer data can be collected in other evaluations. Thus, this preliminary study on the evaluation of mobile end usability should be further deepened.

Acknowledgments

Supported by 2019 The Fundamental Research Funds for the Central Universities of Southwest Minzu University 2019SQN03

References

- [1] AppAnnie . (2019). <https://www.appannie.com/cn/intelligence/>
- [2] QuestMobile.(2019).<https://data.questmobile.com.cn/>
- [3] Cao Shengsheng., (2003) Preliminary study on product usability in the information society .
 Decoration., 135: 8-9.
- [4] Davis F D. (1986)A technology acceptance model of empirically testing new end-user
 information systems: theory and results. Massachusetts: Sloan School of Management,
 Massachusetts Institute of Technology, :78-80.
- [5] Hartson,RH.(1989) Human-Computer Interface Development: ConceptsAnd System for Its
 Management., *ACM Computing Suvreys.*, 1: 5-60.
- [6] Viswanath Venkatesh V. Ramesh.(2006) Web and wireless site Usability: Understanding
 differences and Modeling use,*MIS Quarterly.*30: 81-206.