Expert Systems with Applications 38 (2011) 10292–10302

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A fuzzy group analytical hierarchy process approach for software quality assurance management: Fuzzy logarithmic least squares method

Kevin Kam Fung Yuen [a,*], Henry C.W. Lau [b]

[a] Department of Business Administration, Zirve University, Kizilhisar Campus, Gaziantep 27260, Turkey
[b] School of Management, University of Western Sydney, Sydney, Australia

ARTICLE INFO

Keywords:
Fuzzy group analytical hierarchy process (AHP)
Decision analysis
Software quality assurance
Software engineering
Software project evaluation

ABSTRACT

This paper proposes a Fuzzy Group Analytical Hierarchy Process approach for assessing the quality of software, with judgments by a group of experts at different levels. The international standard of software quality attributes, ISO/IEC9126-1:2001 which comprises of 6 criteria with 27 subcriteria, is applied as the attributes of software quality. Regarding the prioritization method, the modified fuzzy Logarithmic Least Squares Method (LLSM) is applied to derive the importance weight vectors. The Fuzzy Prioritization Programming for Direct Rating Scales (FPP-DRS) on the basis of the modified fuzzy LLSM and the rescaling functions is proposed to design the direct rating scales in fuzzy number. The Fuzzy Synthesis Programming for Absolute Measurement (FSPAM) including Integrated Fuzzy Scores of Individual Expert (IFSIE) and Group Integrated Fuzzy Score (GIFS) is proposed to aggregate the criteria values from various experts. The proposed approach can help a group of various experts including developers, testers and purchasers, to measure the level of the software quality of the in-house development or the third party development.

1. Introduction

A high quality software system is one of the essential attributes of high confidence computer systems. Any malfunctions of the software systems can cause great inconvenience, even a disaster to the users. A high confidence system is a system which meets the standard of software quality. Software quality assurance plays an essential role in the development of such software systems.

Measurement of the software quality includes the measurement of in-house developments and a selection of vendors' products. SQA takes the responsibility to make the "go/not go" decision in this matter. If the quality of the product released or purchased is below standard, the company will suffer a significant loss. If the product development is behind schedule, the company also loses a lot. As it is difficult to build a perfect or error free software system or to purchase highly compatible software components, SQA must apply comprehensive techniques to determine whether the systems reach the right level of quality. These techniques include the clear definition of quality attributes, measurement tools, and the integration framework.

To determine software quality, quality metric models have been studied by many researchers. This research selects six attributes with 27 subcriteria in ISO/IEC 9126-1 (2001), which is the revision of 1991 version (ISO/IEC 9126, 1991), for the development of the proposed fuzzy group analytical hierarchy process.

ISO/IEC 9126-1 defines terms for the software quality characteristics and how these characteristics are decomposed into subcharacteristics. The details of this are given in Section 5. ISO/IEC 9126-1, however, does not describe how any of these subcharacteristics could be measured. To address this issue, three more parts are extended: ISO/IEC 9126-2 (2003), ISO/IEC 9126-3 (2003), ISO/IEC 9126-4 (2003). ISO/IEC 9126-2 (2003) defines external metrics which measure the behaviours of the computer-based system that includes the software. ISO/IEC 9126-3 (2003) defines internal metrics which measure the software itself. ISO/IEC 9126-4 defines quality in use metrics which measure the effects of using the software in a specific context of use. However, a drawback of the existing international standards is that they provide very general quality models and guidelines, but are very difficult to apply specific domain (Bertoa, Troya, & Vallecillo, 2006). Expert judgments using the proposed fuzzy group analytical hierarchy process (FGAHP) can be the ideal solution, especially for some subcriteria which are less computable.

Regarding measurement tools and the integration framework including aggregation techniques and synthesis approaches, the analytic hierarchy process (AHP) (Saaty, 1980, 1990, 2005) is the popular approach which enables the user to make decisions to address these issues. The software development of AHP can be found in various researches (e.g. Ossadnik & Lange, 1999; Zhu, Aurum, Gorton, & Jeffery, 2005). The limitation is that the measurement

* Corresponding author.
E-mail addresses: kevinkf.yuen@gmail.com (K.K.F. Yuen), H.Lau@uws.edu.au (H.C.W. Lau).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.02.057

scales for the value of the utility function, which is basically numerical and probabilistically judgmental, induce difficulties when making the evaluation. This can be solved by applying linguistic labels, represented by fuzzy numbers, as the scales for the software metrics. Studies in fuzzy AHP (e.g. Boender, de Graan, & Lootsma, 1989; Buyukozkan, Kahraman, & Ruan, 2004; Chang, 1996; Chang, Wu, & Lin, 2008; Chen, Tzeng, & Ding, 2008; Laarhoven & Pedrycz, 1983; Mikehailov & Tsvetinov, 2004; Wang, Elhag, & Hua, 2006, 2008) can be applied to address this limitation.

The outline of this paper is as follows. Section 2 reviews the related works. Section 3 presents the foundations of fuzzy theory. Section 4 presents the generic managerial process of FGAHP. Sections 5–9 illustrate the details of the generic process. Section 5 presents the (ISO/IEC 9126-1, 2001) software quality model which consists of six criteria with 27 subcriteria. Section 6 presents the rating scale schema. Section 7 presents the prioritization methods for deriving fuzzy importance and fuzzy metric scales of the hierarchical criteria using a modified fuzzy LLSM model. Section 8 illustrates the synthesis method to integrate the evaluation scores and importance of the hierarchical criteria. Section 9 shows the case study demonstrating the proposed approach and discussion of the applicability of the proposed model. The Conclusion is in Section 10. The notation summary is shown in the Appendix.

2. Related works

Related research about FGAHP in software quality measurement is rarely found in the literature. Two articles (Buyukozkan et al., 2004; Chang et al., 2008) which are identified address the selection problems.

Buyukozkan et al. (2004) developed a fuzzy AHP approach for the selection of software development strategy. The main problem is that the research used the extent analysis method (EAM) (Chang, 1996) in fuzzy AHP, which has been used in many studies as it is regarded as less complex than other methods. However, the recent research from Wang, Luo, & Hua (2008) pointed out that EAM was problematic. Wang et al. (2006, 2008) proposed a Modified Fuzzy Logarithmic Least Squares Method (MF-LLSM) as the appropriate alternative (Boender et al., 1989; Laarhoven & Pedrycz, 1983). The modified fuzzy LLSM is applied in this paper.

Chang et al. (2008) proposed fuzzy AHP for the selection of software projects using 21 subcriteria only (there should be 27 according to the document) in ISO 9126-1:2001. However, firstly, their paper just used ratio relative measurement with its limitations that several projects should be compared, so that the best one among them can be chosen. This paper applies absolute measurement. The advantage of this is that the project is evaluated by itself without any influence from other projects as there is no need to compare among alternatives; but there is a need to understand the performance of the project that is being measured. Secondly, the prioritization method is not determined. Users have to define the two parametric values $\alpha, \beta \in [0,1]$ to convert the fuzzy numbers into crisp values, and next to integrate the crisp inputs. However, the choices of α and β are infinite. The approach proposed in this paper does not take care of this parametric process as fuzzy aggregation by modified fuzzy LLSM is applied. Thirdly, their output values are not of fuzzy number. The output value of a fuzzy number suggests the reasonable interval for the decision maker. Usually the lower boundary of the fuzzy number is a pessimistic estimation of the quality measurement, the upper boundary is the opposite, and the modal value is the adoptive value. This is the beauty of a fuzzy number. Finally, the two studies (Buyukozkan et al., 2004; Chang et al., 2008) did not include group judgment capability as the aggregation methods appeared not to support this function. Conversely, algorithms of this new approach support group judgments.

3. Fuzzy theory foundations

FGAHP applies fuzzy linguistic labels. A fuzzy linguistic label can be represented by a fuzzy number which is represented by a fuzzy set (Zadeh, 1965, 1975, 1996). Fuzzy sets capture the ability to handle uncertainty by approximate methods. A triangular fuzzy number (TFN) is applied mostly in fuzzy theories and applications. The details are as follows:

If X is a universal set (or universal of discourse) of elements x's, then a fuzzy set α in X is a set of ordered pairs, i.e. $\alpha = \{(x, \mu_\alpha(x)) : x \in X\}$. μ_α is called the membership function or the grade of membership which defined as $\mu_\alpha : X \to [0,1]$.

A TFN is represented by 3-tuple (l, m, u), and its membership has the form:

$$\mu_\alpha(x) = \begin{cases} \frac{x-l}{m-l}, & l \leqslant x \leqslant m \\ \frac{u-x}{u-m}, & m \leqslant x \leqslant u\,, \quad l \leqslant m \leqslant u, \\ 0, & otherwise \end{cases} \tag{1}$$

l is the fuzzy up boundary, and u is the fuzzy low boundary and m is the modal value.

Consider two TFNs $\alpha_1 = (l_1, m_1, u_1)$ and $\alpha_2 = (l_2, m_2, u_2)$. Some of the operational axioms are as follows:

Addition:

$$\begin{aligned}\alpha_1 + \alpha_2 &= (l_1, m_1, u_1) + (l_2, m_2, u_2) \\ &= (l_1 + l_2, m_1 + m_2, u_1 + u_2)\end{aligned} \tag{2}$$

Subtraction:

$$\begin{aligned}\alpha_1 - \alpha_2 &= (l_1, m_1, u_1) - (l_2, m_2, u_2) \\ &= (l_1 - l_2, m_1 - m_2, u_1 - u_2)\end{aligned} \tag{3}$$

Multiplication:

$$\alpha_1 \cdot \alpha_2 = (l_1, m_1, u_1) \cdot (l_2, m_2, u_2) = (l_1 \cdot l_2, m_1 \cdot m_2, u_1 \cdot u_2) \tag{4}$$

Division:

$$\alpha_1 / \alpha_2 = (l_1, m_1, u_1) / (l_2, m_2, u_2) = (l_1 / l_2, m_1 / m_2, u_1 / u_2) \tag{5}$$

Inversion:

$$(l_1, m_1, u_1)^{-1} = (1/(u_1 \cup l_1), 1/m_1, 1/(u_1 \cap l_1)). \tag{6}$$

Eq. (6) is used for fuzzy reciprocal scales which will be discussed in Section 6.

Example 1. Let $\alpha_1 = (5, 6, 7), \alpha_2 = (\frac{1}{4}, \frac{1}{3}, \frac{1}{2})$. Then their memberships are:

$$\mu_{\alpha_1}(x) = \begin{cases} x-5, & 5 \leqslant x \leqslant 6 \\ 7-x, & 6 \leqslant x \leqslant 7\,, \\ 0, & otherwise \end{cases} \quad \mu_{\alpha_2}(x) = \begin{cases} 12x-3, & \frac{1}{4} \leqslant x \leqslant \frac{1}{3} \\ 3-6x, & \frac{1}{3} \leqslant x \leqslant \frac{1}{2}\,. \\ 0, & otherwise \end{cases}$$

Addition of both is $\alpha_1 + \alpha_2 = (\frac{21}{4}, \frac{19}{3}, \frac{15}{2})$, subtraction of both is $\alpha_1 - \alpha_2 = (\frac{19}{4}, \frac{17}{3}, \frac{13}{2})$, multiplication of both is $\alpha_1 \cdot \alpha_2 = (\frac{5}{4}, 2, \frac{7}{2})$, and division of both is $\frac{\alpha_1}{\alpha_2} = (20, 18, 14)$. Inversions of both are $\alpha_1^{-1} = (\frac{1}{7}, \frac{1}{6}, \frac{1}{5})$ and $\alpha_2^{-1} = (2, 3, 4)$.

4. Overview of FGAHP

The FGAHP Approach is associated with the following five steps:

Step 1: Problem definition analysis:

In this step a hierarchy structure to solve the domain problem is built. The structure includes an objective, single or multiple level(s) of criteria, rating scales for all criteria, and corresponding experts for the evaluation. More details are illustrated in Sections 5 and 6.

Step 2: Group assessment process:

This step contains three processes. Firstly, group pairwise comparisons are performed to determine the relative scores for the weights of the criteria. Secondary, other group pairwise comparisons are used to derive the relative scores of the rating scales for each criterion. Finally, a direct rating process is conducted for the absolute measurement.

Step 3: Assessment validation process:

The consistency ratio is obtained by comparing the consistency index with the random index (Table 1, Saaty, 2005) which is an average random consistency index derived from a sample of randomly generated reciprocal matrices using the scale in Table 2. CR has the form:

$$\text{Consistency ratio} = \frac{\text{Consistence index}}{\text{Random index}}, \tag{7}$$

where the consistency index is in the form

$$CI = \frac{\lambda_{\max} - n}{n - 1}, \tag{8}$$

$\lambda_{\max}$ is a principal eigenvalue of a pairwise matrix such that $\lambda_{\max} \geqslant n$. This method is good for measuring consistency by using the Eigen system method. An alternative method to get $\lambda_{\max}$ is the form

$$\lambda_{\max} = n + \frac{1}{n} \sum_{1 \leqslant i < j \leqslant n} \frac{\delta_{ij}^2}{1 + \delta_{ij}}, \quad \delta_{ij} = \left(\frac{a_{ij}}{w_i / w_j} - 1 \right) \tag{9}$$

In this paper, the Eigen value method is applied to measure the personal judgment using the modal value of the fuzzy number. To determine the validity, if CR > 0.1, the pairwise matrix is not consistent, then the comparisons should be revised. Otherwise, the pairwise matrix is accepted. In this paper, only the modal value of the fuzzy number is taken into consideration. This means the boundary is irrelevant to the consistency.

Step 4: Prioritization process:

Prioritization means to convert the pairwise matrices of the importance of the criteria classes to the weight priorities vectors, and to convert the pairwise matrices of ordinal rating scales of the criteria to the metric scales vectors. There are various methods. This research applies fuzzy modified LLSM. More details are given in Section 7.

Step 5: Synthesis process:

In this step the direct rating results of the group of experts are aggregated in order to classify the performance. The numerical values of the metric scales of the direct rating process come from the prioritization of the pairwise matrices. A weighted mean of various experts' direct judgments is applied to produce the aggregated result. The important techniques of FGAHP for software quality are addressed in Section 8.

Table 1
Random consistency index (RI) (Saaty, 2005).

N	1	2	3	4	5	6	7	8	9	10
RI	0	0	.52	.89	1.11	1.25	1.35	1.40	1.45	1.49

Table 2
Pairwise comparison scale schema.

Verbal scales	Labels	Fuzzy number
Equally important	EI	(1,1,1)
Weakly important	WI	(1,2,3)
Moderately important	MI	(2,3,4)
Moderately plus	MP	(3,4,5)
Strongly important	SI	(4,5,6)
Strongly plus	SP	(5,6,7)
Very strongly	VS	(6,7,8)
Very, very strongly	VVS	(7,8,9)
Extremely important	EXI	(8,9,9)
Reciprocals of above	Add prefix "I" to the above labels	For a fuzzy number is (l_i, m_i, u_i), the reciprocals of the fuzzy number is the inversion of the fuzzy number, i.e., $(1/(u_i \cup l_i), 1/m_i, 1/(u_i \cap l_i))$

5. Hierarchical software quality model

IEEE (1991) defined software quality as (1) the degree to which a system, component, or process meets specified requirements, and (2) the degree to which a system, component, or process meets customer or user needs or expectations. IEEE (1991) defined software quality assurance as (1) a planned and systematic pattern of all actions necessary to provide adequate confidence that an item or product conforms to established technical requirements, and (2) a set of activities designed to evaluate the process by which the products are developed or manufactured.

There are various quality models. Khan, Mustafa, & Ahson (2006) and Khosravi & Gueheneuc (2004) reviewed the hierarchical and non-hierarchical models of software quality attributes such as Factor–Criteria–Metrics Model, McCall's Model, Boehm's Model, FURPS and Dromey's Model.

This paper chooses the ISO/IEC 9126-1 (2001) model, which is the more generic model, as the evaluation criteria to measure the software quality. ISO/IEC 9126 (1991) has been replaced by ISO/IEC 9126 (software product quality) and ISO/IEC 14598 (software production evaluation). The software product quality characteristics defined in this part of ISO/IEC 9126:2001 can be used to specify both functional and non-functional customer and user requirements (ISO/IEC 9126-1, 2001).

ISO/IEC 9126-1 enables software product quality to be specified and evaluated from different perspectives by those associated with acquisition, requirements, development, use, evaluation, support, maintenance, quality assurance, and audit of software (ISO/IEC 9126-1, 2001). This paper facilitates the ISO/IEC 9126-1 model for software quality assurance. Fig. 1 shows the hierarchy model consisting of six criteria and 27 sub-criteria, which are defined as follows (ISO/IEC 9126-1, 2001):

[1] Functionality (C_1): The capability of the software product to provide functions that meet stated or implied needs when the software is in use under specified conditions.
   - [1.1] Suitability (C_{11}): the capability of the software product to provide an appropriate set of functions for specified tasks and user objectives.
   - [1.2] Accuracy (C_{12}): the capability of the software product to provide the right or agreed results or effects with the needed degree of precision.
   - [1.3] Interoperability (C_{13}): the capability of the software product to interact with one or more specified systems.
   - [1.4] Security (C_{14}): the capability of the software product to protect information and data so that unauthorized persons or systems cannot read or modify them, and authorized persons or systems are not denied access to them.
   - [1.5] Functionality compliance (C_{15}): the capability of the software product to adhere to standards, conventions or regulations in laws and similar prescriptions relating to functionality.

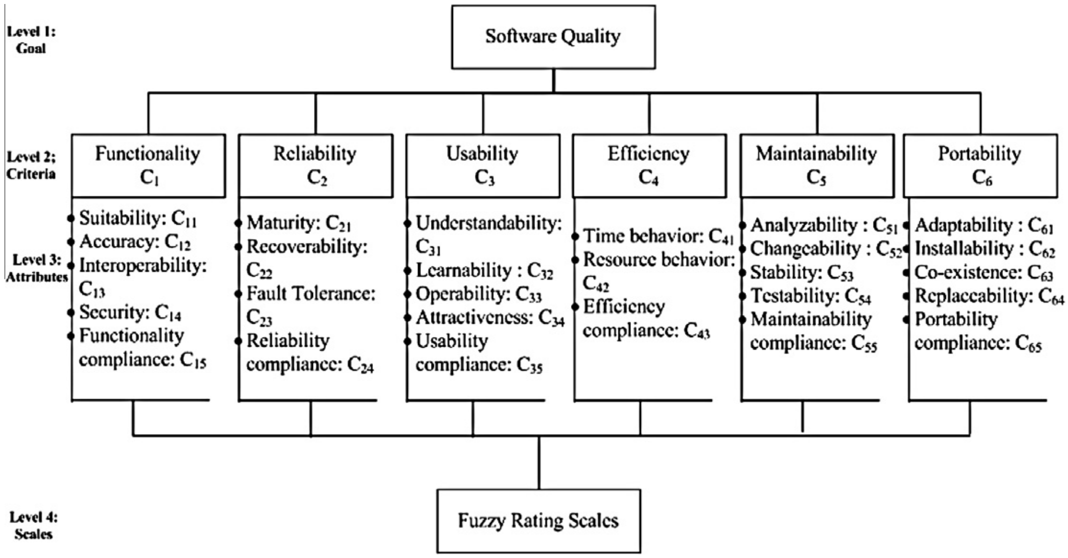

Fig. 1. FGAHP structure for evaluating Software vendor with respect to ISO six criteria.

[2] Reliability (C_2): the capability of the software product to maintain a specified level of performance when used under specified conditions.
- [2.1] Maturity (C_{21}): the capability of the software product to avoid failure as a result of faults in the software.
- [2.2] Recoverability (C_{22}): the capability of the software product to re-establish a specified level of performance and recover the data directly affected in the case of a failure.
- [2.3] Fault tolerance (C_{23}): the capability of the software product to maintain a specified level of performance in cases of software faults or of infringement of its specified interface.
- [2.4] Reliability compliance (C_{24}): the capability of the software product to adhere to standards, conventions or regulations relating to reliability.

[3] Usability (C_3): the capability of the software product to be understood, learned, used, and attractive to the user under specified conditions.
- [3.1] Understandability (C_{31}): the capability of the software product to enable the user to understand whether the software is suitable, and how it can be used for particular tasks and conditions of use.
- [3.2] Learnability (C_{32}): the capability of the software product to enable the user to learn its application.
- [3.3] Operability (C_{33}): the capability of the software product to enable the user to operate and control it.
- [3.4] Attractiveness (C_{34}): the capability of the software product to be attractive to the user.
- [3.5] Usability compliance (C_{35}): the capability of the software product to adhere to standards, conventions, style guides or regulations relating to usability.

[4] Efficiency (C_4): the capability of the software product to provide appropriate performance, relative to the amount of resources used, under stated conditions.
- [4.1] Time behavior (C_{41}): the capability of the software product to provide appropriate response and processing times and throughput rates when performing its function under stated conditions.
- [4.2] Resource behavior (C_{42}): the capability of the software product to use appropriate amounts and types of resources when the software performs its function under stated conditions.
- [4.3] Efficiency compliance (C_{43}): the capability of the software product to adhere to standards or conventions relating to efficiency.

[5] Maintainability (C_5): the capability of the software product to be modified. Modifications may include corrections, improvements or adaptation of the software to changes in environment, and in requirements and functional specifications.
- [5.1] Analyzability (C_{51}): the capability of the software product to be diagnosed for deficiencies or causes of failures in the software, or for the parts to be modified to be identified.
- [5.2] Changeability (C_{52}): the capability of the software product to enable a specified modification to be implemented.
- [5.3] Stability (C_{53}): the capability of the software product to avoid unexpected effects from modifications of the software.
- [5.4] Testability (C_{54}): the capability of the software product to enable modified software to be validated.
- [5.5] Maintainability compliance (C_{55}): the capability of the software product to adhere to standards or conventions relating to maintainability.

[6] Portability (C_6): the capability of the software product to be transferred from one environment to another.
- [6.1] Adaptability (C_{61}): the capability of the software product to be adapted for different specified environments without applying actions or means other than those provided for this purpose for the software considered.
- [6.2] Installability (C_{62}): the capability of the software product to be installed in a specified environment.
- [6.3] Co-existence (C_{63}): the capability of the software product to co-exist with other independent software in a common environment sharing common resources.
- [6.4] Replaceability (C_{64}): the capability of the software product to be used in place of another specified software product for the same purpose in the same environment.
- [6.5] Portability compliance (C_{65}): the capability of the software product to adhere to standards or conventions relating to portability.

There are two difficulties in using the hierarchical criteria. One is to derive the fuzzy importance (or fuzzy weights) among the hierarchical criteria as there are many tradeoffs. For example, reliability competes with efficiency. Usually the more reliable the less efficient, as the longer the processing time, the more resource utilization and the less efficiency compliance. Usability competes with functionality. More functions of a software application usually lead to less user-friendliness. Portability also competes with maintainability. The conclusion is that everyone has a different perception about the criteria. Thus it is necessary to design the grading method to balance their ideas.

The second difficulty is to derive the metric scale values for the subcriteria as the scale values are perceived differently for different subcriteria by different experts. To address this issue, the rating scale schemas are introduced before the prioritization method.

6. Rating scale schemas

Two categories of rating scale schemas are used in this paper. One is the absolute scale, and the other is a relative scale. The relative scales are used for the pairwise comparison matrices (Table 2) to determine the importance or weights of the criteria in different levels of the AHP model, and to determine the fuzzy scale values of obsolete rating scales by the prioritization method of the reciprocal matrix. The relative scales are used for comparison. For example, A is weakly more important than B. This means A is (1,2,3) times better than B. If the modal value of the fuzzy number is considered only, this means A is 2 times better than B.

The direct rating scales (or metric scales), which are the absolute scales, are used for the direct measurement of the criteria (Table 3), which are independent of each other. For example, the sentence "A is satisfactory", implies no comparison between A and others. However, the score for the satisfactory is undetermined at this moment, but will be discussed after the introduction of the prioritization methods. Here, this sentence can be assumed to mean that A is (0.3,0.5,0.7). If a modal value of the fuzzy number is taken, A is 0.5.

The intervals between two linguistic labels such as good and excellent have different score meanings with respect to functionality and reliability respectively, among different people. Table 7 illustrates this issue which will be discussed in the case study.

Table 3
Direct rating scale schema.

Linguistic scales	Labels	Fuzzy scale values
Poor	P	The scale values from 0 to 1 are determined by the normalized prioritization method
Weak	W	
Satisfactory	S	
Good	G	
Excellent	E	

7. Prioritization methods

This paper chooses the recent research, the modified fuzzy Logarithmic Least Squares Method (LLSM) (Wang et al., 2006, 2008), as the prioritization method in FGAHP model. The related comparison with extent fuzzy AHP approach (Chang, 1996) can be referred to (Wang et al., 2008), and the comparison with LLSM approach (Laarhoven & Pedrycz, 1983; Boender et al., 1989) can be referred to (Wang et al., 2006).

Consider a fuzzy comparison matrix expressed by Table 4. a_{ijk} is a triangular fuzzy judgment for comparing criterion i to criterion j from expert k. e_{ij} is the cardinal number of expert judgments. In addition, $a_{ijk} = (l_{ijk}, m_{ijk}, u_{ijk}) = a_{jik}^{-1} = (1/u_{jik}, 1/m_{jik}, 1/l_{jik})$ (by Eq. (6)) for $k = 1,\ldots,e_{ij}$, $i,j = 1,\ldots,n$ and $i \neq j$. $a_{ijk} = (1,1,1)$ if $i = j$.

The pairwise comparison matrices can be used to identify the relative scores for the weights of the criteria, and also can be used to assign the absolute scores of the rating scales for each criterion.

The modified fuzzy LLSM, which derives the priorities of the triangular fuzzy comparison matrix in Table 4, has the form:

$$\text{Min } J = \sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\sum_{k=1}^{e_{ij}}\left(\begin{array}{l}\left(\ln w_i^L - \ln w_j^U - \ln a_{ijk}^L\right)^2\\ +\left(\ln w_i^M - \ln w_j^M - \ln a_{ijk}^M\right)^2\\ +\left(\ln w_i^U - \ln w_j^L - \ln a_{ijk}^U\right)^2\end{array}\right)$$

$$\text{Subject to}\begin{cases} w_i^L + \sum\limits_{j=1,j\neq i}^{n} w_j^U \geqslant 1 \\ w_i^U + \sum\limits_{j=1,j\neq i}^{n} w_j^L \leqslant 1 \\ \sum\limits_{i=1}^{n} w_i^M = 1 & i = 1,\ldots,n \\ \sum\limits_{i=1}^{n}\left(w_i^L + w_i^U\right) = 2 \\ w_i^U \geqslant w_i^M \geqslant w_i^L > 0 \end{cases} \tag{10}$$

The solutions of the above optimum model are the normalized triangular fuzzy weights (fuzzy priority or fuzzy importance of the criteria) which is the form $w_i' = (w_i^U, w_i^M, w_i^L),\ i = 1,\ldots,n$.

For the fuzzy scale values for the direct rating scales (Table 3), this paper proposes Fuzzy Prioritization Programming for Direct Rating Scales (FPP-DRS), which the set of the fuzzy scale values are transformed from a reciprocal matrix by Eq. (10), and an extra step needed to be performed is to rescale $(w_i^L, w_i^M, w_i^U),\ i = 1,\ldots,n$ to $(\bar{w}_i^L, \bar{w}_i^M, \bar{w}_i^U) = \bar{w}_i$ by dividing $\max(w_1^U,\ldots,w_n^U)$, i.e.

Table 4
Group fuzzy pairwise comparison for matrix $\bar{A} = (a_{ijk})$.

Criteria	Criterion 1	Criterion 2	$\cdots$	Criterion n
Criterion 1	(1,1,1)	$\left\{\begin{array}{c}(l_{121}, m_{121}, u_{121})\\ \vdots\\ (l_{12e_{12}}, m_{12e_{12}}, u_{12e_{12}})\end{array}\right\}$	$\cdots$	$\left\{\begin{array}{c}(l_{1n1}, m_{1n1}, u_{1n1})\\ \vdots\\ (l_{1ne_{1n}}, m_{1ne_{1n}}, u_{1ne_{1n}})\end{array}\right\}$
Criterion 2	$\left\{\begin{array}{c}(l_{211}, m_{211}, u_{211})\\ \vdots\\ (l_{21e_{21}}, m_{21e_{21}}, u_{21e_{21}})\end{array}\right\}$	(1,1,1)	$\cdots$	$\left\{\begin{array}{c}(l_{2n1}, m_{2n1}, u_{2n1})\\ \vdots\\ (l_{2ne_{2n}}, m_{2ne_{2n}}, u_{2ne_{2n}})\end{array}\right\}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
Criterion n	$\left\{\begin{array}{c}(l_{n11}, m_{n11}, u_{n11})\\ \vdots\\ (l_{n1e_{n1}}, m_{n1e_{n1}}, u_{n1e_{n1}})\end{array}\right\}$	$\left\{\begin{array}{c}(l_{n21}, m_{n21}, u_{n21})\\ \vdots\\ (l_{n2e_{n2}}, m_{n2e_{n2}}, u_{n2e_{n2}})\end{array}\right\}$	$\cdots$	(1,1,1)

$$\bar{w}_i^L = \frac{w_i^L}{\max\left(w_1^U, \ldots, w_n^U\right)} \tag{11}$$

$$\bar{w}_i^M = \frac{w_i^M}{\max\left(w_1^U, \ldots, w_n^U\right)} \tag{12}$$

$$\bar{w}_i^U = \frac{w_i^U}{\max\left(w_1^U, \ldots, w_n^U\right)} \tag{13}$$

Following example demonstrates the prioritization problem.

Example 2 (*Prioritization problems*). Consider a 3×3 group pairwise matrix for 3 criteria by 2 expert judgments as follows:

$$Ex2 = \begin{pmatrix} (1,1,1) & \left\{\begin{matrix}(3,4,5)\\(2,3,4)\end{matrix}\right\} & \left\{\begin{matrix}(5,6,7)\\(6,7,8)\end{matrix}\right\} \\ \left\{\begin{matrix}\left(\frac{1}{5},\frac{1}{4},\frac{1}{3}\right)\\\left(\frac{1}{4},\frac{1}{3},\frac{1}{2}\right)\end{matrix}\right\} & (1,1,1) & \left\{\begin{matrix}(3,4,5)\\(4,5,6)\end{matrix}\right\} \\ \left\{\begin{matrix}\left(\frac{1}{8},\frac{1}{7},\frac{1}{6}\right)\\\left(\frac{1}{8},\frac{1}{7},\frac{1}{6}\right)\end{matrix}\right\} & \left\{\begin{matrix}\left(\frac{1}{5},\frac{1}{4},\frac{1}{3}\right)\\\left(\frac{1}{6},\frac{1}{5},\frac{1}{4}\right)\end{matrix}\right\} & (1,1,1) \end{pmatrix}$$

The consistence ratios for expert 1 and 2 are 0.073 and 0.062, respectively, which is within the acceptance range.

If the matrix $Ex2$ is treated as a fuzzy priority vector, then it can be solved by the modified fuzzy LLSM shown in Eq. (10), which can be calculated by *Excel*, *Mathlab* or *Mathematica*, and finally the fuzzy priority is

$$WEx2 = \left\{\begin{matrix}\left(w_1^L, w_1^M, w_1^U\right)\\\left(w_2^L, w_2^M, w_2^U\right)\\\left(w_3^L, w_3^M, w_3^U\right)\end{matrix}\right\} = \left\{\begin{matrix}(0.608, 0.673, 0.720)\\(0.208, 0.253, 0.314)\\(0.072, 0.074, 0.078)\end{matrix}\right\}$$

If the matrix $Ex2$ is treated as a Direct Rating Scales, which assign a numerical scale to the metric scale set (good, normal, bad), then three equations (Eqs. (11)–(13)) are applied to above fuzzy priority vector. From the above result, $\max\left(w_1^U, w_2^U, w_3^U\right) = 0.720$. Next the above fuzzy priority vector $WEx2$ is divided by 0.720, and then the fuzzy result $WS2$ of the DRS is:

$$WS2 = \left\{\begin{matrix}\textit{Good}\\\textit{Normal}\\\textit{Bad}\end{matrix}\right\} = \left\{\begin{matrix}(0.844, 0.934, 1)\\(0.289, 0.351, 0.436)\\(0.100, 0.102, 0.108)\end{matrix}\right\}$$

After the prioritization processes are performed, next the synthesis process is discussed.

8. Synthesis method

In the synthesis process, the direct rating scores of the group of experts for all criteria and the fuzzy importance of all criteria are aggregated. The importance of all criteria and the numerical representation of the direct rating scales for each criterion are determined in the prioritization process. Regarding the synthesis method, this research proposed Fuzzy Synthesis Programming for Absolute Measurement (FSPAM).This method is established on the synthesis methods of modified fuzzy LLSM (Wang et al., 2006, 2008), which were used for relative measurement.

FSPAM includes the Integrated Fuzzy Scores of Individual Expert (IFSIE) and the Group Integrated Fuzzy Score (GIFS) which are calculated with the presentation in Table 5.

IFSIE $\left(w_{E_k}^L, w_{E_k}^M, w_{E_k}^U\right)$ can be obtained by solving the following two linear programming models and an equation as follows:

$$w_{E_k}^L = \underset{w \in \Omega w}{\mathrm{Min}} \sum_{j=1}^{m} \bar{w}_{kj}^L w_j, \quad k = 1, \ldots, K, \tag{14}$$

$$w_{E_k}^U = \underset{w \in \Omega w}{\mathrm{Max}} \sum_{j=1}^{m} \bar{w}_{kj}^U w_j, \quad k = 1, \ldots, K, \tag{15}$$

$$w_{E_k}^M = \underset{w \in \Omega w}{\mathrm{Max}} \sum_{j=1}^{m} \bar{w}_{kj}^M w_j^M, \quad k = 1, \ldots, K, \tag{16}$$

where $\Omega_w = \left\{\begin{matrix}W = (w_1, \ldots, w_m)| \\ w_j^U \geqslant w_j^M \geqslant w_j^L, \sum_{j=1}^m w_j = 1,\ j = 1, \ldots, m\end{matrix}\right\}$ is the space of weights. $\left(w_j^L, w_j^M, w_j^U\right)$ is the normalized triangular fuzzy weight of criterion $j(j = 1,\ldots,m)$ and $\bar{w}_{kj} = \left(\bar{w}_{kj}^L, \bar{w}_{kj}^M, \bar{w}_{kj}^U\right)$ is the representative fuzzy number of the metric linguistic label rated by expert E_k with respect to the criterion j ($k = 1,\ldots,K; j = 1,\ldots,m$).

GIFS $\left(\tilde{w}^L, \tilde{w}^M, \tilde{w}^U\right)$ can be derived by the weighted average of $\left\{\left(w_{E_k}^L, w_{E_k}^M, w_{E_k}^U\right)\right\}$, and has the form:

$$\left(\tilde{w}^L, \tilde{w}^M, \tilde{w}^U\right) = \frac{1}{K} \sum_{k=1}^{K} \left(w_{E_k}^L, w_{E_k}^M, w_{E_k}^U\right) \tag{17}$$

Example 3 (*Synthesis problem*). Continue to Example 2. Assume an objective is measured by three criteria, which their fuzzy priority vector is $WEx2$. The set of direct rating scales applies $WS2$. Assume there are three decision makers provide rating scores for each criterion. The details are shown in Table 6. Find the synthesis result.

By using Eqs. (14)–(16), IFSIE of each decision maker is:

$$\left\{\begin{matrix}\left(w_{E_1}^L, w_{E_1}^M, w_{E_1}^U\right)\\\left(w_{E_2}^L, w_{E_2}^M, w_{E_2}^U\right)\\\left(w_{E_3}^L, w_{E_3}^M, w_{E_3}^U\right)\end{matrix}\right\} = \left\{\begin{matrix}(0.669, 0.787, 0.882)\\(0.404, 0.499, 0.613)\\(0.669, 0.787, 0.882)\end{matrix}\right\}$$

By using Eq. (17), then GIFS is (0.581,0.691,0.793). Next section illustrates and discusses the FGAHP solution in details.

Table 6
Rating scores with respective to objective.

E	Criterion 1 (0.608,0.673,0.720)	Criterion 2 (0.208,0.253,0.314)	Criterion 3 (0.071,0.074,0.071)
E_1	(0.844,0.934,1)	(0.289,0.351,0.436)	(0.844,0.934,1)
E_2	(0.289,0.351,0.436)	(0.844,0.934,1)	(0.289,0.351,0.436)
E_3	(0.844,0.934,1)	(0.289,0.351,0.436)	(0.844,0.934,1)

Table 5
Synthesis of local fuzzy weights.

Expert	Criterion 1 $\left(w_1^L, w_1^M, w_1^U\right)$	$\cdots$	Criterion j $\left(w_j^L, w_j^M, w_j^U\right)$	$\cdots$	Criterion m $\left(w_m^L, w_m^M, w_m^U\right)$	IFSIE
E_1	$\left(\bar{w}_{11}^L, \bar{w}_{11}^M, \bar{w}_{11}^U\right)$	$\cdots$	$\left(\bar{w}_{1j}^L, \bar{w}_{1j}^M, \bar{w}_{1j}^U\right)$	$\cdots$	$\left(\bar{w}_{1m}^L, \bar{w}_{1m}^M, \bar{w}_{1m}^U\right)$	$\left(w_{E_1}^L, w_{E_1}^M, w_{E_1}^U\right)$
$\vdots$	$\vdots$	$\cdots$		$\cdots$	$\vdots$	$\vdots$
E_k	$\left(\bar{w}_{k1}^L, \bar{w}_{k1}^M, \bar{w}_{k1}^U\right)$	$\cdots$	$\left(\bar{w}_{kj}^L, \bar{w}_{kj}^M, \bar{w}_{kj}^U\right)$	$\cdots$	$\left(\bar{w}_{km}^L, \bar{w}_{km}^M, \bar{w}_{km}^U\right)$	$\left(w_{E_k}^L, w_{E_k}^M, w_{E_k}^U\right)$
$\vdots$	$\vdots$	$\cdots$		$\cdots$	$\vdots$	$\vdots$
E_n	$\left(\bar{w}_{n1}^L, \bar{w}_{n1}^M, \bar{w}_{n1}^U\right)$	$\cdots$	$\left(\bar{w}_{nj}^L, \bar{w}_{nj}^M, \bar{w}_{nj}^U\right)$	$\cdots$	$\left(\bar{w}_{nm}^L, \bar{w}_{nm}^M, \bar{w}_{nm}^U\right)$	$\left(w_{E_n}^L, w_{E_n}^M, w_{E_n}^U\right)$
GIFS						$\left(\tilde{w}^L, \tilde{w}^M, \tilde{w}^U\right)$

9. Application

9.1. Background

A company designing and manufacturing Smartphone is involved in the software and hardware development. In order to meet the requirement of the customer, the company usually applies ISO/IEC 9126:2001 to measure the software quality. Recently the company has been developing a new model of Smart phone. The department of Software Quality Assurance has received the testing job from the department of Software Engineering. The testing job is to test the quality in use of the new product. The result is critical to the release decision. The SQA department, which collaborates with other departments, has applied the proposed fuzzy group analytic hierarchy process approach to evaluate the quality of the in-house developed product.

9.2. The FGAHP solution

The solution of FGAHO comprises of five steps which are introduced in Section 4, and explained in detail in Sections 5–8.

Step 1: Problem definition analysis: The objective of the FGAHP is to determine the measurement of the quality of the in-house development. The rating scale schemas in Tables 2 and 3 are applied. Six criteria with 27 subcriteria of ISO/IEC 9126-1 2001) are applied as the attributes of software quality. The hierarchy structure is illustrated in Fig. 1. The management board for this project is organized in this stage and will be presented in the next step.

Step 2: Processes of group assessment: The results of group assessment, assessment validation, and prioritization are presented in the same Tables. The processes are described step by step.

The assessment includes three parts. In Part 1 the fuzzy scale values for the linguistic labels with respect to each criterion are defined. In Part 2 the fuzzy importance for the hierarchical criteria is defined. In Part 3 the scores for the hierarchical criteria are defined.

To negotiate the importance of the criteria, the members of the board of decision makers are drawn from different backgrounds. To simplify the complexity of the new approach presentation, three experts are presented in parts 1 and 2. Experts 1, 2, 3 (E_1^w, E_2^w, E_3^w) are the software assurance manager, software development manager, and the marketing manager respectively. In part 3, the direct rating activities are performed only by the three experts (E_1,E_2,E_3) from SQA department.

In part 1, the pairwise measurements for the absolute scales of six criteria in Table 2 are illustrated in Table 7. The policy defines that the subcriteria are applied the scale values according to their parents. In part 2, the pairwise measurements for the importance of the six criteria with 27 subcriteria are presented in Tables 8–14. In part 3, the direct measurements with absolute scales ($S_1,\ldots,S_6$) for the six criteria with 27 subcriteria by three experts are illustrated in Tables 15–20.

Step 3: Assessment validation process: Assessment validations mainly verify the validities of the pairwise measurements by

Table 7
Group fuzzy pairwise comparison for the scale values of six criteria.

	Poor	Weak	Satisfactory	Good	Excellent	FI
P	{EI}	{IWI, IMI, IWI}	{IMP, ISP, IVS}	{IVVS, IEXI, IVVS}	{IEXI, IEXI, IEXI}	(0.065, 0.068, 0.080)
W		{EI}	{ISP, ISP, ISP}	{IVVS, IVS, IVVS}	{IVVS, IVVS, IEXI}	(0.086, 0.097, 0.109)
S			{EI}	{IMP, IMP, ISI}	{ISI, IMP, ISI}	(0.274, 0.299, 0.335)
G				{EI}	{IWI, EI, IWI}	(0.690, 0.757, 0.857)
E					{EI}	(0.856, 0.955, 1)
S_1:$CR(SQA, R\&D, MKT)$ = (0.0660, 0.0703, 0.0893) < 0.1						
P	{EI}	{IWI, IWI, IWI}	{IMI, IMI, IMI}	{ISP, IVS, IVVS}	{IEXI, IEXI, IEXI}	(0.064, 0.071, 0.090)
W		{EI}	{ISI, ISI, ISI}	{IVS, IVS, IVS}	{IVVS, IVVS, IVVS}	(0.074, 0.087, 0.100)
S			{EI}	{ISI, ISI, ISI}	{ISP, ISP, ISP}	(0.181, 0.203, 0.228)
G				{EI}	{IMI, IMI, IMI}	(0.520, 0.560, 0.621)
E					{EI}	(0.919, 0.976, 1)
S_2:$CR(SQA, R\&D, MKT)$ = (0.0880, 0.0856, 0.0845) < 0.1						
P	{EI}	{IMI, IMP, IWI}	{IVS, IVVS, IVS}	{IVVS, IEXI, IVVS}	{IEXI, IEXI, IEXI}	(0.060, 0.061, 0.068)
W		{EI}	{ISP, ISP, ISP}	{IVVS, IVS, IVVS}	{IVVS, IVVS, IEXI}	(0.091, 0.100, 0.112)
S			{EI}	{IMP, IMP, ISI}	{ISI, IMP, ISI}	(0.288, 0.312, 0.349)
G				{EI}	{IWI, EI, IWI}	(0.673, 0.748, 0.857)
E					{EI}	(0.835, 0.944, 1)
S_3:$CR(SQA, R\&D, MKT)$ = (0.0970, 0.0940, 0.0893) < 0.1						
P	{EI}	{IWI, IMI, IWI}	{IMP, ISP, IVS}	{IVVS, IEXI, IEXI}	{IEXI, IEXI, IEXI}	(0.065, 0.068, 0.079)
W		{EI}	{ISP, ISP, ISP}	{IVVS, IVS, IVVS}	{IVVS, IVVS, IEXI}	(0.085, 0.097, 0.109)
S			{EI}	{IMP, IMP, ISI}	{ISI, IMP, ISI}	(0.273, 0.298, 0.335)
G				{EI}	{IWI, EI, IWI}	(0.696, 0.761, 0.854)
E					{EI}	(0.853, 0.952, 1)
S_4:$CR(SQA, R\&D, MKT)$ = (0.0661, 0.0703, 0.0861) < 0.1						
P	{EI}	{IWI, IMI, IWI}	{IVS, ISP, IVS}	{IVVS, IVS, IVS}	{IEXI, IVVS, IVVS}	(0.074, 0.080, 0.092)
W		{EI}	{IVS, ISP, ISP}	{IVVS, IVS, ISP}	{IVVS, IVVS, IVS}	(0.101, 0.115, 0.128)
S			{EI}	{IMP, IMI, IWI}	{ISI, IMP, IWI}	(0.366, 0.417, 0.503)
G				{EI}	{IWI, EI, EI}	(0.728, 0.807, 0.887)
E					{EI}	(0.875, 0.958, 1)
S_5:$CR(SQA, R\&D, MKT)$ = (0.0880, 0.0738, 0.0227) < 0.1						
P	{EI}	{IWI, IMI, IWI}	{IVS, ISP, IVS}	{IVVS, IVS, IVS}	{IEXI, IVVS, IVVS}	(0.073, 0.079, 0.092)
W		{EI}	{IVS, ISP, ISP}	{IVVS, IVS, ISP}	{IVVS, IVVS, IVS}	(0.099, 0.114, 0.128)
S			{EI}	{IMP, IWI, EI}	{ISI, IMP, IWI}	(0.400, 0.444, 0.519)
G				{EI}	{IWI, EI, EI}	(0.692, 0.743, 0.796)
E					{EI}	(0.858, 0.949, 1)
S_6:$CR(SQA, R\&D, MKT)$ = (0.0880, 0.0664, 0.0183) < 0.1						

Table 8
Group fuzzy pairwise comparison for the fuzzy importance of six criteria.

	C_1	C_2	C_3	C_4	C_5	C_6	Fuzzy importance
C_1	{EI}	{ISI, IWI, EI}	{IMI, IWI, WI}	{IVVS, WI, WI}	{MI, IVS, VS}	{SI, IEXI, SI}	(0.116, 0.141, 0.167)
C_2		{EI}	{MI, EI, WI}	{IMI, MP, WI}	{EXI, IMP, VS}	{EXI, ISI, SI}	(0.235, 0.274, 0.297)
C_3			{EI}	{IWI, MP, EI}	{EXI, IMP, SI}	{EXI, ISI, MI}	(0.168, 0.197, 0.232)
C_4				{EI}	{EXI, IEXI, SI}	{EXI, IEXI, MI}	(0.152, 0.173, 0.200)
C_5					{EI}	{MI, IWI, IWI}	(0.093, 0.100, 0.111)
						{EI}	(0.104, 0.114, 0.l25)
$CR(SQA, R\&D, MKT)$ = (0.0812, 0.0171, 0.0043) < 0.1							

Table 9
Group fuzzy pairwise comparison for the fuzzy importance of five subcriteria of functionality.

Criteria	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	Fuzzy importance
C_{11}	{EI}	{WI, IWI, WI}	{IMP, SP, ISP}	{IMI, WI, ISI}	{IVS, VS, IWI}	(0.118, 0.149, 0.184)
C_{12}		{EI}	{IVVS, EXI, IEXI}	{ISP, MP, IEXI}	{IEXI, EXI, IMP}	(0.113, 0.119, 0.128)
C_{13}			{EI}	{IWI, IMI, IWI}	{IWI, EI, MI}	(0.187, 0.206, 0.233)
C_{14}				{EI}	{IWI, MP, MI}	(0.264, 0.339, 0.411)
C_{15}					{EI}	(0.160, 0.186, 0.203)
$CR(SQA, R\&D, MKT)$ = (0.0281, 0.0074, 0.0185) < 0.1						

Table 10
Group fuzzy pairwise comparison for the fuzzy importance of four subcriteria of reliability.

Criteria	C_{21}	C_{22}	C_{23}	C_{24}	Fuzzy importance
C_{21}	{EI}	{SI, SI, VS}	{SP, VVS, MI}	{IMI, SI, IWI}	(0.386, 0.428, 0.484)
C_{22}		{EI}	{WI, WI, IWI}	{IEXI, EI, IEXI}	(0.080, 0.089, 0.101)
C_{23}			{EI}	{IEXI, IWI, ISP}	(0.0646, 0.789, 0.105)
C_{24}				{EI}	(0.365, 0.404, 0.415)
$CR(SQA, R\&D, MKT)$ = (0.0350, 0.0023, 0.0071) < 0.1					

Table 11
Group fuzzy pairwise comparison for the fuzzy importance of five subcriteria of usability.

Criteria	C_{31}	C_{32}	C_{33}	C_{34}	C_{35}	Fuzzy importance
C_{31}	{EI}	{WI, IWI, WI}	{MP, WI, WI}	{MI, WI, IWI}	{IVS, VS, WI}}	(0.195, 0.274, 0.341)
C_{32}		{EI}	{WI, MP, WI}	{IWI, SI, IMP}	{IEXI, EXI, WI}	(0.177, 0.225, 0.270)
C_{33}			{EI}	{IWI, WI, IMP}	{IEXI, MP, WI}	(0.097, 0.120, 0.156)
C_{34}				{EI}	{IEXI, MP, MP}	(0.172, 0.211, 0.246)
C_{35}					{EI}	(0.169, 0.170, 0.175)
$CR(SQA, R\&D, MKT)$ = (0.0504, 0.0222, 0.0307) < 0.1						

Table 12
Group fuzzy pairwise comparison for the fuzzy importance of three subcriteria of efficiency.

Criteria	C_{41}	C_{42}	C_{43}	Fuzzy importance
C_{41}	{EI}	{WI, IMI, WI}	{IWI, MP, MI}	(0.297, 0.405, 0.507)
C_{42}		{EI}	{IMI, EXI, WI}	(0.303, 0.380, 0.464)
C_{43}			{EI}	(0.190, 0.216, 0.239)
$CR(SQA, R\&D, MKT)$ = (0.0885, 0.0885, 0.0885) < 0.1				

consistence ratio. The results are shown in Tables 7, 8, 19, 10, 11, 12, 13, 14 respectively. The judgments are acceptable as CR is less than 0.1 for raters (SQA, R& D, MKT). CR is used for monitor the validity of the input. If CR is larger than 0.1, the raters are required to revise the work again scores until the CR is in the acceptable range.

Step 4: Prioritization process: The prioritization method is used for two purposes in this paper. One is to derive the fuzzy importance by the modified fuzzy LLSM (Eq. (10)), and another is to

Table 13
Group fuzzy pairwise comparison for the fuzzy importance of five subcriteria of maintainability.

Criteria	C_{51}	C_{52}	C_{53}	C_{54}	C_{55}	Fuzzy importance
C_{51}	{EI}	{WI, IWI, WI}	{IMI, IWI, IMP}	{IMI, IWI, ISI}	{IVS, VVS, IWI}	(0.094, 0.118, 0.152)
C_{52}		{EI}	{ISP, WI, IVVS}	{ISP, WI, IEXI}	{IEXI, EXI, IMP}	(0.089, 0.102, 0.113)
C_{53}			{EI}	{IWI, WI, IWI}	{IMI, EXI, WI}	(0.241, 0.311, 0.378)
C_{54}				{EI}	{IWI, EXI, IMI}	(0.231, 0.291, 0.354)
C_{55}					{EI}	(0.163, 0.178, 0.184)
$CR(SQA, R\&D, MKT)$ = (0.0172, 0.0397, 0.0641) < 0.1						

Table 14
Group fuzzy pairwise comparison for the fuzzy importance of five subcriteria of portability.

Criteria	C_{61}	C_{62}	C_{63}	C_{64}	C_{65}	Fuzzy importance
C_{61}	{EI}	{MP, SP, IWI}	{IMI, IWI, ISI}	{IMI, IWI, IVVS}	{IVS, VVS, IWI}	(0.097, 0.111, 0.l37)
C_{62}		{EI}	{IEXI, IEXI, IMI}	{IEXI, IEXI, IMP}	{IEXI, IWI, WI}	(0.052, 0.054, 0.059)
C_{63}			{EI}	{IWI, WI, WI}	{IWI, EXI, MI}	(0.280, 0.356, 0.419)
C_{64}				{EI}	{IWI, EXI, MP}	(0.288, 0.348, 0.408)
C_{65}					{EI}	(0.109, 0.132, 0.151)
$CR(SQA, R\&D, MKT) =$ (0.0357, 0.0383, 0.0442) < 0.1						

Table 15
Synthesis of fuzzy scores for functionality.

Expert	C_{11} (0.118, 0.149, 0.184)	C_{12} (0.113, 0.119, 0.128)	C_{13} (0.187, 0.206, 0.233)	C_{14} (0.264, 0.339, 0.411)	C_{15} (0.160, 0.186, 0.203)	Fuzzy scores
E_1	Good	Excellent	Good	Good	Excellent	(0.736, 0.817, 0.904)
E_2	Good	Excellent	Excellent	Excellent	Good	(0.792, 0.888, 0.960)
E_3	Good	Excellent	Good	Good	Good	(0.709, 0.781, 0.875)

Table 16
Synthesis of fuzzy scores for reliability.

Expert	C_{21} (0.386, 0.428, 0.484)	C_{22} (0.080, 0.089, 0.101)	C_{23} (0.0646, 0.789, 0.105)	C_{24} (0.365, 0.404, 0.415)	Fuzzy scores
E_1	Excellent	Good	Good	Excellent	(0.837, 0.907, 0.945)
E_2	Excellent	Excellent	Good	Good	(0.712, 0.776, 0.837)
E_3	Good	Good	Good	Excellent	(0.666, 0.729, 0.778)

Table 17
Synthesis of fuzzy scores for usability.

Expert	C_{31} (0.195, 0.274, 0.341)	C_{32} (0.177, 0.225, 0.270)	C_{33} (0.097, 0.120, 0.156)	C_{34} (0.172, 0.211, 0.246)	C_{35} (0.169, 0.170, 0.175)	Fuzzy scores
E_1	Excellent	Excellent	Excellent	Excellent	Good	(0.806, 0.910, 0.976)
E_2	Good	Excellent	Excellent	Excellent	Excellent	(0.780, 0.890, 0.972)
E_3	Good	Good	Good	Excellent	Excellent	(0.729, 0.823, 0.917)

Table 18
Synthesis of fuzzy scores for efficiency.

Expert	C_{41} (0.297, 0.405, 0.507)	C_{42} (0.303, 0.380, 0.464)	C_{43} (0.190, 0.216, 0.239)	Fuzzy scores
E_1	Excellent	Good	Good	(0.743, 0.838, 0.928)
E_2	Good	Good	Good	(0.696, 0.761, 0.854)
E_3	Excellent	Good	Excellent	(0.780, 0.880, 0.956)

Table 19
Synthesis of fuzzy scores for maintainability.

Expert	C_{51} (0.094, 0.118, 0.152)	C_{52} (0.089, 0.102, 0.113)	C_{53} (0.241, 0.311, 0.378)	C_{54} (0.231, 0.291, 0.354)	C_{55} (0.163, 0.178, 0.184)	Fuzzy scores
E_1	Good	Excellent	Good	Good	Excellent	(0.765, 0.850, 0.921)
E_2	Good	Excellent	Excellent	Excellent	Excellent	(0.853, 0.941, 0.989)
E_3	Good	Excellent	Good	Good	Excellent	(0.765, 0.850, 0.921)

Table 20
Synthesis of fuzzy scores for portability.

Expert	C_{61} (0.097, 0.111, 0.l37)	C_{62} (0.052, 0.054, 0.059)	C_{63} (0.280, 0.356, 0.419)	C_{64} (0.288, 0.348, 0.408)	C_{65} (0.109, 0.132, 0.151)	Fuzzy scores
E_1	Good	Excellent	Good	Excellent	Excellent	(0.767, 0.853, 0.922)
E_2	Excellent	Good	Good	Good	Good	(0.708, 0.766, 0.824)
E_3	Good	Excellent	Good	Good	Excellent	(0.719, 0.781, 0.839)

Table 21
Synthesis of overall score.

	C_1 (0.116, 0.141, 0.167)	C_2 (0.235, 0.274, 0.297)	C_3 (0.168, 0.197, 0.232)	C_4 (0.152, 0.173, 0.200)	C_5 (0.093, 0.100, 0.111)	C_6 (0.104, 0.114, 0.125)	Fuzzy scores
E_1	$\begin{pmatrix}0.736\\0.817\\0.904\end{pmatrix}^T$	$\begin{pmatrix}0.837\\0.907\\0.945\end{pmatrix}^T$	$\begin{pmatrix}0.806\\0.910\\0.976\end{pmatrix}^T$	$\begin{pmatrix}0.743\\0.838\\0.928\end{pmatrix}^T$	$\begin{pmatrix}0.765\\0.850\\0.921\end{pmatrix}^T$	$\begin{pmatrix}0.765\\0.853\\0.922\end{pmatrix}^T$	$\begin{pmatrix}0.780\\0.871\\0.940\end{pmatrix}^T$
E_2	$\begin{pmatrix}0.792\\0.888\\0.960\end{pmatrix}^T$	$\begin{pmatrix}0.712\\0.776\\0.837\end{pmatrix}^T$	$\begin{pmatrix}0.780\\0.890\\0.972\end{pmatrix}^T$	$\begin{pmatrix}0.696\\0.761\\0.854\end{pmatrix}^T$	$\begin{pmatrix}0.853\\0.941\\0.989\end{pmatrix}^T$	$\begin{pmatrix}0.708\\0.766\\0.824\end{pmatrix}^T$	$\begin{pmatrix}0.742\\0.827\\0.907\end{pmatrix}^T$
E_3	$\begin{pmatrix}0.709\\0.781\\0.875\end{pmatrix}^T$	$\begin{pmatrix}0.666\\0.729\\0.778\end{pmatrix}^T$	$\begin{pmatrix}0.729\\0.823\\0.917\end{pmatrix}^T$	$\begin{pmatrix}0.780\\0.880\\0.956\end{pmatrix}^T$	$\begin{pmatrix}0.765\\0.850\\0.921\end{pmatrix}^T$	$\begin{pmatrix}0.719\\0.781\\0.839\end{pmatrix}^T$	$\begin{pmatrix}0.717\\0.780\\0.880\end{pmatrix}^T$
Overall performance (by Eq. (17)): (0.746, 0.832, 0.909)							

derive the fuzzy scale values for direct rating scales by FPP-DRS (Eqs. (10)–(13)). The results of the fuzzy importance of the hierarchical criteria are presented in Tables 8–14, and the fuzzy values of the direct rating scales of the hierarchical criteria are shown in Table 7.

Step 5: Synthesis process: The synthesis is performed by Fuzzy Synthesis Programming for Absolute Measurement (FSPAM) (Eqs. (14)–(17)). Tables 15–20 combine the results for the six criteria with 27 subcriteria with respect to each SQA expert. Table 21 combines the results of Tables 7 and 14, 15, 16, 17, 18, 19 to get the overall performance.

The final fuzzy value *(0.746, 0.832, 0.909)* is derived in Table 21 as the level of the quality. Different company policies and scenarios have different opinions as to what quality is acceptable. In this result, the quality of the new development is more than acceptable. Thus the SQA can give permission for a product of this quality to be used.

10. Conclusion

This research proposes a fuzzy group analytical hierarchy process approach for software quality evaluation under conditions of uncertainty. Six criteria with 27 subcriteria from ISO/IEC 9126:2001 are chosen for building the model. To illustrate the usability and validity of this model, a case study has been conducted. The advantages of the FGAHP approach includes following three issues.

Firstly, as fuzzy importance vectors among the hierarchical criteria are perceived differently by different people, this paper applied modified fuzzy LLSM to derive the consensus fuzzy importance or fuzzy weight. The modified fuzzy Logarithmic Least Squares Method is used. This is better than other methods as was discussed in the related works.

Secondly, as fuzzy intervals of the linguistic labels, and distances among any adjacent linguistic labels are perceived differently for each criteria, and differently by each expert, this paper proposes Fuzzy Prioritization Programming for Direct Rating Scales (FPP-DRS), in which three more equations are applied to the modified fuzzy LLSM and rescale the original interval values, to address this issue.

Thirdly, each expert perceives the quality of the attributes of the software products differently. Thus Fuzzy Synthesis Programming for Absolute Measurement (FSPAM) including the Integrated Fuzzy Scores of Individual Expert (IFSIE) and the Group Integrated Fuzzy Score (GIFS) is proposed to address this issue and to come to an agreement.

The scope of this paper does not include the methods or frameworks of how to give the direct rating scores for each criterion as these methods are different from various software applications such as web applications, window applications, distributed applications, mobile applications, and database applications. The relevant issue was discussed by ISO/IEC 9126–2,3,4 (2003), Xenos and Christodoulakis (1997), Bevan (1999) and Losavio, Chirinos, Matteo, Levy, and Ramdane-Cherif (2004). Future research will address this issue.

Future research will also compare the fuzzy group AHP approach with the fuzzy group Cognitive Network Process (CNP) (Yuen, 2009, 2011a, 2011b) in the areas of Software Quality Assurance management.

Appendix A

The notations are summarized as follows:

i, j, k	indices
$\mu_A(x)$	the membership value of the fuzzy set A with respective to x such that $x \in \Re$
(l, m, u)	l and $u$are the fuzzy boundary and m is the modal value such that $l \leqslant m \leqslant u$
a_1, a_2	a fuzzy set with index 1 or 2
$WEx2$	the fuzzy priority vector in Examples 2 and 3
$WS2$	the set of The direct rating scales in Examples 2 and 3
CR	consistence ratio
CI	consistence index
RI	random index
$\lambda_{\max}$	a principal eigenvalue of a pairwise matrix
n	the cardinal number of the criteria
C_i	the ith criterion
C_{ij}	the jth sub-criterion in the ith criterion
a_{ijk}	a triangular fuzzy judgment $(l_{ijk}, m_{ijk}, u_{ijk})$ for comparing criterion i to criterion j from expert k
e_{ij}	the cardinal number of expert judgments for C_{ij}
J	objective function of the modified fuzzy LLSM to be minimized
$w_i' = (w_i^U, w_i^M, w_i^L)$	the normalized triangular fuzzy weights (fuzzy priority or fuzzy importance)
$\bar{w}_i = (\bar{w}_i^L, \bar{w}_i^M, \bar{w}_i^U)$	the fuzzy scale value for rating scale i
$\bar{w}_{kj} = \left(\bar{w}_{kj}^L, \bar{w}_{kj}^M, \bar{w}_{kj}^U\right)$	the representative fuzzy number of the metric linguistic label rated by expert E_k with respect to the criterion$j$.
$\left(w_{E_k}^L, w_{E_k}^M, w_{E_k}^U\right)$	Integrated Fuzzy Scores of Individual Expert (IFSIE) of expert k

(*continued on next page*)

Ω_w	the space of weights
$(\tilde{w}^L, \tilde{w}^M, \tilde{w}^U)$	Group Integrated Fuzzy Score (GIFS) derived from the mean of the set of IFSIE $\left\{\left(w^L_{E_k}, w^M_{E_k}, w^U_{E_k}\right)\right\}$

References

Bertoa, M. F., Troya, J. M., & Vallecillo, A. (2006). Measuring the usability of software components. *The Journal of Systems and Software, 79*, 427–439.

Bevan, N. (1999). Quality in use: Meeting user needs for quality. *The Journal of Systems and Software, 49*, 89–96.

Boender, C. G. E., de Graan, J. G., & Lootsma, F. A. (1989). Multi-criteria decision analysis with fuzzy pairwise comparisons. *Fuzzy Sets and Systems, 29*, 133–143.

Buyukozkan, G., Kahraman, C., & Ruan, D. (2004). A fuzzy multi-criteria decision approach for software development strategy selection. *International Journal of General Systems, 33*(2-3), 259–280.

Chang, D. Y. (1996). Applications of extend analysis method on fuzzy AHP. *European Journal of Operative Research, 95*, 649–655.

Chang, C. W., Wu, C. R., & Lin, H. L. (2008). Integrating fuzzy theory and hierarchy concepts to evaluate software quality. *Software Quality Journal, 16*(2), 263–276.

Chen, M. F., Tzeng, G. H., & Ding, C. G. (2008). Combining fuzzy AHP with MDS in identifying the preference similarity of alternatives. *Applied Soft Computing, 8*(1), 110–117.

IEEE. (1991). IEEE std 610.12-1990 – IEEE standard glossary of software engineering terminology, correct edition, Feb. 1991. In *IEEE software engineering standards collection*. New York: The Institute of Electrical and Electronics Engineers.

ISO/IEC 9126. (1991). *Software product evaluation – Quality characteristics and guidelines for use*. ISO.

ISO/IEC 9126-1. (2001). *Software engineering-product quality – Part 1: Quality model*. ISO.

ISO/IEC 9126-2. (2003). *Software engineering-product quality – Part 2: External metrics*. ISO.

ISO/IEC 9126-3. (2003). *Software engineering-product quality – Part 3: Internal metrics*, 1st ed. ISO.

ISO/IEC 9126-4. (2003). *Software engineering-product quality – Part 4: Quality in use metrics*. ISO.

Khan, R. A., Mustafa, K., & Ahson, S. I. (2006). *Software quality: Concepts and practices*. Oxford, U.K.: Alpha Science.

Khosravi, K., & Gueheneuc, Y. G. (2004). A quality model for design patterns. *Summer*.

Laarhoven, P. J. M. V., & Pedrycz, W. (1983). A fuzzy extension of Saaty's priority theory. *Fuzzy Sets and Systems, 11*, 229–241.

Losavio, F., Chirinos, L., Matteo, A., Levy, N., & Ramdane-Cherif, A. (2004). ISO quality standards for measuring architectures. *The Journal of Systems and Software, 72*, 209–223.

Mikehailov, L., & Tsvetinov, P. (2004). Evaluation of services using fuzzy analytic hierarchy process. *Applied Soft Computing, 5*, 23–33.

Ossadnik, W., & Lange, O. (1999). AHP-based evaluation of AHP-software. *European Journal of Operational Research, 118*(2), 578–588.

Saaty, T. L. (1980). *The analytic hierarchy process*. New York: McGraw-Hill.

Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research, 48*, 9–26.

Saaty, T. L. (2005). *Theory and applications of the analytic network process: decision making with benefits, opportunities, costs, and risks*. RWS Publications.

Wang, Y. M., Elhag, T. M. S., & Hua, Z. (2006). A modified fuzzy logarithmic least squares method of fuzzy analytic hierarchy process. *Fuzzy Sets and Systems, 157*, 3055–3071.

Wang, Y. M., Luo, Y., & Hua, Z. (2008). On the extent analysis method for fuzzy AHP and its application. *European Journal of Operational Research, 186*, 735–747.

Xenos, M., & Christodoulakis, D. (1997). Measuring perceived software quality. *Information and Software Technology, 39*, 417–424.

Yuen, K.K.F. (2009). Cognitive network process with fuzzy soft computing technique for collective decision aiding, The Hong Kong Polytechnic University, Phd thesis.

Yuen, K.K.F. (2011a). The pairwise opposite matrix and its cognitive prioritization operators: the ideal alternatives of the pairwise reciprocal matrix and analytic prioritization operators, Journal of the Operational Research Society, Accepted for publication.

Yuen, K.K.F. (2011b). The primitive cognitive network process: comparisons with the analytic hierarchy process, International Journal of Information Technology & Decision Making, Accepted for publication.

Zadeh, L. A. (1965). Fuzzy sets. *Information Control, 8*, 338–353.

Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning. *Information Science, part I*(8), 199–249. part II (8), 301–357; part III (9), 43–86.

Zadeh, L. A. (1996). Fuzzy logic = computing with words. *IEEE Transactions on Fuzzy Systems, 4*(2), 103–111.

Zhu, L., Aurum, A., Gorton, I., & Jeffery, R. (2005). Tradeoff and sensitivity analysis in software architecture evaluation using analytic hierarchy process. *Software Quality Journal, 13*(4), 357–375.