Inf Syst Front (2017) 19:301–319
DOI 10.1007/s10796-016-9674-6

Estimating participants for knowledge-intensive tasks in a network of crowdsourcing marketplaces

Yiwei Gong[1]

Published online: 18 July 2016

Abstract Crowdsourcing has become an increasingly attractive practice for companies to abstain on-demand workforce and higher level of flexibility in open contexts. While knowledge-intensive crowdsourcing is expected to be prosperous, most current crowdsourcing calls are still about general and low-priced tasks. An obstacle of conducing knowledge-intensive crowdsourcing is the lack of diversity of expertise and the small scale of crowd in isolated crowdsourcing marketplaces. In this paper, a network of crowdsourcing marketplaces is envisioned for efficient knowledge-intensive crowdsourcing and engagement of massive and diverse participants across different marketplaces. Based on an algorithm for estimating participants for knowledge-intensive crowdsourcing tasks, an experiment with 100 simulations indicates that conducting crowdsourcing tasks in a network of crowdsourcing marketplaces results in higher customer satisfaction than doing that in isolated marketplaces. This finding advocates the development of a network of crowdsourcing marketplaces to open up the potential of knowledge-intensive crowdsourcing.

Keywords Knowledge-intensive crowdsourcing · Flexibility · Search friction · Estimation algorithm

✉ Yiwei Gong
yiweigong@whu.edu.cn

1 School of Information Management, Wuhan University, Wuhan, Hubei 430072, People's Republic of China

1 Introduction

The concept of crowdsourcing is first coined in 2006 and simply means outsourcing certain tasks and problem formulations to an undefined (and generally large) network of people in the form of open calls (Howe, 2006). In a dynamic and globalized business environment where companies have to constantly advance and satisfy the needs of the customer and to remain competitive, it is not enough to respond the growing complexity of their environment with just their own capabilities or those of their partners. As a result, an increasing number of companies, ranging from small startups to those listed in Fortune 500, are trying to make use of crowdsourcing to access knowledge and skills that previously unavailable to them and to solve parts of business processes formerly executed in-house (Afuah and Tucci, 2012; Satzger et al., 2013). In this way, companies can have an on-demand workforce and higher knowledge absorptive capacity and business processes can be adapted, which results in higher level of flexibility (Gong and Janssen, 2012).

A typical form of crowdsourcing is publishing the request for proposals through an online marketplace with the details of the needed service and its expected duration and (a range of) cost. Then potential participants bid on the task by submitting their proposals. If more than one proposals were received for a task, only one candidate would be selected to carry out the task. At the end, company can decide to accept and pay for the work, or refuse it if it does not fulfil the expectation. There are many crowdsourcing marketplaces isolated to each other on the Internet, such as Amazon Mechanical Turk (AMT, www.mturk.com), Crowd Flower (www.crowdflower.com), Upwork (www.upwork.com), etc. AMT is a popular example of online crowdsourcing marketplaces. Most tasks on AMT typically involve basic computer and language skills, such as tagging photos according to their contents, rewriting sections of prose, transcribing audio, choosing representative screenshots from a short video clip, or responding to survey questions (Abascal-Mena et al., 2014). Those tasks are often priced between 0 and 0.1 US dollar (Difallah et al., 2015). In fact, today's most tasks in crowdsourcing are very simple tasks (Satzger et al., 2013) which are also called 'micro-tasks' in literature.

Crowdsourcing can be considered as an online and distributed problem-solving model (Brabham, 2008). It is suggested that engaging crowds can help companies develop solutions to a variety of business challenges. As the business challenges and tasks vary, so do the knowledge and skills that crowdsourcing participants have. Unlike simple and low-priced tasks that commonly require general skills, challenging and valuable tasks, such as IT services (Ford et al., 2015; Nevo and Kotlarsky, 2014; Lu et al., 2015), are often knowledge intensive and require crowdsourcing participants with special skills and knowledge. Knowledge-intensive crowdsourcing is regarded as one of the most promising areas of crowdsourcing in the future, given the critical role it can play in today's knowledge-based economy (Kittur et al., 2013). Despite its importance, the technology that has turbocharged knowledge-intensive crowdsourcing's potency and application is still relatively new (Lakhani, 2013), and limited work so far optimizes knowledge-intensive crowdsourcing (Roy et al., 2015; Yang et al., 2015). This impedes the implementation of knowledge intensive-crowdsourcing and the understanding of the depth and reach of knowledge crowds across the economy.

On the way to fully unlock the potential of knowledge-intensive crowdsourcing, a challenge is the identification of suitable participants considering their skills and knowledge (Satzger et al., 2013; Yang et al., 2015; Roy et al., 2015), especially across different isolated crowdsourcing platforms. Matching skills to tasks often relays on sophisticated online crowdsourcing platforms to manage distributed workers and support task providers (Lakhani, 2013). However, prior research on crowdsourcing takes the online crowdsourcing platforms and their functionalities as given (Majchrzak and Malhotra, 2013), not even considering the isolated platform issue. In this paper, a network of marketplaces for knowledge-intensive crowdsourcing is envisioned. An algorithm of AHP-TOPSIS based on Grey Relation Analysis (GRA) is employed to enable the estimation of crowdsourcing participants in both isolated and networked situations. AHP is an abbreviation of Analytical Hierarchical Process, while TOPSIS stands for Technique for Order Preference by Similarity to Ideal Solution. In the previous work of this study which is presented in (Gong, 2015), the accuracy and efficiency of this algorithm on identifying best-fit knowledge-intensive crowdsourcing participant has been validated by using real-world log data of 348 completed crowdsourcing tasks. This paper extends the previous work by involving an environment of networked crowdsourcing marketplaces. An experiment with 100 simulations is then used to create another dataset which allows for a comparison between the customer satisfaction rates of tasks deployed in a network of crowdsourcing marketplace and the customer satisfaction rates of tasks deployed in isolated marketplaces. The main goal of this paper is to prove that knowledge-intensive crowdsourcing performs better in an open and networked environment than on isolated platforms and to advocate the development of a network of crowdsourcing marketplaces to open up the potential of knowledge-intensive crowdsourcing across different online crowdsourcing marketplaces.

The remainder of this paper is structured as follows. Section 2 envisions a network of crowdsourcing marketplaces. In Section 3 related work is discussed. Section 4 describes the design of an algorithm of AHP-TOPSIS based on GRA, including its validation. Then, Section 5 presents the conducted simulation and discusses their results. Section 6 concludes the paper, discusses its limitations and points to future work.

2 Envisioning a network of crowdsourcing marketplaces

Crowdsourcing is one of the emerging Web 2.0 based phenomenon and a connective and collaborative social-technical environment that enables individuals to get involved in internet-mediated social participation, communication, and collaboration (Zhao and Zhu, 2014). With a rapid development recently, online crowdsourcing marketplaces have become more and more sophisticated, making it simpler to manage and support distributed general tasks. Some crowdsourcing marketplaces provide simple text searching, ranking or commenting applications for helping tasks providers or participants. Given this, online marketplaces are foreseen to expand continually, unlocking an incredible number of opportunities for careers and skilled work (Kittur et al., 2013). While the demand for using crowdsourcing marketplaces for knowledge-intensive tasks is climbing, there are some obstacles that hinder task providers from expected results. Matching between knowledge-intensive tasks and crowdsourcing participants is a long-standing challenge of this kind. Implementing a mechanism for a task-participant matching concerning the requirement of knowledge and skills is not an easy job. For example, 10 years after its launch, it is still difficult to conduct knowledge-intensive crowdsourcing on AMT because of the lack of worker profiles indicating skills or experience, inability to post worker or employer ratings and reviews, minimal infrastructure for effectively managing workers or collecting analytics, etc. (Vakharia and Lease, 2015). Many scholars and experts have tried to address this challenge by the design and development of appropriate crowdsourcing systems (eg Geiger and Schader, 2014; Roy et al., 2015; Yuen et al., 2015). Although those studies are respectably realistic and contribute to current practice and research on crowdsourcing marketplaces, they focus on standalone crowdsourcing platform.

However, a closed single crowdsourcing marketplace environment often fails in fulfilling the expectation of companies which would like to access a large crowd for their problem-solving. It has been reported that most crowdsourcing platforms for conducting knowledge-intensive tasks obtain relatively few submissions from potential solvers and the majority of these submissions do not meet the expectation of task providers (Mahr et al., 2015). A limitation of current crowdsourcing marketplaces is the lack of domain-specific expertise (Heimerl et al., 2012). Many crowdsourcing marketplaces do not provide much skill variety (Kittur et al., 2013). When a task requires specific (level of) knowledge and skills, it becomes relevantly difficult to access the required expertise by only one marketplace. It is reported that a company had to deploy similar tasks on different crowdsourcing platforms in order to reach different online communities (Guittard et al., 2015). Furthermore, some crowdsourcing marketplaces have a domestic focus, for example, local crowdsourcing marketplaces in China and Russia or other non-English speaking countries. And some others have a preference on different kinds of tasks, for example the majority of tasks on AMT is micro-task, while the majority of tasks on InnoCentive (www.innocentive.com) is innovation-oriented. From a task providers' perspective, the above evidences indicate that conducting knowledge-intensive crowdsourcing in a standalone marketplace results in limited customer satisfaction.

Although current crowdsourcing marketplaces are isolated with each other, this isolation is expected to be broken up along with the development of crowdsourcing technology and economy. The advancement of new information and communication technologies and the evolution of the internet have a profound impact on the structure of business (Garrigos-Simon et al., 2012), and companies are responding by implementing more and more flexible business models (Sommer, 2003). While firstly in the past century the integration of multinational companies, and later in recent decades the business model of outsourcing, the advances in new technologies and the globalized environment keep shaping the horizon of business and the boundary of companies (Zenger et al., 2011). For knowledge-intensive business, the creation of multinational companies and outsourcing can be explained by the knowledge-based view (KBV) of the firm. The KBV considers knowledge as the most strategically significant resource of a firm and the source of competitive advantage resides (Alavi and Leidner, 2001). In the past, a critical success factor of many companies consist of the creation of multinational partnerships network in the form of outsourcing or co-investment was that they could work together to achieve or retain competitive advantage by exchange and complementation of knowledge and skills. However, the consolidation of the networks of companies means the consolidation of their own rigidity. In the new socioeconomic arena where the crowd has a great diversity of knowledge and skills, crowdsourcing can overcome the advantages of the networks of companies by its higher level of flexibility (Garrigos-Simon and Narangajavana, 2015). From multinational companies to outsourcing and to crowdsourcing, it is a process of decentralization and the expansion of the boundary of companies. As innovative companies that have a high knowledge absorptive capacity will keep looking for better mechanisms for efficient information and knowledge aggregation and the boundary of companies will be dynamic, given high-powered incentives from the market and appropriate value-creating coordination (Zenger et al., 2011). We can therefore envision that the decentralization will continue in the development of crowdsourcing and the business model of crowdsourcing will evolve from relying on a single marketplace to deploying knowledge-intensive tasks in a network of crowdsourcing marketplaces.

Another support of this envisioning is the theory of Search Friction which is a term developed in the economics of the labor market (Pissarides, 2011). The theory of Search Friction is typically used to explain the frictional unemployment phenomenon where unemployment and vacant jobs exist at the same time because both workers and jobs are heterogeneous, and a mismatch can happen between the characteristics of supply and demand, because of different expectations on prices and wages, staggered contracts, location, time etc. (Diamond, 1982). The presence of search and matching cost has important and general implication for how markets perform their function of allocating supply and demand. The cost of search and matching is regarded as a hindrance to the process of efficient allocation (Mortensen, 2011). A satisfied job-worker matching is the outcome of an information gathering process in which workers look for the content of available jobs while employers engage in complementary recruiting activities to broadcast job information. The limitations of cognitive as well as the available resource and time put bounds on the information that the workers or employers can access and process at one time (Williamson, 2002). Crowdsourcing marketplace can play an important role in reduce search and matching cost in such a job-worker matching process by facilitating information gathering and processing for both sides. For this reason crowdsourcing has drawn the attention of economic and management science research on search including the studies of broadcast search (Jeppesen and Lakhani, 2010; Corvello and Iazzolino, 2013) and distant search (Afuah and Tucci, 2012). The study on using crowdsourcing for conducting broadcast search for problem-solving indicates that the greater diversity of the crowd expertise—the increasing distance between the problem and the solvers' field of expertise leads to a higher probability of developing of a winning solution (Jeppesen and Lakhani, 2010). At the same time, the study on using crowdsourcing for conducting distant search for problem-solving indicates that the larger scale of the crowd—the more pervasive the problem-solving know-how in a crowd, the higher the likelihood there will be someone in the crowd who will participate in solving the problem (Afuah and Tucci, 2012). In summary, the research on Search Friction implies that diversity and scale of crowd influence the probability of having a satisfying match between knowledge-intensive tasks and their participants. A network of crowdsourcing marketplaces can have a greater diversity of expertise and a larger scale of crowd. It is therefore expected to have the potential of a higher satisfaction in conducting knowledge-intensive crowdsourcing.

From the above discussion on the limitations of current standalone crowdsourcing marketplaces, the demand of flexibility and dynamic boundary of companies, and the implications of search literature, the following hypothesis can be made.

Hypothesis Conducting knowledge-intensive tasks in a network of crowdsourcing marketplaces results in higher customer satisfaction than doing so in isolated marketplaces.

This hypothesis is made from a task providers' perspective. Based on the finding that companies gain limited satisfaction in their knowledge-intensive crowdsourcing in a standalone crowdsourcing marketplace, this hypothesis uses customer satisfaction as a criteria to reflect the efficiency of knowledge-intensive crowdsourcing. Investigating this hypothesis is the first step to reveal the value of a network of crowdsourcing marketplaces.

3 Related works

In literature, there has been a considerable amount of work towards developing appropriate platforms and suggesting frameworks for efficient crowdsourcing, but only limited studies focus on mechanisms that facilitate matching between tasks and participants. There are mainly two directions in those studies on matching: either from a crowdsourcing participant's perspective and recommending tasks to participants, or from a task provider's perspective and estimating participants for tasks.

From a participant's perspective, task recommendation is the main topic under research. For example, Geiger and Schader (2014) proposed a systematic design of personalized task recommendation approaches based on literature study. Yuen et al. (2015) proposed a task recommendation framework based on a unified probabilistic matrix factorization, aiming to recommend tasks to workers in dynamic scenarios with a focus on solving the cold start problem.

From a task provider's perspective, a little bit more studies can be found in literature. For example, Roy et al. (2015) presented a framework for optimizing participant-to-task assignment in knowledge-intensive crowdsourcing. This framework relies on a set of pre-computed indexes and maintains them adaptively to enable effective task assignment. Human factors, such as expertise, minimum wage requirements, and acceptance ratio, are integrated into the assignment process. Other scholars discuss crowdsourcing from a business process management point of view. Mechanisms for estimating crowdsourcing participants for business process execution can be found in the research of BPEL4People in social networks (Schall et al., 2014), in which a ranking method based on Hyperlink-Induced Topic Search (HITS) algorithm is provided to estimate the expertise of crowdsourcing works in a social network. In this method, a certain skill, its expected level and the importance of a task are used as input, and the ranking result presents a list of all the suitable crowdsourcing works in a social network. The underlying concept of BPEL4People is that the flexibility of traditional SOA-based business process systems can be enhanced by enabling human-based services with very the same API used by software-based Web services. In this way, tasks would be able to match to suitable workers that are registered and active on the crowdsourcing social platform (Satzger et al., 2013). The interesting idea of social network crowdsourcing introduced by Schall et al. (2014) is a different vision of the future of crowdsourcing. The authors assumed that, on a global scale, "a clique in a social network is more likely to efficiently work on collaborative tasks than a group of random workers" (Schall et al., 2014, p. 13). In a comparison with the social network vision, the one presented in this paper is based on the commonest marketplace form of crowdsourcing where the size of tasks is well designed for assigning them to individuals. In a knowledge-intensive crowdsourcing situation, the consideration on selecting participants with an appropriate level of knowledge and skills has a higher priority than the consideration on the collaboration between crowdsourcing participants.

Selecting the best-fit crowdsourcing participant from a group of candidates is a typical Multiple Attribute Decision Making (MADM) problem (Hwang and Yoon, 1985) in which decision-making is for selecting the most appropriate one from many feasible solutions. Algorithms for decision-making have a rich literature in management and operations research, economics, and several areas of computer science including machine learning, theory of algorithms, and artificial intelligence. Analytical Hierarchical Process (AHP) (Saaty, 1980) is one of the most outstanding MADM methods, which first estimates the relationship among criteria weight and then the total value of each choice based on the obtained weight (Ngai and Chan, 2005). Essentially, the calculation process using AHP comprises of three stages: (Sekhar et al., 2015)

1. Hierarchical structure creation of the decision problem.
2. Pair wise judgments via a structured questionnaire that yield relative priorities (local weights) on the identified criteria.
3. Synthesis of the relative priorities (local weights) into the global priorities (global weights) that lead to the selection of the final decision.

The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is another outstanding MADM method, which is based on the concept that the best choice should have the shortest Euclidean distances from the positive ideal and the farthest from the negative ideal (Hwang and Yoon, 1985). The ranking of alternatives is based on the shortest distance from the positive ideal solution (PIS) and the farthest from the negative ideal solution (NIS). The calculation process of TOPSIS comprises of the following steps: (Jamali and Tooranloo, 2009)

1. Compute the normalized decision matrix.
2. Calculate the weighted normalized decision matrix.
3. Determine the PIS and NIS.
4. Calculate the separation measures, using the n-dimensional Euclidean distance.
5. Calculate the relative closeness to the ideal solution.
6. Rank the preference order.

AHP and TOPSIS can be used in combination (Patil and Kant, 2014; Lin et al., 2008) where AHP is used to calculate the weights of the parameters and these weights are later used in TOPSIS. A drawback of TOPSIS method is its linear variation of each alternatives, which cannot provide an accurate ranking between two alternatives that have the same distances to the ideal. This problem can be solved by Grey Relational Analysis (GRA) which is an effective method to solve decision making problems by generalizing estimates under limited samples and uncertain conditions (Deng, 1989). GRA is a kind of flexible measurement of curve similarity. By using GRA, the nonlinear relationship between the sequences of each alternative can be well reflected, which can compensate the inaccuracy problem of TOPSIS method. AHP-TOPSIS based on GRA has been proved to be useful in solving MADM problems (eg Chen, 2004; Oztaysi, 2014).

Despite the vast scope of the existing work on algorithms for decision-making, crowdsourcing brings an array of domain-specific challenges that require novel solutions. Formulating an algorithm with well-specified objectives, allowing researchers to propose novel solutions and techniques that can be easily compared, leading to a deeper understanding of the underlying issues (Slivkins and Vaughan, 2013). In this study, an AHP-TOPSIS based on GRA algorithm is employed for estimating crowdsourcing participants for given tasks on both isolated platforms and a network of marketplaces to allow for a comparison on their efficiency.

4 Estimation algorithm for knowledge-intensive crowdsourcing

In this section an algorithm of AHP-TOPSIS based on GRA is proposed for estimating crowdsourcing participants. This algorithm is specified for the decision-making in selecting participants for knowledge-intensive tasks. The final output of this algorithm is a ranking of participants for a certain tasks. The accuracy and efficiency of this algorithm is validated by testing it on a Chinese crowdsourcing marketplace's log data of 5 years.

4.1 Overview of the estimation algorithm

This algorithm has the following three phases.

4.1.1 Phases 1: identifying estimation parameters

The estimation parameters used by an algorithm should be quantitative, otherwise they cannot be calculated. In addition, the data of parameters should be public accessible, otherwise the information aggregation will have a high cost. In the use of this estimation algorithm, there are two underlying assumptions. The first one is that task providers and candidate participants do not know each other in their actual life. This means that a task provider makes his or her decision of selection only based on the related candidate participants' information that is publicly available online. The second one is that task providers will insist on looking for the best-fit participant rather than shifting to other strategy like choosing the first acceptable candidate. This means that all related candidate participants should be involved in the consideration during the decision- making. In a typical marketplace crowdsourcing model, participants' online information comes from either their online profiles or the proposals that they submitted to the task. Both these two sources of information are involved in the formulation of the estimation parameters in this study. Afterwards, the parameters that cannot be quantitated has to be ignored, and the parameters that reflect the same aspect should be merged. At the end, parameters are categorized into benefit parameters (the larger the value is, the better the solution is) and cost parameters (the smaller the value is, the better the solution is).

4.1.2 Phase 2: using AHP to calculate the weight of parameters

In AHP, the multi-attribute weight measurement is calculated via pair-wise comparison of the relative importance of two factors. Assuming that there are N number of decision parameters, denoted as $(P_1, P_2, \cdots, P_n)$, its judgment matrix would be $A = [a_n]$, in which a_n represents the relative importance of P_1 and P_2. Using the row vector average normalization proposed by Saaty (1980), the weight of P_i is calculated as:

$$W_i = \frac{\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}\right)^{\frac{1}{n}}} i, j–1, 2, \cdots, n$$

4.1.3 Phase 3: using GRA-based TOPSIS to estimate participants

In this phase the algorithm has the following ten steps.

1. Normalizing of Initial decision matrix $X = (x_{ij})_{m \times n}$, get the normalization matrix $Z = (z_{ij})_{m \times n}$. $(i = 1, 2, \cdots, m; j = 1, 2, \cdots n)$For benefit parameters:

$$Z_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \tag{1}$$

For cost parameters:

$$Z_{ij} = \frac{\max_i x_{ij} - x_{ij}}{\max_i x_{ij} - \min_i x_{ij}} \tag{2}$$

2. Calculating the weighted decision matrix $S = (s_{ij})_{m \times n}$, $(i = 1, 2, \cdots, m; j = 1, 2, \cdots n)$.

$$S_{ij} = w_{ij} z_{ij}$$

3. Calculating the positive ideal solution S^+ and negative ideal solution S^-

$$S^+ = \left(s_1^+, s_2^+, \cdots, s_n^+\right); \quad S^- = \left(s_1^-, s_2^-, \cdots, s_n^-\right)$$

Where $s_j^+ = \max_i s_{ij} = w_j; s_j^- = \min_i s_{ij} = 0, i = 1, 2, \cdots, m; j = 1, 2, \cdots, n.$

4. Calculating the Euclidean distance between each solution and positive/negative ideal solution d_i^+, d_i^-

$$d_i^+ = \sqrt{\sum_{j=1}^{n}\left(s_{ij} - s_j^+\right)^2}, \quad d_i^- = \sqrt{\sum_{j=1}^{n}\left(s_{ij} - s_j^-\right)^2}$$

Where $i = 1, 2, \cdots, m; j = 1, 2, \cdots, n$.

5. Calculating the grey relation coefficient matrix of each solution and positive/negative ideal solution L^+, L^-:

$$L^+ = \left(l_{ij}^+\right)_{m \times n}, \quad L^- = \left(l_{ij}^-\right)_{m \times n}$$

$$\text{Where } l_{ij}^+ = \frac{\min_i \min_j \left|s_j^+ - s_{ij}\right| + \theta \max_i \max_j \left|s_j^+ - s_{ij}\right|}{\left|s_j^+ - s_{ij}\right| + \theta \max_i \max_j \left|s_j^+ - s_{ij}\right|}$$

$$l_{ij}^- = \frac{\min_i \min_j \left|s_j^- - s_{ij}\right| + \theta \max_i \max_j \left|s_j^- - s_{ij}\right|}{\left|s_j^- - s_{ij}\right| + \theta \max_i \max_j \left|s_j^- - s_{ij}\right|}$$

Where $\theta \in (0, 1)$, is distinguishing coefficient, Here the value of θ is set to be 0.5.Then the formulas can be simplified as $L_{ij}^+ = \frac{\theta}{|z_{ij} - 1|} + \theta$; $L_{ij}^- = \frac{\theta}{|z_{ij} + 1|} + \theta$

6. Calculating the grey relation grade of each solution and positive/negative ideal solution l_i^+, l_i^-:

$$l_i^+ = \frac{1}{n}\sum_{j=1}^{n} l_{ij}^+; \quad l_i^- = \frac{1}{n}\sum_{j=1}^{n} l_{ij}^-$$

7. Applying nondimensionalization to d_i^+, d_i^-, l_i^+ and l_i^-

$$D_i^+ = \frac{d_i^+}{\max_i d_i^+}, \quad D_i^- = \frac{d_i^-}{\max_i d_i^-}$$

$$L_i^+ = \frac{l_i^+}{\max_i l_i^+}, \quad L_i^- = \frac{l_i^-}{\max_i l_i^-}$$

Where $i = 1, 2, \cdots, m$

8. Calculating the relative closeness degree:

$$P_i^+ = \frac{D_i^+}{D_i^+ + D_i^-}, \quad U_i^+ = \frac{L_i^+}{L_i^+ + L_i^-}$$

9. Combining P_i^+ and U_i^+: $Q_i^+ = v_1 P_i^+ + v_2 U_i^+$Where v_1 and v_2 reflect the degree of preference of decision makers, $v_1 + v_2 = 1$, and $v_1 = v_2 = 0.5$

Sorting solutions by the value of Q_i^+. The better Q_i^+ is, the better the solution is, and vice versa. $\max(Q_i^+)$ is the final decision.

4.2 Identifying and weighing parameters

Phase 1 in applying the proposed estimation algorithm is to identify the parameters that are used in the algorithm. This study extends the previous work presented in (Gong, 2015) by taken the divers settings of many crowdsourcing marketplaces into account. Crowdsourcing marketplaces on the Internet are countless. There is no a global ranking or index of online crowdsourcing marketplaces, and it is not possible to exhaustively investigate every crowdsourcing marketplace around the world. Therefore, it is very difficult to capture all pertinent aspects of participants and their proposals in a single coherent model (Slivkins and Vaughan, 2013). Instead of proposing a novel model, an analysis on a list of sustainable crowdsourcing marketplaces allows for the identification of a core subset of those aspects. This study chooses 8 sustainable crowdsourcing marketplaces as a sample, including 4 marketplaces with an English user interface and the rest with a Chinese user interface. Those marketplaces are AMT, InnoCentive, Crowd Flower, Upwork (merged from oDesk and Elance), TaskCN (www.taskcn.com), Epweike (www.epweike.com), Zhubajie (www.zhubajie.com), and 680 (www.680.com). They are considered to be sustainable as they have bee existing for a considerable long period of time and they are still running at the time of writing this paper.

An analysis on what information each of the 8 crowdsourcing marketplaces use in the description of candidates and their proposals to a task results in 12 distinct aspects. This analysis also indicates the common information that is available to task providers in their decision-making for selecting crowdsourcing participants. The following table provides an overview of these 12 aspects and the number of marketplaces which use them in the description of participants and proposals.

In Table 1, 7 common aspects are identified, as they are used by more than half of the 8 marketplaces. These aspects are more likely to be found in the public assessable information provided by a crowdsourcing marketplace. Focusing on these aspects would ease the information aggregation among a network of crowdsourcing marketplaces. At the same time, the rest 5 aspects are considered to be less representative. The 7 common aspects are therefore selected for formulating the parameters in the application of the proposed algorithm.

However, not all the 7 aspects can be directly used as parameters for the estimation algorithm. 4 of them need to be operationalized and preprocessed. First of all, task providers are able to give a customer satisfaction rate to each completed task. A customer satisfaction rate can only indicate the quality of a certain completed task. To reflect the general level of quality of tasks that are completed by a crowd worker, a 'mean value of customer satisfaction rate' should be used. Secondly, the total volume of trade cannot reflect the actual capability in completing a knowledge-intensive task in some situations. For example, a candidate who has completed 1000 general tasks of 1 US dollar will have a higher total volume of trade than someone who has finished 5 knowledge-intensive tasks of 100 US dollar. Therefore, the 'mean value of trade' would be a more precise indicator of a crowd worker's experience in completing knowledge-intensive tasks. It is easy to calculate the mean value of trade, given the total volume of trade and the number of completed tasks are available. Thirdly, the proposed price only reflects the expectation of a candidate. Research on biding in and knowledge-intensive crowdsourcing has indicated that proposed price itself did not contribute to crowdsourcing task success (Gefen et al., 2015). A decision-making on selecting the participant for a complex and valuable knowledge-intensive task has a concern on the quality assurance. For example, the task provider might not choose a candidate whose bidding price is very much lower than his expectation, as such a low proposed price causes his worry on the task delivery. It is meaningless to directly compare the proposed prices from different candidates without a consideration on the expected price from the task provider. Instead, the difference between the proposed price from a candidate and the expected price given by the task provider is more important in the decision-making. Therefore, the proposed price should be processed into 'discrepancy rate between the proposed price and task provider's expectation'. Finally, for the same reason, proposed duration should also be processed into the 'discrepancy rate between the proposed duration and task provider's expectation'.

Table 1 Aspects found in crowdsourcing participant and proposal description

Aspects	Number of Marketplaces	Aspects	Number of Marketplaces
Skill matching rate	6	Credit	2
Skill test score	6	Level of membership	1
Customer satisfaction rate	6	Views of personal page	3
Number of completed tasks	6	VIP or not	1
Total volume of trade	7	Proposed price	6
Integrity authentication	1	Proposed duration	6

An explanation of all the parameters and their units of measurement (if necessary) is given in the following list.

- P1—skill matching rate (in percentage): a matching rate between the required skills for a task and the skills hold by a candidate.
- P2—skill test score: the result (in centesimal system) of test on a set of skills required for completing a certain kind of tasks, eg website development.
- P3—mean value of customer satisfaction rate (in percentage): The average percentage of the customer satisfaction rate received by a crowd worker.
- P4—number of completed tasks: the number of tasks that were completed by a crowd worker.
- P5—mean value of trade (in US dollar): The average payment a crowd worker received from his or her completed tasks.
- P6—discrepancy rate between the proposed price and the task provider's expectation (in percentage): The

discrepancy between the proposed price from a candidate and the expected price from the task provider in a form of percentage.

- P7—discrepancy rate between the proposed duration and the task provider's expectation (in percentage): The discrepancy between the proposed duration from a candidate and the expected duration from the task provider in a form of percentage.

Within these 7 parameters, it is easy to recognize that P1 to P5 are benefit parameters and P6 and P7 are cost parameters.

To be able to weight those identified parameters with the AHP method in the next phase, a hieratical relationship between the parameters needs to be built. The 7 parameters are derived from 7 most common aspects mentioned in Table 1 and they reflect different attributes of a candidate. 'Skill matching rate' and 'skill test score' are the indicators to the level of required skills of a candidate, while 'mean value of trade', 'number of completed tasks' and 'mean value of customer satisfaction rate' actually reflect the crowdsourcing experience of a candidate. 'Discrepancy rate between the proposed price and the task provider's expectation' and 'discrepancy rate between the proposed duration and the task provider's expectation' are apparently about the proposal that is submitted by a candidate. A hieratical model of the parameters is able to be created (see Fig. 1).

With the above hieratical relationship of the 7 parameters, it is able to move to the second phase: weighting those identified parameters with the AHP method. To simplify the discussion, the three parameters on the second level in the above figure are assumed to be equally important in the decision-making for selecting a candidate. This assumption can be formulated as:

$$\mathrm{Weight(Skill)} : \mathrm{Weight(Experience)} : \mathrm{Weight(Proposal)} = 1 : 1 : 1$$

For the same reason, the leaf parameters that are under the same patent are also assumed to be equally important with each other. This assumption can be formulated as:

$$\begin{aligned}&\mathrm{Weight(P1)} : \mathrm{Weight(P2)} = 1 : 1;\ \mathrm{Weight(P3)}\\&: \mathrm{Weight(P4)} : \mathrm{Weight(P5)} = 1 : 1 : 1;\ \mathrm{Weight(P6)}\\&: \mathrm{Weight(P7)} = 1 : 1\end{aligned}$$

Supporting these two assumptions requires further investigation and scientific research on the process and impact factors of decision-making, and it is out of the scope of this study. It is worth to mention that this subjective setting on the importance of the second and third level of parameters will not influence the algorithm itself but the accuracy of the computation result. An evaluation of the algorithm in a later stage indicates that these assumptions are reasonably acceptable.

Based on the above setting of importance, we are able to build the initial AHP judgment matrix of each parameter (see Table 2).

By using the above matrix, the weight of P1 to P7 can be achieved:

$$\omega i = (0.159776593,\ 0.159776593,\ 0.120297876,\ 0.120297876,\ 0.120297876,\ 0.159776593,\ 0.159776593)$$

4.3 Validation of the algorithm

In order to evaluate the proposed algorithm on its accuracy and efficiency in estimating crowdsourcing participants, an experiment was carried out on data from a popular Chinese crowdsourcing marketplace, Epweike. This experiment used the proposed algorithm to estimate crowdsourcing participants of certain tasks, and then the estimation result was compared with the actual selection decision made by the task provider. This comparison allows for an analysis on the accuracy of the method and the impact of the number of candidates on the actual decision-making which reflects the efficiency of this method.

In this experiment, a dataset of 348 valid and completed tasks between 2010 and 2015 for IT services crowdsourcing is used. The content of those task includes software/mobile application development, website construction, database and system design, server maintenance, etc. In those tasks the number of openly visible proposals is more than 3. Tasks that have less than 4 proposals were ignored, because the GRA-based TOPSIS method does not return a meaningful ranking result when the number of candidates is less than 4.

To demonstrate how the third phase of this algorithm is implemented, the calculation of Task X is given as an example. This task has 23 candidate participants and 18 openly available proposals. The other 5 proposals are closed and only visible to the task provider, and therefore those 5 participants are not taken into account in the experiment. The best one is identified from the rest 18 candidates by using the GRA-based TOPSIS method. Through the calculation, the final value of Q_i^+ of Task X is presented in Table 3.

By descending the order of Q_i^+, a ranking of participants of Task X can be achieved.

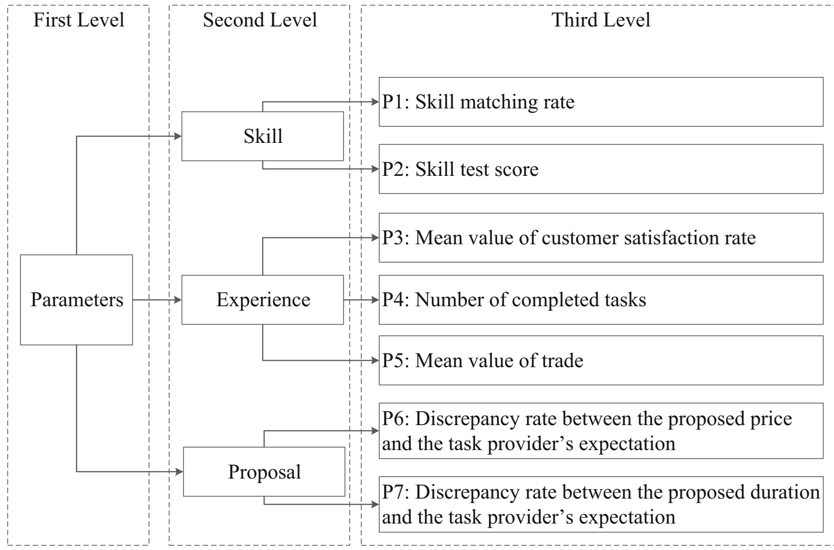

Fig. 1 A hieratical relationship model of the 7 parameters

According to the result presented in Table 4, the best candidate for Task X, by the proposed method of estimation, is Participant 10. The values of each of its parameters are:

$$P_{10} = (0.93,\ 100,\ 0.86,\ 10,\ 8000,\ 0\%,\ 0\%)$$

But the subjective choice made by the task provider is Participant 1, and the values of its parameters are:

$$P_1 = (0.82,\ 100,\ 0.6,\ 2,\ 4500,\ 1\%,\ 5\%)$$

In this example, the estimation result is not matchable with the actual selection result.

In this experiment, Phase 3 was repeated on all the 348 tasks. Then the result of the estimation was compared with the actual selection made by the task provider to find out whether they are matchable. A metric of matching rate is used to indicate the percentages of matchable tasks. In this experiment, the overall matching rate is 87 % (303 out of 348 tasks). The matching rates under different number of participants are presented in Table 5.

The matching rates presented in Table 5 indicate that the proposed method has a high accuracy when the number of participants is between 4 and 10. There are 248 tasks (71.26 % of the total 348 tasks) fall into this range and the matching rates are above 90 %. Specially, in the simplest situation where only 4 participants competing a task, the estimation result is 100 % matched with the manual decision-making. This means the proposed method can achieve an estimation result that is very similar with the manual decision-making, when the manual decision-making is simple and the information overload problem does not appear.

Technically speaking, the accuracy of this method will not be influenced by the number of participants. However, the matching rates generally declines along with the increase of participants (see Fig. 2). A possible explanation is that when the number of participants increase, it is more difficult to manually identify the best-fit participant. The task provider might have to rely on its own experience or even instinct in the decision-making for participant selection rather than an objective comparison between candidates. When the number of participants is more than 60, manually identifying the best-fit participant becomes very difficult and the actual selection result deviates very much from the estimation result, which results in matching rates below 33 %. The following curve demonstrates the decline of the matching rate. Under the assumption that the accuracy of manual decision-making is mainly and negatively

Table 2 Initial AHP judgment matrix of each parameter

	P1	P2	P3	P4	P5	P6	P7
P1	1	1	3/2	3/2	3/2	1	1
P2	1	1	3/2	3/2	3/2	1	1
P3	2/3	2/3	1	1	1	2/3	2/3
P4	2/3	2/3	1	1	1	2/3	2/3
P5	2/3	2/3	1	1	1	2/3	2/3
P6	3/2	3/2	3/2	3/2	3/2	1	1
P7	3/2	3/2	3/2	3/2	3/2	1	1

Table 3 Value of Q_i^+ for each participant of Task X

Participants	Q_i^+	Participants	Q_i^+
1	0.504601337	6	0.459536143
2	0.507368043	7	0.487222997
3	0.463986991	8	0.499194489
4	0.521388498	9	0.575854617
5	0.501450868	10	0.613623483

Table 4 A ranking of participants of Task X

Order	Participants	Order	Participants
1	10	6	5
2	9	7	8
3	4	8	7
4	2	9	3
5	1	10	6

influenced by information overload, the proposed method could solve this problem and improve the decision-making for crowdsourcing tasks with a large number of participants. To solidly prove this statement, investigation on the factors that impact manual decision-making for selecting crowdsourcing participants is desired in the future.

By using real-world data, the experiment on the proposed algorithm of AHP-TOPSIS based on GRA demonstrate its effectiveness and accuracy in the decision-making for selecting the best-fit crowdsourcing participant. This algorithm is useful especially in a situation with a large number of candidates. Although this algorithm is tested on in a standalone marketplace environment, it can also be used in a network of crowdsourcing marketplaces environment, if the information of candidates can be aggregated from different marketplaces. This algorithm provides a technical tool that allows for a further study on the proposed hypothesis in a network of marketplaces environment.

However, accessing data of the 7 parameters on different crowdsourcing marketplace platforms require a sufficient level of interoperability between those marketplaces. IEEE defines interoperability as “the ability of two or more systems or components to exchange information and to use the information that has been exchanged” (IEEE, 1990, p. 114). The interoperability in a network of crowdsourcing marketplace is essential to ensure the openness of data and operation across platforms. According to the maturity levels of interoperability that is defined by Janssen et al. (2014), aggregation of information of the 7 parameters requires a semantic interoperability—interpreting information of those parameters in the same way even if they named differently in different marketplace, while the distribution and manage of tasks (including commenting and evaluating completed tasks) on different marketplaces requires pragmatic interoperability—organizational and collaborative aspects of quality and trust.

5 Participants estimation in a network of crowdsourcing marketplaces

This study aims at investigating the hypothesis that “*conducting knowledge-intensive tasks in a network of crowdsourcing marketplaces results in higher customer satisfaction than doing so in isolated marketplaces*”. At the time when conducting this study, a network of interoperable crowdsourcing marketplaces was not yet available to perform tests with real data. Therefore, a synthetic dataset that represents a network of crowdsourcing marketplaces situation is generated for the study on the proposed hypothesis. Based on generated dataset, situations of conducting tasks in isolated crowdsourcing marketplaces and situation of conducting tasks in a network of crowdsourcing marketplaces can be simulated. In the experiment presented in this section, the algorithm of AHP-TOPSIS based on GRA is used as a tool to explore different situation. And then the comparison of the results from different situation can verify the hypothesis.

Table 5 Matching rates under different number of participants

Number of Participants	Number of Tasks	Number of Matched tasks	Matching Rate	Number of Participants	Number of Tasks	Number of Matched tasks	Matching Rate
4	65	65	100 %	31	11	7	64 %
5	55	55	100 %	32	11	8	73 %
6	24	23	96 %	37	8	5	63 %
7	18	16	89 %	38	3	2	67 %
8	61	58	95 %	39	5	3	60 %
9	15	14	93 %	40	4	2	50 %
10	10	9	90 %	42	2	1	50 %
11	5	4	80 %	44	2	1	50 %
12	7	5	71 %	47	2	1	50 %
15	4	3	75 %	52	3	2	67 %
17	5	4	80 %	57	3	1	33 %
22	5	4	80 %	60	1	0	0 %
23	4	3	75 %	62	3	1	33 %
29	8	5	63 %	63	4	1	25 %

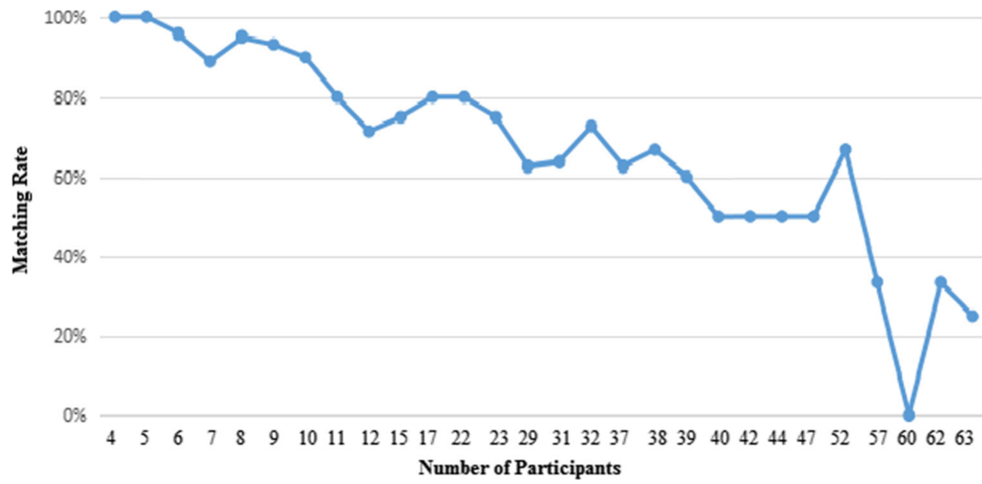

Fig. 2 Decline of matching rates along with the increase of participants

5.1 A dataset for simulating a network of crowdsourcing marketplaces

In order to simulate a network of crowdsourcing marketplaces that allows for an application of the proposed algorithm, a dataset is generated using Matlab according to the following arrangements.

- Assume that there is a network consisting 10 crowdsourcing marketplaces.
- In each simulation, the number of tasks within the network is randomly generated between 1 and 500.
- In each task, the number of candidates is randomly generated between 1 and 500.
- For each candidate, the value of P1, P2, P6 and P7 is randomly generated between 0 and 100 %.
- For each candidate, the initiate value of P3 is randomly generated between 0 and 100 %.
- For each candidate, the initiate value of P4 is randomly generated between 1 and 100.
- For each candidate, the initiate value of P5 is randomly generated between 0 and 1000.
- For each task, an instant customer satisfaction rate (SCR) is randomly generated between 0 and 100 %.

In total, this experiment contains 100 simulations. It is worth to mention that although the value of P1, P2, P6 and P7 is randomly generated in each simulation, starting from the second simulation, the value of P3, P4, and P5 of the participant selected in the last simulation will be changed by the superposition of the current value and the data generated by the last simulation. For example, if a participant Y was selected to complete the task and the task has its SCR, then in the next simulation, the value of P3 of Y is a superposition of the initiate value of P3 and the SCR that Y received in the last simulation. This allows for a simulation of the development and growing up of crowd workers. The more simulations are conducted, the better the dataset is.

5.2 Estimation of participant

5.2.1 Initial matrix of isolated marketplaces

In this experiment, an assumption is that all participants are working in only one marketplace. This means a participant will not submit more than one proposal. In reality, a candidate might try to submit multiple proposal (often by using different IDs) to increase the chance of win. This situation is not taken into account in this experiment. By the AHP-TOPSIS based on GRA algorithm, the local best-fit participant of each task on each isolated marketplace can be estimated and selected, and the SCR for the task can be generated. Through the calculation, a matrix of crowdsourcing participants on all the marketplaces can be created:

$$M_i T_j = \left(P_{ij}, {Q^+}_{ij}, SCR_{ij}\right)$$

Where

Mi	is the marketplace *i*, *i* = 1,......,10.
Tj	is the task *j*, *j* = 1,......,n.
Pij	is the selected participant in task *j*, marketplace *i*.
Q^+ij	is the value Q^+ of the selected participant in task *j*, marketplace *i*.
SCRij	is the customer satisfaction rate of the selected participant in task *j*, marketplace *i*.

```
%For matrix MTj, first find the maximal value of Q+ij in different
marketplaces, and choose it as the Q+j in the new matrix; then find the
matrix which has the maxQ+ij, and choose the value 'Pij' and 'SCRij' as the
value 'Pj' and 'SCRj' in the new matrix.%

For j=1:n
  MTj=(Pj, Q+j, SCRj)
  % MTj is the matrix of network marketplace.%
  For i=1:10
      Q+j=maxQ+ij
  % Get the maximal Q+ij among different marketplaces and assign its value
to the Q+j in the network marketplace.%
  Find  MiTj=(Pij,maxQ+ij,SCRij)  %Find the matrix which has the maxQ+ij
      If Q+j=maxQ+ij
      Pj=Pij
      SCRj=SCRij
      End
```

5.2.2 Matrix of a network of marketplaces

In order to simulate the situation of a network of marketplaces, this experiment transforms the data from the matrix of isolated marketplaces into the one of a networked marketplaces. The transformation progress is described as the following:

Based on the matrix transformation above, the matrix of a network of marketplaces is:

$$MT_j = \left(P_J, MaxQ_j^+, SCR_j\right)$$

For example, for Task1 in 5 different marketplaces ($i = 1,\ldots,5$, $j = 1$), the initial matrix of isolated marketplaces are:

$M_1T_1 = (11, 0.504601337, 73.33\ \%)$
$M_2T_1 = (8, 0.863986991, 72\ \%)$
$M_3T_1 = (23, 0.507368043, 65.25\ \%)$
$M_4T_1 = (5, 0.501450868, 64\ \%)$
$M_5T_1 = (32, 0.621388498, 61.11\ \%)$

The $\max Q^+_{ij}$ is 0.863986991 in M_2T_1. So the $Pij = 8$ and $SCRij = 72\ \%$ are chosen from M_2T_1 and combined the $\max Q^+_{ij}$ as the new matrix representing a network of those marketplaces:

$MT_1 = (8, 0.863986991, 72\ \%)$

5.3 Result analysis

There are in total 100 the simulation in this experiment. The number of tasks in a simulation are between 10 and 500. In each simulation there is a comparison between the situation of isolated marketplaces and the situation of networked marketplace. For space reason, only the details in Simulation 1 is given as an example.

In Simulation 1, there are 10 marketplaces from M_1 to M_{10} and 500 tasks from T_1 to T_{500}.

Thus, $M_iT_j = (P_{ij}, Q^+_{ij}, SCR_{ij})$, $i = 1, \ldots \ldots, 10$; $j = 1, \ldots \ldots, 500$

The best participant of each task on every marketplace is known by applying the estimation algorithm. Based on the initial matrix of isolated marketplaces and the matrix of a network of marketplaces, by selecting the $\max Q^+_{ij}$, the best participant of each task in a network of marketplaces can also be identified. In order to verify the hypothesis that deploying task in a network of marketplaces results in higher customer satisfaction than in isolate marketplaces, the SCR is employed to measure the efficiency. SCR reflects the customer satisfaction rate of task provider, and it indicates to what extend a crowdsourcing task has been successfully completed. A high SCR also means the task provider made a right decision in selecting the crowdsourcing participate.

```
For i=1,......,10
While(j=1,......,500)
{
  Ri=mean(SCRij)
  % calculate the mean value of SCR in each marketplace
  R1=mean(Ri)
  % calculate the mean value of every mean value of SCR in each isolated
  marketplace
  R2=mean(SCRj)
  % calculate the mean value of SCR in a network of marketplaces

  Vi=var(SCRij)
  % calculate the variance of SCR in each marketplace
  V1=mean(Vi)
  % calculate the mean value of variance of SCR in each isolated
  marketplaces
  V2=var(SCRj)
  % calculate the variance of SCR in a network of marketplaces
}
```

For the whole 500 tasks in this example, the higher the mean value of SCR is, the better the quality of selection is, while the lower the variance of SCR is, the higher the stability of selection is.

To allow for a focus on SCR, the matrix MT_j can be simplified by the following operation.

For this example experiment, the calculation results are listed as the following:

$R_i = (0.493651418\ 0.480351667\ 0.489786131\ 0.514644481\ 0.488365109\ 0.500576404\ 0.485454964\ 0.479718389\ 0.505378207\ 0.507681448)$
$R1 = 0.494560822$
$R2 = 0.500565276$
$V_i = (0.080410458\ 0.079033747\ 0.085378209\ 0.079256614\ 0.085445299\ 0.082321288\ 0.086333093\ 0.08560183\ 0.082304576\ 0.084790464)$
$V1 = 0.083087558$
$V2 = 0.07958054$

In a comparison between R1 and R2, and between V1 and V2:

R1 < R2, it means the quality of selection in a network of marketplaces is better than that in isolated marketplaces;
V1 > V2, it means the stability of selection in a network of marketplaces is better than that in isolated marketplaces.

The above description is the data analysis for one simulation. The same process of calculation is repeated in the rest simulation, and the matrix of results from all the 100 simulations is created (see Table 6).

To be more intuitional on the mean difference between R1 and R2, a statistical chart of it is given in Fig. 3. Out of the 100 simulations, positive difference between R1 and R2 (R2 > R1) is accounted for 89 %. This means that the participant who is selected from a network of marketplaces is more likely to get a higher customer satisfaction rate than the one who is selected from isolated marketplaces.

Another statistical chart can also be created for showing the difference between V1 and V2 (see Fig. 4). Negative difference between V1 and V2 (V2 < V1) is accounted for 93 % out of the 100 simulations. It means that the participant who is selected from a network of marketplaces is more likely to have a higher stability of customer satisfaction than the one who is selected from isolated marketplaces.

The above statistical analysis on the mean and variance difference indicates that estimating and selecting participants for knowledge-intensive tasks from a network of crowdsourcing marketplaces is more likely to have a higher customer satisfaction in a comparison with doing that in isolated crowdsourcing marketplaces by using the same algorithm and parameters. This finding support the hypothesis

Table 6 Simulation results

Number of Tasks	Isolated		Networked		Mean Difference (R2-R1)	VAR Difference (V2-V1)
	MEAN (R1)	VAR (V1)	MEAN (R2)	VAR (V2)		
409	0.4358	0.0866	0.7688	0.0624	0.333	−0.0242
454	0.5961	0.0605	0.8092	0.0587	0.2131	−0.0018
72	0.6775	0.0937	0.4255	0.0883	−0.252	−0.0054
458	0.4429	0.0727	0.6963	0.0338	0.2534	−0.0389
320	0.578	0.0625	0.713	0.0386	0.135	−0.0239
58	0.458	0.0834	0.9604	0.0207	0.5024	−0.0627
146	0.8909	0.0946	0.9194	0.0891	0.0285	−0.0055
278	0.7076	0.0681	0.8905	0.099	0.1829	0.0309
479	0.7546	0.0532	0.8335	0.01	0.0789	−0.0432
483	0.4899	0.0879	0.7316	0.0675	0.2417	−0.0204
87	0.485	0.0651	0.7958	0.059	0.3108	−0.0061
486	0.5458	0.0991	0.7112	0.0683	0.1654	−0.0308
479	0.6561	0.0589	0.9838	0.0575	0.3277	−0.0014
248	0.4998	0.0749	0.7894	0.0532	0.2896	−0.0217
402	0.8988	0.0821	0.8802	0.057	−0.0186	−0.0251
80	0.4766	0.0994	0.6723	0.0305	0.1957	−0.0689
217	0.4968	0.0548	0.6594	0.0297	0.1626	−0.0251
459	0.461	0.0911	0.8952	0.0195	0.4342	−0.0716
398	0.498	0.0617	0.4501	0.0199	−0.0479	−0.0418
480	0.4799	0.0861	0.6332	0.0157	0.1533	−0.0704
331	0.504	0.0765	0.5522	0.0464	0.0482	−0.0301
27	0.6346	0.0627	0.9502	0.0504	0.3156	−0.0123
426	0.6296	0.0429	0.8988	0.0322	0.2692	−0.0107
468	0.4701	0.0787	0.882	0.07	0.4119	−0.0087
343	0.4312	0.0665	0.4363	0.0175	0.0051	−0.049
381	0.6396	0.0795	0.9868	0.0663	0.3472	−0.0132
374	0.6353	0.094	0.7161	0.0695	0.0808	−0.0245
202	0.4222	0.0975	0.6501	0.0757	0.2279	−0.0218
331	0.5177	0.0902	0.7941	0.0273	0.2764	−0.0629
94	0.6157	0.0984	0.7768	0.0225	0.1611	−0.0759
356	0.7367	0.0792	0.5752	0.0727	−0.1615	−0.0065
26	0.5204	0.0623	0.659	0.0184	0.1386	−0.0439
146	0.4093	0.0935	0.7418	0.0573	0.3325	−0.0362
33	0.8123	0.0622	0.9904	0.0577	0.1781	−0.0045
58	0.4941	0.0875	0.5003	0.0115	0.0062	−0.076
413	0.4263	0.0536	0.4637	0.0209	0.0374	−0.0327
350	0.5428	0.0876	0.6234	0.0454	0.0806	−0.0422
165	0.5572	0.0704	0.5189	0.0536	−0.0383	−0.0168
476	0.6257	0.0767	0.6938	0.046	0.0681	−0.0307
27	0.6037	0.0568	0.6801	0.0288	0.0764	−0.028
225	0.4056	0.0413	0.971	0.0397	0.5654	−0.0016
197	0.5206	0.0435	0.9522	0.0667	0.4316	0.0232
385	0.4316	0.0627	0.8707	0.0129	0.4391	−0.0498
400	0.369	0.0336	0.8427	0.0153	0.4737	−0.0183
102	0.7538	0.064	0.5615	0.0426	−0.1923	−0.0214
250	0.8247	0.0779	0.6537	0.0145	−0.171	−0.0634
228	0.6676	0.0319	0.7287	0.0241	0.0611	−0.0078
10	0.726	0.0498	0.9656	0.0273	0.2396	−0.0225
358	0.5042	0.0719	0.6506	0.0211	0.1464	−0.0508
500	0.4945	0.0831	0.5006	0.0796	0.0061	−0.0035
145	0.676	0.0763	0.5809	0.0232	−0.0951	−0.0531
343	0.7054	0.0455	0.8207	0.027	0.1153	−0.0185
331	0.6887	0.0715	0.7998	0.0138	0.1111	−0.0577
90	0.5005	0.0734	0.7235	0.0672	0.223	−0.0062
68	0.6678	0.0498	0.8189	0.0354	0.1511	−0.0144
254	0.7556	0.0585	0.7999	0.0618	0.0443	0.0033
480	0.5069	0.0726	0.7914	0.0398	0.2845	−0.0328
177	0.4768	0.0549	0.6071	0.0482	0.1303	−0.0067
297	0.5888	0.0582	0.8934	0.0343	0.3046	−0.0239
120	0.6922	0.0501	0.9021	0.0277	0.2099	−0.0224
378	0.3745	0.084	0.6994	0.0212	0.3249	−0.0628
135	0.5254	0.0687	0.8721	0.0541	0.3467	−0.0146
258	0.7436	0.0899	0.8251	0.0868	0.0815	−0.0031

Table 6 (continued)

Number of Tasks	Isolated		Networked		Mean Difference (R2-R1)	VAR Difference (V2-V1)
	MEAN (R1)	VAR (V1)	MEAN (R2)	VAR (V2)		
353	0.8675	0.0452	0.8015	0.0387	−0.066	−0.0065
447	0.4142	0.0792	0.5143	0.0343	0.1001	−0.0449
480	0.5202	0.0457	0.6213	0.0288	0.1011	−0.0169
278	0.568	0.0828	0.6764	0.0608	0.1084	−0.022
78	0.7918	0.078	0.989	0.0676	0.1972	−0.0104
83	0.4388	0.044	0.4938	0.0175	0.055	−0.0265
136	0.8188	0.0294	0.9133	0.0385	0.0945	0.0091
422	0.4194	0.0811	0.7869	0.0253	0.3675	−0.0558
135	0.6258	0.0954	0.7749	0.0174	0.1491	−0.078
409	0.5146	0.0395	0.6712	0.0195	0.1566	−0.02
129	0.657	0.0704	0.8564	0.0228	0.1994	−0.0476
465	0.6892	0.0495	0.8148	0.025	0.1256	−0.0245
181	0.4724	0.085	0.9374	0.0659	0.465	−0.0191
106	0.7537	0.0792	0.9291	0.0616	0.1754	−0.0176
133	0.5357	0.0251	0.5672	0.0147	0.0315	−0.0104
312	0.6308	0.0876	0.8042	0.0638	0.1734	−0.0238
242	0.4786	0.0991	0.7498	0.0756	0.2712	−0.0235
182	0.3699	0.0563	0.5511	0.0764	0.1812	0.0201
417	0.8336	0.0896	0.5743	0.0157	−0.2593	−0.0739
297	0.675	0.0874	0.7703	0.0629	0.0953	−0.0245
279	0.5592	0.0941	0.6619	0.0239	0.1027	−0.0702
459	0.8946	0.0986	0.9381	0.028	0.0435	−0.0706
150	0.7464	0.0873	0.9896	0.0466	0.2432	−0.0407
381	0.7515	0.0807	0.8381	0.0774	0.0866	−0.0033
379	0.6063	0.0843	0.9086	0.0562	0.3023	−0.0281
196	0.7504	0.0811	0.8736	0.026	0.1232	−0.0551
288	0.4647	0.0387	0.7249	0.0259	0.2602	−0.0128
47	0.4689	0.0581	0.9438	0.0721	0.4749	0.014
36	0.506	0.0181	0.9278	0.0128	0.4218	−0.0053
270	0.9262	0.0201	0.8907	0.0145	−0.0355	−0.0056
392	0.3686	0.0523	0.5564	0.0371	0.1878	−0.0152
468	0.6684	0.0711	0.7566	0.0366	0.0882	−0.0345
74	0.4135	0.0546	0.4592	0.04	0.0457	−0.0146
289	0.6552	0.0271	0.9861	0.022	0.3309	−0.0051
240	0.5876	0.0546	0.8133	0.0483	0.2257	−0.0063
16	0.4969	0.0233	0.6753	0.0123	0.1784	−0.011
175	0.5073	0.0149	0.6562	0.0358	0.1489	0.0209

proposed in Section 2 that "deploying knowledge-intensive tasks in a network of crowdsourcing marketplaces results in higher customer satisfaction than doing so in isolated marketplaces".

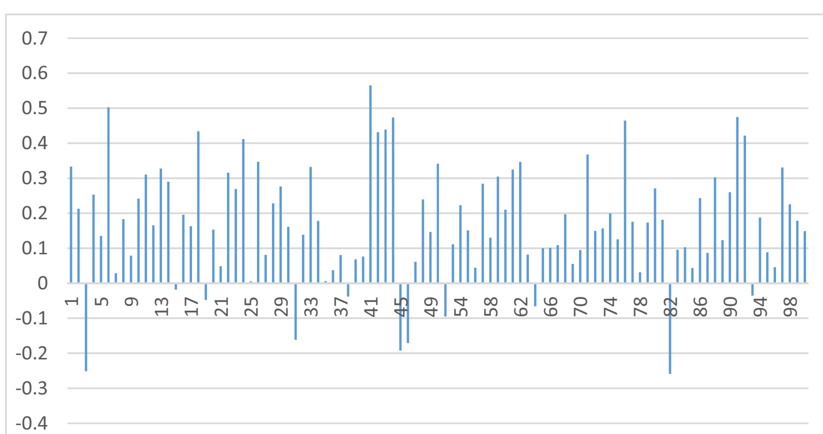

Fig. 3 Difference between R1 and R2

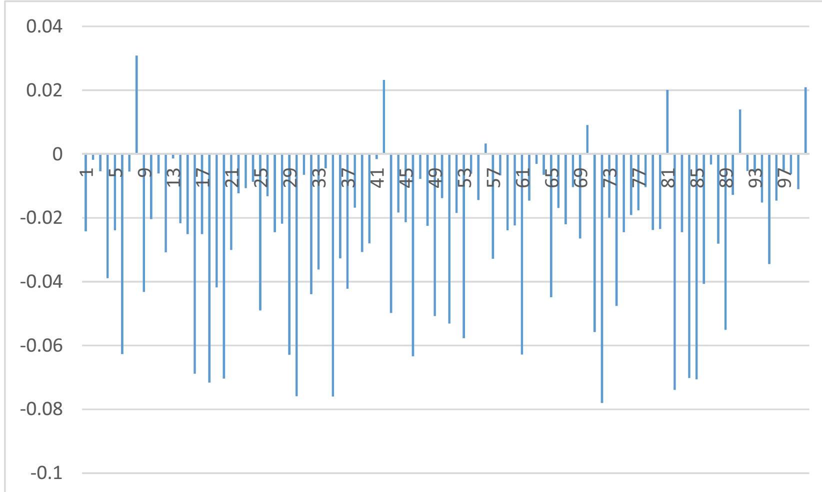

Fig. 4 Difference between V1 and V2

6 Conclusion and discussion

The focal point of this study is a network of crowdsourcing marketplaces that enables prosperous knowledge-intensive crowdsourcing in the future. Networked crowdsourcing marketplaces are expected to provide a greater diversity of expertise and a larger scale of crowd. Although such a network of crowdsourcing marketplaces is not currently existing at the time of doing this research, a hypothesis that it is better in supporting knowledge-intensive crowdsourcing than current isolated marketplaces environment has led to a discussion of the limitations of current knowledge-intensive crowdsourcing and a envision of the future crowdsourcing marketplaces. Based on prior research on a standalone crowdsourcing marketplace, an algorithm for estimating crowdsourcing task participants has been redesigned to adapt to networked marketplaces environment. A synthetic dataset that represents the situation of a network of crowdsourcing marketplaces is generated to allow for an application of the algorithm to explore the efficiency of conducting knowledge-intensive crowdsourcing tasks in this network. The analysis on the experiment results supports the hypothesis and indicates that conducing knowledge-intensive tasks in a network of marketplaces would have a higher customer satisfaction. This study is novel, and to the knowledge of the author, it is the first one towards to the situation of a network of crowdsourcing marketplaces. The finding of this study advocates the development of a network of crowdsourcing marketplaces to fully open up the potential of knowledge-intensive crowdsourcing.

6.1 Practical and theoretical implications

Conducting knowledge-intensive crowdsourcing often relays on sophisticated online crowdsourcing platforms, no matter in a standalone marketplace or in a network of marketplaces. Current online platforms provide general and basic functions and services for task providers to manage request for proposals, selection of task participants and the communication with selected participants in later stages. In a network of crowdsourcing marketplaces, task providers still need the similar functions and services from their crowdsourcing platforms. This study indicates that manual estimation of candidates for knowledge-intensive tasks will encounter the problem of information overload. In a network of crowdsourcing marketplaces environment where the number of candidate participants is expected to be larger, pure manual estimation would hardly work. Therefore, software agents that automatically rank the participants of a task would be indispensable for task providers in selecting suitable participants. In this study the same algorithm is used in both isolated and networked environment. In addition to the verification of hypothesis, the approach of this study also suggests that well designed estimation algorithms that are useful in single crowdsourcing marketplace could be reused in a networked marketplaces environment. From this point of view, sophisticated online crowdsourcing platforms and their intelligent decision support tools and services are the enabler of creating a network of crowdsourcing marketplaces. Crowdsourcing platform is an important factor influencing companies' decision to adopt crowdsourcing (Thuan et al., 2016). From practical view, this study provides insight for online crowdsourcing marketplace designers and developers to create more intelligent tools and services for task providers.

This research is also valuable to companies interested in expanding their business boundaries by seeking the help of outside experts. The finding indicates that employing multiple crowdsourcing marketplaces is more likely to have task success and consequently higher satisfaction. However, broadcasting request of proposal does not guarantee the success of the task. The statistical analysis of the experiment results also indicates the chance of dissatisfaction to appear. Extant research suggests that many crowdsourcing participants do not know the identity of task providers and underinvest the time needed to complete a successful task (Mahr et al., 2015). This calls for further investigation on the mechanisms for successful task deployment from both task provider and participant perspectives.

From theoretical view, online crowdsourcing marketplace platforms, and more specifically, intelligent software agents for estimating crowdsourcing participants play a role of mediator. In economy theories of Search Friction, the presence of search and matching cost has important and general implication for how online marketplaces perform their function of allocating supply and demand (Mortensen, 2011). This search and matching cost in online crowdsourcing is considered to be caused by information accessing and processing that is required in decision-making (Afuah and Tucci, 2012), such as selecting suitable crowdsourcing participants. Another contribution of this study is its attempts to reduce the information accessing cost by the creation of networked marketplaces and to limit information processing cost by using estimation algorithm. In essential, both of these two attempts require a sufficient level of interoperability between marketplaces.

6.2 Limitations

Coordinating the knowledge-intensive crowdsourcing work in a network of crowdsourcing marketplaces is undoubtedly very complicated. To allow for the simulations of such an environment, this study has greatly simplified the situation. It is inevitable to have limitations caused by this simplification. First of all, this study only discusses crowdsourcing from a task provider's perspective. A network of crowdsourcing marketplaces can also be studied from a crowd worker's perspective and have different insights to this environment. Second, only a one-to-one manner of task-participant matching is taken into account in this study. The search and matching process between tasks and the crowd can have many other forms. For example, matching the work of a whole knowledge-intensive business process to a group of crowd workers will be far more challenging. Third, in the application of the AHP method, this study equally treats the parameters on the same level in the hieratical relationship model. Whether it is optimal to equally assign the weights of parameters turns out to be a need in further investigation into the decision-making process of selecting task participants. Last but not least, this study relies on an underlying assumption that a network of crowdsourcing marketplaces has a sufficient level of interoperability between those marketplaces. There is a great need in addressing the interoperability issues between marketplaces to allow a network of marketplaces to be created. Future research is desired to have more insights into the issues mentioned above.

6.3 Future research

This study discusses a network of crowdsourcing marketplaces from a task provider's perspective. The main concerns of a task provider are how to disseminate task information to as much potential crowdsourcing participants as possible, and how to evaluate and select suitable participants or even the best-fit one from the candidates who are willing to carry out the task by submitting a proposal. To be addressed in future research, there are many challenges in coordinating the deployment of knowledge-intensive tasks in a network of crowdsourcing marketplaces, including:

- Mechanisms that allow for distributing task information among different marketplaces and manage the distribution conveniently;
- Identification of a crowd worker in a network of crowdsourcing marketplaces;
- Collection of the submission and also the profile (such as skills and experience) of candidates from different marketplaces;
- Mechanisms that enable the breakdown of a knowledge-intensive business process into interrelated crowdsourcing tasks and coordinate them in a network of crowdsourcing marketplaces, and so on.

Alternatively, future research could be conducted from a crowdsourcing participant's perspective. The main concerns of a knowledge worker in a network of crowdsourcing marketplaces are how to find out suitable tasks and how to manage their submissions in different marketplaces. The challenges to be addressed in future research can include the following aspects:

- Mechanisms that allow for searching tasks in a network of crowdsourcing marketplaces;
- Coordination of different submissions distributed in a network of crowdsourcing marketplaces; especially, those tasks exclusive with each other due to the limited availability of the crowdsourcing participant;
- Mechanisms that enable/stimulate knowledge workers to participate cross-domain tasks to enable higher level of innovation, and so on.

Overcoming the above challenges need advanced mediation and analytical tools, while creating openness and interoperability of each marketplace to share and distribute data is another key issue. We are in a very far distance from achieving an open and frictionless network of crowdsourcing marketplaces. The study presented in this paper makes the first attempt to point out research directions of a network of crowdsourcing marketplaces and share the ideas with the Information Systems research community.

Acknowledgments This work is supported by the National Natural Science Foundation of China (Grand No. 71501145).

References

~~Abascal-Mena, R., López-Ornelas, E., Zepeda-Hernández, J. S., Gómez-Torrero, B. E., León-Martagón, G., & Morales-Franco, H. (2014). *Worker community: using crowdsourcing to link informal workers with potential clients.* Paper presented at the the 6th International Conference on Social Computing and Social Media, Heraklion, Crete, Greece.~~

~~Afuah, A., & Tucci, C. L. (2012). Crowdsourcing as a solution to distant search. *Academy of Management Review, 37*(3), 355–375.~~

Alavi2001 Alavi, M., & Leidner, D. E. (2001). Review: knowledge management and knowledge management systems: conceptual foundations and research issues. *MIS Quarterly, 25*(1), 107–136.

~~Brabham, D. C. (2008). Crowdsourcing as a model for problem solving. *Convergence: The International Journal of Research into New Media Technologies, 14*(1), 75–90.~~

~~Chen, M. F. (2004). Combining grey relation and TOPSIS concepts for selecting an expatriate host country. *Mathematical and Computer Modelling, 40*(13), 1473–1490.~~

~~Cervello, V., & Iazzolino, G. (2013). Factors affecting the practices of external problem solvers in broadcast search. *Journal of Technology Management & Innovation, 8*(2), 166–177.~~

~~Deng, J. (1989). Introduction to grey system. *Journal of Grey System, 1*(1), 1–24.~~

~~Diamond, P. A. (1982). Aggregate demand management in search equilibrium. *Journal of Political Economy, 90*(5), 881–894.~~

~~Difallah, D. E., Catasta, M., Demartini, G., Ipeirotis, P. G., & Cudré-Mauroux, P. (2015). *The dynamics of micro-task crowdsourcing: the case of Amazon MTurk.* Paper presented at the the 24th International Conference on World Wide Web, Florence, Italy~~

~~Ford, R. C., Richard, B., & Ciuchta, M. P. (2015). Crowdsourcing: a new way of employing non-employees? *Business Horizons, 58*(4), 377–388.~~

~~Garrigos-Simon, F. J., Alcamí, R. L., & Ribera, T. B. (2012). Social networks and Web 3.0: their impact on the management and marketing of organizations. *Management Decision, 50*(10), 1880–1890.~~

~~Garrigos-Simon, F. J., & Narangajavana, Y. (2015). From crowdsourcing to the use of masscapital: the common perspective of the success of Apple, Facebook, Google, Lego, TripAdvisor, and Zara. In F. J. Garrigos-Simon, I. Gil-Pechuán, & S. Estelles-Miguel (Eds.), *Advances in crowdsourcing* (pp. 1–13). Switzerland: Springer.~~

~~Gefen, D., Gefen, G., & Carmel, E. (2015). How project description length and expected duration affect bidding and project success in crowdsourcing software development. *Journal of Systems and Software.*~~

Geiger2014 Geiger, D., & Schader, M. (2014). Personalized task recommendation in crowdsourcing information systems — current state of the art. *Decision Support Systems, 65*, 3–16.

~~Gong, Y. (2015). Enabling flexible IT services by crowdsourcing: a method for estimating crowdsourcing participants. In M. Janssen (Ed.), *Open and big data management and innovation* (Lecture Notes in Computer Science, Vol. 9373): Springer.~~

~~Gong, Y., & Janssen, M. (2012). From policy implementation to business process management: principles for creating flexibility and agility. *Government Information Quarterly, 29*(1), 61–71.~~

~~Guittard, C., Schenk, E., & Burger-Helmchen, T. (2015). Crowdsourcing and the evolution of a business ecosystem. In F. J. Garrigos-Simon, I. Gil-Pechuán, & S. Estelles-Miguel (Eds.), *Advances in crowdsourcing* (pp. 49–60). Switzerland: Springer.~~

~~Heimerl, K., Gawalt, B., Chen, K., Parikh, T. S., & Hartmann, B. (2012). *Communitysourcing: engaging local crowds to perform expert work via physical kiosks.* Paper presented at the The ACM SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, USA.~~

~~Howe, J. (2006). The rise of crowdsourcing. *Wired Magazine.*~~

~~Hwang, C. L., & Yoon, K. (1985). *Multiple attribute decision making: methods and applications a state-of-the-art survey* (Lecture Notes in Economics and Mathematical Systems): Springer-Verlag.~~

~~IEEE (1990). Interoperability. *IEEE standard computer dictionary: a compilation of IEEE standard computer glossaries.* New York, USA: Institute of Electrical and Electronics Engineers.~~

~~Jamali, R., & Tooranloo, H. S. (2009). Prioritizing academic library service quality indicators using fuzzy approach: case study: libraries of Ferdowsi university. *Library Management, 30*(4/5), 319–333.~~

~~Janssen, M., Estevez, E., & Janowski, T. (2014). Interoperability in big, open, and linked data–organizational maturity, capabilities, and data portfolios. *Computer, 47*(10), 44–49.~~

~~Jeppesen, L. B., & Lakhani, K. R. (2010). Marginality and problem-solving effectiveness in broadcast search. *Organization Science, 21*(5), 1016–1033.~~

~~Kittur, A., Nickerson, J. V., Bernstein, M. S., Gerber, E. M., Shaw, A., Zimmerman, J., et al. (2013). *The Future of Crowd Work.* Paper presented at the The 16th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW 2013), San Antonio, Texas, USA.~~

~~Lakhani, K. R. (2013). Using the crowd as an innovation partner. *Harvard Business Review, 91*(4), 60–69.~~

~~Lin, M. C., Wang, C. C., Chen, M. S., & Chang, C. A. (2008). Using AHP and TOPSIS approaches in customer-driven product design process. *Computers in Industry, 59*(1), 17–31.~~

Lu, B., Hirschheim, R., & Schwarz, A. (2015). Examining the antecedent factors of online microsourcing. *Information Systems Frontiers, 17*(3), 601–617. Lu2015a

~~Mahr, D., Rindfleisch, A., & Slotegraaf, R. J. (2015). Enhancing crowdsourcing success: the role of creative and deliberate problem-solving styles. *Customer Needs and Solutions, 2*(3), 209–221.~~

Majchrzak, A., & Malhotra, A. (2013). Towards an information systems perspective and research agenda on crowdsourcing for innovation. *Journal of Strategic Information Systems, 22*(4), 257–268. Majchrzak2013a

~~Mortensen, D. T. (2011). Markets with search friction and the DMP model. *The American Economic Review, 101*(4), 1073–1091.~~

Nevo, D., & Kotlarsky, J. (2014). Primary vendor capabilities in a mediated outsourcing model: can IT service providers leverage crowdsourcing? *Decision Support Systems, 65*, 17–27. Nevo2014

~~Ngai, E. W. T., & Chan, E. W. C. (2005). Evaluation of knowledge management tools using AHP. *Expert Systems with Applications, 29*(4), 889–899.~~

~~Oztaysi, B. (2014). A decision model for information technology selection using AHP integrated TOPSIS-Grey: The case of content management systems. *Knowledge-Based Systems, 70*, 44–54.~~

~~Patil, S. K., & Kant, R. (2014). A fuzzy AHP-TOPSIS framework for ranking the solutions of Knowledge Management adoption in Supply Chain to overcome its barriers. *Expert Systems with Applications, 41*(2), 679–693.~~

~~Pissarides, C. A. (2011). Equilibrium in the labor market with search frictions. *The American Economic Review, 101*(4), 1092–1105.~~

~~Roy, S. B., Lykourentzou, I., Thirumuruganathan, S., Amer-Yahia, S., & Das, G. (2015). Task assignment optimization in knowledge-intensive crowdsourcing. *The VLDB Journal, 24*(4), 467–491.~~

~~Saaty, T. L. (1980). *The analytic hierarchy process: planning, priority setting, resource allocation* (Decision Making Series): Mcgraw-Hill.~~

~~Satzger, B., Psaier, H., Schall, D., & Dustdar, S. (2013). Auction-based crowdsourcing supporting skill management. *Information Systems, 38*(4), 547–560.~~

~~Schall, D., Satzger, B., & Psaier, H. (2014). Crowdsourcing tasks to social networks in BPEL4People. *World Wide Web, 17*(1), 1–32. doi:10.1007/s11280-012-0180-6.~~

~~Sekhar, C., Patwardhan, M., & Vyas, V. (2015). A delphi-AHP-TOPSIS based framework for the prioritization of intellectual capital indicators: a SMEs perspective. *Procedia - Social and Behavioral Sciences, 189*, 275–284.~~

~~Slivkins, A., & Vaughan, J. W. (2013). Online decision making in crowdsourcing markets: theoretical challenges. *ACM SIGecom Exchanges, 12*(2), 4–23.~~

~~Sommer, R. A. (2003). Business process flexibility: a driver for outsourcing. *Industrial Management & Data Systems, 103*(3), 177–183.~~

Thuan2016 Thuan, N. H., Antunes, P., & Johnstone, D. (2016). Factors influencing the decision to crowdsource: a systematic literature review. *Information Systems Frontiers, 18*(1), 47–68.

~~Vakharia, D., & Lease, M. (2015). *Beyond AMT: an analysis of crowd work platforms*. Paper presented at the iConference 2015, Newport Beach, CA, USA.~~

~~Williamson, O. E. (2002). The theory of the firm as governance structure: from choice to contract. *The Journal of Economic Perspectives, 16*(3), 171–195.~~

~~Yang, J., Bozzon, A., & Houben, G. J. (2015). *Knowledge crowdsourcing acceleration*. Paper presented at the the 15th International Conference on Web Engineering (ICWE 2015), Rotterdam, the Netherlands~~

~~Yuen, M. C., King, I., & Leung, K. S. (2015). TaskRec: a task recommendation framework in crowdsourcing systems. *Neural Processing Letters, 41*, 223–238.~~

~~Zenger, T., Felin, T., & Bigelow, L. S. (2011). Theories of the firm-market boundary. *Academy of Management Annals, 5*(1), 89–133.~~

Zhao, Y., & Zhu, Q. (2014). Evaluation on crowdsourcing research: current status and future direction. *Information Systems Frontiers, 16*(3), 417–434. Zhao2014a

Dr. Yiwei Gong is an associate professor at the School of Information Management at Wuhan University and a guest researcher at the Faculty of Technology, Policy, and Management at Delft University of Technology. Yiwei serves as a reviewer for prestigious international journals and conferences in the field of Information Systems, and he is an editorial board member of Government Information Quarterly. His expertise and research interests include Business Process Management, Enterprise Architecture and Service Innovation.