Summer 5-28-2021

Task Decomposition for Knowledge-intensive Crowdsourcing: Managing Dependency and Structural Complexity

Xiaojie SHI
School of Information Management, Wuhan University, Wuhan, 430072, China, sprinashi@whu.edu.cn

Full Research Paper

Task Decomposition for Knowledge-intensive Crowdsourcing: Managing Dependency and Structural Complexity

Xiaojie SHI[1*]

[1]School of Information Management, Wuhan University, Wuhan, 430072, China

Abstract: Knowledge-intensive crowdsourcing (KIC) is expected to provide flexible access to knowledge and expertise. In this research, we proposed a task decomposition method to support the design decisions for KIC tasks decomposition and investigated how the level of granularity affects the crowdsourcing performance. To address the structural complexity, we employed the business process lens and coordination theory to describe the components of KIC tasks and their structural relationships. Afterwards, seven arithmetic tasks with different levels of structural complexity were designed and decomposed into subtasks with different levels of granularity by the proposed decomposition method. A laboratory experiment including 1960 groups of tests to simulate a real crowdsourcing environment was conducted to explore the relationship between different levels of granularity and the crowdsourcing performance. The results suggest that moderate decomposition helps to reduce the completion time and improve the quality of outcomes. A critical point of the level of granularity at which the completion time achieves the minimum is identified.

Keywords: knowledge-intensive crowdsourcing, business process lens, task decomposition, coordination theory

1. INTRODUCTION

Knowledge-intensive crowdsourcing refers to accomplishing knowledge-intensive tasks through the collaborative creation of knowledge content and requires special skills obtained by the means of crowdsourcing . In the dynamic and globalized environment, an increasing number of companies are attempting to make use of KIC to access knowledge and expertise. Despite its great potential, KIC often encounters a high risk of failure such as timeouts and poor quality of outcomes[1,2]. Task decomposition is one of the classical approaches to resisting project failure risks, and it has been explored and verified in many domains, such as service systems , cloud manufacturing and product development [3].

However, research on the KIC task decomposition is still in its infancy by far. Existing research on the management of KIC projects often focuses on task assignments or collaboration between crowd workers, e.g. . Task decomposition is often regarded as a design decision in the preparation of a crowdsourcing project with limited attention and is conducted by task providers or crowd workers based on their experience in practice, leaving the method of task decomposition in ambiguity. A key issue in task decomposition is identifying the proper level of granularity, which describes the grain size of the subtasks [4]. To the best of our knowledge, an in-depth investigation of KIC tasks decomposition to address the level of granularity and its impact on crowdsourcing performance has not been reported in KIC literature. A generic task decomposition method that could address the complex interdependency between subtasks and achieve maximum benefits from the decomposition is desired. To address this knowledge gap, this study proposes a KIC task decomposition method and aims to reveal how level of granularity affects the performance of crowdsourcing in terms of the completion time and quality.

To manage the structural complexity of KIC tasks, a business process lens and the coordination theory have been employed. Through a business process perspective, the decomposition of KIC tasks is an undertaking in which interdependent subtasks are constructed with a logical sequence. It allows the transformation of a complex

* Corresponding author. Email: sprinashi@whu.edu.cn

task decomposition problem into a business process coordination problem that aims to manage the dependencies between subtasks. The dependencies of the subtasks are clarified by the coordination theory in this study. The proposed task decomposition method was demonstrated through a laboratory experiment that consisted of 1960 groups of tests. The results indicate that a moderate decomposition of KIC tasks results in less completion time and higher-quality crowdsourcing. With regard to the level of granularity, a critical point at which the completion time of the crowdsourcing reaches the minimum was discovered from the experiments. The results of our research give decomposition guidance, which will help in identifying an appropriate level of granularity for supporting the design decision for KIC task decomposition and improving the crowdsourcing performance.

The rest of this paper is structured as follows. In Section 2, we introduce the business process lens and the coordination theory which serve as the theory background of our study. In Section 3, seven types of tasks are designed according to the coordination theory. Based on this, a task decomposition method is proposed to decompose these tasks into suitably sized subtasks. Experiments exploring the relationship between the level of granularity and the performance of crowdsourcing are presented in Section 4. Finally, in Section 5 we discuss the results of our study to provide both the theoretical and practical implications. Finally, we conclude the paper by proposing the limitation and the future research direction in Section 6.

2. BACKGROUND AND RELATED WOKRS

2.1 Business process lens and coordination theory

Decomposing a KIC task into a set of subtasks has to address the interdependencies among them. Dependencies arise when actions taken by one referent affect the actions or outcomes of another referent [5]. To analyze and manage these dependencies, a business process lens has been adopted in this study. A business process, according to van der Aalst and Hee [6], can be defined as a combination of individual activities and a workflow describing their logical order. The business process lens allows investigating the independent crowdsourcing elements and linking them into an integrated process.

The business process lens views the subtasks and their logical relationships within a knowledge-intensive task as a set of independent activities to be coordinated. Furthermore, the business process lens provides an organic view through which to understand the elements of a KIC task and their structures. Analyzing a KIC task from the business process lens provides insight into how its structure impacts its execution. In this way, comparing different business process designs enables the comparison and analysis of different decomposition solutions.

To understand and manage the dependencies within knowledge-intensive tasks, this research employed the coordination theory. The coordination theory, proposed by Malone and Crowston [7], involves how activities can be coordinated to work together harmoniously. According to the coordination theory, there are three kinds of basic dependence relationships between activities that bring the coordination problems (Figure 1): flow, share and fit. The relationship of flow means that the output of one activity is the exclusive input of another activity. Share means that two different activities require the same limited resources, such as human resources and money. The fit relationship indicates that the former two activities jointly form the latter activity. These three basic relationships and their combinations constitute dependencies in knowledge-intensive activities.

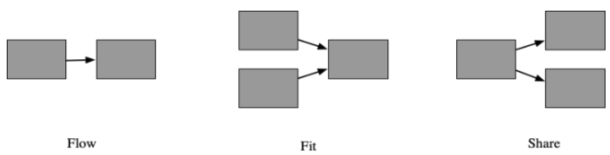


Figure 1. Three basic dependencies in processes [7]

The above three basic relationships and their combinations can enable the design of different processes that represent a variety of knowledge-intensive tasks. If these three kinds of dependencies and their combinations are well-managed within the task, the task decomposition can be addressed. Therefore, investigating and analyzing common dependencies and their related coordination mechanisms facilitate the development of a generic decomposition method for different KIC tasks.

2.2 Task decomposition

While there is limited research on the decomposition of KIC tasks, task decomposition has been studied in many other domains, such as robot systems, cloud manufacturing and product development. Relative studies of task decomposition contain decomposition approaches from different aspects. As a widely adopted methodology, task modular decomposition [8] divides a task into multiple subtask modules with certain associations among those subtasks. It attaches great importance to the interdependencies among subtasks. Quantitative criteria such as the degree of coupling [3,9] and the degree of equilibrium [10] are often used in task modular decomposition to evaluate the decomposition outcomes.

The degree of coupling measures the amount of information interactions between subtasks [11]. The greater the amount of information interaction between subtasks, the higher is the degree of coupling. Frequent information interactions between subtasks demand collaboration, which will increase the difficulties of crowdsourcing these subtasks to different crowd workers. The greater the amount of the information interactions between the subtasks, the more dependencies there are between them. It is vital to enable a suitably low coupling degree between two subtasks.

The dependency structure matrix (DSM) is a common method used to depict the correlation between two subtasks [12]. It divides a task into n subtasks and distributes it into an $n \times n$ matrix. As shown in Eq.（1）

, the design structure matrix is an n-th-order square matrix that stands for all of the subtasks, and n represents the number of subtasks. a_{ij} in the DSM represents the information interaction between the subtasks in the row and the corresponding subtasks in the column. The value of a_{ij} is shown in Eq.( ). If there is a correlation between Subtask i and Subtask j, then the combination between them is marked as 1 in the matrix. Otherwise, the value of their combination is 0. By using the DSM, the correlation between two subtasks can be clearly reflected.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{1}$$

$$a_{ij} = \begin{cases} cond(t_i, t_j) \ , \ i \neq j \\ 0 \qquad\qquad , \ i = j \end{cases} \tag{2}$$

$$R = \begin{cases} \sum_{s,t \in T} \frac{cond(t_1,t_2)}{|T| \times |T-1|}, |T| > 1 \\ 0, |T| \leq 1 \end{cases} \tag{3}$$

The formula used to calculate the coupling coefficient is given in Eq.(3) [13]. In this formula, R represents the coupling coefficient. t_i and t_j are two independent subtasks, $\sum_{t_1,t_2 \in T} cond(s,t)$ stands for the amount of information interaction between the subtasks, and $|T|$ presents the number of subtasks.

The degree of equilibrium reflects a subtask's uniformity in size [10]. An appropriate equilibrium degree of subtasks helps balance the completion time of the subtasks and consequently improve the overall completion time of the crowdsourcing project [14]. In the context of crowdsourcing, the standard deviation of a subtask execution time is used to quantify its degree of equilibrium [10]. The equilibrium degree of a subtask can be evaluated by Eq.(4). D represents the value of the standard deviation of the task execution time, and n is the number of subtasks. $\overline{T}$ is the average of the task execution time.

$$D = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (t_i - \bar{t})^2} \tag{4}$$

2.3 Related works addressing the task decomposition of KIC

Knowledge-intensive tasks are often large in scale and high in structural complexity and cannot be accomplished by an individual worker . The key to performing knowledge-intensive tasks is to decompose them into several subtasks and coordinate among these subtasks [15]. In related studies, there are two main research perspectives: technical and business perspectives. The technical perspective mainly focuses on how to supply crowdsourcing using technologies and emphasizes technical methods, techniques, and frameworks for solving problems in crowdsourcing . The business perspective focuses on how to complete certain business goals efficiently and effectively [16], involving crowdsourcing outcomes in terms of time, quality or cost.

Though existing studies have promoted the development of KIC, neither technical nor the business perspective studies well address the following two issues. First, current studies on decomposition limit their vision to specific types of tasks, such as article writing [17] and proofreading [18], leading to a low generalizability of the conclusions. Generic methods that can guide decompositions of different types of KIC tasks are still lacking. Second, how to achieve maximum benefits from the decomposition of KIC tasks with a proper level of granularity is still unclear. Many KIC studies do not focus on the decomposition of tasks, leaving this undertaking up to the task providers or crowd workers. Inappropriate levels of granularity will have an adverse effect on the crowdsourcing performance and bring problems such as low completion quality, high cost in terms of time and money or even failure of the crowdsourcing project [19]. We are therefore motivated to investigate a generic task decomposition method with predictable effects on the crowdsourcing performance.

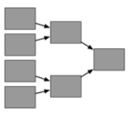
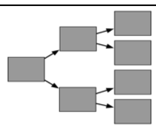
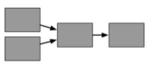
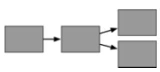
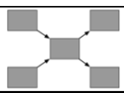
3. THE DECOMPOSITION METHOD FOR KIC TASKS

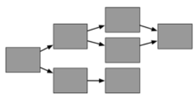
3.1 Task structural complexity

To study the structural complexity of KIC tasks, in this section, seven types of business processes with different levels of structural complexity are designed. As the business process lens helps transfer a KIC decomposition problem into a business process problem, the design of KIC tasks is actually the design of business processes. As defined in Section 2.2, flow, fit and share are the three basic dependencies. Table 1 demonstrates how single or multiple types of dependencies constitute tasks with different levels of structural complexity. In the figures listed in the second column, the rectangle represents a basic activity that is a minimum-sized building block constituting a KIC task, and the arrow represents the sequence relationships between the activities. Based on the three types of dependencies, seven knowledge-intensive crowdsourcing tasks with different levels of structural complexity are obtained.

Type 1 is made up of the flow relationships, which indicate that there are only flow relationships inside the

Table 1. Task structural complexity demonstrations

Type	Demonstration	Level of structural complexity	Constitution
1		low	flow
2		low	fit
3		low	share
4		medium	flow, fit
5		medium	flow, share
6		medium	share, fit

7		high	flow, fit, share

business process. The category is regarded as the business processes with the least structural complexity. Type 2 includes business processes that only possess fit relationships. Similarly, Task 3 contains only share relationships. Types 1, 2 and 3 of business processes represent KIC tasks with a low level of structural complexity. Combining any two types of dependencies could produce tasks with a higher level of structural complexity. We therefore have the flow and fit relationships with which to form Type 4, the flow and share relationships to form Type 5, and the share and fit relationships to form Type 6. These 3 types of business processes are regarded as tasks with a medium level of structural complexity. Type 7 contains all of the three basic relationships—fit, share, and flow—and represents KIC tasks with a high level of structural complexity.

3.2 The task decomposition method

In previous studies, the decomposition is dominated by crowd workers or task providers based on their own experience [1,15]. Decomposition solutions differ when deduced by different people. In addition, if the sizes of the subtasks are not balanced, this will lead to a large difference between the completion times of the subtasks, which will affect the overall completion progress of the whole task [10]. In consideration of the above two conditions, a task decomposition method that helps control the decomposition process by utilizing the concepts of the task coupling degree and task equilibrium degree is proposed in this section. The method can be divided into 3 steps, and it is introduced as follows.

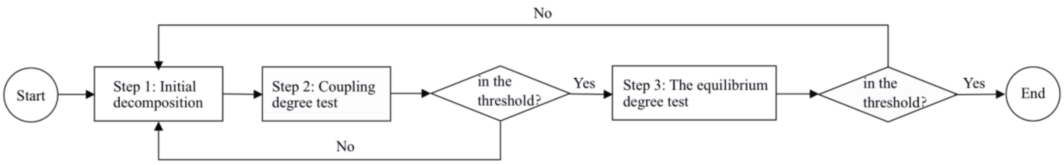


Figure 2. The steps in task decomposition

Step 1: Initial decomposition. In this step, a preliminary decomposition based on the structure of the KIC task is conducted. Tasks with a low level of structural complexity that consist of a single kind of relationship, such as Types 1, 2, and 3, can be evenly divided. For task types with higher structural complexity, we try to divide a task in as balanced a way as possible during the decomposition. This means that the number of flow, fit, and share relationships in each subtask should be approximately the same. At the end of this step, the temporary subtask Set T is obtained.

Step 2: Coupling degree test. This step aims to ensure the independency of subtasks after the initial decomposition. The coupling degree is controlled within a reasonable range, which might vary for different types of tasks and decomposition goals. To maintain the proper coupling degree of subtasks, it is necessary to set a reasonable decomposition threshold in advance. The appropriate decomposition threshold can be determined according to the actual decomposition experience or repeated experiments. In this step, if the degree of coupling of the task meets the requirements of the threshold, the subtasks in Set T are independent of each other, and then the decomposition can be continued with Step 3. If not, there are too many information interactions between the subtasks. In this case, a return to step 1 is needed. The re-decomposition requires merging subtasks with a higher degree of coupling.

Step 3: The equilibrium degree test. This step uses the completion time of a subtask to measure the scale of a subtask. A balanced completion time means that the subtasks are of balanced sizes to avoid the situation in which a few subtasks take much more time than do the others, resulting in an unnecessary long project duration. The completion time of each subtask is evaluated by experts in this step. Similar to the case for Step 2, a reasonable threshold of the equilibrium degree is given in advance. If the equilibrium degree of the subtasks is within the threshold, that the subtasks are balanced and the decomposition is effective. Then, an ultimate decomposition result Set U is obtained. If the equilibrium degree of the subtasks exceeds the threshold, the scale of the subtasks is unbalanced, and need to go back to Step 1 for re-decomposition. The subtasks that have an overly long completion time need to be split, and subtasks with a completion time that is too short should be merged.

4. EXPERIMENT

In this chapter, a knowledge-intensive task is designed and decomposed into a number of subtasks according to the decomposition method proposed in Section 3.2. After the decomposition, a crowdsourcing experiment was conducted to explore how the level of granularity of the tasks will affect the performance of the crowdsourcing. In this study, conducting a laboratory experiment is chosen rather than conducting experiments in a real-world crowdsourcing environment. This is because there are many factors that could impact the performance of crowdsourcing in a real crowdsourcing environment. For example, crowd workers may be interrupted by personal affairs. In such a case, the performance of the crowdsourcing would be greatly affected and could not well reflect the effect of the task decomposition. What's more, since the background and the ability of the crowd workers various, the participants of the laboratory experiment are randomly selected. Therefore, this study uses a laboratory experiment to explore the effects of the task decomposition to avoid influences from other factors.

The experiments were conducted in February 2019. The participants of the experiments were undergraduate students of our university in various majors. Recruiting college students to conduct lab experiments for crowdsourcing or process design research is a strategy has been adopted by previous studies, and its validity has been verified [20].

4.1 Design of the knowledge-intensive tasks

The performance of crowdsourcing can be measured in many ways, such as according to time, price, and quality. In the experiment, we judge the performance of crowdsourcing by two parameters: the completion time of the task and the quality of the results. Existing KIC studies often employ experiments to investigate KIC tasks. The experiments simulate the design and execution of KIC tasks including arithmetic tasks, article editing, article writing, software developing, etc. This research chose arithmetic tasks as the knowledge-intensive tasks for the following two reasons. The first is that the quality of the outcomes can be easily judged by the accuracy of the calculation results. The other reason is that arithmetic tasks are relatively short in terms of completion time and are low in cost, therefore allowing the feasibility of conducting a large number of crowdsourcing tasks in the experiment.

Seven different kinds of KIC tasks of arithmetic were designed. For example, Task 7 (Table 2) is an arithmetic task that consists of 11 activities. There are several dependencies between these subtasks. The result of subtask f relies on the calculation output of subtask c. The calculation processes of subtask b and subtasks c depend on the

Table 2. The task of Type 7

activity	Please calculate the value of k.
activity 1	a=12+4-6
activity 2	b=a+9+7
activity 3	c=a-15+7
activity 4	d=b-5+1
activity 5	e=b-5+8
activity 6	f=c+1+4
activity 7	g=d-e+2
activity 8	h=f-3+5

activity 9	i=f+1+3
activity 10	j=h+i-6
activity 11	k=g+j-3

result of subtask a. Subtask g depends on the results of d and e. According to the concept of the relationships of fit, share and flow, the relationship between subtask f and subtask c is a flow relationship. Subtask a, subtask b and subtask c are in a share relationship. Subtasks d, e and g are in a fit relationship. Likewise, the arithmetic task can be translated into a business process problem, as shown in Figure 3. In a similar way, the other tasks can be translated into business processes.

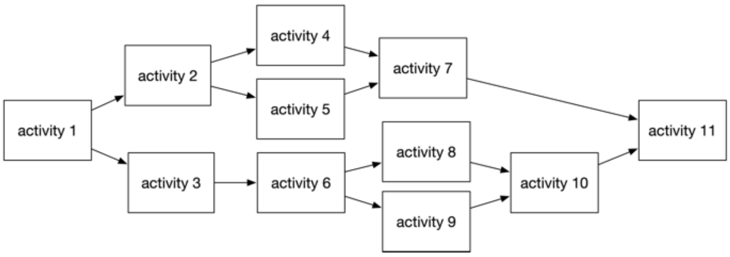


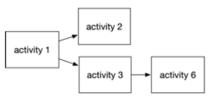
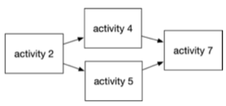
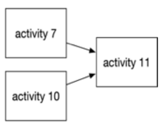
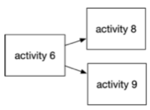
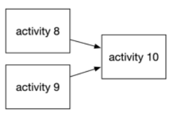
Figure 3. The task of Type 7 as expressed through the business process lens

4.2 Decomposition of the tasks

The decomposition method presented in Section 3.2 is applied in this section to decompose the knowledge-intensive tasks into subtasks to below a certain level of granularity. Decomposing the Type 7 tasks into 5 subtasks serves as an example demonstrating the decomposition process.

In Step 1, a preliminary decomposition of the task according to the structure is conducted. Overall, the flow, share, and fit relationships in each subtask are evenly distributed. The decomposition results, as well as the dependency types and activities, are presented in Table 3. They meet the requirements of Step 1, and the decomposition continues with Step 2.

Table 3. The task decomposition for Type 7

Subtasks	Demonstration	Dependency types
Subtask 1		share, flow
Subtask 2		share, fit
Subtask 3		fit
Subtask 4		share
Subtask 5		fit

In the second step, we need to judge and analyze the coupling degree of the decomposition results from Step 1. Based on experience, we set the task coupling degree threshold to 0.6. According to Eq.(2) and Eq.(3), we can get the task information association matrix and the degree of coupling of the task is within the defined threshold. The decomposition proceeds to the next step. In Step 3, the job is to analyze and control the degree of equilibrium of the subtasks. Based on experience, we set the threshold value to 4 s. We evaluate the time required for each subtask based on previous crowdsourcing experiments. According to Eq.(4), the calculated equilibrium degree of the subtasks is 2.76 s, which is less than the threshold of 4 s and therefore meets the conditions. At the end of the process, the task decomposition solution is valid.

4.3 Data collection and analysis

There are 7 kinds of knowledge-intensive tasks, which are of different levels of structural complexity. They are decomposed into 7 seven different levels of granularity, respectively. Thus, 196 different kinds of subtasks are obtained. Each kind of subtask requires a participant to complete it independently in our laboratory room under monitoring. To eliminate or reduce the impact of different calculation speeds of individuals on the experiment, the test for each subtask was performed repeatedly by different participants 10 times. Furthermore, to avoid the impact of different academic background on the calculation speed, we averagely assigned participants in different majors to each repeating group. At the end, in total, 1960 sets of results were collected in the experiment. As the performance of crowdsourcing is reflected by the completion time and accuracy, the results and completion times of the subtasks are recorded in the experiment. The experimental platform is powered by www.wjx.cn, which can assign the tasks randomly to the participants and record the execution times of the subtasks.

By adding the completion times of each group of subtasks, the overall completion time of the seven complex types of crowdsourcing tasks with different levels of granularity is obtained. The results are presented in Figure 4. We find that at the same level of granularity, the more complex is the knowledge-intensive task, the longer is the time it takes during crowdsourcing. By breaking down these knowledge-intensive tasks, the crowdsourcing time can be reduced. To further analyze the changes of the completion time with different levels of granularity, we calculate the rate of time change by Δt/t. We assume that when the value of |Δt/t| is below 0.06, the time change can be neglected.

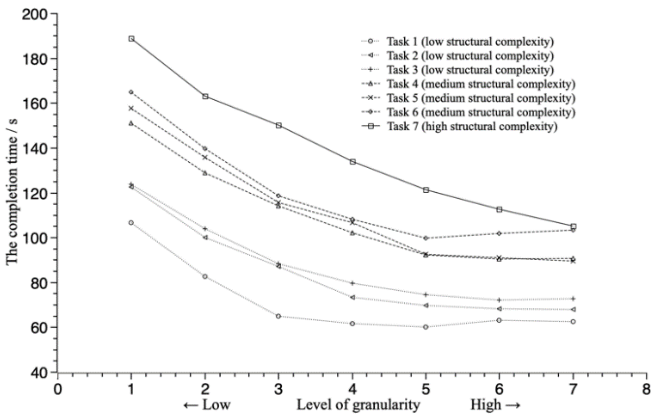


Figure 4. The completion times of the tasks with different levels of granularity for different types of tasks

For Type 1, the values of $|(t_3-t_4)/t_3|$, $|(t_4-t_5)/t_4|$, $|(t_5-t_6)/t_5|$ and $|(t_6-t_7)/t_6|$ are below 0.06. This indicates that when the level of granularity is less than 3, the crowdsourcing time decreases as the decomposition increases. However, when the task is decomposed into more than 3 subtasks, the overall completion time no longer reduces as the decomposition proceeds further. In contrast, the completion time stays the same and even increases. Therefore, we define the level of granularity of 3 as a critical point for Type 1, at which point the overall completion time reaches its minimum. For tasks with low levels of structural complexity, the critical point occurs when the level of granularity is 4. For tasks with medium levels of structural complexity, such a point occurs when the level of granularity is 5. For Type 7, the structural complexity of which is the highest, such a point still does not appear when the task is decomposed to a level of granularity of 7. Therefore, where the critical point appears is related to the level of structural complexity of the tasks. The more complex is the task, the higher is the level of granularity at which the critical point appears. By calculating the accuracy of each task, we observe the relationship between the level of granularity and the accuracy of tasks with different levels of structural complexity. Figure 5 presents the accuracies of the calculation results at different levels of granularity for the 7 types of tasks.

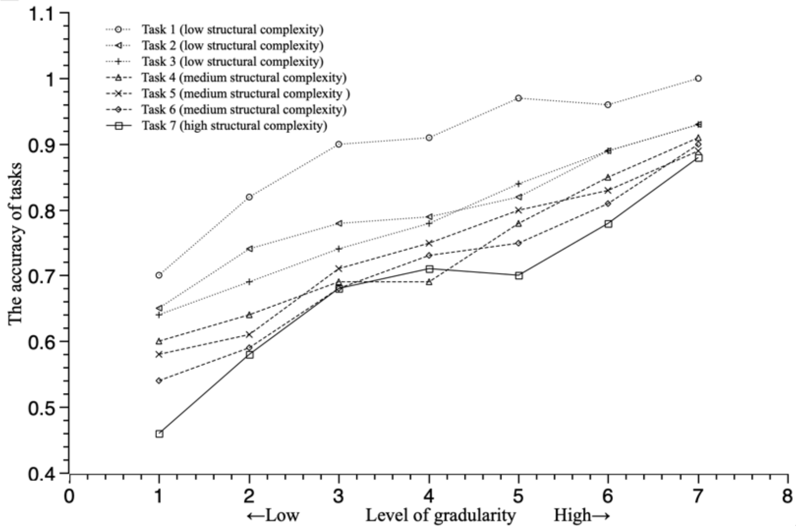


Figure 5. The accuracy of the results at different levels of granularity for each type of task

Based on the results, we found that the accuracy of each task stays the lowest when the task is not decomposed. The most complex task 7 has an accuracy of only 0.45, while task 1, which has a low level of structural complexity, has an accuracy of 0.7. For the same KIC task, as the level of level of granularity increases, the accuracy increases gradually. To compare the accuracies of tasks with different levels of structural complexity, we averaged the accuracies of tasks 1-3 and 4-6, respectively, and obtained the average accuracy for different levels of structural complexity. By comparing the accuracies of the tasks with these three different levels of structural complexity, we find that at the same level of granularity, the less complex is the KIC task, the higher-quality will be the crowdsourcing outcome.

For each decomposition solution, 10 repeated experiments were performed to ensure the credibility of the data. To examine the degree of dispersion of the completion times of each of these 10 groups of tasks, the standard deviation of the completion times for different types of tasks at different levels of granularity are calculated. The findings show that for each type of task, as the level of granularity increases, the standard deviation of the completion time decreases, indicating that a high level of granularity degree helps improve the predictability of the completion time of the tasks. The structural complexity of knowledge-intensive tasks brings instability to the completion time of the crowdsourcing. The results also indicate that as the level of granularity increases, the deviation often reduces gradually. This means that higher level of granularity results in less uncertainty in the completion time.

5. DISCUSSION

The theoretical model (Figure 6) summarizes the key findings of the experiment. This study investigates how the level of granularity effects KIC performance by the two dimensions of quality and time, where the level of granularity is an independent variable and KIC performance is a dependent variable. Structural complexity exerts an influence on the relationship between the level of granularity and KIC performance, and thus serves as the moderate variable in the research. In order to avoid the influence of other unrelated variables, we control the external disturbances that may impact the results of the experiment, such as the interruption from personal affairs and their different calculation speeds.

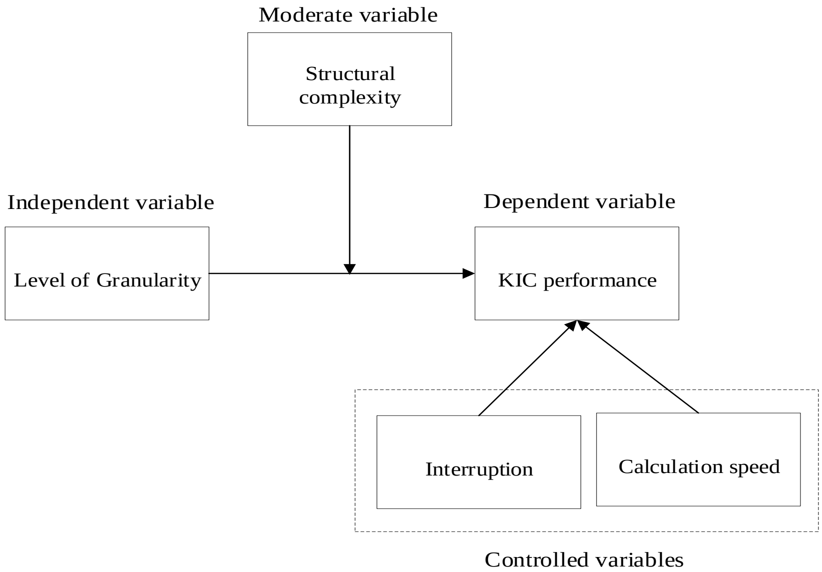


Figure 6. The theoretical model

The research discussed in this paper can have potential impacts on KIC research. Previous research regarded task decomposition as a given preparation stage of KIC projects [15]. In contrast, our study investigates and explains the relationships between the level of granularity and the completion time and quality of KIC tasks. The results in this study allow KIC researchers to include and control task decomposition in their research design.

In addition, this study indicates a new approach to the research of KIC task decomposition by applying business process methods and the coordination theory. We use the three basic dependence relationships to model the relationships between subtasks. Thus, an abstract decomposition problem is transferred into a concrete business process modeling problem. Business process methods are applied as the coordination mechanism to enable task decomposition. According to the coordination theory, dependencies and the mechanisms used for managing them are generic. Although further investigation of the different aspects influencing crowdsourcing performance is required to develop a more comprehensive method, this approach is promising in terms of its generalizability. As the business process can serve as a template of for creating multiple, real-life instances of the same process, it enables the development of generic process patterns to decompose KIC tasks, removing the different types and features of the tasks.

The results of our research also provide useful practical insights for use in the task decomposition of KIC. The results demonstrate the benefits of task decomposition in crowdsourcing. The structural complexity of the tasks influences the relationship between the level of granularity and the crowdsourcing performance. Thus, KIC tasks with more complex structures should be decomposed into subtasks with a smaller size. An important finding of this research is the existence of a critical point in the relationship between the level of granularity and the completion time. As the task is reaches a higher level of granularity, the completion time of the crowdsourcing continually declines. However, once the critical point has been reached, the completion time stays the same or even inclines when the level of granularity becomes higher. Thus, to complete KIC tasks within a minimum amount of time, it is vital to identify the critical point. Furthermore, the decrease of the completion time becomes slow when the level of granularity nears the critical point. Employing a low level of granularity could result in the reduction of the completion time of task decomposition. The decision of whether to continue with a higher level of granularity to reach the critical point is a tradeoff between the cost of task decomposition and the savings from further reducing the completion time of the crowdsourcing task. In addition, increasing the level of granularity could make the completion time of a task more predictable. These findings allow for reasonable decision-making with regard to KIC task decomposition.

6. CONCLUSIONS

To fill the knowledge gap in managing KIC task decomposition, we start from the dependency relationships of KIC subtasks and investigate them through a business process lens. In the research, seven types of KIC tasks with different levels of structural complexity were designed according to the coordination theory. A decomposition based on the structure of the tasks was conducted to obtain 196 groups of subtasks that represent different level of granularity. To explore how the level of granularity will affect the performance of a KIC, a series of laboratory experiments was conducted. The results indicate that the performance of the KIC is closely related to the level of granularity. Task decomposition helps reduce the completion time of the crowdsourcing and improve the predictability of the completion time, which further helps reduce the timeout risk. The quality of the crowdsourcing also improves as a result of task decomposition. A critical point exists at which the level of granularity enables the minimum completion time of the crowdsourcing. The proposed decomposition method and the findings of the experiment provide guidelines for supporting the design decision for effective KIC task decomposition which could improve the crowdsourcing performance at the end.

In this study, there are several limitations that should be viewed as the starting points for further research. First, our analysis is based on experiments using arithmetic tasks. More kinds of knowledge-intensive tasks should be explored to further test the generalizability of the conclusion. Second, to constrain the scale of the KIC tasks in the experiment, tasks with the different structural complexities consist of the same number of subtasks. Whether and how the scale of a KIC task influences the crowdsourcing performance remains to be investigated. Third, a critical point at which the task completion time will reach its minimum has been identified in the research. Knowing the position of this point can enable more precise decision-making in preparing crowdsourcing projects. However, currently, this point can only be obtained through a large number of crowdsourcing experiments and is time- and labor-consuming. Whether there is a method that can be used conveniently to locate or estimate this point is worth exploring.

REFERENCES

[1] Zhao, Y., & Zhu, Q. (2014). Evaluation on crowdsourcing research: Current status and future direction. Information Systems Frontiers, 16(3), 417-434.

[2] Daniel, F., Kucherbaev, P., Cappiello, C., Benatallah, B., & Allahbakhsh, M. (2018). Quality control in crowdsourcing: A survey of quality attributes, assessment techniques, and assurance actions. Acm Computing Surveys, 51(1), 7.

[3] Zhang, X., Yang, Y., & Bao, B. (2016). Task decomposition and grouping for customer collaboration in product development. Journal of Intelligent Systems, 25(3), 361-375.

[4] Chiriac, N., Hölttä-Otto, K., Lysy, D., & Suh, E. S. (2011). Level of modularity and different levels of system granularity. Journal of Mechanical Design, 133(10), 101-117.

[5] Mccann, J. E., & Ferry, D. L. (1979). An Approach for Assessing and Managing Inter-Unit Interdependence. Academy of Management Review, 4(1), 113-119.

[6] Van Der Aalst, W., & Van Hee, K. (2004). Workflow management: models, methods and systems. Boston: MIT Press.

[7] Malone, T. W., & Crowston, K. (1994). The Interdisciplinary Study of Coordination. Acm Computing Surveys, 26(1), 87-119.

[8] Habib, M., & Paul, C. (2010). A survey of the algorithmic aspects of modular decomposition. Computer Science Review, 4(1), 41-59.

[9] Liu, W., Ji, J., Yang, Y., & Zhang, L. (2018). Capability-Based Design Task Decomposition in Heavy Military Vehicle Collaborative Development Process. The International Academy for Production Engineering, 70(1), 13-18.

[10] Bao, B., Yang, Y., Li, F., & Xue, C. (2014). Decomposition model in product customization collaborative development task. Computer Intergrated Manufacturing Systems, 20(7), 1537-1545.

[11] Smith, R. P., & Eppinger, S. D. (1997). A predictive model of sequential iteration in engineering design. Management Science, 43(8), 1104-1120.

[12] Maheswari, J. U., & Varghese, K. (2005). Project scheduling using dependency structure matrix. International Journal of Project Management, 23(3), 223-230.

[13] Pang, H., & Fang, Z. (2008). Task decomposition strategy and granularity design in networked collaborative environment. Computer Intergrated Manufacturing Systems, 14(3), 425.

[14] Mizusawa, K., Tajima, K., Matsubara, M., Amagasa, T., & Morishima, A. (2018). Efficient pipeline processing of crowdsourcing workflows. Paper presented at the Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Torino, Italy, October 22 - 26, 2018

[15] Jiang, J., An, B., Jiang, Y., Lin, D., Bu, Z., Cao, J., et al. (2018). Understanding crowdsourcing systems from a multiagent perspective and approach. ACM Transactions on Autonomous and Adaptive Systems (TAAS), 13(2), 8.

[16] Luz, N., Silva, N., & Novais, P. (2015). A survey of task-oriented crowdsourcing. Artificial Intelligence Review, 44(2), 187-213.

[17] Kittur, A. (2011). CrowdForge: crowdsourcing complex work. Paper presented at the Chi 11 Extended Abstracts on Human Factors in Computing Systems, Santa Barbara, USA, October 16 - 19, 2011

[18] Jiang, H., & Matsubara, S. (2014). Efficient task decomposition in crowdsourcing. Paper presented at the International Conference on Principles and Practice of Multi-Agent Systems, Gold Coast, Queensland, Australia, December 1-5, 2014

[19] Allahbakhsh, M., Benatallah, B., Ignjatovic, A., Motahari-Nezhad, H. R., Bertino, E., & Dustdar, S. (2013). Quality control in crowdsourcing systems: Issues and directions. IEEE Internet Computing, 17(2), 76-81.

[20] Thuan, N. H., Antunes, P., & Johnstone, D. (2018). A Decision Tool for Business Process Crowdsourcing: Ontology, Design, and Evaluation. Group Decision and Negotiation, 27(2), 285-312.