http://doi.org/10.17576/apjitm-2022-1101-02
Asia-Pacific Journal of Information Technology and Multimedia
Jurnal Teknologi Maklumat dan Multimedia Asia-Pasifik
Vol. 11 No. 1, June 2022: 12 - 25
e-ISSN: 2289-2192

Received: 29 August 2021
Accepted: 29 November 2021
Published: 30 June 2022

COMPOSING MULTI-RELATIONS ASSOCIATION RULES FROM CROWDSOURCING REMUNERATION DATA

SITI SALWA SALLEH
NURHAYATI ZAKARIA
NORJANSALIKA JANOM
SYARIPAH RUZAINI SYED ARIS
NOOR HABIBAH ARSHAD

ABSTRACT

In crowdsourcing, requesters are companies that require external workers to execute specific tasks, whereas a platform acts as a mediator to match and allocate the tasks to digital workers. To assign it to a worker, the platform must first identify the types of tasks and match them to the appropriate workers based on their level of competency. Each worker has different ICT competencies which affect work quality and remuneration. However, general practise frequently assumes a single level of worker's capability for all tasks, hence the categorisation of difficulty of tasks is unclear and inconsistent. Apart from causing dissatisfaction among workers, this also implies an absence of incentive standardisation. Therefore, this study explores this matter and which aims to identify and visualise the parameters that affect remuneration determination. To gather the data, focus group discussions and interviews with crowdsourcing players have been conducted. The data comprise a lot of redundancies, therefore an apriori algorithm is used to normalise it by removing redundancies and then extracting significant patterns. Next, an association rule is used to uncover correlations between parameters. To gain a more understandable insight, the data relationship is visualised using an alluvial chart that manages to illustrate the flow. Findings show that task type, outcome variation, and competency requirements demonstrate a degree of interdependence. It is suggested that there is a significant pattern showing that the remuneration scheme is determined by five levels of DW, which are expert, advanced, intermediate, novice, and basic. Advance workers are most likely to participate in the crowdsourcing, and the remuneration scale is suggested to be wider compared to others. The study's findings provide input for remuneration strategy in future work.

Keyword: crowdsourcing, framework, multidimensional, digital worker, digital task, remuneration

INTRODUCTION

Crowdsourcing is a concept where a company outsources a set of digital tasks to third party (Kietzmann, 2017). According to the World Economic Forum's Future of Jobs' report, 84 percent of organisations expect to rapidly digitalize their work processes, including a significant increase in remote work—with the potential to shift 44 percent of their employees to work remotely (World Economic Forum, 2020). In Malaysia, crowdsourcing began in 2012 when the Malaysia Digital Economy Corporation (MDEC) launched a strategic study with the goal of training individuals from the B40 group to do online digital jobs. Despite its primary goal of uplifting the B40 household income, MDEC enhanced the sector through a number of international collaborations, where the implementation of business model and digital economy framework has been evolved. In the local context, the e-Rezeki program, owned by MDEC, is an example of a crowdsourcing platform. In the international arena, there are many platforms such as 99designs and Amazon Mechanical Turk. Businesses outsource some jobs to crowdsourcing to save money on office space rentals, staff training and development, and to avoid paying permanent employees to perform routine or odd jobs (Pilloni, 2018; Zakariah *et al*., 2016).

The crowdsourcing process begins with a requester forwarding it to a mediator (platform), who then assigns it to a digital worker (DW). Typically, the advertisement is shared on the online marketplace where it describes the skill needed, the duration, the anticipated completion date, and the rate of pay. Allocating it to the DW requires the platform to identify the types of tasks and match them to the suitable workers based on their competency level, which is the most critical aspect in the crowdsourcing process (Zhao *et al.,* 2019; Thuan *et al.*, 2013). Tasks are given to DW, who are distant and sometimes not thoroughly vetted (Qiu et al., 2018). For some specific quality and professional tasks, the platform wants them to participate multiple times (Wang & Yu, 2020). However, the level of competency affects work quality and remuneration (Corney et al., 2009; Rouse, 2010). Thus, it is important for DW to possess distinct ICT competency as well as the quality of their past work (Schenk & Guittard, 2011; Ye & Kankanhalli, 2013; Nakatsu *et al.,* 2014). A common issue with existing task allocation methods is that a platform often assumes a single level of worker's ability for all tasks (Luo & Jennings, 2021). Apart from creating issues in the incentives, the effects of this are losing high-quality workers, and ways to engage DW need to be formed (Pilloni, 2018). In other perspective, Oppenlaender *et al.*, (2020) highlighted that fair compensation is the priority in creative work as workers complaint about creative work being underpaid. The workers insist that requesters or platforms have to be transparent and accurate about the time and effort required to allow workers to make an informed decision about self-selecting to participate.

Every DW aims to maximise his/her own payment, while the requester intends to achieve high-quality final solutions to tasks with minimal cost (Xu *et al.*, 2019). Pricing tasks improperly can disincentive workers from performing them. This trade-off between quality and worker's incentives complicates pricing decisions in crowdsourcing markets, necessitating the development of a new scheme that considers both the platform and the requester (Qiao et al., 2019). Naturally, the primary goal of any crowdsourcing business model is to determine the true cost at the lowest feasible price (Mridha & Bhattacharyya, 2019). However, if rewards are not well quoted, the workers will degrade the overall quality of the outcomes (Luo & Jennings, 2021). Another remuneration-related issue is the classification of easy and complex tasks, which is also unclear and inconsistent. Some researchers divide tasks into three categories: invention, evaluation, and organization (Corney *et al.*, 2009). Others characterise it as simple, complex and, creative (Schenk & Guittard, 2011; Thuan *et al.,* 2013). It was also divided into well-structured and unstructured parts (Nakatsu *et al.,* 2014; Buettner, 2015). Simple versus complex tasks have no clear definition, and this implicates no standardisation in terms of incentive.

To look into this matter, this study was conducted based on the following research questions; 1) Which parameters have a significant effect on remuneration? 2) How are these parameters related to one another? 3) How should digital workers be classified, and does classification affect remuneration? The study's aim was to identify, classify, and group the parameters, and to aid comprehension, the relationships have been visualized. The outcomes of this study will form a basis for future research on payment-related matters as a design of incentive scheme is critically important (Qiu *et al.,* 2019). This article is organized as follows. It begins with an introduction and then continues on to a review of the literature in Section Two. Section Three discusses the methods, while Section Four presents the analysis, then followed with findings and discussion. Finally, in the conclusion section, the concluding remark and future work are presented.

RELATED WORK

Requesters, platform and digital workers (DW) are the three key components of the crowdsourcing ecosystem. The basic flow of crowdsourcing process is depicted in Figure 1, and the description of each component is as follows:

1. Requesters are companies that outsource the jobs or tasks.
2. Platform is an operator that connects, passes, and offers tasks.
3. A digital worker (DW) is an individual who pull the tasks.

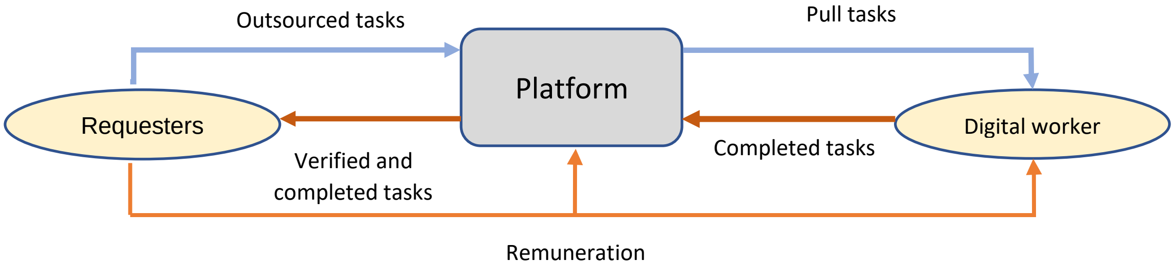

FIGURE 1. Crowdsourcing Business Model

Requesters are companies that need external workers to perform some tasks and a platform serves as a middleman to administer the matching, allocating and offering certain incentives for accomplishing the assignment (Hosseini *et al.*, 2014; Aldahari *et al.*, 2018; Zakariah *et al.*, 2016). Platform advertises the tasks and typically has a mechanism for raising both the requester's and the DW's rewards and ratings (Qiu *et al.*, 2019; Qiu, Squicciarini & Hanrahan, 2019). Low rewards may not attract enough workers, resulting in insufficient responses to a task (Luo & Jennings, 2021). In some models, tasks are offered through auctions, where remuneration is established either by bidding, prices quoted or advertised beforehand (cited in Bhattacharyya & Mridha, 2018). Normally, the allocation strategy is determined by the level of difficulty of tasks and the capabilities of workers to complete it. The level of DW competency varies due to their heterogeneity as they may be homemakers, retirees, unemployed grads, gig workers, students, or new graduates depending on their circumstances (Hu & Zhang, 2017). Among them, the DW could also be a professional who works full-time elsewhere, such as an engineer or a software developer.

According to Goncalves *et al.* (2017), the primary challenge in crowdsourcing is appropriately routing and assigning workers to suitable tasks. As a response, that research studied a routing and assignment mechanism that would enable tasks to be assigned to DW based on their cognitive capabilities. They assessed respondents' visual and fluency in cognitive abilities using a kit of Factor-Referenced Cognitive Tests. The findings indicate that the cognitive abilities of participants are highly correlated with their performance which relates to level of competency. Chen *et al.*,(2015) proposed an algorithm called Opt-KG to address the budget allocation problem with imperfect workers by assuming that the costs of workers are the same. Similar research was conducted in another study, that also focused on allocation tasks based on certain budget and experimented on dynamic worker pool (John & Bhatnagar, 2019). In the study, they proposed a Markov mathematical framework for modelling the problem. The study outcome showed that the allocation process is comparable to conventional way but can be performed in a shorter time frame.

The other vein of evidence suggests that social incentives can improve the performance of collaborative work even if the task criteria are higher (Feyisetan & Simperl, 2017). Meanwhile, Mizusawa et al. (2018) proposed a pipeline processing system that allows the price of a task to fluctuate dynamically in the platform marketplace system. They compared the new methods against the batch processing scheme, and the results indicated that fluctuating prices demonstrate significant reductions in matchmaking time. Other than that, Qiu *et al.* (2019) designed incentive mechanisms to reduce the information asymmetry between requesters and DW. They suggested to rate the DW in order to motivate them to obtain higher rating in order to attract requesters, which in turn bolsters platform sustainability. Yu *et al.* (2019) proposed a reputation-based incentive mechanism, which motivates DW to improve tasks quality in a location-based participatory sensing application for environmental monitoring. The higher the reputation of a DW, the greater the chance the participant will gain the reward in the auction. In addition, there is also a model that encompasses a series of proper reward-penalty function pairs and workers' personal order values to align the interests of different requesters (Xu *et al.*, 2019). To provide matchmaking autonomy, a multi-objective recommendation model has been proposed where it allows every worker and requester to set the parameters that meet their goals (Aldahari *et al.*, 2018).

The allocation mechanism, selection of DW based on task type, remuneration, and incentives have all been thoroughly studied. To elicit additional perspectives on these issues, this study employs an approach that solicitate input from platform operators and related agencies in the local crowdsourcing business.

MATERIAL AND METHODS

This research includes theoretical and empirical activities that are divided into three stages. The overall work is presented in Figure 2, and the three stages comprise knowledge acquisition and data gathering, analysis, and visualization. Previous work and documents were reviewed as part of the knowledge acquisition process. Documents from MDEC and local operating platforms were also reviewed. Two platform operators gave DW profiles and task assignment data samples.

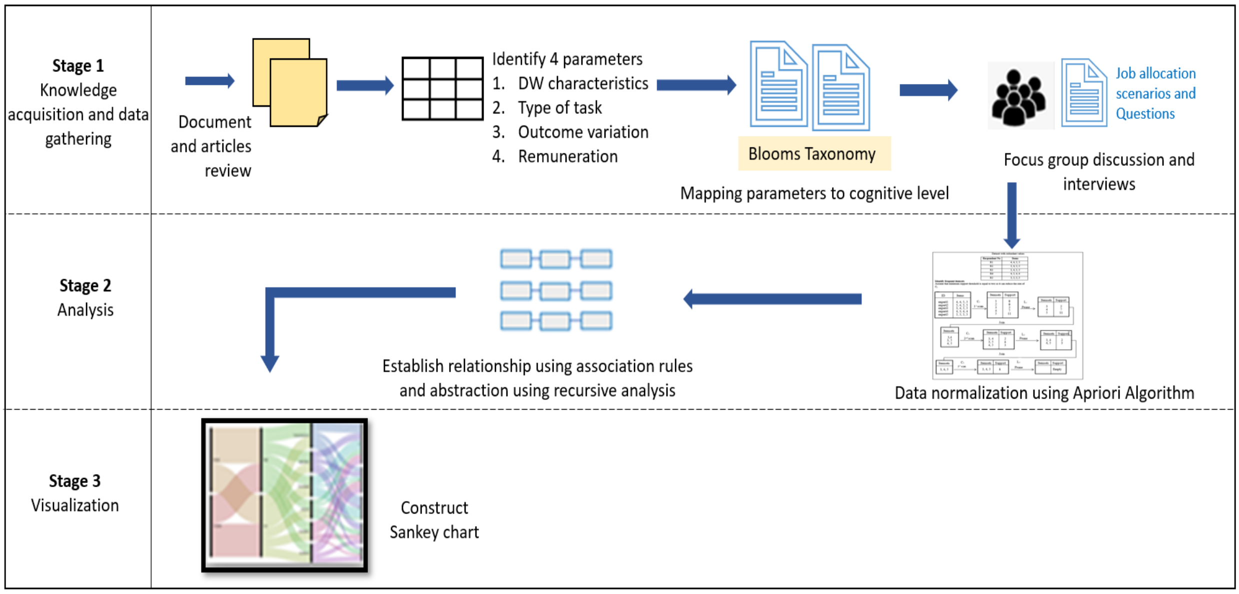

FIGURE 2. Flow of study

Table 1 summarises past research on the four characteristics of job type, DW, task result, and payment. It demonstrates that crowdsourcing encompasses a wide variety of tasks. However, this research proposes categorising tasks into two categories: simple and complex. What distinguish the two categories are the outcome variations. For instance, while performing simple tasks such as image tagging, the result is a low variation, which is the tagged photos, while simple creative tasks such as drawing a sheep resulted in a large variety of sheep drawings.

TABLE 1. Parameters discussed by previous researchers

	Researcher	Task	DW	Task Outcome	Remuneration
1	Luo & Jennings, 2021	Many different tasks	Many different types of workers	High variation	Mechanism for budget-limited task allocation
2	Oppenlaender *et al.*, 2020	Creative work	Professional, casual Novelty, self-developer	High variation	Different levels of worker aim for different incentives
3	Wang & Yu, 2020	Diverse and complex Novel, creative, innovative	Knowledgeable Unknown Heterogenous	High variation	Incentive based on work quality
4	Jonas *et al.*, 2020	Simple task Creative and ideation task	Individual or casual worker. Professional work with high rewards.	High variation	Very low Paid per task
5	Yu *et al.*, 2019	Simple task but require reliable worker	Reputable workers	Low variation	Auction based Tasks given to the lowest bidder and Reputation-based
6	Haralabopoulos *et al.*, 2019	Micro task (Searching, labelling and clicking)	Individual or casual worker	Low variation	Monetary incentive results low quality of work Volunteer produced better quality of work
7	Aldahari *et al.*, 2018	Macro task (Web design)	Individual	High variation	Quality of work and trustworthiness proficiency level of the worker in each skill or type of task
8	Bhattacharyya & Mridha, 2018	Simple task	Individual Non-expert	Low variation	Auction or prices. Price powered by requester
9	Zakariah *et al.*, 2016)	Simple task	Individual with ICT knowledge, skills, and experience	Simple task with low variation	Financial incentives, coupon, bonus, free product, free service
10	Nakatsu *et al.*, 2014	Well-structured task (simple), unstructured task (creative),	Non-expert with low commitment. Experts require high commitment	Easy to complete	-

After that, the task types were mapped using Bloom's digital taxonomy and compiled a collection of task examples in order to connect them to the outcome variation. The taxonomy was also used to map digital tasks to cognitive or skill levels, as summarised in Table 2.

TABLE 2. Competency level and type of tasks

Competency level	Example of instructions / actions (Job scenario)	Task type and outcome		
		Type	Low variant	High variant
Knowledge	Label, name, describe, list, match, identify, outline, reproduce, select, state	Simple	Yes	-
		Complex	-	Yes
Comprehension	Convert, defend, distinguish, estimate, explain, extend, generalize, infer, summarize, predict	Simple	Yes	
		Complex	-	Yes

Application	Change, computes, demonstrate, discover, operate, predict, prepare, produce, relate, show, solve, use	Simple Complex	Yes	- Yes
Analysis	Differentiate, distinguish, identify, illustrate, infer, outline, point out, relate, select, separate, breakdown, categorize, diagram, inventory	Simple Complex	Yes -	- Yes
Evaluation	Appraise, compare, conclude, contrast, criticize, describe, discriminate, explain, justify, interpret, relate, summarizes, support	Simple Complex	- -	Yes Yes
Synthesis	Categorize, combine, compile, compose, create, devise, design, explain, generate, modifies, organize, plan, rearrange, reconstruct, relate, reorganize, revise, rewrite, summarize, tell, write	Simple Complex	- -	Yes Yes

A low variant of outcome means that the task does not produce a variety of outputs, while a high variant means that the task produces "no specific form" of output, as aforementioned sheep drawing. A simple and repetitive task that produces a low variant of an outcome does not require high skill. On the other hand, non-repetitive tasks that produce high variants of outcomes, like logo design, require creative skills as the outcome is novel. The same context goes to complex tasks with low outcome variation, that demands specific talent. Those with large outcome variations necessitate professional skills, as well as a variety of solutions based on imagination, reasoning, and experience.

The data gathering included a series of focus group discussions (FGDs) and interviews with operators of local crowdsourcing platforms and MDEC project managers. Five of them participated in the focus group, while the remaining six, who represent various platform operators, participated in in-depth interviews. Table 3 summarises the compilations of parameters obtained from the FGDs.

TABLE 3. Parameters involved in remuneration

Parameters	Digital worker Type	Digital worker Competency	Task Type	Task Outcome
Classification	*Individual* Integrative Independent	*Non-expert* Limited skill and knowledge, moderate comprehension level and applications familiarity *Rating scheme:* Basic and novice *Moderate skill* Possess some experiences and the skill are progressing *Rating scheme:* Intermediate and advance	*Simple* Unstructured Structured	Low variant of outcome
	Group Interdependent	*Expert* Possess specific skills and able to do analysis, evaluation and synthesis or creation *Rating scheme:* Expert	*Complex* Unstructured Structured	High variant of outcome

Workers are classified into five levels by platform operators but are aggregated into three as a result of the FGD agreement to widen the range of remuneration. The first level is non-experts, who are rated as basic workers and possess basic ICT skills. While a novice has limited experience and needs more training, the second level is moderate. It comprises intermediates who can do simple ICT tasks and can accomplish them independently, but close supervision may be required. At the same level is an advanced (or self-developer) worker, who progresses consistently with advanced ICT knowledge. They are intrinsically motivated and seek tasks that will make them learn or gain knowledge, as they strive for continuous improvement (Oppenlaender *et al.*, 2020). The third level is an expert or professional who can solve difficulties related to his/her expertise and they are well experienced. The experts complete the tasks full-time and work for long hours in pursuit of maximum productivity and income (Oppenlaender *et al.*, 2020). Apart from simple and complex, each entry in the table is a form of structured or unstructured action done to complete the task.

Interviews with respondents were conducted in five cycles. In the interview, respondents need to specify the type of task, level of competency needed, and variant of outcome in connection to the job for each scenario given. Additionally, respondents explained how they used star ratings to aggregate DW remuneration. Level of DW competency is transcribed into basic = 1, novice = 2, intermediate = 3, advance = 4, and expert = 5. While the DW type is denoted nominally as either an individual worker or a group of workers. Table 4 shows an example of transcribed data. Other than that, pertinent data about the level of commitment and the type of payment, which may be a voluntary contribution, a flat fee, or a prize, were also acquired. Additionally, assignment history and ideas on how attitude should be assessed were also shared by the respondents.

TABLE 4. Example of transcribed data

Work Scenario	Tasks S= Simple C = Complex	Out come L = low H = high	DW rate and competency level																								
			Respondet 1					R2					R3					R4					R5				
			Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5
Synthe sis	S	L	1	2				4	5				4	3				5	4				5	5			
	S	H	5	5	4			5	4	4			5	5	4			5	5	3			5	5	5		
	C	L	4	4	5			5	4	3			5	4	3			4	4	5			5	5	4		
	C	H	5	4	4	4		3	3	3	3		5	5	4	5		5	5	4	5		5	5	4	4	
Evalua tion	S	L	4	5				2	3				4	5				5	2				3	4			
	S	H	4	3	5			4	2	5			5	3	5			5	2	5			5	4	5		
	C	L	5	5	5	5		5	3	5	4		5	3	5	5		5	4	5	5		5	4	5	5	
	C	H	5	4	5	5	4	5	3	4	5	4	5	5	4	5	5	5	5	5	5	4	5	5	5	4	5
Application	S	L	2	2				1	2				1	1				1	2				1	1			
	S	H	2	3	3			4	4	3			4	4	3			2	2	3			4	4	3		
	C	L	3	3	4	4		4	4	3	5		3	3	3	3		4	4	3	3		4	4	4	4	
	C	H	5	5	4	4	4	5	5	5	5	5	5	5	4	4	5	5	4	3	3	4	5	5	5	5	5
Analysis	S	L	1	1				2	2				2	1				1	1				2	1			
	S	H	1	1	2			2	1	1			1	2	2			3	3	2			2	2	1		
	C	L	4	4	5	5		5	4	5	5		3	4	5	5		4	3	4	4		5	5	5	5	
	C	H	5	5	5	5	4	5	4	5	5	5	5	5	5	4	4	5	4	5	4	4	5	5	5	5	5
Compre hension	S	L	1	2				1	1				2	2				1	2				1	2			
	S	H	2	3	3			3	3	3			2	2	3			2	2	3			2	2	2		
	C	L	3	4	4	4		4	4	3	3		3	3	4	3		3	3	3	3		4	4	4	4	
	C	H	3	4	4	4	4	5	5	4	4	4	5	5	5	5	4	4	3	3	3	4	5	5	5	5	5
Knowled ge	S	L	1	2				2	1				2	3				3	4				1	2			
	S	H	2	2	1			1	1	2			2	1	3			4	3	4			1	1	1		
	C	L	3	4	4	3		3	3	3	3		4	5	5	4		4	3	2	4		5	5	5	5	5
	C	High	5	4	4	4	3	4	4	4	4	3	5	5	4	4	4	4	3	4	3	4	5	5	5	5	5

ANALYSIS OF PARAMETERS AND RELATIONSHIP ESTABLISHMENT

The data analysis process began with exercises aimed at removing redundancy. Appriori algorithm was used, which calculated the similarity of the most often occurring items, counted them, and deleted them. The works are summarised in Figure 3.

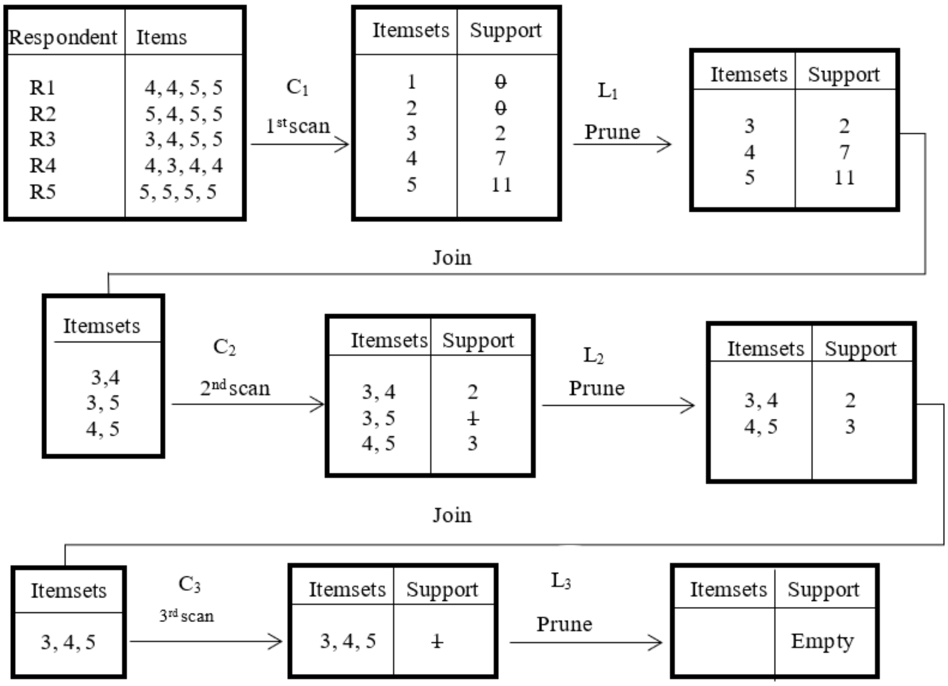

FIGURE 3. Steps in apriori algorithm

In the diagram, C_n indicates the list of candidate items to be counted. While "Support" counts up the number of occurrences of each item, for example (C_1: {1} – 0, {2} – 0, {3} -2, {4} – 7, {5} – 11). It iterates within two steps of joining and pruning. Join generates all feasible C_k candidates, whereas prune removes those that are superfluous from C_k. It ran iteratively until redundant data had been removed and only significant patterns of parameters were counted in the end. Table 5 shows examples of the 24 pattern statements that were obtained.

TABLE 5. Examples of pattern obtained using Apriori algorithm

Dimension			
1 Tasks	2 Outcome Variation	3 Competency Level	DW competency level
Complex	Low	Synthesis	3 => 4, 4 => 5, 3, 4 => 5
Complex	High	Synthesis	4 => 5
Complex	Low	Evaluation	4 => 5, 3 => 5, 3, 4 => 5
Complex	High	Evaluation	4 => 5
Complex	Low	Application	3 => 4
Complex	High	Application	4 => 5
Complex	Low	Analysis	3 => 4
Complex	High	Analysis	4 => 5
Complex	Low	Knowledge	3 => 4
Complex	High	Knowledge	3 => 4
Complex	Low	Comprehension	3 => 4
Complex	High	Comprehension	3 => 4
Simple	Low	Synthesis	4 => 5
Simple	High	Synthesis	4 => 5
Simple	Low	Comprehension Evaluation	4 => 5
Simple	High	Comprehension Evaluation	2 => 5, 3 => 5, 4 => 5, 3, 4 => 5

Simple	Low	Application	1 => 2
Simple	High	Application	4 => 5, 3 => 4
Simple	Low	Analysis	1 => 2
Simple	High	Analysis	1 => 2
Simple	Low	Knowledge	1 => 2
Simple	High	Knowledge	1 => 2
Simple	Low	Comprehension	1 => 2
Simple	High	Comprehension	2 => 3

Following that, an association rule was used to identify the relationship between each of the pattern statements' itemset. The association rules operate by calculating and comparing each rule's confidence value. The weakest rules were removed. For example, for statement "A, B", means "what is in A is also in B." A fragment of the pseudocode to construct the relationship is presented in Figure 4. In this example, {3 ⇒ 4} has been set as Rule 1 and {4 ⇒ 5} as Rule 2.

Example: Frequent itemset {3, 4} {4, 5}

Rule 1: {3 ⇒ 4}
Confidence = support{3} ∪ support{4}/support{3} * 100
= (2/2) * 100 = 100%
Rule 1 is selected.

Rule 2: {4 ⇒ 5}
Confidence = support{4} ∪ support{5}/support{4} * 100
= (3/7) * 100 = 42.8571%
Rule 2 is rejected.

If the **min_conf = 50%** then only the first rule will be generated.

FIGURE 4. A fragment of pseudocode for an association rule

Iterative simulations were run to assess and confirm the confidence threshold value. Based on simulation and expert validation, the minimum confidence threshold was 50%. As a consequence, rules with values less than or equal to the threshold values were removed. The association rule was repeated until all relationships attained the abstraction level. Thus, in the preceding example, any rules that were less than 50% were considered weak and eliminated. Table 6 illustrates the established relationship. It can be seen in the table that relationship number 1 shows that for complex tasks, DW at intermediate, advanced, and expert levels will be assigned to do the tasks.

TABLE 6. Example of relationships established using association rule

Relationship No	Dimension			
	1 Type of tasks	2 Outcome variation	3 Competency level	DW category
1	Complex	Low	Synthesis	3, 4 and 5
2	Complex	High	Synthesis	4 and 5
3	Simple	Low	Synthesis	4 and 5
4	Simple	High	Synthesis	4 and 5
5	Complex	Low	Evaluation	3, 4 and 5
6	Complex	High	Evaluation	4 and 5
7	Simple	Low	Evaluation	4 and 5
8	Simple	High	Evaluation	4 and 5
9	Complex	Low	Application	3 and 4

10	Complex	High	Application	4 and 5
11	Simple	Low	Application	1 and 2
12	Simple	High	Application	3 and 4
13	Complex	Low	Analysis	3 and 4
14	Complex	High	Analysis	4 and 5
15	Simple	Low	Analysis	1 and 2
16	Simple	High	Analysis	1 and 2
17	Complex	Low	Knowledge	3 and 4
18	Complex	High	Knowledge	4 and 4
19	Simple	Low	Knowledge	1 and 2
20	Simple	High	Knowledge	1 and 2
21	Complex	Low	Comprehension	3 and 4
22	Complex	High	Comprehension	3 and 4
23	Simple	Low	Comprehension	1 and 2
24	Simple	High	Comprehension	1 and 2

VISUALIZATION

A visual has been developed to aid comprehension of the relationship. The alluvial charts were finally created and will be discussed in the next section. To validate the understanding of the created visuals, a series of expert reviews and critiques were implemented. Experts are distinct from individuals who participated in data collection. They are local crowdsourcing platform workers from Crowdsourcing SIG researchers, Crowdsourcing Malaysia Association members, and MDEC's point person. Figure 5 presents the flow of relationships represented by alluvial charts.

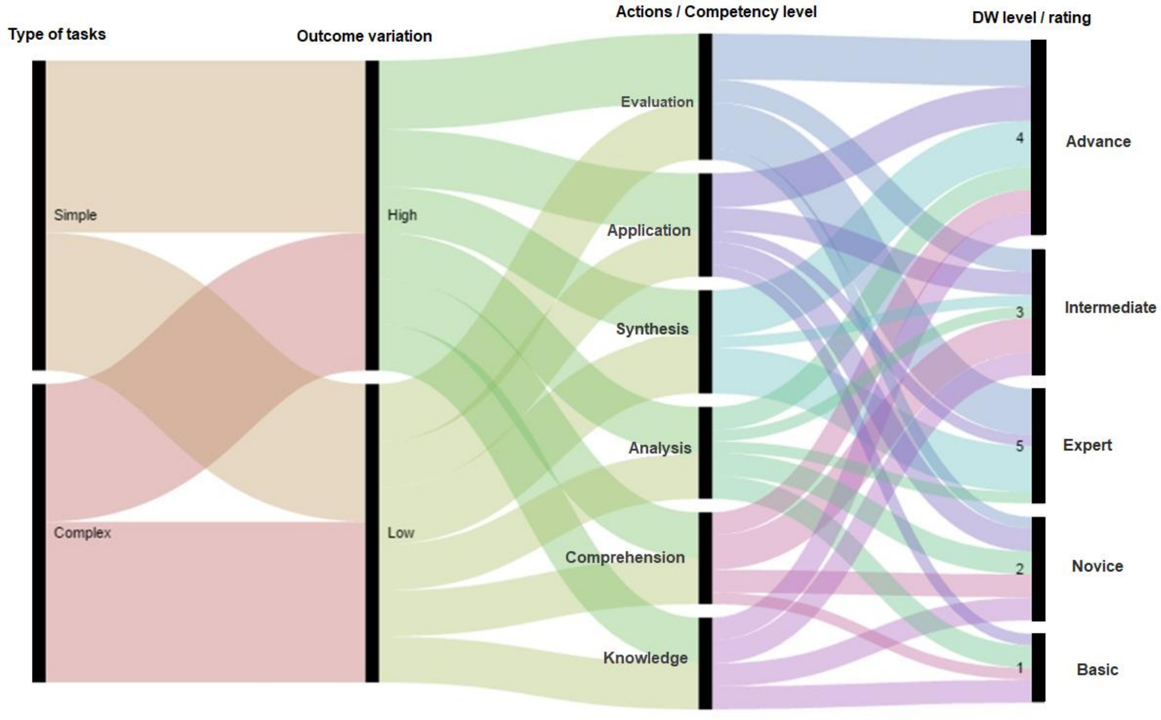

FIGURE 5. Relationships represented by Alluvial Charts

FINDINGS AND DISCUSSION

The flow graph in the alluvial chart comprises four vertical axes. The axes represent the type of task, outcome variation, competency level, and DW rate. The size of a data cluster is represented by the height of the axes, on which the strongest consensus is gained from respondents. The same goes for the thickness of the stream, which also represents the size of the consensus contained. The chart excludes DW types because all workers are viewed as individuals in the platform’s perspectives. Some occupations may require a team of DW with specific skills, knowledge, and experience, but group work is the result of an individual's integration (Samini, 2017). The first axis "type of task" shows a similar sized streams crossed each other to connect into the second axis. This suggests that each task has the potential to produce high or low variance of outcome. The second, which is S "outcome variation" axis shows the stream spread to all action types (competency level). As can be observed, the heights of the blocks on both axes are nearly same, indicating that the types of work outsourced by requesters are distributed almost evenly.

The third axis denotes the six activities or competencies necessary to properly complete the job scenario. Each skill level is similar in height, indicating that all respondents (platform operators) have similar comprehension. The stream from the axis spreads and crosses in various ways, indicating that platform operators delegate the tasks in a variety of different ways, but yet the most difficult tasks that require evaluation and synthesising ability are given to experts. Tasks that require knowledge are given to basic, novice, and intermediate. On the other hand, tasks requiring comprehension are spread between beginners, novices, and experts. It is shown that simple tasks with high outcome variation that fall under "application" spread to all five levels of DW. Complex analysis tasks with low and high outcome variation are distributed across all levels as well. After checking, it was noticed that tasks that were given to basic and novices were related to their areas of expertise, like accounting and bookkeeping, but were considered lack of ICT knowledge.

The fourth axis specifies the level and ratings, which are related to the remuneration scale range. It is suggested that the remuneration and incentives must be fixed within a particular range. Repetitive tasks, such as validating links and surfing websites, pay the least as it takes the shortest amount of time to complete one cycle of task. Difficult work that has a significant degree of variance in outcome and requires more time and dedication should receive the highest compensation. It is found that the vertical axes for advanced types of DW are the highest in height, implying that this type of DW is the largest group and has been offered most tasks. Platform operators justified this by stating that they are fresh graduates who are serious about increasing their earnings and developing their talents. Since advance workers are most likely to participate in the crowdsourcing, thus the remuneration scale is suggested to be wider compared to the others. It can also be seen that intermediate and expert are almost at par. Consensus from the validation session agreed that the most important criteria in selecting DW and calculating remuneration is the competency level. In addition, star ratings should cover willingness and overall attitude. Even if the transaction is simply virtual, one's passion and attitude can be detected during the communication and transaction process, thus, remuneration can be adjusted based on both expertise and commitments. The interrelation of each parameter can form a cubic. Each surface represents a single parameter and is split into multiple segments. The cubic will be defective if one of the six surfaces of type of task, outcome variant, competency level, competency star value rating, type of DW, and remuneration is absent. It is visualised in Figure 6.

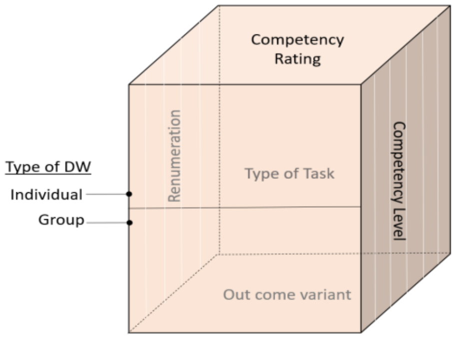

FIGURE 6. The cubic

CONCLUSION

This research discovered four important parameters affecting remuneration determination and visualised their relationship. Data have been gathered through FGDs and interviews with crowdsourcing agencies and platform operators. Significant patterns were extracted using apriori algorithms, and relationships between parameters were revealed using an association rule. The alluvial chart, which depicts the visual, explains the data story. According to the findings, five levels of DW determined an important part of remuneration. This outcome provides recommendations to industry participants, such as members of the Malaysia Crowdsourcing Innovation Association on best practises, classification criteria, and how to specify their remuneration budget. The findings of this study enable us in designing a remuneration strategy for future work.

REFERENCES

Aldahari, E., Shandilya, V., & Shiva, S. (2018). Crowdsourcing Multi-Objective Recommendation System*. *Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*. https://doi.org/10.1145/3184558.3191579

Bhattacharyya, M., & Mridha, S. K. (2018, January 11). Studying the influence of requesters in posted-price crowdsourcing. *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*. https://doi.org/10.1145/3152494.3152495

Buettner, R. (2015, January). A Systematic Literature Review of Crowdsourcing Research from a Human Resource Management Perspective. *2015 48th Hawaii International Conference on System Sciences*. https://doi.org/10.1109/HICSS.2015.549

Chen, X., Lin, Q., & Zhou, D. (2015). Statistical decision making for optimal budget allocation in crowd labeling. *Journal of Machine Learning Research*, *16*. https://doi.org/10.2139/ssrn.2408163

Corney, J. R., Torres-Sánchez, C., Jagadeesan, A. P., & Regli, W. C. (2009). Outsourcing labour to the cloud. *International Journal of Innovation and Sustainable Development*, *4*(4). https://doi.org/10.1504/IJISD.2009.033083

Feyisetan, O., & Simperl, E. (2017). Social Incentives in Paid Collaborative Crowdsourcing. *ACM Transactions on Intelligent Systems and Technology*, *8*(6). https://doi.org/10.1145/3078852

Goncalves, J., Feldman, M., Hu, S., Kostakos, V., & Bernstein, A. (2017). Task Routing and Assignment in Crowdsourcing based on Cognitive Abilities. *Proceedings of the 26th International Conference on World Wide Web Companion - WWW '17 Companion*. https://doi.org/10.1145/3041021.3055128

Haralabopoulos, G., Wagner, C., McAuley, D., & Anagnostopoulos, I. (2019). Paid Crowdsourcing, Low Income Contributors, and Subjectivity. *IFIP Advances in Information and Communication Technology*, *560*. https://doi.org/10.1007/978-3-030-19909-8_20

Hosseini, M., Phalp, K., Taylor, J., & Ali, R. (2014, May). The four pillars of crowdsourcing: A reference model. *2014 IEEE Eighth International Conference on Research Challenges in Information Science (RCIS)*. https://doi.org/10.1109/RCIS.2014.6861072

Hu, Z., & Zhang, J. (2017). Optimal posted-price mechanism in microtask crowdsourcing. *IJCAI International Joint Conference on Artificial Intelligence*, *0*. https://doi.org/10.24963/ijcai.2017/33

John, I., & Bhatnagar, S. (2019, January 3). Efficient Budget Allocation and Task Assignment in Crowdsourcing. *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*. https://doi.org/10.1145/3297001.3297050

Kietzmann, J. H. (2017). Crowdsourcing: A revised definition and introduction to new research. *Business Horizons*, *60*(2). https://doi.org/10.1016/j.bushor.2016.10.001

Luo, Y., & Jennings, N. R. (2021). A budget-limited mechanism for category-aware crowdsourcing of multiple-choice tasks. *Artificial Intelligence*, *299*. https://doi.org/10.1016/j.artint.2021.103538

Mizusawa, K., Tajima, K., Matsubara, M., Amagasa, T., & Morishima, A. (2018). Efficient pipeline processing of crowdsourcing workflows. *International Conference on Information and Knowledge Management, Proceedings*. https://doi.org/10.1145/3269206.3269292

Mridha, S. K., & Bhattacharyya, M. (2019). Introducing Collaboration in Competitive Crowdsourcing Markets. *IEEE Intelligent Systems*, *34*(1). https://doi.org/10.1109/MIS.2019.2898184

Nakatsu, R. T., Grossman, E. B., & Iacovou, C. L. (2014). A taxonomy of crowdsourcing based on task complexity. *Journal of Information Science*, *40*(6). https://doi.org/10.1177/0165551514550140

Oppenlaender, J., Milland, K., Visuri, A., Ipeirotis, P., & Hosio, S. (2020, April 21). Creativity on Paid Crowdsourcing Platforms. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3313831.3376677

Pilloni, V. (2018). How Data Will Transform Industrial Processes: Crowdsensing, Crowdsourcing and Big Data as Pillars of Industry 4.0. *Future Internet*, *10*(3). https://doi.org/10.3390/fi10030024

Qiao, Y., Wu, J., Zhang, L., & Wang, C. (2019, June 24). Mechanism design for cross-market task crowdsourcing. *Proceedings of the International Symposium on Quality of Service*. https://doi.org/10.1145/3326285.3329044

Qiu, C., Squicciarini, A., & Hanrahan, B. (2019). Incentivizing distributive fairness for crowdsourcing workers. *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, *1*.

Qiu, C., Squicciarini, A., Khare, D. R., Carminati, B., & Caverlee, J. (2018). CrowdEval: A cost-efficient strategy to evaluate crowdsourced worker's reliability. *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, *2*.

Qiu, C., Squicciarini, A., & Rajtmajer, S. (2019, November 3). Rating Mechanisms for Sustainability of Crowdsourcing Platforms. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. https://doi.org/10.1145/3357384.3357933

Rouse, A. C. (2010). A Preliminary taxonomy of crowdsourcing. *ACIS 2010 Proceedings - 21st Australasian Conference on Information Systems*.

Samimi, P., Ravana, S. D., Webber, W., & Koh, Y. S. (2017). Effects of objective and subjective competence on the reliability of crowdsourced relevance judgments. *Information Research*, *22*(1).

Schenk, E., & Guittard, C. (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics*, *7*(1). https://doi.org/10.3917/jie.007.0093

Thuan, N. H., Antunes, P., & Johnstone, D. (2013). *Factors Influencing the Decision to Crowdsource*. https://doi.org/10.1007/978-3-642-41347-6_9

Wang, G., & Yu, L. (2020). Analysis of enterprise sustainable crowdsourcing incentive mechanism based on principal-agent model. *Sustainability (Switzerland)*, *12*(8). https://doi.org/10.3390/SU12083238

World Economic Forum. (2020). The Future of Jobs Report 2020 | World Economic Forum. *The Future of Jobs Report, October*.

Xu, J., Wang, S., Zhang, N., Yang, F., & Shen, X. (2019). Reward or penalty: Aligning incentives of stakeholders in crowdsourcing. *IEEE Transactions on Mobile Computing*, *18*(4). https://doi.org/10.1109/TMC.2018.2847350

Ye, H., & Kankanhalli, A. (2013). Leveraging crowdsourcing for organizational value co-creation. *Communications of the Association for Information Systems*, *33*(1). https://doi.org/10.17705/1cais.03313

Yu, R., Cao, J., Liu, R., Gao, W., Wang, X., & Liang, J. (2019). Participant incentive mechanism toward quality-oriented sensing: Understanding and application. *ACM Transactions on Sensor Networks*, *15*(2). https://doi.org/10.1145/3303703

Zakariah, Z., Janom, N., & Arshad, N. H. (2016). Crowdsourcing business model for low income community: Conceptual model. *Proceedings of 2016 SAI Computing Conference, SAI 2016*. https://doi.org/10.1109/SAI.2016.7556126

Zhao, B., Wang, Y., Li, Y., Gao, Y., & Tong, X. (2019). Task allocation model based on worker friend relationship for mobile crowdsourcing. *Sensors (Switzerland)*, *19*(4). https://doi.org/10.3390/s19040921

Siti Salwa Salleh[1,3]
[1]Fakulti Sains Komputer dan Matematik, Universiti Teknologi MARA,
Cawangan Negeri Sembilan, Kampus Seremban, Negeri Sembilan
ssalwa@uitm.edu.my

Nurhayati Zakaria[2]
Norjansalika Janom[2]
Syaripah Ruzaini Syed Aris[2]
Noor Habibah Arshad[2]
[2]Fakulti Sains Komputer dan Matematik, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia
[3]Institut Pengangkutan Malaysia (MITRANS) , Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia
norjan@fskm.uitm.edu.my, ruzaini@fskm.uitm.edu.my, habibah@fskm.uitm.edu.my