A Profile-Aware Microtasking Approach for Improving Task Assignment in Crowdsourcing Services

Jabu MTSWENI1, Ernest Ketcha NGASSAM2, Legand BURGE III3

1Defence, Peace, Safety, and Security (DPSS), CSIR, Pretoria, South Africa
Tel: +27128414319, Email: mtswenij@gmail.com

2University of South Africa, Florida, Johannesburg, South Africa
Tel: +27116709126, Email: eketcha@gmail.com

3Howard University, Washington, DC, United States of America
Tel: +2028064852, Email: blegand@scs.howard.edu

Abstract: Microtasking as one of the crowdsourcing models has found penetration mostly in developed nations and is slowly making inroads in developing nations. It has been used for quickly solving a myriad of social and business challenges by tapping into the diversity of the unknown crowd. Nevertheless, there are still a number of research challenges that are found in various microtasking initiatives, such as lack of user profiling, quality of task submissions, poor task design and assignment, ambiguous task assessments, and lack of diverse platforms to cater for the needs of differently skilled crowds. Most prominent microtasking platforms do not consider or even exploit the profile of the crowd in improving task design, assignment, and eventually quality of completed microtasks. In this paper, the objective was to explore the design and evaluation of a conceptual approach that could be implemented in microtasking environments for purposes of improving task assignments and eventually quality. A design science research methodology was adopted to build the proposed artefact, which was evaluated for relevance using comparative analysis and crowdsourcing metrics. The results suggest that exploiting micro workers’ profiles could improve task design, assignment, assessment, and ultimately quality. However, stakeholders’ awareness of effective microtasking approaches still requires research attention.

Keywords: crowdsourcing, microwork, microtasks, microtasking, developing nations, mobile services, task assignment

1. Introduction

The notion of microtasking as one of the crowdsourcing model that is widely adopted in developed digital economies for tapping into human intelligence or micro-workers in solving specific digital tasks (e.g. ad-hoc information gathering, tagging, and classification) that cannot be easily completed by computers [1][2]. A few micro-work platforms are also emerging in developing countries [3], particularly for minimizing rising unemployment in disadvantaged communities [4].

In essence, microtasking promotes the completion of digital tasks by dividing a complex task (e.g. translation of a book) into simple and smaller tasks that can be completed by diverse skilled micro-workers located across various locations. The small tasks, also referred to as microtasks could be generated from a large task owned by a requester or micro-employer [5].

The apparent benefits of microtasking are that organisations can tap into the diversity of micro-workers who have the ability to perform posted tasks quickly and cheaper compared
to internal work-force [6][7]. Evidence of the popularity of microtasking can be noted in platforms such as Mechanical Turk [8][9], where over 200,000 human intelligent tasks (HITs) in different categories, scale, and from different parts of the world are usually available for completion by micro-workers for a small fee. Additional benefits that are commonly cited for microtasking include: innovation and entrepreneurship stimulation, equitable access to employment opportunities, skills and work-profile development, diverse participation leading to diverse contributions, increased productivity and value co-creation [4], [10] [11]. Despite the various opportunities presented by microtasking, a number of limitations are still evident in various microtasking platforms. Some of the challenges include quality control of contributions by the diverse and unknown crowd of micro-workers, micro-worker profiling, efficient task design and assignment, transparent and fair task assessment, seamless and contextual payment of tasks completed, sustained participation, ethical and security considerations [4]. Other critical research issues stem from the fact that current microtasking solutions mostly do not exploit the expertise of micro-workers for task assignment and eventually improving task quality [12].

Thus, the main objective of this paper was to explore the design and evaluation of a profile-aware microtasking approach that could be implemented in diverse microwork solutions for improving task assignment.

The next section (Section 2) discusses the key background concepts on crowdsourcing and microtasking. Thereafter, research methodology adopted for the research presented in this paper is discussed (Section 3). This is followed by the design and development of a profile-aware microtasking framework that could be implemented in an effort to improve task assignment and quality (Section 4). The proposed artefact as guided by the design science research methodology is evaluated using different approaches in Section 5. The paper is concluded with a summary of contributions in Section 6.

2. State-of-the-Art: Crowdsourcing and Microtasking

There are different interpretations in different domains about what is meant by crowdsourcing [7][13]. Although the origin of the term is often credited to Jeff Howe [14], the concept of crowdsourcing has long existed prior to 2006, albeit in various modes (e.g. Wikipedia). For purposes of this paper, the definition of Jeff Howe will be used as reference [14]. In brief, Howe [14] defines crowdsourcing as an act by an organisation or individual of taking a function that could be performed by targeted resources (e.g. employees) and outsourcing it for solutions through an open call to a diverse and unknown group of people referred to as the crowd [14]. In essence, crowdsourcing is a distributed problem-solving business model focusing on participative online activities [15][7][16]. It has found use in various business domains, including governments, global corporations, non-profit organizations, schools, and social environments. As discussed in [3], various implementations of crowdsourcing exists ranging from crowdfunding to crowd voting. Most forms adopt a voluntary-based approach, where users participate without financial benefits [17]. The crowdsourcing process as stated in [7], could be linear or iterative depending on the scope of the tasks.

Most crowdsourcing implementations can be realized using any of the four models [18], that is: (1) peer production, (2) investments (3) competitions, and (4) microtasking. The peer production model sources contributions for a project or solution from a large network of knowledgeable individuals. Typical examples that adopt this model of crowdsourcing are open source projects (e.g. Linux OS), online collaborative projects such as Wikipedia and Stack Overflow. Users generally contribute to such initiatives without any financial reward, and decide on the scope of their contributions.

The investments model involves the idea or function initiator working in collaboration with external parties or sponsors to achieve their goal. The initiator might have an idea that
sponsors are interested in funding. At the successful completion of the project, the collaborator or sponsor might get the product or solution for free. This type of a model is common in crowdfunding initiatives.

The *competitions model* is another way that crowdsourcing could be implemented. This model has seen wide adoption in businesses, particularly in software development and designs of innovative products or solutions. In this instance, an organisation would either closely or openly advertise a function that would need to be completed by competing contestants with the best solution winning large sums of money and/or prizes. The focus in this paper is *microtasking*, which is another commonly applied crowdsourcing model. In this model, large tasks are decomposed into a “set of self-contained microtasks that could be completed by a diverse group of individuals for a small fee or incentive” [18]. It is sometimes referred to as paid crowdsourcing [17], and “users contribute at the time scale of minutes” [6].

As with the other different crowdsourcing models, microtasking is grounded on four pillars of the reference model as discussed in [16] and depicted in Figure 1.

![Figure 1: Pillars of Crowdsourcing – Reference Model](image)

The *crowd* is the foundation of all microtasking platforms as they form part of the contributors of solutions sourced by decomposed tasks. In microtasking, the crowd is also referred to as a micro-worker. According to [16], the *crowd* is characterized by: diversity, unknown-ness, largeness, undefined-ness, and suitability. The profile-aware proposition stems from this pillar, which is often neglected by various crowdsourcing platforms. In [7], it is noted that in most cases the crowd participation in completing the tasks is voluntary, but may be incentivized.

The *crowdsourcer* or micro-employer is generally an individual or organisation who commences the crowdsourcing process [7]. Basically, it is the one who puts out an open call through some platform for completion of tasks by the crowd. In microtasking terms, the crowdsourcer is called a requester. According to [16], the *crowdsourcer* is typified by task design, incentives provision, ethicality provision, open call, and privacy provision. The *crowdsourcer* plays a integral part in the crowdsourcing process, particularly in task design and evaluation.

The *crowdsourcing task* is the actual task that needs a solution from the *crowd*. The scope of the crowdsourcing task differs depending on the implementation model. In microtasking, tasks are generally of minute scale; hence labelled microtasks. A *crowdsourcing task* would be modular, varying degrees of complexity, solvable, user-driven, and amenable to individual or collaborative contributions [16]. In this paper, we specifically focused on task assignment.

The final pillar in the reference model is the *crowdsourcing platform*. This is the actual environment where the task could be advertised, completed, assessed and/or incentivized. This could be an online or physical platform. In [16], four distinct features of a *crowdsourcing platform* are highlighted: support crowd-related interactions, crowdsourcing-related interactions, provides task-related facilities, and provides platform-related facilities, such as managing platform misuse and providing a payment mechanism. In the microtasking model, the platform is commonly called microtasking platform. A number of
common microtasking platforms are summarized in [3], [4]. For purposes of the study presented in this paper, we focused only on the platforms that are purported for developing communities, such as Money for Jam (M4JAM) [3], MobileWorks [19], and mClerk [20].

3. Research Methodology

A design science research (DSR) methodology [21][22] was adopted for this study, mainly because it provides a systematic approach for designing, developing, and evaluating research artefacts, such as a reference framework proposed in Section 4. The methodology is also widely used in information systems research for the creation of successful artefacts. The methodology suggests a step-by-step approach [23] starting with (1) the identification and understanding of the research problem, (2) defining the objectives of the proposition, (3) design and development of an artefact, (4) demonstrating the proposed artefact within a relevant context, (5) evaluating the solution, and finally (6) communicating the results to the scientific community. For purposes of this paper, step 1 – 2 are already addressed in the previous sections, and in Section 4, step 3 will be addressed. In section 5, step 4-5 will be covered.

Since various research methods can be exploited to address the steps alluded above, for this paper, a conceptual modelling [24] method was opted to build the profile-aware microtasking approach. The requirements were determined based on the systematic literature reviews and the research challenges reported in [3], [4]. A conceptual model was considered relevant for the identified problem since it can easily represent concepts, including relationships between them, and is commonly used for designing valuable information system artefacts. In evaluating the proposed artefact, qualitative comparative analysis (QCA) [25] and crowdsourcing metrics [26], chiefly in relation to task assignment, were used. These methods provided a good foundation for widely evaluating the proposed approach.

For comparative analysis, the microtasking platforms noted in Section 2 and briefly discussed in Section 5 were evaluated using the proposed profile-aware microtasking artefact as elaborated in Section 4. Finally, in order to measure if the proposed approach is of any significance, it was evaluated against the crowdsourcing metrics proposed in [7].

4. A Profile-aware Microtasking Approach

Since microtasking follows an open call approach, is time-driven, and the crowd is generally unknown to the crowdsourcer or micro-employer [16], the design, assignment, and assessment of tasks become critical for motivation, correct task match, and sustained participation. However, since the crowd is assumed to be diverse in size, location, age, and expertise, it is impractical to expect the crowdsourcer to manually assign tasks to a diverse crowd and manually evaluate all submissions by the crowd for the desired solution. Thus, task design is essential, however, without understanding the profile of the potential crowd, incorrectly designed tasks could solicit incorrectly completed tasks. For example, a crowdsourcing platform such as CrowdFlower [12][15] enables the crowdsourcer to design and simultaneously distribute one task to multiple platforms used by a diverse group of individuals and organisations. As may be noted, such an approach cannot adopt a rigid task design for the different platforms targeted. For instance, a task may be completed by an intelligent crowd in one platform, distributed to a well-trained crowd in another platform, or contributions might also be received from a lowly skilled crowd using microtasking platforms, such as MobileWorks [27]. This diversity calls for profile-aware task design, assignment, and eventually assessment leading to improved quality, organisational performance, and crowd satisfaction [17]. Profile-aware refers to the diverse profile (e.g. language, qualification, reputation, skills, task history, etc.) of the crowd or micro-workers.
of a specific platform that need to be taken into considerations when designing and assigning microtasks.

Relevant approaches such as the cross-platform Curriculum Vitae (CV) are suggested in [12] to deal with the issue of profiling users for task assignment. However, balancing the protection of the crowdsourcer’s intellectually property (IP) (i.e. during task design) and crowd’s privacy (during task assignment and assessment) are some of the issues that need attention in existing profile-aware microtasking solutions.

Based on a systematic literature review [7], [12], [15], [17], [28], [29], the following high-level requirements (see Table I) were identified as important for a profile-aware microtasking approach. These are categorized based on the reference model [16]. The proposed artefact is shown in Figure 2.

<table>
<thead>
<tr>
<th>Table I: High-level Requirements for a Profile-Aware Microtasking Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Micro-Worker Profile</td>
</tr>
<tr>
<td>Location</td>
</tr>
<tr>
<td>Language</td>
</tr>
<tr>
<td>Experience</td>
</tr>
<tr>
<td>Core Skills</td>
</tr>
<tr>
<td>Track record</td>
</tr>
<tr>
<td>Self-assessments</td>
</tr>
<tr>
<td>Machine-processable</td>
</tr>
<tr>
<td>Anonymity</td>
</tr>
</tbody>
</table>

In order to improve task design, assignment, assessment, and ultimately quality, the micro-worker need to at least have a profile that exhibits the features shown in Table I. It is important that the location and language of the micro-worker is known since some microtasks are targeted at specific audiences. Secondly, the experience and core competencies of the micro-worker should be transparent for purposes of improving task assessment. The track record (e.g. performance history, earned rewards, etc.) and self-assessments are a good indicator on what tasks should be assigned or not to a specific crowd.

Because crowdsourcing attracts large crowds, it is important that the profile of the micro-worker is machine-processable for automation purposes. As maybe be noted, the micro-worker profile need not to include sensitive personal information, such as names, emails, or identification numbers. These are not necessary for task design and assignment, thus it is key that the anonymity of the crowd is also preserved to sustain trust and participation.
Task design plays a vital role in receiving useful submissions from the crowd, as such in moving towards a profile-aware microtasking approach, a microtask needs to be simple, modular (decomposable) so that various contributions can be easily solicited, specific (provide enough details), solvable, measurable, and contextual in order to match targeted crowd as per available profiles.

The platform plays an important in ensuring that the designed microtask can be assigned to the relevant crowd. As such, in addition to the requirements noted in [16], the platform need to be transparent on the target audience and include task-related support, such as task design, assignment, and assessment. The platform should at least provide diverse interactions to both the crowd and the crowdsourcer. For instance, the crowd need to be able to create and monitor their profile over time, and the crowdsourcer (micro-employer) should be able to also monitor the effectiveness of their tasks over time. In addition, the micro-employer should at least provide microtasks that are designed according to the target crowd, and considering ethical issues (e.g. is the task in line with the general ethical norms?). Privacy and security are important considerations when developing profile-aware tasks. In order to balance the need for transparent crowd profiles, it is also important for the micro-employer to adhere to high standards of ethics to ensure that the privacy and anonymity of the crowd profile is preserved at all times. This could be ensured that personal information such as names and email addresses are not shared with the micro-employer by the intermediary crowdsourcing platform, and these are anonymised when making the crowd profile transparent to the micro-employer. By design, in crowdsourcing, the crowd is generally not known to the crowdsourcer, and at the same time the micro-employer is not overly exposed to the crowd. Thus, the profiling process should focus on the expertise and context rather than on personal information. This is important to prevent unfair collusions and negative business practices. The focus after all is mostly on the solutions provided by the crowd than the personal details of the crowd.

Lastly, feedback is seen as essential towards sustaining positive participation in microtasking platforms, thus platforms should, where possible, provide automated feedback to micro-workers via micro-employers.
5. Evaluation

In determining the relevance of the proposed approach in Section 4, we opted for a comparative analysis using existing microtasking platforms. These are briefly explained below:

- **Money for Jam (M4JAM):** it is a South African microtasking initiative that is deployed via the WeChat platform, and can be accessed using smartphones. Its target market is mainly what they called jobbers and follows a micro-jobbing model, which in essence is microtasking. It boasted of a 300% growth in 2015 and payments in excess of R854 000 [30]. The platform provides a variety of tasks, such as taking a photo of products at specific shops; registered workers have limited time to complete the advertised tasks. Incentives range from R15 to R45 [31].

- **mClerk:** it is a mobile-based microtasking platform targeting low-income users. It is SMS-based, although it is also capable of sending images via SMS to enable design of graphical tasks [20]. Most of the tasks catered for by mClerk are simple, such as digitizing local-language documents, and workers are paid using mobile. The tasks are distributed to micro-workers.

- **MobileWorks:** this is also a mobile phone-based crowdsourcing platform intended to provide employment to micro-workers in developing regions [19]. Human optical recognition tasks are dominant and mostly are completed via a mobile web-application. The platform has seen wide adoption in India.

5.1 Qualitative Comparative Analysis

Referring to the proposed profile-aware microtasking approach, Table II shows the comparative analysis between these platforms. The focus of the evaluation for this paper was only limited to the profile-aware elements. As may be noted, the majority of the platforms can be considered as limited in terms of being profile-aware. For instance, M4JAM does not consider the profile of the micro-worker as recommended in Section 4.1. Micro-workers only provide their phone numbers, and location, language, qualification, and experience information is not captured by the platform during registration. The platform relies heavily on the good-will of the micro-worker for task quality. Micro-workers anonymity is also not guaranteed in M4JAM since their terms and conditions clearly state that users’ information can be shared with third-parties. We were unable to identity how the other two platforms manage privacy and anonymity aspects.

<table>
<thead>
<tr>
<th>Elements</th>
<th>Platforms</th>
<th>M4JAM</th>
<th>mClerk</th>
<th>MobileWorks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locality</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Automation</td>
<td></td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>History</td>
<td></td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Core Competencies</td>
<td></td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

Keys:

- ✓: fully considered
- ×: not considered
- ✓: partially considered

As may be noted in Table II, the core skills and qualifications of the micro-worker are not considered when assigning tasks by all the evaluated platforms. In fact, most tasks in...
these platforms can be completed by anyone irrespective of their skills or qualifications. The track record and/or performance of micro-workers are not considered when a micro-worker opts to participate in a task by M4JAM and mClerk. MobileWorks tracks the performance history, since a participating worker may be deactivated from the platform due to continuous low tasks’ quality.

Although, most microtasking platforms are based on open calls, it is our position that not all tasks within a specific platform can be completed by all users of that platform, especially in developing regions where skill levels vary greatly. Thus, a profile-aware approach is vital in improving task design, assignment, and assessment.

5.2 Crowdsourcing Metrics

Using the crowd membership metrics in [7], the proposed profile-aware microtasking approach was also evaluated for relevance and significance. Table III below shows the summary of the evaluation.

<table>
<thead>
<tr>
<th>Crowd size</th>
<th>Age</th>
<th>Gender</th>
<th>Nationality/Residency</th>
<th>Skill, Knowledge, expertise</th>
<th>Individual or corporation</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>

From Table III, it is clear that the proposed profile-aware platform does not consider personal information (i.e. Age or Gender) as key to task assignment. Our proposition is that the identity of participants needs to be anonymised to avoid unintended consequences such as collusion between the micro-employer and micro-worker. At the same time, the nationality and residency of micro-workers is vital, since some tasks might be governed by certain IP and copyright laws, and others might be location-bound. The skills, knowledge, and expertise play a critical role in relation to task assignment and completion, and as such this element is part of the proposed profile-aware microtasking solution. With the context of the proposed artefact, the micro-worker is considered to be an individual, thus the possibility of one organisation participating in completing certain tasks is ignored since individuals will ultimately be the ones completing the work. However, the challenge that might emerge in such situations is when micro-workers within one organisation are sharing profiles, which could affect task quality.

Although the presented profile-aware approach has not been technically tested, the theoretical evaluation discussed above suggests that the design and development of profile-aware tasks could potentially improve task design, assignment, and consequently task quality.

6. Conclusion

Microtasking platforms are diverse, and useful for developing nations. They have exhibited the potential of availing employment opportunities to a number of casual digital workers around the world. However, in spite of the successes, these initiatives are still faced with different technical and research challenges. The issues of task design, assignment, assessment, and quality remain partially addressed in many microtasking platforms, and in our view the root cause being the neglect of the user profile. Thus, in this paper we have motivated the design and development of a profile-aware microtasking approach that could be exploited to improve task assignment in multiple platforms, and indirectly improve task quality, productivity, and sustained participation. The suggested approach considered micro-workers’ locality, work history, core skills, and automation of these profiles by the platform as key towards the improvement of task assignment in microtasking platforms.
The presented approach was evaluated using comparative analysis and crowdsourcing metrics. The focus on comparative analysis was on three common microtasking platforms mainly used in developing nations. From the analysis, it was evident that these existing platforms are not profile-ware. However, the proposed approach also did not satisfy all the crowdsourcing metrics for the crowd. Nevertheless, its significance for academia and businesses lies in the improvement of task design, task assignment and quality. This could lead to benefits such as improvement in the adoption of crowdsourcing platforms. Further research would involve the technical implementation of the suggested approach, especially in matured crowdsourcing platforms.

Acknowledgement
This work is based on the research supported in part by the National Research Foundation of South Africa for the Grant, Unique Grant no: 876914. Any opinion, finding and conclusion or recommendation expressed in this material is that of the authors and the NRF does not accept any liability in this regard.

References


