ISSN 2186-7437

NII Shonan Meeting Report

No. 2013-15

Intelligent Information Processing Chances of Crowdsourcing

Wolf-Tilo Balke, TU Braunschweig, Germany
Takahiro Hara, University of Osaka, Japan
Seung-won Hwang, POSTECH University, South Korea
Christoph Lofi, NII Tokyo, Japan

November 18 - 21 , 2013

National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-Ku, Tokyo, Japan

Intelligent Information Processing Chances of Crowdsourcing

Organizers:
Wolf-Tilo Balke, TU Braunschweig, Germany
Takahiro Hara, University of Osaka, Japan
Seung-won Hwang, POSTECH University, South Korea
Christoph Lofi, NII Tokyo, Japan

November 18 - 21 , 2013

Introduction

Currently a variety of platforms like Amazon Mechanical Turk, CrowdFlower, or SamaSource are offering frameworks with different degrees of sophistication where (usually relatively simple) cognitive tasks can be dynamically posed to a large and readily available workforce. This ability of cheaply distributing simple jobs via the Web allows for new modes of labor and information processing. In fact, the knowledge society has already brought substantial changes to business processes in todays economy. This is especially true for the basic question of what and where people work.

Here the ubiquity of sophisticated mobile devices and communication services allow for almost unlimited flexibility and freedom in negotiating and outsourcing short-term work contracts and delivering results. Currently, mobile crowdsourcing by smartphone users is a hot research area. In any case, in the industrialized world there is a clear transition from traditional production of goods or processing of raw materials towards the provisioning of services and the flexibility with respect to the place where such services are actually physically provided has dramatically increased. Still, although services could in principle be offered flexibly from virtually anywhere in the world, typical constraints like the local cost of labor or easy access to an educated workforce, remain valid. Crowd-Sourcing promises to break with these traditional work models, by offering a dynamic global information-processing workforce which is available 24/7 with close to no overhead. This shift paves the way for approaching large-scale information task which were previously infeasible for both algorithmic and traditional human-based approaches.

The central challenge in the current knowledge society is to efficiently and intelligently deal with an overwhelming amount of information, a daunting task for computer systems and humans alike. To this aim the data management and data mining communities consider a wide variety of operators, algorithms, and workflows.

For some information-heavy areas like for example customer relationship management, where everyday services like ordering procedures, customer data management, complaint handling, etc. have to be performed, out-sourcing the work to specialized workers has become a commonly accepted solution for increasing efficiency. Although such services do not produce anything in the traditional material sense, they are critical for company goals like efficient sales handling, customer satisfaction and retention, etc. Whereas such tasks used to be done on-site, nowadays call centers all over the world centrally provide such services at considerably reduced costs for a large number of customers. These services are quite basic and easy to provide in terms of education. On an educationally higher level, business intelligence services can serve as a good example: extracting relevant information from company data and using it to recognize or design value-adding areas like new products, promising customer segments, or better business processes for a company is a profitable business. Indeed "infopreneur" is a term coined for the growing number of persons whose primary business is gathering and selling electronic information. However, this current form of out-sourcing information-centric tasks is still quite static (i.e. a fixed team of specialists is contracted for a larger task). In contrast, crowdsourcing as understood in our workshop dynamically assigns small intelligence tasks to workers from a large pool in a demand-driven fashion. The advantages are obvious: if at creation time each process can be effectively broken down to manageable tasks and a viable time plan, it can be fulfilled very efficiently. The main factor is elasticity: peaks and slumps in activity can be dynamically handled and missing expertise or competences can be contracted. Thus, the efficiency of the overall process is hard to beat.

The main purpose of this Shonan meeting is to bring together researchers from the field of data management, information processing, HCI, and mobile computing to discuss the technical challenges, possible societal impact, as well as promising industrial applications for on-demand crowdsourcing techniques in vast information management challenges. The seminar puts a clear focus on operations in data management and data processing workflows. Indeed there are many open questions to discuss: How can operators/workflows benefit from crowdsourcing? Can the resulting quality be controlled? Which workers should be selected? How to determine expected response times? How to deal with privacy risks?

As stated above, a special focus should be paid to crowd-sourceable operators for applications for data and information management, information organization, and information access. In recent years algorithms aimed at these tasks have raised a lot of attention and indeed, methods have grown quite powerful even over huge and largely unstructured information repositories like the Web. Applications are almost limitless ranging from basic information extraction over knowledge management to complex business intelligence. However, with more complex information processing or information mining capabilities also the complexity, susceptibility for errors and danger of overspecialization of these algorithms increases. Since most failings can be traced back to limited cognitive abilities, missing contextual knowledge or heuristics gone wrong, the idea of direct human supervision and intervention at processing time is currently pursued in many domains. But also the quality of the work delivered by workers raises concerns: todays platforms are facing spam and individual workers work quality, skill, and reliability have to be measured for effective quality control. While for spam detection simple methods like gold questions or majority vote may work well, more complex quality assessment need new and more powerful models. Actually, ranking schemes based on reputation mechanisms already play a vital role in Web platforms, where matchings or transactions between anonymous parties are brokered. Hence their applicability for crowd-sourcing scenarios should be discussed. In fact, the need for human assistance in bridging the final semantic gap for todays information processing has already given rise to information systems that rely on hybrid architectures. Such hybrid architectures transparently combine the efficiency of current algorithms with the cognitive power and flexibility of humans.

Here, generally two design directions are popular:

- Using human input for improving the steps performed by information processing algorithms by providing training samples, answering questions about ambiguous results, or by providing relevance feedback.
- Involving humans directly into the information processing process, explicitly out-sourcing some of the required tasks or operators within the process.

Both general approaches are still very new, and no established research community has yet developed for crowd-assisted information processing algorithms. This Shonan meeting can provide a significant stimulus to the research community in order to advance this still new field of interest.

Topics of Interest

The meeting is primarily intended to focus on topics and problems related to information and knowledge processing. In this area, there are many tasks for which basic algorithmic approaches exist, but fall short because they often cannot grasp the semantics of the data they operate on correctly. Here, we envision that crowd-sourcing techniques are running in parallel in a hybrid system, and supplementing the algorithms when necessary. Especially, operators and algorithms of the following areas shall be discussed with their potential synergy with crowd-sourcing in mind:

- Complex databases operators like cognitive comparison and similarity functions, as for example sorting or joining images, ambiguous labels, descriptions, etc.
- Information and knowledge mining tasks, as for example entity and relation detection, entity reconciliation, or improving typical extraction pattern
- Improving data or knowledge representation, as for example schema matching, ontology cleaning, or data cleaning
- Sensor data stream processing (e.g., energy efficient stream join, uncertain stream processing)
- Obtaining cognitive meta-data from natural-language, as for example sentiment or emotion analysis, intention detection, sarcasm detection, etc.
- Semantic querying and retrieval, as for example question answering techniques or semantically-aware information retrieval algorithms

- Mobile-Crowd Sourcing platform, harnessing the special features of mobile devices as e.g., sensors
- Privacy issues, especially for mobile participants (e.g., location, trajectory, POI)
- Ethics of crowd-computing: discussions and insights on how the large-scale application of crowd-sourcing affects both workers and information management systems from an ethical perspective

Participants

- Wolf-Tilo Balke, Technische Universitt Braunschweig, Germany
- Seung-won Hwang, POSTECH University, South Korea
- Takahiro Hara, University of Osaka, Japan
- Christoph Lofi, National Institute of Informatics, Japan
- Yukino Baba, University of Tokyo, Japan
- Sozo Inoue, Kyusyu Institute of Technology, Japan
- Koichi Kise, Osaka Prefecture University, Japan
- Kai Kunze, Osaka Prefecture University, Japan
- Kinda El Maarry, Technische Universitt Braunschweig, Germany
- Yoshifumi Masunaga, Ochanomizu Univeristy, Japan
- Shigeo Matsubara, Kyoto University, Japan
- Atsuyuki Morishima, University of Tsukuba, Japan
- Stephan Sigg, Technische Universitt Braunschweig, Germany
- Yoshito Tobe, Aoyama Gakuin University, Japan
- Sanjay Madria, Missouri University of Science and Technology, United States
- Jiyin He, Centrum Wiskunde en Informatica, Netherlands
- Xuan Zhou, Renmin University of China, China
- Koji Zettsu, NICT, Japan
- Nestor Alvaro, National Institute of Informatics, Japan
- Hyunsouk Cho, POSTECH, South Korea
- Victor Munts-Mulero, CA Technologies, Spain

Program

This seminar followed a different organizational structure than most other Shonan meetings, as its goal was not to present individual research results, but to bring together experts from the field which freely discuss the current state-of-the art during the meeting itself. Then, potential problem areas are identified and respective research challenges and potential solutions are developed and presented. The meeting was started with an introductory talk by the organizers, providing some general overview of the topic, the goal of the workshop, and the schedule.

But basically, the first day was dedicated to helping the participants get to know each other as well as their respective research. For this purpose, each participant prepared a brief presentation of roughly 10 minutes, outlining herself with her research interests. The participants explained briefly in a general way how their research currently involves crowd-sourcing in intelligent information processing, or how crowd-sourcing techniques might help to address problems currently found in their research area. These introductions provided interesting views or fruitful insights on problems and solutions, some of which have been used to start further discussions on the following day.

The core of the seminar are the breakout sessions on the second day. Here, in a group discussion, five main problem areas have been identified. Based on this, the whole group was split into five subgroups, each discussing one topic in detail over the course of the day. On the next day, each group presented their discussion results with a brief talk. The last day was used to summarize the meeting and discuss future initiatives and tasks.

Arrival Day (Sunday, 17th November)

15:00-18:30	Check-In in Shonan Center
19:00-20:30	Welcome Dinner
21:00-	Free Time

Day 1 (Monday, 18th November)

07:30-09:00	Breakfast
09:00-09:10	Shonan Introduction by Staff
09:10-12:00	Seminar Session with one Coffee Break Opening briefing from organizers Position talks from participants
12:00-14:00	Lunch with Photo Shooting
14:00-18:00	Seminar Session with one Coffee Break Position talks from participants (continued) Discussion to categorize the issues addressed by the participants
18:30-19:30	Dinner
19:30-	Free Time

Day 2 (Tuesday, 19th November))

07:30-09:00	Breakfast
09:00-12:00	Seminar Session with one Coffee Break Break-out Sessions discussing important issues and find new research directions on the target topic
12:00-13:30	Lunch
13:30-18:00	Seminar Session with one Coffee Break Break-out Sessions (continued) Prepare for group presentation
18:30-19:30	Dinner
19:30-	Free Time

Day 3 (Wednesday, 20th November)

07:30-09:00	Breakfast
09:00-12:00	Seminar Session with one Coffee Break Presentation from each group and discussion
12:00-13:30	Lunch
13:30-18:00	Excursion to Kamakura
19:00-21:30	Banquet Dinner
21:30-	Free Time

Day 4 (Thursday, 21th November)

07:30-09:00	Breakfast
09:00-12:00	Seminar Session with one Coffee Break Idea marketplace and future collaborations Final organizer presentation and wrap up
12:00-13:30	Lunch

Summary of Breakout Sessions

In the remainder of this document, we summarize the results of the breakout sessions. Three of the sessions have been compiled into papers submitted to the UnCrowd 2014 workshop, while the other two are summarized using bullet lists.

Crowdsourcing: the Art of Involving the Community in Social Computing

Victor Muntés-Mulero[1], Yoshifumi Masunaga, Takahiro Hara[2], and Sozo Inoue[3]

[1] CA Labs, CA Technologies, Barcelona, Catalonia
[2] Department of Multimedia Engineering, Graduate School of Infomation Science and Technology, Osaka University
[3] Graduate School of Engineering, Kyushu Institute of Technology

Abstract. The number of organizations adopting crowdsourcing strategies is growing at an amazing speed. While the term *crowdsourcing* is being increasingly used in different contexts, it is still ambiguous and lacks a clear definition. Crowdsourcing implicitely includes a social dimension. Social computing was coined in 1994 by Doug Schuler, and its current meaning was come along around the middle of 2000 due to the wisdom of crowds by James Surowiecki, referring to systems that support the gathering, representation, processing, use, and dissemination of information that is distributed across social collectivities. In this paper, we discuss the relationship between crowdsourcing and social computing and its potential social impact. We define the main components of a crowdsourcing platform and we provide a classification of challenges that can be solved through crowdsourcing.

1 Introduction

Due to the tremendous popularization of the high-speed Internet and smart devices with wireless communication capabilities, a vast number of computers are connected and communicated with each other anytime and anywhere. This situation has drastically changed our life style, and various types of new services have been launched. A typical example of such services is *social networking services* such as Facebook, Flicker, and Twitter. Through social networking services, people form on-line communities and generate an extremely large amount of user-generated content, which have been taking a significant role as information sources on human intelligence.

The term crowdsourcing was coined in 2006 by Jeff Howe in his artcle[6] published in Wired maganize, and he has provided some perspectives and definition of crowdsourcing, which we can find in [6]. In a blog article[4], he stated that *"crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form*

[4] Crowdsourcing, http://www.crowdsourcing.com/cs/2006/06/index.html

of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers." As shown in this definition, crowdsourcing was originally the term of a business structure, which simply means outcourcing a function (or task) to the crowd.

The number of solutions using crowdsourcing strategies is growing quickly. Human brain-guided computation is able to efficiently perform tasks that computers cannot solve. Besides, the current economic recession affects millions of families worldwide[5]. In other words, unemployment affects a large spectrum of the population in many countries including highly qualified professionals.

Just as an example, Amazon Mechanical Turk (mturk.com), CrowdFlower (crowdflower.com) or ClickWorker (clickworker.com) provide a platform that enables companies or individuals to utilise human intelligence to perform tasks that present complex challenges for computers. In some cases, like CrowdFlower or ClickWorker, these platforms offer a variety of crowdsourcing services. They improve quality by using several reviews of each data unit, gold standards, etc. However, the limits between a crowdsourcing digital platform and yet another website collecting some information are blurry. Many different platforms may be labelled as crowdsourcing ranging from Wikipedia to a sophisticated industrial platform for high-quality translation. All of them share one idea: people collaborate remotely to achieve a common goal. However, what do all these platform have in common? May a simple form on a website be considered a crowdsourcing platform? Where are the limits of crowdsourcing?

In this paper, we explore the concept of crowdsourcing. We discuss the relationship between crowdsourcing and social computing and we make a first proposal about the organization of a generic crowdsourcing platform. We also discuss the potential social impact that the generalization of crowdsourcing may have. Finally, we analyze the different types of tasks that are currently being solved using crowdsourcing strategies and technologies.

This paper is organised as follows. Section 2 introduces previous work including some basic properties of aggregation functions. Then, Section 3 revisits the concept of social computing, analyzes some basic principles of crowdsourcing and proposes a definition. Section 4 extends the definition of Section 3 by presenting a generic organization of a crowdsourcing platform including the main components of such a platform. Section 5 describes the most common types of tasks solved through crowdsourcing. In Section 6, we discuss the potential social impact of crowdsourcing. Finally, Section 7 concludes and draws some future research lines.

2 Previous Work

One of the main issue related to crowdsourcing is output quality [1, 5, 8, 10, 17, 18]. Incentive mechanisms are also tightly linked to quality [7]. Different sce-

[5] For example, the unemployment rate in Spain and Portugal was 26.7 and 15.5, respectively (November 2013) (see "Seasonally adjusted unemployment". Eurostat)

narios may require different incentive mechanisms. For instance, an industrial environment may require to provide economic incentives to motivate the crowd to perform certain tasks. The work presented in [15] presents some results that confirm the importance of money compared to other motivations in certain domains. Also from an industrial perspective, first steps have been done to establish the basis for crowd coordination and create rewarding mechanisms that are based on involving human beings in the evaluation of the quality of other workers through the so-called AV-Units [12]. Some platforms use a combination of the following extrinsic incentives to keep a working community engaged: economical rewards, gamification, *e.g.*, public scoreboards, and free training. For instance, while Mechanical Turk rewards workers with economic incentives, Duolingo[6] provides a strategy based on gamification, offering language training for free while users actually translate real strings from websites.

Another important aspect for crowdsourcing platforms is worker management. Venetis and Garcia-Molina propose *Gold Standard Performance* to detect the performance of a worker before the crowdsourcing task starts [17]. Other characteristics of a worker such as demographics or personality traits may be relevant to the quality of their work under specific task conditions [8]. In general, it is considered that inaccurate evaluation of tasks may encourage other fraudsters to misbehave in the platform or discourage good workers. For example, Hirth et al. raise "Majority Decision Approach" [5] to judge whether worker's submission is correct in simple tasks, and using "Control Group Approach" method in complicated cases. Crowdsourcing the quality evaluation of the jobs performed by the crowd is also proposed in [4, 9].

A definition and analysis of previous work on social computing is provided in Section 3.

3 Crowdsourcing and Social Computing

The term social computing was coined in 1994 by Doug Schuler [14] referring to any type of computing application in which software serves as an intermediary or a focus for a social relation, which includes newsgroups such as email, community computing, groupware, technological systems for better integration of social needs, accountability for the impact of computing, and social nature of software. However the meaning of social computing has been changed according to the development of the Web. A detailed analysis about the use of the term social computing in the web society is performed in [11]. The investigation shows that the first or the oldest Wikipedia article on social computing was written on January 21, 2005, and the opening section of the article states that *"Social computing refers to the use of social software, and thus represents a growing trend of ICT usage concerned with tools that support social interaction and communication. Social computing is rather based on existing social conventions or related to specific social contexts than characterized by its technological attributes. Examples of social computing is the use of e-mail for maintaining social relationships,*

[6] http://www.duolingo.com/

instant messaging for daily microcoordination at one's workplace, or weblogs as a community building tool." Note that the definition is seen as an extension of that of Schuler. This work shows the constant evolution of the term social computing: *"In the weaker sense of the term, social computing has to do with supporting any sort of social behavior in or through computational systems. It is based on creating or recreating social conventions and social contexts through the use of software and technology. Thus, blogs, email, instant messaging, social network services, wikis, social bookmarking and other instances of what is often called social software illustrate ideas from social computing, but also other kinds of software applications where people interact socially. In the stronger sense of the term, social computing has to do with supporting computations that are carried out by groups of people, an idea that has been popularized in James Surowiecki's book, The Wisdom of Crowds [16]. Examples of social computing in this sense include collaborative filtering, online auctions, prediction markets, reputation systems, computational social choice, tagging, and verification games."* It should also be mentioned that the spread of the social computing in the stronger sense is partly thanks to the Web 2.0 report by Tim OReilly [13] which was published in 2005, where the utilization of the collective intelligence in the Surowieckis sense should be one of the seven principles which is essential for Web to be Web 2.0. Based on the stronger sense of the social computing definition, a formal model of social computing was introduced. That is, since the stronger sense is based on the wisdom of crowds, it reveals a formal computational model of Surowiecki's idea. A formal model of social computing is depicted in Figure 1.

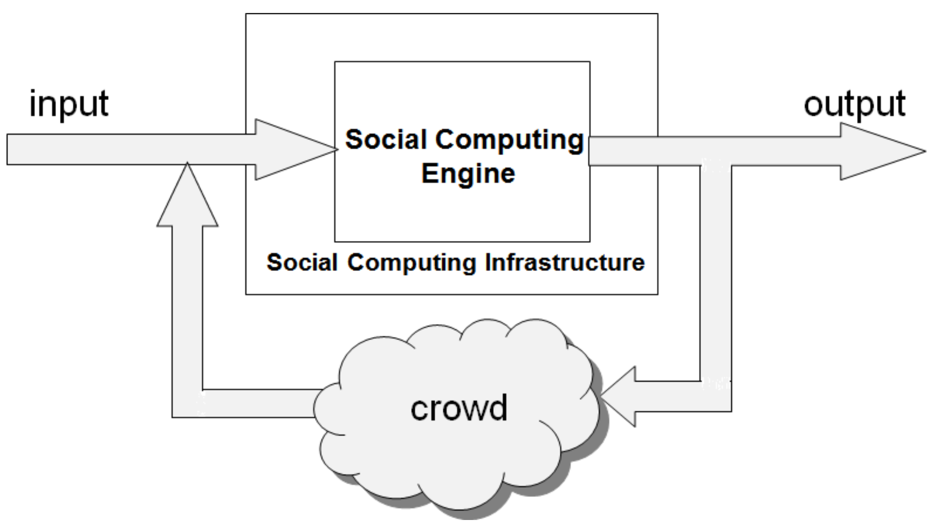


Fig. 1. Formal model of social computing (extracted from [11])

Note that the essential difference between the traditional computing and social computing is that there is a *social feedback loop* from the output to the input in the social computing model, which represents people's direct participation in the computing process. As Surowiecki mentioned in his book, this participation tends to yield an exact decision if a group of people satisfies the following three conditions: diversity, independence, and decentralization. Briefly speaking, diversity of opinion means that each person should have private information even if it is just an eccentric interpretation of the known facts. Independence means that peoples opinions are not determined by the opinions of those around them. Decentralization means that people are able to specialize and draw on local knowledge. As he also mentioned, the aggregation function is needed to weave a variety of opinions into a single outcome. The social computing engine shown in Figure 1 represents the aggregation function. If we take Wikipedia as an example of social computing, a set of MediaWiki, Wiki edit conflict resolver, and the 3RR (three-revert rule) constitutes the social computing engine. If we take the stock market as an example of social computing, NASDAQ system, for example, is the social computing engine. The social computing infrastructure represents the Web as a platform.

3.1 Discussion about the definition for crowdsourcing

Before presenting a definition for crowdsourcing, we re-examine some basic principles of the wisdom of crowds of Surowiecki that should be linked to crowdsourcing.

- **Diversity**: any crowdsourcing platform should promote diversity. It is only through diversity of backgrounds, knowledge, opinions, roles and experience that we can leverage the full potential of the crowd. Any crowdsourcing platform may ensure that workers and users are diverse.
- **Independence**: independence among workers is a desirable property in most crowdsourcing platforms. Complete independence might not be possible in an increasingly connected world. However, crowdsourcing platforms must promote workers and contributors to act based on their own experience, knowledge and information. In this way, we maximize the capacity to provide collective emerging mechanisms, based on the opinion and capacity of independent remote Internet users.
- **Openness**: crowdsourcing platforms must be as open as possible and promote participation. We consider open calls a necessary part of a crowdsourcing platform, where remote workers must freely decide to participate. Although profiles and skills might be important in order to solve certain tasks, a crowdsourcing platform should not be based on looking for specific and well-known service providers. Identity should not be important during the open call in the recruitment process.
- **Decentralization**: crowdsourcing tends to descentralize work. Depending on the platform the degree of democracy might be higher (i.e. Wikipedia) or lower (a crowdsourcing platform created by a company), however, the

community should play an important role in assessing quality, collaborating and even deciding the objectives of the crowd.

Based on these basic principles and taking into account previous definitions such as [2] and [6], we define crowdsourcing as:

Definition 1. *Crowdsourcing is a problem-solving and production model consisting in involving independent internet or intranet users through an open call to contribute to the mission of a social computing system*

Note that with this definition, we remark the role of crowdsourcing as one of the possible mechanisms to engage remote users in a social computing system. Besides, we have deliberately avoided using *internet users* exclusively in the definition as we understand that it is possible to use crowdsourcing in other restricted environments such as for instance inside a large enterprise, through the corporate intranet.

4 Crowdsourcing platform components

In this section, we describe those components that are essential in a crowdsourcing platform. We understand that the subcomponents of a crowdsourcing platform are also part of the definition of crowdsourcing. Figure 2 depicts the essential components of a crowdsourcing platform. First of all, we remark the importance of defining a clear goal. Social computing involves human beings by definition and crowdsourcing tackles the challenge of involving them into the social computing system. While, human intelligence brings a huge number of opportunities, it also requires dealing with ambiguity. Human potential resides in part in our ability to return solutions even when a problem is ambiguously defined. However, results might not be as accurate as necessary. In order to reduce noise in the answer caused by a bad definition of the problem to be solved and the task to be performed, the goal needs to be clear to remote workers from the beginning. Once the goal is clear, we divide a crowdsourcing platform in four subcomponents:

- **Task Creation and Management Mechanisms**: any crowdsourcing platform must create efficient mechanisms to interact with remote users. This implies a subcomponent to create and manage subtasks, a digital infrastructure to make these tasks public through open calls, mechanisms to ensure efficient machine-human interaction and the definition of usable UIs. We also consider an essential part of a crowdsourcing system a task monitoring mechanism that allows collecting information about the work progress and, optionally, learn from those processes. Tasks may require human beings or they may be automatic. Crowdsourcing platforms need to combine these two types of tasks when necessary. An important and usual type of task will be that task devised to merge results from different crowdsourced tasks performed by human beings. A possible mechanism to combine tasks through collaboration patterns is presented in [12].

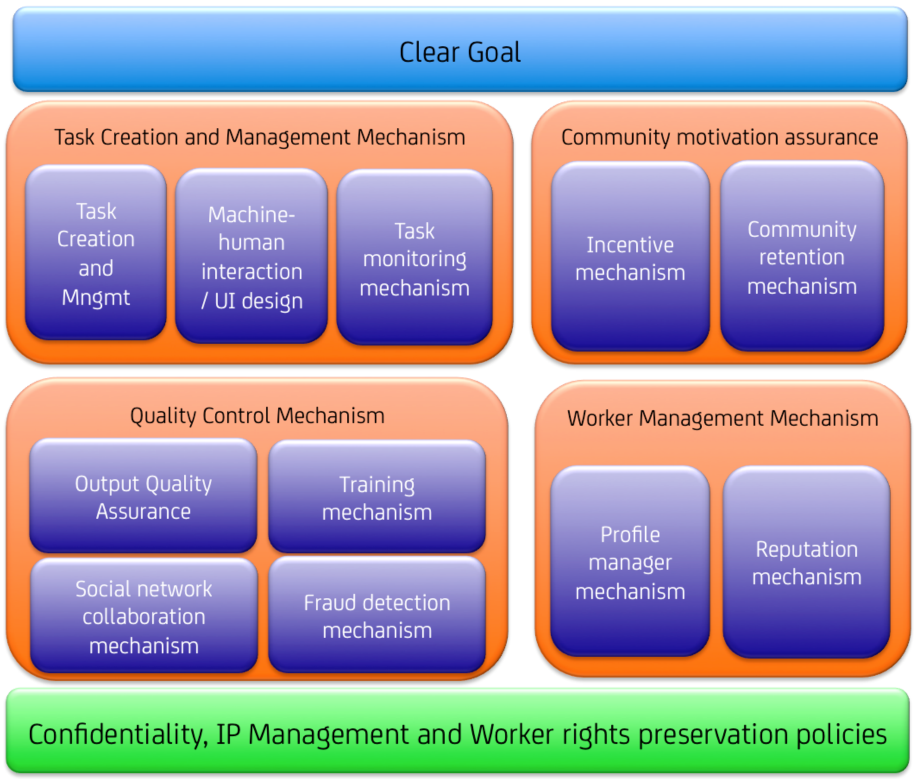


Fig. 2. Components of a crowdsourcing platform

– **Community Motivation Assurance Mechanisms**: another essential part of a crowdsourcing mechanism is to provide mechanisms and a estrategy for assuring community motivation. This may differ significantly from one type of platform to another, ranging from economic incentives to other incentives such as social reputation. In any case, it is essential to design a strategy to incentivate and grow a community and, also very important, to retain actual users of the platform.

– **Quality Control Mechanisms**: quality assurance is essential in many scenarios such as in industry. A crowdsourcing platform must provide mechanisms to control the output quality, to train users so that they can grow and learn through their participation in the platform and mechanisms to enable the collaboration between the workers of the platform when necessary. Besides, quality is usually related to well-intentioned workers. Unfortunately, it is usual to find internet users that misbehave in order to be rewarded without being compliant with the expected output quality. Because of this, it is necessary to develop sophisticated fraud detection mechanisms to identify fraudsters and organized attacks.

– **Worker Management Mechanisms**: managing workers or remote users is also essential in any crowdsourcing system. Mechanisms for worker management may range from simple registration mechanisms to identify users, to sophisticated platforms that allow workers to create complex profiles and reputation indicators. In some scenarios, certain tasks may be restricted to certain profiles and skills, as a quality measure.

Finally, legal frameworks are still ambiguous with respect to digital platforms. The complexity of a multijurisdictional environment poses a serious challenge when it comes to leverage the potential of billions of workers located in any part of the world. Because of this, it is crucial to define policies to rule IP management and confidentiality and to preserve workers rights.

5 Crowdsourcing task classification

In this section, we briefly discuss about a categorization of tasks or crowdsourcing application types. Our discussion is based on the categorization presented by Daren Brabham in [3]. However, we extend the current proposal and enumerate some examples for each category:

– ***Knowledge discovery and management*** tasks are tipically issued when an organization requires information from usually unknown external sources. The task consist on asking internet users to provide it. Practical examples of this type of task may be represented by Peertopatent.org, SeeClickFix.com or BeMyEye.com platforms. IBM Innovation Jam[7], Kaggle[8], or X Prize Foundation[9].
– ***Broadcast search*** is used when a complex problem is defined and an open call is created to find empically provable solutions to that problems. An example of this type of platform would be Innocentive.com or the Red Ballon Challenge issued by DARPA.
– ***Peer-vetted creative production*** is another type of task where there is not an objective mechanism to evaluate the quality of the output and this depends on the opinons of people or on market support. Examples of these types of tasks may be Threadless.com, iStockphoto.com or Local Motors[10].
– ***Distributed human intelligence tasking*** is used when an organization owns data that needs to be analyzed or processed. Usually the volumes of information are large and tasks are subdivided in smaller tasks. Examples of platforms based on these types of tasks are Amazon Mechanical Turk[11] or ODesk.com.

[7] collaborationjam.com
[8] kaggle.com/
[9] xprize.org
[10] localmotors.com
[11] mturk.com

- ***Crowdfunding*** would be a fifth type of task which is not listed in Brabhams list of task categories. A project that requires funding is published and the crowd freely decides to make donations to make the project possible. Typical examples would be Kickstarter.com, CyberAgent Crowdfunding[12] or GoFundMe.com.

6 Social Impact of Crowdsourcing

Finally, we analyze the potential social impact of crowdsourcing. Beyond being able to characterize a crowdsourcing system, we aim at understanding and explaining how the use of crowdsourcing may impact society worldwide.

- **Crowdsourcing enables real-time data collection in new environments**: the fact that billions of people already have highly sophisticated mobile devices with increasing sensoring capabilities is opening a whole world of new opportunities to collect real-time data. This allows analyzing real environments in real time, by collecting massive amounts of information from individual contributors that may be automatically or manually submitting information through crowdsourcing mechanisms. Therefore, the crowd acts as a humongous sensor network in real time. Some application examples are target-tracking (e.g., person or animal) in the real world, high-granularity air pollution monitoring, and realtime event/accident detection.
- **Crowdsourcing allows modelling human behaviour**: with crowdsourcing platforms we can precisely monitor human beings behaviour and decisions. Some application examples are crowd-behavior analysis, emergency evacuation monitoring/recommendation, tour recommendation, and traffic control by congestion monitoring. While there have been some existing services (not based on crowdsourcing) that address this issue (e.g., Citysense[13]), crowdsourcing could contribute to further enhance the movitation and incentive people's participation. With this we open a new research area to improve automatic mechanisms based on human behaviour mining. Besides, new opportunities arise to improve machine learning from collective this massive information.
- **Increasing the competitivity of companies**: companies face serious problems to manage massive amounts of information within dynamic and international markets. Companies also face other types of challenges. Four competing needs are pushing industry to find new solutions: the need for managing massive amounts of information, the need for including humans in the loop to guarantee quality, the need for elastic systems that allow adaptability to dynamic environments and workloads, and the need for reducing time-to-market for their products and being agile in order to exploit new emerging markets. With the evolution of new technologies and the capacity

12 ca-crowdfunding.com
13 Sense Networks, https://www.sensenetworks.com/products/macrosense-technology-platform/citysense/

to produce and access humongous amounts of digital information, processing and analyzing data becomes critical. A natural way to handle this information would be using automatic processes. However, the quality provided by automatic processes is not yet sufficient and the need for including humans in the loop is unquestionable. Although applications require human intervention, it is unfeasible to keep an in-house team of professionals to cope with massive amounts of information, dynamic workloads and the time constraints in the production chain. The concept of elasticity has been generalized with the adoption of cloud computing solutions. An essential competitive advantage for many organizations is their capacity to allocate and free resources depending on the workload and this collides with the static characteristics of in-house human teams. The intervention of human beings may also represent an important bottleneck in terms of performance and it may be a factor to increase the time-to-market of many products that require humans to be involved in its process of development. Besides, international companies face important challenges when it is necessary to commercialize and localize products in new regions to open new markets. This makes it necessary to outsource many industrial processes increasing the costs and diminishing the control over quality. Crowdsourcing will provide solutions for all these challenges allowing to parallelize processes involving human beings connected remotely in an international environment, enabling elasticity and reducing costs.

- **Democratization and transparency**: crowdsourcing allows to democratize the use of the Internet workforce and the distribution of private funds. This may reduce the dependence on public funds and the social impact of many organizations such as NGOs.
- **Crowdsourcing enables the resolution of problems using unprecendented mechanisms**: thanks to crowdsoucing, we can solve problems that could not be solve before. Just as an example see *fold.it*[14] or the *GoldCorp Challenge* case[15].
- **Crowdsourcing creates new business models for both enterprises and workers**: through crowdsourcing new telework opportunities arise and, with this, new opportunities to create innovative digital enterprises, replacing traditional business models. This will also have an important impact on workers. On the one side, it may have negative consequences, creating a lack of stability for workers, since crowdsourcing represents a new trend towards a freelancer-based marketplace. Besides, it may also represent the exclusion of many non-digital natives. On the other side, it represents better opportunities for multiemployment, remote and free professional training and a good opportunity for work-life balance. Besides, it may democratize the access to job opportunities independently of gender, age or geographic location.

[14] solveIt: Solve puzzles for science: http://fold.it

[15] http://www.ideaconnection.com/open-innovation-success/Open-Innovation-Goldcorp-Challenge-00031.html

7 Conclusions and future work

Crowdsourcing is still an emerging area of research. Althought the number of platforms using different flavors of crowdsourcing is increasing very fast, there exists a lack of conceptualization of the main challenges related to this type of technology. Quality assurance, motivation and reputation mechanism, and fraud detection are just examples of three important issues that need to be studied. Our knowledge on the potential impact that crowdsourcing will have in the near future is still unclear. Crowdsourcing may become a key trigger to leverage the ability of social network users to create content and ideas jointly using emerging intelligence techniques, to exploit smart cities using mobile devices as sensors and to increase the competitivenss of the industrial sector.

Acknowledgement

The paper is based on the discussion in the Shonan Meeting for “Intelligent Information Processing - Chances of Crowdsourcing” held on November 13-21, 2013. We thank for the organizers and participants of the meeting.

References

1. Bentivogli, L., Federico, M., Moretti, G., Paul, M.: Getting expert quality from the crowd for machine translation evaluation. In: MT Summit XIII. pp. 521–528 (2011)
2. Brabham, D.C.: Crowdsourcing. The MIT Press essential knowledge series, MIT Press (2013)
3. Brabham, D.: Crowdsourcing. MIT Press (2013)
4. Gao, Q., Vogel, S.: Consensus versus expertise: a case study of word alignment with mechanical turk. In: Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon’s Mechanical Turk. pp. 30–34. CSLDAMT ’10, Association for Computational Linguistics, Stroudsburg, PA, USA (2010), http://dl.acm.org/citation.cfm?id=1866696.1866700
5. Hirth, M., Hossfeld, T., Tran-Gia, P.: Cost-optimal validation mechanisms and cheat-detection for crowdsourcing platforms. In: Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2011 Fifth International Conference on. pp. 316–321 (2011)
6. Howe, J.: Wired 14.06: The Rise of Crowdsourcing, http://www.wired.com/wired/archive/14.06/crowds.html
7. Kamar, E., Horvitz, E.: Incentives for truthful reporting in crowdsourcing. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 3. pp. 1329–1330. AAMAS ’12, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2012), http://dl.acm.org/citation.cfm?id=2343896.2343988
8. Kazai, G., Kamps, J., Milic-Frayling, N.: The face of quality in crowdsourcing relevance labels: Demographics, personality and labeling accuracy. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management. pp. 2583–2586. CIKM ’12, ACM, New York, NY, USA (2012), http://doi.acm.org/10.1145/2396761.2398697

9. Kontokostas, D., Zaveri, A., Auer, S., Lehmann, J.: Triplecheckmate: A tool for crowdsourcing the quality assessment of linked data. In: Knowledge Engineering and the Semantic Web. Communications in Computer and Information Science, vol. 394, pp. 265–272 (2013)
10. Mashhadi, A.J., Capra, L.: Quality control for real-time ubiquitous crowdsourcing. In: Proceedings of the 2Nd International Workshop on Ubiquitous Crowdsouring. pp. 5–8. UbiCrowd '11, ACM, New York, NY, USA (2011), http://doi.acm.org/10.1145/2030100.2030103
11. Masunaga, Y.: Social computing: Its evolving definition and modeling in the context of collective intelligence. In: SocialInformatics. pp. 314–319 (2012)
12. Muntés-Mulero, V., Paladini, P., Manzoor, J., Gritti, A., Larriba-Pey, J.L., Mijnhardt, F.: Crowdsourcing for industrial problems. In: Nin, J., Villatoro, D. (eds.) CitiSens. Lecture Notes in Computer Science, vol. 7685, pp. 6–18. Springer (2012)
13. O'Reilly, T.: What is web 2.0: Design patterns and business models for the next generation of software. (2005), http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html
14. Schuler, D.: Social computing. Commun. ACM 37(1), 28–29 (Jan 1994), http://doi.acm.org/10.1145/175222.175223
15. Silberman, M.S., Irani, L., Ross, J.: Ethics and tactics of professional crowdwork. XRDS 17(2), 39–43 (Dec 2010), http://doi.acm.org/10.1145/1869086.1869100
16. Surowiecki, J.: The Wisdom of Crowds. Anchor (2005)
17. Venetis, P., Garcia-Molina, H.: Quality control for comparison microtasks. In: Proceedings of the First International Workshop on Crowdsourcing and Data Mining. pp. 15–21. CrowdKDD '12, ACM, New York, NY, USA (2012), http://doi.acm.org/10.1145/2442657.2442660
18. Zhu, S., ShaunKane, Feng, J., Sears, A.: A crowdsourcing quality control model for tasks distributed in parallel. In: Extended Abstracts on Human Factors in Computing Systems. pp. 2501–2506 (2012)

Skill ontology-based model for Quality Assurance in Crowdsourcing

Kinda El Maarry[1], Wolf-Tilo Balke[1], Hyunsouk Cho[2], Seung-won Hwang[2], Yukino Baba[3]

[1]Institut für Informationssysteme, TU Braunschweig
{elmaarry, balke}@ifis.cs.tu-bs.de
[2]Department of Computer Science and Engineering, POSTECH, Korea
{prory, swhwang}@postech.ac.kr
[3]The University of Tokyo, Japan
yukino_baba@mist.i.u-tokyo.ac.jp

Abstract.
Crowdsourcing continues to gain more momentum as its potential becomes more recognized. Nevertheless, the associated quality aspect remains a valid concern, which introduces uncertainty in the results obtained from the crowd. We identify the different aspects that dynamically affect the overall quality of a crowdsourcing task. Accordingly, we propose a skill ontology-based model that caters for these aspects, as a management technique to be adopted by crowdsourcing platforms. The model maintains a dynamically evolving ontology of skills, with libraries of standardized and personalized assessments for awarding workers skills. Aligning a worker's set of skills to that required by a task, boosts the ultimate resulting quality. We visualize the model's components and workflow, and consider how to guard it against malicious or unqualified workers, whose responses introduce this uncertainty and degrade the overall quality.

Keywords: crowdsourcing, quality assurance, skill ontology, uncertain data.

1 Introduction

The hope of being able to somehow benefit from "the wisdom of the crowd" [1] is the main driver for the rising popularity of crowdsourcing [2], coupled with the information flood and the flexible and relatively cheap solution today's crowdsourcing platforms offer. Thus, more and more companies and organizations are turning to crowdsourcing. Some notable names include: NASA, Threadless, iStockphoto, InnoCentive, etc [3]. Yet, the question of automatically assuring the returned quality of results [4] and the uncertainty that is associated with it, remains an unsolved question [5]. This is because checking every single submitted response is costly, time consuming and threatens to invalidate most of the crowdsourcing gains. This in turn encourages unethical workers to submit low quality results. According to [6], crowdsourcing

adfa, p. 1, 2011.

follows Sturgeon's law, i.e. as much as 90% of crowdsourcing contributions may actually be fraudulent.

Most requestors end up relying on redundancy or repeated labeling as means of verification of user performance. A common approach tests the reliability of users by blending a set of questions for which the answers are known, so-called gold questions, into the workload. Another possibility is the assignment of multiple workers to the same task and then aggregating their responses. While both approaches pose problems for tasks where the comparison between individual workers' results is difficult (see [7] for a detailed discussion), the redundancy approach usually incurs monetary costs and places the costs for quality control at the task provider's doorstep. Moreover, popular techniques for aggregation like e.g., majority voting, have been shown to suffer from severe limitations [8].

To correctly tackle the issue of uncertainty, various factors affecting the quality of the responses were investigated [9]. While results show monetary incentives having an effect on quality in contrast to the experimental results in [10], it's still rather tricky; low paid jobs yield sloppy work, and highly paid jobs attract unethical workers. Another investigated factor was workers' qualification, where not only was it shown that qualified workers produce better quality and strive to maintain their qualification level, but in a setup that relies on qualifications for task assignments, unqualified workers are pushed to diligently work on improving their own qualifications.

To that end, we investigate a skill ontology-based model to be adopted by crowdsourcing platforms, which aims at identifying those qualified workers and assigning them to the tasks they're eligible to. This can be realized through identifying the skills required to adequately work on a task, and aligning it to the skills a certain worker has. Consequently, by excluding non-qualified workers or non-ethical workers who falsely try to build up their qualifications, the model would be practically excluding the sources of uncertainty introduced to the data altogether.

The rest of the paper is organized as follows. We start off by reviewing the current related work. Next, we define what quality stands for in a crowdsourcing setup and identify the different types of quality that our model needs to realize. Section 4 presents in details the proposed skill-ontology model. This is followed in section 5 by an overview of the model's workflow. Finally, in the last section, we provide a summary and an outlook on future work.

2 Related Work

In recent years, many web-based collaboration platforms and marketplaces are relying on that same "wisdom of the crowd" ideolody, where anonymous users' contributions are in some way combined to provide innovative and diverse services. Threadless (online t-shirt design contest) [11] and istockPhoto, are two prominent examples exploiting that ideology. [12] presents an audio document retrieval service "PodCastle", which collects anonymous transcriptions of podcast speech data to train an acoustic model. This was followed two years later by an alternative crowdsourcing-based approach [13]. These examples support the main argument in [14], i.e. that the way people collaborate and interact on the web has been so far poorly leveraged through the existing service-oriented computing architectures.

So instead, a mixed service-oriented system i.e. service-oriented crowdsourcing, is desirable, enabling a more seamless approach, which would also exploit the on-demand allocation of flexible workforces. This steers the trend ever more towards crowdsourcing now being offered by many platforms: Amazon Mechanical Turk, Samasource, Crowdflower, etc. However, every chance needs to overcome challenges, and the main challenge here is that crowdsourcing results are often questionable in terms of their quality, and the associated uncertainty introduced in aggregated results becomes an issue.

This is actually very similar to the missing confidence in third-party services which posed serious issues in the web services community, see e.g., [15]. One solution here was to adopt credentials proving the eligibility of each discovered service. Simply put, a service is eligible if it meets certain quality requirements (in functionality, as well as typical QoS parameters like availability or response time). When composing complex workflows out of individual services these quality requirements can be interpreted as mutual agreements. Such agreements can be expressed for example by Web Service Level Agreement Language (WSLA) [16], or Web Service Management Language (WSML) [17]. In our context, a service provider is none other than a worker who has some skills and a task provider's confidence in results would be based upon the worker's credentialed skills. These credentials can be attained by passing a standardized test or a personalized test that the provider designs for that particular task. An agreement is reached, when a worker's credentialed skills matches those listed by the task provider (requestor) as the exact skills required for the corresponding task.

A lot of work in crowdsourcing literature has already been devoted to mitigate such quality concerns. The solution of redundancy and repeated labeling was first expanded by Dawid and Skene [22], who took into consideration the response's quality based on the workers. Through applying an expectation maximization algorithm, the overall error rate for each worker can be computed. Other approaches that estimate these error rates includes: a Bayesian version of the expectation maximization algorithm approach [23] and a probabilistic approach that takes into account both the worker's skill and the difficulty of the task at hand [24]. A further step was taken in [25] with an algorithm separating the unrecoverable error rates from recoverable bias.

Here, rather than looking at the worker's error rates, we aim at identifying the workers who are a good match for the corresponding task. Each worker has a skill profile, and every skill in the ontology is associated with a library of assessments. These assessments validate whether a worker indeed possesses the necessary skill or not. Both can be viewed analogously to the competence profiles provided by learning objects – entities that are used for task-focused training or learning in the IEEE 1484.12.1 – 2002 Standard for Learning Object Metadata.

These skills can be managed and referred to in a skill ontology. This fortunately leads us to a rich literature to derive and adapt from, which has been devoted to building competencies models, see [26] and [27]. Competency covers: knowledge, experience, skill and willingness to achieve a task. Such models have been used for quite a long time in organizations to help identify and attract suitable workers, as well as to help the workers acquire the needed skills. In order to identify the skills required for a task, skill gap analysis can be used to create the task's corresponding competency map [28]. Workers having the corresponding skills in a task's competency map could be then identified through competency matching. [29] Formalizes another approach that focuses on Ontology-based semantic matchmaking between demanded skills (skills required by a task in our case) and supply (the workers possessing that skill). However, competency models still have their own challenges. Given their complexity, competencies have to be precisely defined within the different specific domains. Moreover, developing assessments that can truly capture one worker's competency level is unfortunately very often underestimated [30].

But of course, assigning the right worker for a task involves much more than just choosing the workers based on their skills. A worker maybe be highly competent relative to the task he's assigned to, yet his work ethics may earn him a bad reputation This might simply boil down to wanting to finish a task as fast as possible and with the least effort incurred. So the overall quality is in fact affected by both the workers' skills and reputation. This elicits the need for deploying quality control measures, whether in design time, run time or both, see [18] for a more comprehensive list of these measures. Computing workers' reputations poses a real challenge, and many reputation approaches have been investigated whether it's based on a reputation model [19], on feedback and overall satisfaction [20], or on deterministic approaches [21], etc.

3 Types of Quality

Upon addressing the data uncertainty that arise with crowdsourcing tasks, different aspects of quality can be identified. This breakdown allows us to identify the corresponding quality assurance mechanisms, which needs to be addressed by the proposed model. A detailed description of each of those quality aspects follows next.

3.1 Result's quality

Comes first to mind, and covers both the requester's expectations and the usefulness of the results. In terms of requestor's expectations, the structure of the returned results will be heavily influenced by what the requester wants and expects. Accordingly, the crowdsourcing task should be designed in a way that elicit that specific structure in the returned results (factual correctness in the form of a yes/no answer, consensus, opinion diversity, opinion quality, etc.). In terms of usefulness, requestors may also measure the quality in terms of how the results are consistent or abiding to the task description, or whether they are transparent and traceable .i.e. there exists a logical pattern the worker followed to give that response.

3.2 Platform's quality

Refers to the usability of the platform, where a platform's interface and offered tools should equally support both workers and requesters. For workers, the platform should promote a fair working environment. Fairness encompasses: 1) guaranteed payments, 2) nondiscriminatory conduct, 3) payments matching the corresponding load of work. For the requesters, the platform should offer an adequate set of tools to easily and efficiently: 1) upload data and download results, 2) design tasks, 3) automatically assign qualified workers, 4) block spammers, 5) train workers.

3.3 Task's quality

At a lower granularity, the quality of the task directly affects the results' quality. A requestor should: 1) identify the set of skills required to accomplish a task, 2) describe the task clearly, 3) define the expected effort in terms of complexity or time required to finish, 4) design the task's interface to support an easier workflow for the worker.

3.4 Worker's quality

Refers to how fit a work is for the task at hand. Namely, how qualified and prepared they are to do the task. On one hand, qualified can be mapped to skill levels and how relevant these skills are to the task. On the other hand, prepared can be translated into willingness to complete the task to the best of ones skills. Other contributing factors are: 1) workers' availability, 2) flexibility of working hours (Both can be easily monitored through activity logs), 3) workers' reputation. (Can be based on history and average satisfaction score attained upon the completion of a task).

These different aspects will often in reality be interleaved. For instance, the clarity of a task is not only related to a task's quality, but might also fall under the platform's quality, where a platform ensures that the workers get clear task description that helps them avoid getting penalized if they do the task incorrectly due to vague guidelines.

4 Skill Ontology-based model

Following the quality aspects we identified in section 3, we propose a skill ontology-based model to be adopted by crowdsourcing platforms. The model aims to capture the different aspects of quality that helps diminish the resulting uncertainty by eliminating one of its major sources: unqualified workers. The skill ontology-based model roughly comprises of: 1) skill ontology, 2) ontology merger 3) skill's library of assessments, 3) Skill aligner, 4) reputation system and a 5) task assigner.

4.1 Basic and temporary skill ontologies

At the model's core lies the skill ontology. The model maintains a dynamic ontology, which evolves with the crowdsourcing platform's demands. While some skills will be often required for many tasks e.g. language skills for translating tasks, other skills will be highly specific and tailored for a specific task e.g. identifying the family, genus and species a fish belongs to. Accordingly, two ontologies are maintained: a basic and a temporary one. The basic skill ontology retains those skills that are highly demanded by many tasks. The temporary skill ontology retains newly added skills. Later on, only those skills that were frequently required by many tasks are transferred from the temporary ontology to the basic one.

A requestor is always presented with a single consolidated ontology, in which he can browse the skills required for the task he's designing. When the required skill isn't available in the ontology, the requestor can define a new skill.

4.2 Ontology merger

A new skill, which has been newly defined by a requestor is initially added to the temporary skill ontology. Every defined skill must be associated with at least one assessment. The new skill resides in the temporary ontology until it: 1) has proven to be popular 2) has been verified. Popular skills are skills that were required not only by many tasks, but also by many different requestors. Verified skills are skills that are associated with at least one verified assessment as will be further explained next.

4.3 Skill's library of assessments

Identifying whether a worker has a certain skill or not, can be ascertained through an assessment. A skill's library of assessments may comprise two types of assessments: standardized and personalized assessments. For standardized assessments like: TOEFL for the English language, or MOOC (Massive Open Online Course) certificates, most requestors will approve and conclusively trust them. However, when a requestor doesn't, or when there is simply no standardized test for the required skill, the requestor can create a personalized assessment.

Standardized assessments are inherently verified, since their legitimacy are already proven. On the other hand, personalized assessments, require further investigation for verification. Consider the following scenario: A worker posing as a requestor, creates a new personalized assessment and uploads it for the skill he\she wants to attain. Providing the perfect answers for these personalized assessments becomes then trivial, and the worker can accumulate endless skills in this manner. Assessments' verification can be done manually or automatically (platform-wise or crowd-wise).

1. Manual verification: entails hiring an expert to look over the assessment, this however costs both time and money. Accordingly, as a rule of thumb, this should be limited to cases where a popular skill has only one personalized assessment or multiple personalized assessments from the same requestor.
2. Automatic verification: serves as an alternative to manual verification, when the skill has: at least one verified personalized assessment, or one standardized assessment.

   - Automatic platform-wise verification: The platform creates a new personalized assessment, merging the original questions with those from different verified assessments available in the corresponding skill's library of assessments. If workers can also answer the newly merged questions, the assessment is verified and can be later on used on its own.
   - Automatic crowd-wise verification: Workers who have the corresponding skill in their skill profile, can verify the assessment and earn a higher reputation. Note that, extra measures need to be taken, to avoid workers who maliciously aim at boosting their reputation by creating spam assessments and reporting them later as spam. Accordingly, unlike the task assignment, the workers are automatically assigned a random assessment, rather than choosing one.

Until a personalized assessment is verified, workers are allowed to take. If the workers suspect the assessment to be a spam, they must report it. If not, they may take it and acquire a pending-verification skill in their profile upon passing the assessment. When the assessment is verified, all the corresponding pending-verifications skills are updated. If the assessment was merely spam, workers' who failed to report the assessment as such are penalized, and the corresponding pending-verification skill is revoked.

4.4 Skill aligner

Upon creating a task with a set of prerequisite skills, a requestor can choose to either use one of the available skills in the ontology or define a new one. Choosing one of the skills in the ontology can be a tiresome job, especially since the ontology grows with the needs of the crowdsourcing platform. Ideally, a requestor should be able to quickly see whether the required skill is available in the ontology or not. To that end, a taxonomy can be maintained on top of the ontology, which the requester can quickly traverse. This taxonomy can be automatically built from the skill description and keywords the requestor inputs upon adding the new skill [31], and validated by the crowd.

4.5 Reputation system

To ensure high quality, only qualified workers should be assigned to the corresponding task. Qualified workers are those workers who: 1) have the required skills, 2) are willing and motivated to complete the tasks, 3) are available, and 4) are highly reputable. Each of those can be respectively measured as follows.

1. **Skill profile:** The skill profile holds the worker's list of skills. The profile acts as a primary filter, where only those workers having a task's prerequisite set of skills are considered. A Skill profile can hold two types of skills: 1) verified skills, 2) pending-verification skills. For every skill in the worker's profile, a list of all the tasks that the worker utilized the corresponding skill in are compiled. Furthermore, an accompanying score is attached, reflecting this experience. This score can be derived from the compiled lists of tasks. Only completed tasks with a positive feedback are listed i.e. requestor was satisfied with the worker. Completed tasks with a negative feedback, are only reflected in the skill's score. This gives a chance for the worker to improve his skill, without having a permanent black spot in their skill profile.
2. **Willingness:** A worker's willingness can be captured from his crowdsourcing platform activity, the following can be observed: 1) time needed to finish a job versus that set by the requestor as the optimal processing time for the task to be done, 2) ratio of completed to aborted tasks.
3. **Availability:** A worker's availability can also be captured from the worker's activity log on the crowdsourcing platform, by specifically noting the number of hours the worker logs per day or month. The time zone a worker is in plays an important role, for urgent tasks i.e. assigning workers with different time zones to the requestor's saves time, where requested tasks can be simply finished overnight.
4. **Reputation:** A worker's reputation can be derived from the average requestor's satisfaction. Moreover, the worker's reputation is penalized, when a pending-

verification skill proved to be spam. In addition to such a penalizing system, a reward system can also be in place e.g. Workers contributing in the automatic crowd-wise verification of assessments.

4.6 Task assigner

Initially only those workers with the required skills are considered for a task. A ranking based on the combination of the willingness, availability and reputation measures is then provided. The three measures are by default equally weighted. The requestor can however choose to give higher weight for any of those measures. E.g. availability is more critical than willingness. A requestor may also choose to completely disregard any of the measures e.g. availability is of no importance. Ultimately, workers exceeding the quality threshold defined by the requestor are assigned to the task. Furthermore, responses of workers with higher ranking are given a higher weight.

5 Workflow of the skill-ontology based model

The skill-ontology based model's workflow can be functionally broken down into: requestor-side, platform-side and worker-side for ease of illustration as follows. Figure 1 gives a graphical overview of the various components of the model as well as the system's interactions.

1. **Requestor-side:** After the requestor designs the task according to his needs, the list of skills required for that task has to be specified. To that end, the requestor checks the taxonomy of skills provided by the platform. When the required skill is found, the requestor simply adds it in the task's list of required skills. Checking the skill's library of assessments, the requestor chooses the assessments he approves and deems eligible for the task's requirements. If no such assessment is found, the requestor is free to design an assessment of his own, which is then added to the skill's library of assessments as an unverified assessment. On the other hand, if the requestor never finds the required skill from the start, he can add a new one along with at least one assessment. The new skill is initially added to the temporary Ontology. If the defined assessment is a standard assessment, no verification is needed, otherwise it's added as an unverified assessment. In addition to the list of required skills, the requestor defines a threshold for the worker's quality to be employed, as well as the measures of quality (willingness, availability, reputation) he wants to consider and their corresponding weights of importance.
2. **Platform-side:** The platform maintains at the back-end two ontologies: Temporary and Basic ontology. On the requestor's front end, a view that combines both ontologies is provided. The front-end ontology may or may not reflect the basic ontology at a given time, and may include both verified and unverified skills. Popular unverified skills that are in the temporary ontology are merged with the basic ontology upon verification. Every skill is associated with a library of assessments that holds either standardized assessments and/or personalized assessments. Furthermore, the platform maintains a database of workers, associating each work with a

profile of skills (verified, pending verification) along with their computed measures of quality.

3. **Worker-side:** A worker is free to choose the tasks he wants to be considered for. Only when his skill profile contains the required skills for the corresponding task is he considered for the task. A worker can at any time expand his skill profile, by sitting assessments and attaining new skills. Workers may also boost their reputation by: 1) verifying personalized assessments 2) validating the platform's generated skill taxonomy.

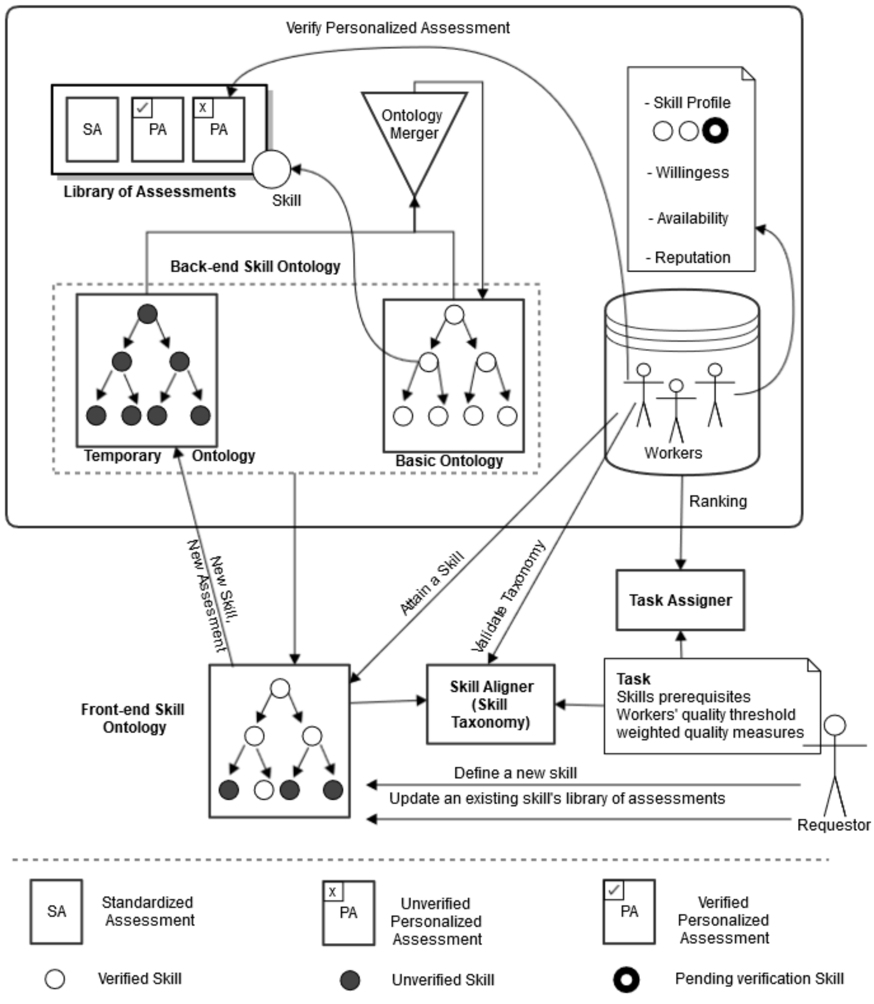


Fig. 1. Skill-Ontology based model workflow

6 Summary & Outlook

Uncertainty is inevitable when dealing with crowdsourcing results. We defined different aspects of quality to identify the corresponding quality assurance measures that should be present. Next, we proposed a skill ontology-based model to be adopted by crowdsourcing platform as a management technique. At its core, the model diminishes the existing uncertainty by eliminating unqualified workers. This is attained by maintaining a dynamically evolving ontology of skills, with libraries of standardized and personalized assessments for awarding credentialed skills. After aligning a worker's set of skills to that required by a task, the resulting quality is improved, where only qualified workers are assigned to the task. Furthermore, in such a setup, qualified workers strive to maintain their qualification level, and unqualified workers are pushed to diligently work on improving their own qualifications. We investigated the model and its workflow on a top level, however, the feasibility of maintaining such a model needs to be further investigated. As examined in the related work section, our model is closely related to web services, reputation-based systems and competency models. Further literature needs to be thoroughly examined, and accordingly adapted to leverage the current model. Furthermore, the proposed workers' quality measures that's to be computed should be formally defined.

7 References

[1] J. Surowiecki, *The Wisdom of Crowds*. Anchor, 2005, p. 336.

[2] J. Howe, "The Rise of Crowdsourcing," *North*, vol. 14, no. 14, pp. 1–5, 2006.

[3] D. C. Brabham, "Crowdsourcing as a Model for Problem Solving: An Introduction and Cases," *Convergence: The International Journal of Research into New Media Technologies*, vol. 14, no. 1. pp. 75–90, 2008.

[4] J. Kamps, S. Geva, C. Peters, T. Sakai, A. Trotman, and E. Voorhees, "Report on the SIGIR 2009 workshop on the future of IR evaluation," *ACM SIGIR Forum*, vol. 43, no. 2. p. 13, 2009.

[5] D. Zhu and B. Carterette, "An Analysis of Assessor Behavior in Crowdsourced Preference Judgments," in *SIGIR 2010 Workshop on Crowdsourcing for Search Evaluation*, 2010, no. Cse, pp. 17–20.

[6] P. Marsden, "Crowdsourcing," *Contagious Magazine issue No.18*, 2009. .

[7] C. Lofi, J. Selke, and W.-T. Balke, "Information Extraction Meets Crowdsourcing: A Promising Couple," *Datenbank-Spektrum*, vol. 12, no. 1, 2012.

[8] L. I. Kuncheva, C. J. Whitaker, C. A. Shipp, and R. P. W. Duin, “Limits on the majority vote accuracy in classifier fusion,” *Pattern Anal. Appl.*, vol. 6, no. 1, pp. 22–31, Apr. 2003.

[9] G. Kazai, “In Search of Quality in Crowdsourcing for Search Engine Evaluation,” *SIGIR Forum*, vol. 44, no. 2, pp. 165–176, 2011.

[10] W. Mason and D. J. Watts, “Financial incentives and the ‘performance of crowds’,” *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 2. p. 100, 2010.

[11] D. C. Brabham, “MOVING THE CROWD AT THREADLESS,” *Information, Communication & Society*, vol. 13, no. 8. pp. 1122–1145, 2010.

[12] “PodCastle: Collaborative Training of Acoustic Models on the Basis of Wisdom of Crowds for Podcast Transcription - INTERSPEECH2009

[13] M. Goto and J. Ogata, “Podcastle: recent advances of a spoken document retrieval service improved by anonymous user contributions,” in *Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech 2011)*, 2011, no. August, pp. 3073–3076.

[14] D. Schall, *Service-Oriented Crowdsourcing: Architecture, Protocols and Algorithms*. Springer, 2012, p. 105.

[15] C. Lai, “Endorsements, Licensing, and Insurance for Distributed System Services,” *The Journal of Electronic Publishing*, vol. 2, no. 1. 1996.

[16] H. Ludwig, A. Keller, A. Dan, and R. King, “A service level agreement language for dynamic electronic services,” *Proc. Fourth IEEE Int. Work. Adv. Issues E-Commerce Web-Based Inf. Syst. (WECWIS 2002)*, 2002.

[17] V. M. Akhil Sahai, Anna Durante, “Towards Automated SLA Management for Web Services - HPL-2001-310R1,” 2002.

[18] M. Allahbakhsh, B. Benatallah, A. Ignjatovic, H. R. Motahari-Nezhad, E. Bertino, and S. Dustdar, “Quality Control in Crowdsourcing Systems: Issues and Directions,” *IEEE Internet Comput.*, vol. 17, no. 2, pp. 76–81, 2013.

[19] N. F. Mohammad Allahbakhsh, Aleksandar Ignjatovic, Boualem Benatallah, Seyed-Mehdi-Reza Beheshti, Elisa Bertino, “Reputation management in crowdsourcing systems.” 2012.

[20] A. Ignjatovic, N. Foo, and C. T. L. C. T. Lee, “An Analytic Approach to Reputation Ranking of Participants in Online Transactions,” *2008 IEEE/WIC/ACM Int. Conf. Web Intell. Intell. Agent Technol.*, vol. 1, 2008.

[21] Z. Noorian and M. Ulieru, “The State of the Art in Trust and Reputation Systems: A Framework for Comparison,” *Journal of theoretical and applied electronic commerce research*, vol. 5, no. 2. 2010.

[22] A. P. Dawid and A. M. Skene, “Maximum likelihood estimation of observer error-rates using the EM algorithm,” *J. R. Stat. Soc. Ser. C Appl. Stat.*, vol. 28, no. 1, pp. 20–28, 1979.

[23] V. C. Raykar, S. Yu, L. H. Zhao, G. H. Valadez, C. Florin, L. Bogoni, and L. Moy, “Learning From Crowds,” *J. Mach. Learn. Res.*, vol. 11, pp. 1297–1322, 2010.

[24] J. Whitehill, P. Ruvolo, T. Wu, J. Bergsma, and J. Movellan, “Whose Vote Should Count More: Optimal Integration of Labels from Labelers of Unknown Expertise,” *Adv. Neural Inf. Process. Syst.*, vol. 22, no. 1, pp. 1–9, 2009.

[25] P. G. Ipeirotis, F. Provost, and J. Wang, “Quality Management on Amazon Mechanical Turk,” in *New York*, 2010, pp. 0–3.

[26] M. a. Campion, A. a. Fink, B. J. Ruggeberg, L. Carr, G. M. Phillips, and R. B. Odman, “Doing Competencies Well: Best Practices in Competency Modeling,” *Pers. Psychol.*, vol. 64, no. 1, pp. 225–262, 2011.

[27] J. S. Shippmann, R. A. Ash, M. Battista, L. Carr, L. D. Eyde, B. Hesketh, J. Kehoe, K. Pearlman, E. P. Prien, and J. I. Sanchez, “The practice of competency modeling,” *Pers. Psychol.*, vol. 53, pp. 703–740, 2000.

[28] J. L. De Coi, E. Herder, A. Koesling, C. Lofi, D. Olmedilla, O. Papapetrou, and W. Siberski, “A model for competence gap analysis,” in *WEBIST 2007: Proceedings of the Third International Conference on Web Information Systems and Technologies*, 2007, pp. 304–312.

[29] S. Colucci, T. Di Noia, E. Di Sciascio, F. M. Donini, M. Mongiello, and M. Mottola, “A formal approach to ontology-based semantic match of skills descriptions,” *J. Univers. Comput. Sci.*, vol. 9, no. 12, pp. 1437–1454, 2003.

[30] K. Koeppen, J. Hartig, E. Klieme, and D. Leutner, “Current Issues in Competence Modeling and Assessment,” *Zeitschrift für Psychologie / Journal of Psychology*, vol. 216, no. 2. pp. 61–73, 2008.

[31] X. Liu, Y. Song, S. Liu, and H. Wang, “Automatic taxonomy construction from keywords,” *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, pp. 1433–1441, 2012.

Towards Mobile Sensor-Aware Crowdsourcing: Architecture, Opportunities and Challenges

Jiyin He[1], Kai Kunze[2], Christoph Lofi[3], Sanjay. K. Madria[4], and Stephan Sigg[5]

[1] Centrum Wiskunde & Informatica, Netherlands
[2] Osaka Prefecture University, Japan
[3] National Institue of Informatics, Japan
[4] Missouri University of Science and Technology, USA
[5] Georg-August University Goettingen, Germany
Emails: { jiyinhe@gmail.com, kai.kunze@gmail.com, lofi@nii.ac.jp, madrias@mst.edu, stephan.sigg@uni-dortmund.de }

Abstract. The recent success of general purpose crowdsourcing platform like Amazon Mechanical Turk paved the way for a plethora of crow-enabled applications and workflows. However, the variety of tasks which can be approached via such crowdsourcing platforms is limited by constraints of the web-based interface. Therefore, in this paper, we propose mobile user interface clients. Switching to mobile clients has the potential to radically change the way crowdsourcing is performed, and allows for a new breed of crowdsourcing tasks. Here, especially the ability to tap into the wealth of precision sensors embedded in modern mobile hardware is a game changer. In this paper, we will discuss opportunities and challenges resulting from such a platform, and discuss a reference architecture.

Keywords: Mobile Platforms, Sensor-Enabled Crowdsourcing, Location-Aware Crowdsourcing

1 Introduction

Crowdsourcing has become a popular approach to many problems that cannot be easily addressed by automated methods and algorithms, or problems that explicitly require significant amounts of human intelligence or human feedback. Crowdsourcing can often be found in knowledge processing tasks such as data or media classification [8], data acquisition tasks such as data completion [6] or information extraction [15], as well as in providing training data for machine-learning-based approaches [17]. Furthermore, crowdsourcing has proven to be useful to the research community for performing large-scale user studies for evaluating new prototype implementations [11], or performing surveys with a large and diverse number of participants for investigating general human behavior or preferences [1]. Instead of laboriously growing own custom crowdsourcing platforms, these tasks mostly rely on general purpose crowdsourcing platforms such as Amazon Mechanical Turk, CrowdFlower, or SamaSource. These platforms allow a complex task to be executed by dividing it into many smaller and simpler sub-tasks, i.e., HITs (Human Intelligence Tasks) – the smallest unit of crowdsourceable work, which are the distributed to a human worker pool. Workers are recruited and retained with payment. Hence, in theory such platforms can be used to perform any dividable tasks that require human intelligence. However, most of these services only offer a web-based interface for workers, and therefore tasks are limited to those that can be displayed and solved within a web browser.

In this paper, we propose an alternative architecture for a general-purpose crowdsourcing platform based on mobile as well as PC devices to interact with the worker pool, referred to as “hybrid crowdsourcing platform”. This will increase not only the ease of use and acceptance of workers in an ever more mobile society, but also the utility and the range of possible crowdsourcing tasks for research applications as well as practical application. In particular, the access to GPS locations and mobile sensors will allow novel crowd-based application that have not been possible before. Our contributions in this paper are as follows:

– We motivate and discuss the need and benefits of mobile sensor-enabled crowdsourcing platforms.
– We highlight use cases of our platform, especially in the area of locality-sensitive services and ubiquitous computing.
– We present the design space and the generic architecture of such a platform, and discuss the impact of certain decisions on the system features and usability.

2 Background

Crowdsourcing can lead to significant cost savings [9, 14, 18], improved product quality [2] and acceleration of time to market [3, 4].

However, crowdsourcing also has the potential to mitigate regional differences in the distribution of labor and human resources. Therefore, most previous work on mobile crowdsourcing platforms as for example [5], [7], or [16] focused on societal aspects of crowdsourcing. These approaches have been tailored for developing countries as an alternative source of labor and income. In developing countries, the spread of personal computers and wired internet connectivity is low. However, still many may have access to mobile phones or even mobile internet service. Therefore, the core challenge discussed in these works is how crowdsourcing can be adapted to the low-end hardware commonly available in developing countries, and how gaps in internet connectivity could be covered using SMS or alternative messaging methods.

In contrast, mobile crowdsourcing as discussed in this paper especially focused on exploiting the capabilities of modern, powerful mobile hardware to offer new functionality to crowdsourcing services. Especially the ability to tap into the user’s geo-location or access to high-quality sensor data allows for completely new applications.

3 General Design

We envision a crowdsourcing platform that can be used in a stationary as well as mobile setting. The various instances in the private devices of users are interconnected through a server in the cloud that takes care of the aggregation of responses, ranking, evaluation

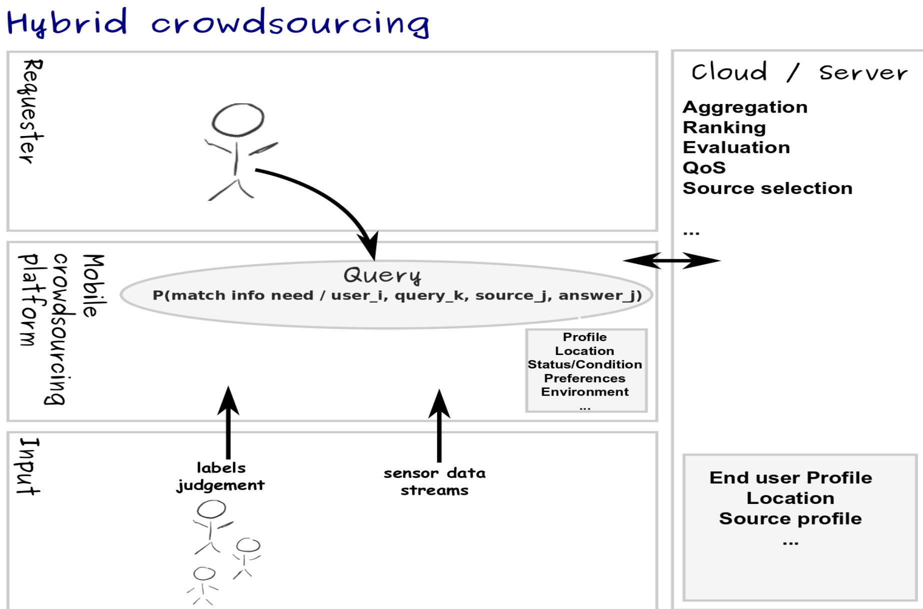


Fig. 1. A proposal of a mobile crowdsourcing platform

and source-selection for a given request. For this purpose, the server stores locations, end user profiles and source profiles among other information (cf. figure 1).

Manual labels and judgements can be harvested as well as sensor data on mobile devices. The requester is likewise part of the crowd as any user or service may issue a query for input of users, services or sensors.

By the combination of mobile, pervasive and crowdsourcing concepts, we will be able to provide crowdsourcing for the masses: A more democratic crowdsourcing usage pattern in which everybody can be crowdsourced or equally state own queries. Mobile crowdsourcing will be seamlessly integrated into daily life with constantly up-to-date, personalised queries that can be completed anytime, anywhere. Instead of playing pointless mobile games to bridge waiting times, people can instead solve interesting queries and even earn money by completing these tasks. Through the integration of context sources in addition to human sources for content provisioning, queries can be highly personalised (e.g. location, environment, condition) and in addition be automatically evaluated for their quality (e.g. fatigue). Such a platform can be exploited to collect huge amounts of labelled sensor data (by asking users to perform certain tasks while being recorded by sensors on the mobile device) from a tightly controlled target population. In addition, it might change the nature of crowdsourcing by empowering ordinary people to set up simple queries that might even reach into their real world (ask people to buy/bring something somewhere). Furthermore, a such a platform might replace traditional data-bases in applications that rely on data which is changing at a high pace. For instance, imagine a dating service, in which a query for a potential partner is not stated to a database of registered users but instead towards the crowd.

4 Opportunities and challenges

State-of-the-art croudsourcing platforms are implemented through web-based services by international players such as Amazon. These platforms require explicit input and reach a maximally diverse population of possible content providers regardless of their location, gender, age, condition or further preferences. However, the result of a request is typically of medium or low quality and requires significant effort to filter out meaningful and quality responses [10]. The integration of crowdsourcing principles with mobile and Pervasive Computing has the potential to disruptively extend the possibilities underlying current crowdsourcing towards, among others, new applications, new classes of data and new possibilities to automatically evaluate quality of responses. We envision a platform with access to implicit information on, for instance, location, condition or further preferences that could restrict a given query to the most intended audience and also utilise sensor information (e.g. fatigue, crowd, loudness level) during the completion of a query in order to automatically estimate the quality of a response. Figure 2 illustrates this concept.

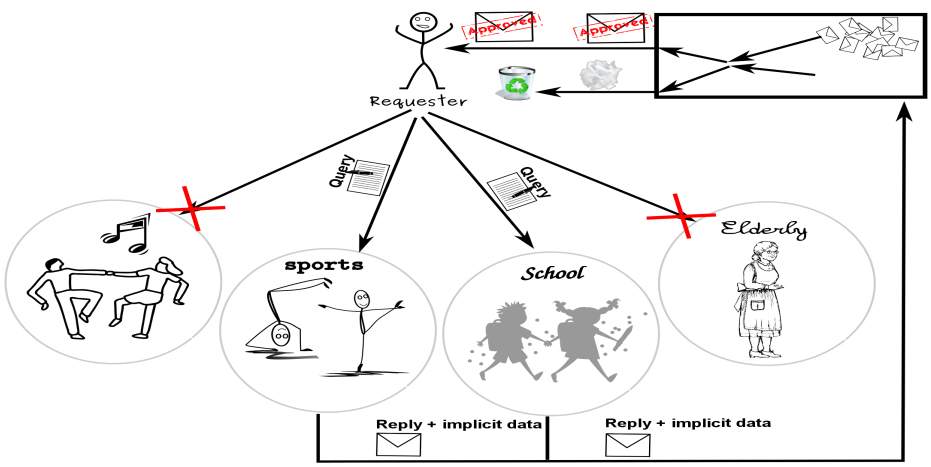


Fig. 2. Concept of mobile crowdsourcing

Expected advantages of a hybrid mobile sensor-aware crowdsourcing paradigm include: (1) improved task performance and efficiency; (2) enabling new crowdsourcing process; and (3) enabling new types of applications. Below, we discuss these aspects, as well as the challenges involved in realising the proposed platform.

4.1 Improved task performance and efficiency

Improved personalisation of request allocation and response aggregation Hybrid mobile sensor-aware crowdsourcing would enable personalised requests filtered by preferences set on mobile devices as well as by dynamic contextual parameters such as location, situation or condition.

Specifically, by maintaining worker profiles including a history of the tasks they have participated, their task performance, as well as the sensor inputs accompanied with this profile, e.g., their location and environment, the system can learn to predict the expertise of the worker, and under which conditions a task may be suitable for that particular worker. For instance, a worker may be able to accomplishing a translating task with high quality in the morning when transiting from home to work, while this performance may decrease in the evening when he/she transits from work to home due to low energy level. Using such information, the system can (1)selectively recommend tasks to target workers, and (2) to selectively return or aggregate worker response to the requester.

Crowdsourcing spontaneous feedback. With a mobile-based platform, both requesters and workers will have less constraints in terms of working locations. This may greatly reduce the time from stating a query to the reception of responses. As a result, responses can be very up-to-date and may include real-time assistance, for instance, in searching/recommendation for pol locations/navigation or spontaneous translation of foreign sentences (e.g. while ordering a menu at a restaurant).

New quality control mechanisms. With sensor data available alongside user input, this data may be utilised to estimate the quality of the provided input. For instance, by analysing the eye-gaze-movement, the platform can estimate fatigue or, reasoning from the loudness level or amount of other people around, which can be used to judge whether the user is impaired in answering questions that require considerable concentration.

Information about situational impacts on cognitive performance. By utilising contextual information, a requester can gain knowledge about the performance of users in various environmental conditions. For instance, by stating a request to several groups of users in various contextual situations, the requester may learn about impacts on cognitive performance. Similarly, by controlling also the situational impacts for a series of queries to several sets of users, the requester can exclude side-effects on the result of a query.

4.2 New crowdsourcing processes

Crowdsourcing for the masses. A crowdsourcing platform on a mobile device, available anytime and anywhere at the convenience of users will change the principle nature of crowdsourcing. Constantly updated, up-to-date and personalised queries can be completed on-demand, interrupted and continued seamlessly. Another aspect is that mobile-based crowdsourcing mitigates hierarchies. Requester and source fall together to the same person as everybody is in the position to state a query. Consequently, quantity of queries will increase while their complexity will fall.

Weakening the strong correlation between labour and human resources. There is a strong relation between the physical location of labour and human resources. While crowdsourcing in general is capable of weakening this correlation, mobile crowdsourcing will further foster this development. In particular, since queries can be more personalised, companies are capable of stating more complex queries also for well-educated workers. This will open new possibilities for workers to offer their workforce without the necessity to relocate.

Participatory Sensing. The envisaged crowdsourcing platform provides access not only to manual input provided by users completing tasks, but also to sensors attached to the mobile platform (Gyroscope, Camera, GPS, Magnetometer, etc.). This might enable, e.g. quick requests for survey purposes even without manual user intervention. Devices and services might extend their contextual perception by harvesting (via automatically answered queries) for sensor information from devices in proximity. Similarly, a mobile crowdsourcing platform may be utilised to acquire labelled sensor data by requesting users to perform specific actions which are then recorded.

4.3 New applications

Mobile crowdsourcing enables new applications for crowdsourcing. For instance, crowdsourcing can replace a database when sensor-based or non-time critical manual feedback is required. There are new challenges introduced by this paradigm as data might then fluctuate in quality and quantity. In addition, crowdsourcing may partly leave the virtual space through a mobile platform. We envision, for instance, an event-hosting company that crowdsources actual manpower on demand. Also, crowdsourcing for educational purposes may serve the need of companies completing actual business-related tasks as well as the need of learners. For instance, a company active in language translations may provide users with text to be translated and later, after collection of all responses, with the corrected aggregated results for educational purposes.

Crowdsourcing as an anonymised customer information system. Mobile crowdsourcing can lower the burden and improve security and privacy in customer information systems. Instead of collecting and maintaining customer-related information for personalised interaction and product design, companies can reach a desired sub-set of customers on demand through mobile crowdsourcing platforms. This will significantly reduce cost and release companies from the burden to maintain huge databases of privacy-critical customer-related information.

Enabling technology for smart cities. A city is defined as smart when investments fuel sustainable economic growth in the respective aspects 'economy', 'mobility', 'environment', 'people', 'living' and 'governance' [12, 13]. A hyrid crowdsourcing platform connects people, government, industry and the environment as all can state queries or provide input to requests stated. Mobile crowdsourcing can therefore serve as an interaction principle in such environments and constitute the backbone of a smart city, interconnecting all major entities.

Mobile crowdsourcing for energy management and smart buildings. Mobile crowdsourcing platform integrates environmental sensors and services. Humans and services acquire maintenance information from infrastructure and surrounding sensors via queries limited by proximity or belonging to a specific entity (building, room, etc). In addition, services can serve as actuators, completing queries designed to control smart buildings and automation. In particular, the controlled entity might change relative to the location of the requester.

4.4 Challenges

High performance data processing and analysis mechanism With a mobile sensor-enabled crowdsourcing platform, we need to be able to process the vast amounts continuously generated explicit user inputs (requests and responses) as well as implicit sensor inputs in real time, e.g., in order to realise the above mentioned personalised request allocation and response aggregation. This requires high performance computational power as well as sophisticated data mining and machine learning algorithms that can scale to this type of data and give spontaneous responses. Further, sensor data as well as user inputs may be noisy. It is non-trivial to extract meaningful features from the raw sensor data as input for machine learning algorithms, or to derive human interpretable results.

Limitations of mobile devices While mobile devices provide great flexibility for people to perform tasks, there is also limitations. These include: the small screen, the limitation of battery life, and the limited types of interactions allowed. For instance, it is less convenient for people to type long sentences in a mobile device compared to that on a PC. With these limitations in mind, dedicated user experience studies need to be carried out while designing and implementing mobile based HITs.

Data security and privacy issues The proposed platform involves collecting data such as a user's location, activities, as well as other personal information measured by the sensors. A major concern is therefore data security and privacy issues. These personal information and mobile users' activities may be disclosed or abused by malicious users, which will threaten the well-being of normal users.

5 Conclusion

In this vision paper, we have discussed a mobile crowdsourcing paradigm which can augment the current web-based crowdsourcing platforms to provide real-time location based query response using mobile devices. Towards this goal, we have provided a hybrid-crowshourcing architecture and discussed several facets to realize this vision. The mobile crowd-sourcing can filter and target the workers who more closely matches not only the queries, but also the location and context requirements. In addition, some of the processing can be done in the centralized web-based part of the proposed architecture, which reduces the burden of processing queries on the mobile devices.

6 References

[1] T. S. Behrend, D. J. Sharek, A. W. Meade, and E. N. Wiebe. The viability of crowdsourcing for survey research. *Behavior research methods*, 43(3):800–813, 2011.

[2] E. Bonabeau. Decisions 2.0: the power of collective intelligence. *MIT Sloan Management Review*, pages 45–52, 2009.

[3] I. Borst. *Understanding Crowdsourcing – effects of motivation and rewards on participation and performance in vountary online activities*. PhD thesis, Erasmus

research institute of Management, Rotterdam School of Management, Erasmus School of Economics, Erasmus University Rotterdam, 2010.

[4] J. V. den Ende, A. Villarroel, and C. Tucci. Strategic crowdsourcing, orchestrating innovation through the cream of the crowd, 2009. Panel symposium, Academy of Management Conference.

[5] N. Eagle. txteagle: Mobile crowdsourcing. In N. Aykin, editor, *Internationalization, Design and Global Development*, volume 5623 of *Lecture Notes in Computer Science*, pages 447–456. Springer Berlin Heidelberg, 2009. ISBN 978-3-642-02766-6. doi: 10.1007/978-3-642-02767-3_50. URL http://dx.doi.org/10.1007/978-3-642-02767-3_50.

[6] M. J. Franklin, D. Kossmann, T. Kraska, S. Ramesh, and R. Xin. Crowddb: answering queries with crowdsourcing. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, pages 61–72. ACM, 2011.

[7] A. Gupta, W. Thies, E. Cutrell, and R. Balakrishnan. mclerk: Enabling mobile crowdsourcing in developing regions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1843–1852, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1015-4. doi: 10.1145/2207676.2208320. URL http://doi.acm.org/10.1145/2207676.2208320.

[8] J. He, J. van Ossenbruggen, and A. P. de Vries. Do you need experts in the crowd?: a case study in image annotation for marine biology. In *Proceedings of the 10th Conference on Open Research Areas in Information Retrieval*, pages 57–60. LE CENTRE DE HAUTES ETUDES INTERNATIONALES D'INFORMATIQUE DOCUMENTAIRE, 2013.

[9] J. Howe. The rise of crowdsourcing. 14(6), September 2009. http://www.wired.com/wired/archive/14.06/crowds.html.

[10] G. Jouret. Inside cisco's search for the next big idea. *Harvard Business Review*, pages 43–45, September 2009.

[11] A. Kittur, E. H. Chi, and B. Suh. Crowdsourcing user studies with mechanical turk. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 453–456. ACM, 2008.

[12] N. Komninos. *Intelligent cities: innovation, knowledge systems and digital spaces*. Spon Press, 2002.

[13] N. Komninos. Intelligent cities: towards interactive and global innovation environments. *International Journal of Innovation and Regional Development*, 1(4): 337–355, September 2009.

[14] J. Lampel and A. Bhalla. The role of status seeking in online communities: Giving the gift of experience. *Journal of Computer-Mediated Communication*, 122, 2007. http://jcmc.indiana.edu/vol12/issue2/lampel.html.

[15] C. Lofi, J. Selke, and W.-T. Balke. Information extraction meets crowdsourcing: A promising couple. *Datenbank-Spektrum*, 12(2):109–120, 2012.

[16] P. Narula, P. Gutheim, D. Rolnitzky, A. Kulkarni, and B. Hartmann. Mobileworks: A mobile crowdsourcing platform for workers at the bottom of the pyramid. In *Human Computation*, 2011.

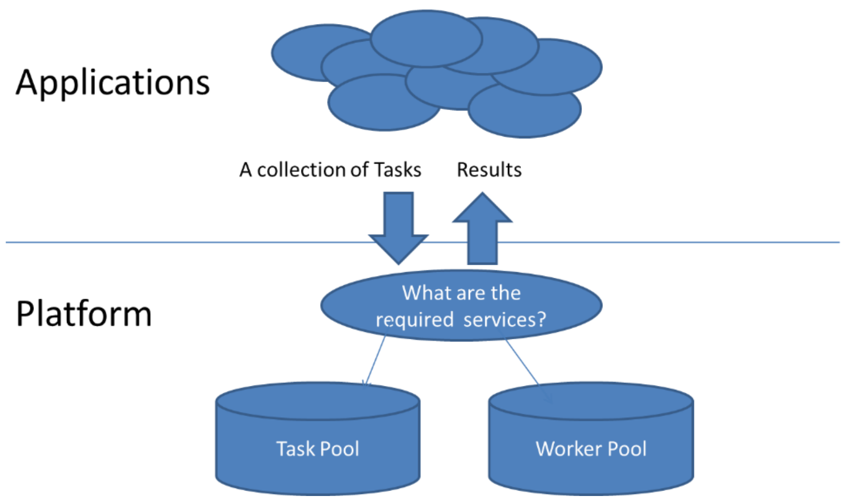
[17] J. Selke, C. Lofi, and W.-T. Balke. Pushing the boundaries of crowd-enabled databases with query-driven schema expansion. *Proceedings of the VLDB Endowment*, 5(6):538–549, 2012.

[18] C. Wu, J. Gerlach, and C. Young. An empirical analysis of open source software developers motivations and continuance intentions. *Information and Management*, 44:253–262, 2007.

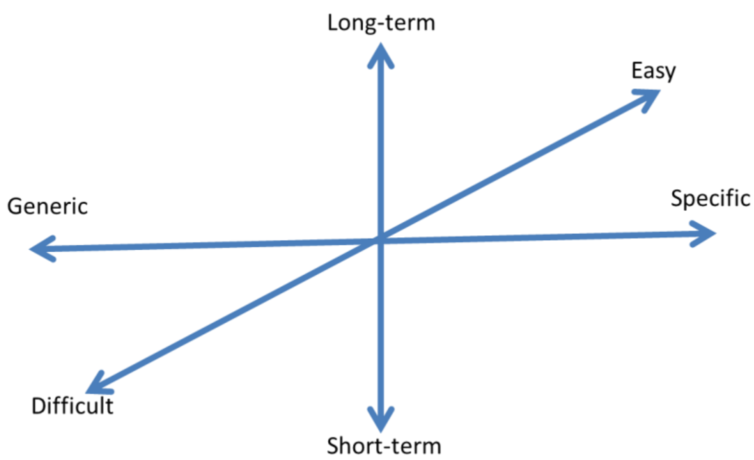
Breakout Group: Applications & Platform Functionality

Koji Zettsu, Koichi Kise, Atsuyuki Morihsima

In crowd-sourcing system, platforms connect worker pools and task pools with the applications.



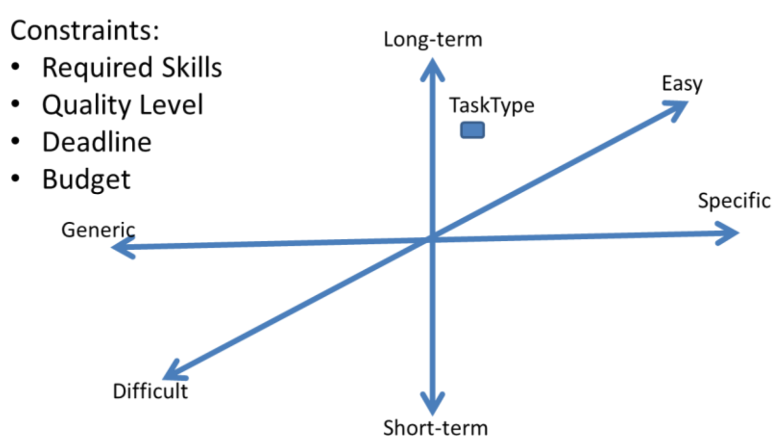
Crowd-sourcing tasks can be identified along 3 major dimensions:



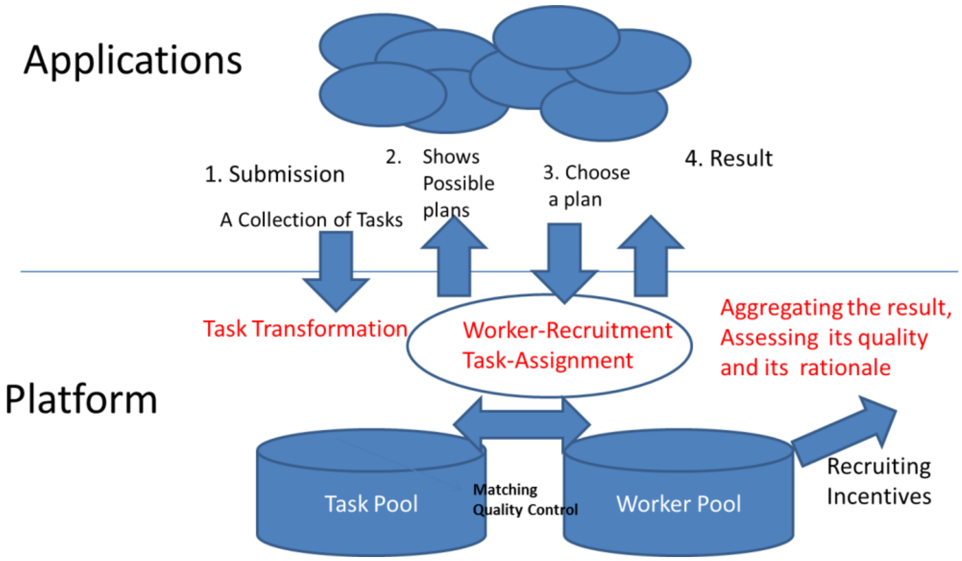
Applications on the other hand can be classified as follows:

	Easy		**Difficult**	
	Generic	**Specific**	**Generic**	**Specific**
Short-term	- Collection of Personal Data - Tagging (generic terms) - Sheep Market - reCapcha - ESP game	- Context-dependent data collection (locality, gender, etc.)	-Easy but real-time processing	- Matching two sentences (in different languages) - Solve mathematical problems - Translation
Long-term	- Logging Behavior - Constructing a large set of data	- Logging Behavior and constructing a set of data (Context-dependent)	-Long-term real-time processing - Searching Balloons	- Research (e.g., innocentive)

A Possible Abstraction: Task Types, Constraints, and the position:



A simple **execution model** in four phases and three functions was developed:



Here, important functions are:
Task Transformation: Function to replace the given microtasks with different set of microtasks that are more tractable
Worker-Management and Task-Assignment: generating plans combining a variety of methods to recruit and motivate workers, to assign tasks to them, and to improve quality of the results
Aggregating the results with the explanation: Function to explain how the system aggregated the result assessed its quality and its rationale

Task Transformation: Replacing given tasks with more tractable ones

Difficult to Easy

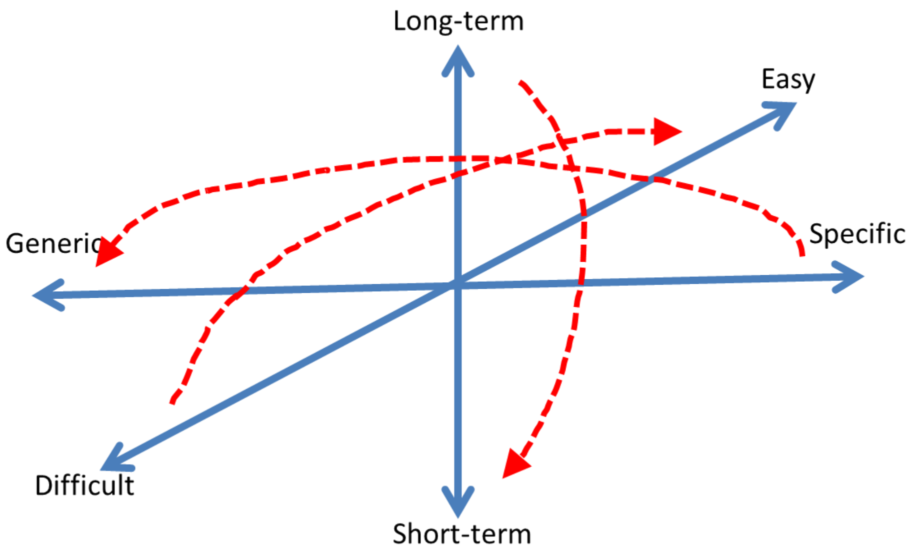
- Assign the same task to many people (achieve real-time processing by distributing the task to many people)
- Changing data-entry tasks to tasks that asks the worker to choose appropriate one from a set of candidates.

Long-term to Short-term

- Divide the problem to small ones (finding a person in a large map, collecting restaurants in a large city, translation)

Specific to Generic

- Remove contexts
- Translate the problem to isomorphic one.



Worker Management and Task Assignment:
Choosing appropriate methods for the given tasks

Matching => can be supported by platforms based on the required skills and deadline

- Worker-driven
    - Workers search for tasks
- System-driven
    - profile-based matching
    - Recommendations by workers (crowdsourcing recruiting)
    - Advertisement
    - Push for urgent tasks

Incentives

- Explaining the significance of tasks: (for some tasks)
- Gamification
    - Ranking, scores
    - Game structure for particular types of tasks
- Paid (fixed payment, merit-based reward)
    - System can give a function to help determine how to pay based on the difficulty and the deadline
- Piggyback functions for particular classes of tasks
    - System can incorporate tasks into other systems based on their types

Worker Management and Task Assignment:
Choosing appropriate methods for the given tasks

Before performing tasks

- Determine the number of duplicate tasks
- Determine appropriate task design and incentives
- Choosing an appropriate set of workers
    - Reliability
    - Skill-based, context-based
    - Qualification Test
    - Gold standard

After receiving the results of tasks

- Removing spammers
- Aggregating the results
- Choosing good results

Aggregating the Results with the explanation:
Explaining how the System Aggregated the Results and Assessed the Quality

- Returns aggregated results and quality-assessment values and its rationale.
- Provides the requesters with the means to verify the quality of the results.

Breakout Group 3: Incentives of Crowdsourcing

Nestor Alvaro, Shigeo Matsubara, Yoshito Tobe, Xuan Zhou

Types of Incentives

Dimension 1

- Financial, e.g. payment
- Social, e.g. reputation
- Moral, e.g. helping others in need

Dimension 2

- Extrinsic, e.g. money, social approval
- Intrinsic, e.g. fun, knowledge, moral satisfaction

What is specific in crowdsourcing?

- Human behaviors coordinated by computer
    - No / Less human interaction or relationship
- Splitting complex tasks into small pieces
    - Less expenses for requesters
    - Less requirements on skills for workers
- Small amount of money for each task
    - Different significance for people from different regions, e.g. US vs developing countries

Factors to be considered in Incentive Mechanism of CS

- Characteristics of Workers
    - Region, Skills
- Types of Tasks
    - Short-run vs. long-run
    - Extrinsic reward vs. Intrinsic reward
    - Simple tasks vs. Tasks requiring skills
        - Requester needs to trust the skills of the workers

Incentive Techniques

- Get more workers
  - Good formation of tasks (clear description, easier to do)
  - Pay more money, fairly
  - Acquirement of knowledge and skills
- Get tasks done faster with better quality
  - How to pay? E.g. bonus
  - Reputation – Turk uses approval rate
- Get better retention
  - How to pay? E.g. increase the payment gradually
  - Reputation. E.g. retention rate / gold members

Incentive Techniques based on Payment

- Bonus Mechanism
  - e.g., Split payment into two parts X+Y, suppose N tasks, pay X/N to each task, and pay Y to the best workers as bonus
  - Conditions
    - Workers care about bonus. (US vs. developing countries)
    - Transparency – workers know who get bonus and why
- Incremental Payment Mechanism
  - E.g., Split tasks into multiple rounds, increase the payment slightly after each round
  - Improving Retention – provide incentives for workers to stay

Incentive Techniques based on Social Recognition

- Reputation
  - Approval Rates for workers
  - Rejection Rates for requesters
  - Retention Rates for workers
  - Bonus Rates for workers
  - Other history information …
- Contest / Ranking
  - For each type of tasks (e.g. translation), maintain a rank list of best workers
  - Endorsement Mechanism, e.g. Best translator
  - Fosters competition among workers
- Connection of worker ID to person ID
  - Inter-person influence
  - Related to real-life profile
  - Issues of privacy

Issues to be Investigated

- Different incentives work for different people.
  - How to find the best mix of incentives?
  - How to find out the types of workers?
- Standard ways to post the tasks, so that workers know what they get.
- How to help requesters find right incentive mechanisms?
  - In the payment related mechanisms, how to determine the parameters, such as % of bonus and % of increase rate?
  - A worker can excel in one type of tasks but sucks in another. How to deal with his reputation?
- Recommendation
  - (workers < > tasks)
  - (requesters < > workers)