The Knowledge Engineering Review, Vol. 33, e2, 1–19.
doi:10.1017/S0269888918000061

A survey of incentive engineering for crowdsourcing

CONOR MULDOON[1], MICHAEL J. O'GRADY[2] and GREGORY M. P. O'HARE[3]

[1]*School of Computer Science, University College Dublin, Belfield, Dublin 4, Ireland;*
e-mail: conor.muldoon@ucd.ie;
[2]*School of Computer Science, University College Dublin, Belfield, Dublin 4, Ireland;*
e-mail: michael.j.ogrady@ucd.ie;
[3]*School of Computer Science, University College Dublin, Belfield, Dublin 4, Ireland;*
e-mail: gregory.ohare@ucd.ie

Abstract

With the growth of the Internet, crowdsourcing has become a popular way to perform intelligence tasks that hitherto would be either performed internally within an organization or not undertaken due to prohibitive costs and the lack of an appropriate communications infrastructure. In crowdsourcing systems, whereby multiple agents are not under the direct control of a system designer, it cannot be assumed that agents will act in a manner that is consistent with the objectives of the system designer or principal agent. In situations whereby agents' goals are to maximize their return in crowdsourcing systems that offer financial or other rewards, strategies will be adopted by agents to game the system if appropriate mitigating measures are not put in place. The motivational and incentivization research space is quite large; it incorporates diverse techniques from a variety of different disciplines including behavioural economics, incentive theory, and game theory. This paper specifically focusses on game theoretic approaches to the problem in the crowdsourcing domain and places it in the context of the wider research landscape. It provides a survey of incentive engineering techniques that enable the creation of apt incentive structures in a range of different scenarios.

1 Introduction

Crowdsourcing platforms (Higgins *et al.*, 2016) enable the performance of tasks by large pools of agents through the use and promotion of incentives of various forms, such as financial rewards, recognition for altruistic behaviour, or aiding a cause deemed worthy by the contributors. Numerous challenges exist around crowdsourcing, as described by O'Grady and O'Hare (O'Grady *et al.*, 2016), not least incentivising citizens. Incentive Engineering is the process by which system developers design incentive structures such that certain desirable properties of system behaviour or outcomes are more likely; the underlying assumption is that agents participating within the system are rational and will thus behave in a manner such as to maximize their utility (see Section 3). There is a rich body of research in the motivational and incentivization sphere that incorporates techniques from several disciplines including behavioural economics, incentive theory, and game theory. In this paper, the focus is specifically on crowdsourcing systems that adopt a game theoretic approach. The reason for focussing on game theory is that, in contrast to alternative disciplines, it enables the creation of Dominant Strategy Incentive Compatible (DSIC) games (see Section 3.2) that align user incentives with system objectives. With alternative behavioural or motivational approaches, for instance, the incentives will be such that the system will be susceptible to manipulation or gaming by users or agents in maximizing their rewards. The paper represents a focussed review (Cooper, 1988) rather than a systematic mapping (Petersen *et al.*, 2015) that identifies the number of papers on subtopics and how such papers might, for example, be categorized. The aim is to provide a critical analysis of the techniques covered, rather than to undertake a statistical analysis to identify gaps in the literature that warrant further research. Systems covered in the paper were chosen as representative, but not exhaustive, applications of game theoretic research within crowdsourcing.

An early use of the term incentive engineering can be found in Drexler and Miller (1988), although adopting an engineering approach to designing incentives in game theory predates this (Myerson, 2013). There are several different vistas through which the problem of incentive engineering may be viewed. For large or complex crowdsourcing tasks that cannot be completed by a single contributor, incentive structures that promote collaboration and the revelation of private information in an incentive compatible or truthful scheme are necessary. Orthogonal to incentive compatibility, are procedures that assume information regarding agents' preferences is known and that modify incentives given the agents' preferences so as to perturb behaviour (Endriss *et al.*, 2011).

Much work within incentive engineering draws from game theory and mechanism design (Myerson, 1981). Game theory is the mathematical study of strategic interactions between multiple rational decision-makers (see Section 3), whereas mechanism design is a sub-field of game theory that may be viewed as the science of designing the rules of games such that agents' behaviour is aligned with the system designers' objectives and desirable outcomes are achieved. Online algorithms are relevant to mechanism design as, in situations where incentives to collaborate or act in a manner consistent with the system designer's objectives exist, certain parameters will invariably need to be determined at run-time. Reasons for this include coordination problems, such as the unavailability of knowledge about the future or current state of the world, or delays regarding information exchange between distributed decision-makers.

In addition to game theory mechanisms, several heuristic and qualitative approaches have been developed that exhibit favourable behaviour within restricted domains. Often, problems within this space are computationally hard and, as such, simplifying assumptions must be adopted to create algorithms that have bounded memory and computational requirements. The complexity problem is further exacerbated in situations whereby incentive structures or parameters of the adopted model vary over time and algorithms must operate in an online manner. For certain classes of problem, approximate algorithms can be created with analytic bounds on the run-time/memory requirements whilst maintaining the desirable properties of designed games.

Current proprietary crowdsourcing frameworks, such as Amazon Mechanical Turk (MTurk), do not use game theory mechanisms in considering the strategic implications of providing incentives for performing tasks. In this survey, results from the literature are identified that could potentially be applied to commercial platforms to improve their performance with respect to the global utility or social surplus of agents or the utility of the framework owner or operator.

The remainder of the paper is structured as follows. Section 2 gives a broad overview of crowdsourcing and provides a definition of how the term is used in the context of this paper. Much of the research within this area is based on game theory and a brief overview of the key concepts is presented in Section 3. A taxonomy that classifies the incentivization mechanisms and enables the comparison and contrast of the various techniques is given in Section 4. Offline applications of game theory and mechanism design within the crowdsourcing domain are discussed in Section 5. Online algorithms that are used in situations whereby parameters of the models must be determined, or must vary over time, as the system is in operation are discussed in Section 6. A discussion is provided in Section 7 and Section 8 concludes the paper.

2 Crowdsourcing

Crowdsourcing may be regarded as a practical implementation of distributed problem solving. Traditionally, its use was primarily restricted to Internet-based platforms such as the Amazon MTurk. More recent platforms explore the potential of mobile computing, typified by smartphones, resulting in Mobile CrowdSensing and Computing (Guo *et al.*, 2015) where both sensing and computing (e.g. interpretation) occur on the device. In See *et al.* (2016b), other interpretations and variations of crowdsourcing or collective sensing are identified.

Crowdsourcing has been demonstrated in a variety of domains; in the environmental monitoring domain alone, examples include generating landcover maps (See *et al.*, 2015), climatology (Muller *et al.*, 2015), invasive species (Adriaens *et al.*, 2015), and earth observation (See *et al.*, 2016a). Notwithstanding the popularity of the field, the issue of data quality is frequently cited by critics as a major limitation (Senaratne *et al*., 2017). Ghezzi *et al*. (2017) provide a recent snapshot of the state-of-the-art and future directions. Aligned with applications in the environmental monitoring domain are spatial and context-aware crowd-sourcing approaches. The underlying concept here is that spatial, temporal, and other constraints are defined with regard to the tasks undertaken by the crowd. Assigning time-constrained spatial tasks to mobile workers such that the tasks can be accomplished with high reliability and spatial/temporal diversity is a problem that is non-deterministic polynomial (NP) time hard; approximate solutions must therefore be considered (Zhao & Han, 2016). Alternatively, a sparse mobile crowdsensing approach could be adopted that leverages the spatial and temporal correlation among the data sensed in different sub-areas to significantly reduce the required number of sensing tasks allocated. This approach reduces the overall sensing cost, while ensuring data quality (Wang *et al*., 2016).

The term crowdsourcing was first coined by Howe in 2006 (Howe, 2006b) in discussing how companies and corporations were increasingly using the crowd, through Internet-based platforms and intelligent networks, for cheap labour in creating content, solving problems, and undertaking corporate research and development. Several examples illuminate this; these include the use of iStockPhoto for crowdsourcing images, Colgate-Palmolive and Procter and Gamble's use of InnoCentive to find research and development solutions for problems that could not be solved with in-house research teams, and the use of MTurk for solving human intelligence tasks (HITs) and technical support problems.

In Howe (2006a), a distinction between crowdsourcing and commons-based peer production (Benkler & Nissenbaum, 2006) is made, although it is noted that there is some common overlap between the concepts. Crowdsourcing can take the form of peer production when the work is performed collaboratively by a group of users. Within crowdsourcing, however, the scope is boarder, and users also often carry out work on their own on tasks that do not require collaboration. In terms of collaborative work and commons-based peer production projects, investment by companies in open-source software, such as IBM's investment of over $1 billion to support the Linux kernel and the Apache Web Server, could be viewed as a form of crowdsourcing, but to a commons project. Ownership of the product is not sought in a commons project; rather value may be gained indirectly through improving the prospects of other products. With crowdsourcing, ownership of the product may be sought. In Howe (2006a), crowdsourcing is defined as:

> *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.*

Further, it is noted that the term is often used without reference to the original article (Howe, 2006b) and its meaning is largely determined or defined by the crowd. This is true of all terms if we are to take lexical or dictionary definitions as these represent how a population commonly uses a term. Lexical definitions can only be considered correct or incorrect regarding the degree to which they reflect usage.

One of the limitations with Howe's definition is that it states the function must have been once performed by employees. This limits the term from being used within the context of new applications that would not have been feasible with pre-existing technology; an exemplar application domain being citizen science. In the context of this paper, the stipulative definition proposed by Brabham (2013) is adopted, which defines crowdsourcing as:

> *an online, distributed problem-solving and production model that leverages the collective intelligence of online communities to serve specific organizational goals.*

In David *et al*. (2012), four fundamental types of crowdsourcing information system are identified, namely crowd processing systems, crowd rating systems, crowd solving systems, and crowd creation systems.

Crowd processing systems require large volumes of homogeneous contributions. All valid contributions are considered qualitatively equal and deliver the same value. Such systems often adopt a divide and conquer approach whereby large jobs are divided up into equal chucks of work or microtasks. Examples of crowd processing systems include Camclickr, Galaxy Zoo, Recaptcha, and VizWiz.

Crowd rating systems also require large volumes of contributions, but value is derived from the collective value that arises from the totality of contributions made and their relationships. Individual contributions represent votes of the participants. Such systems are used to collect online ratings, for example, in opinion panels and prediction markets. Examples of crowd rating systems include the eBay reputation system, eRewards, Hollywood Stock Exchange, and Crowdcast.

Crowd solving systems seek value from heterogeneous contributions of a diverse group of participants. The value of each individual contribution is determined by a given criteria that can vary greatly with the task at hand. Crowd solving systems benefit from larger and more diverse crowds as each contribution potentially increases the value. This is due to a phenomenon that is often referred to as the 'wisdom of the crowd'. The Netflix Prize, FoldIt, Kaggle, InnoCentive, and Idea Bounty are examples of crowd solving systems.

Crowd creation systems also rely on contributions from a diverse group of participants, but seek value from the accumulation of the contributions and their relationships. With crowd creations systems, a variety of contributions are aggregated into a comprehensive artefact. Examples include the Yahoo! Contributor Network, iStockphoto, and Wikipedia.

3 Game theory and mechanism design

From an incentive engineering perspective, there are many potential applications of game theory and mechanism design in the crowdsourcing domain. This section provides a brief overview of key concepts within game theory and mechanism design in a general context.

3.1 Game theory

Game theory is concerned with the analysis of situations whereby multiple agents interact and their actions affect each others' outcomes. With Game Theory, agents evaluate their potential choices considering the potential choices and utilities of other agents within a given scenario. Myerson (1981) defines game theory as follows:

> *Game theory can be defined as the study of mathematical models of conflict and cooperation between intelligent rational decision-makers. Game theory provides general mathematical techniques for analysing situations in which two or more individuals make decisions that will influence one another's welfare.*

Often, intelligent rational decision-makers, within the game theory literature, are referred to as self-interested agents. Self-interest does not imply that agents are necessarily selfish or are only concerned about actions that improve their own circumstances and that their actions are harmful to others; rather, it implies that the agents have beliefs about what are favourable states in the world and their actions are motivated by the intention to bring about those states. This notion is captured by representing the agents' desired states with utility functions that map states of the world to numeric values that quantify the degree of preferences of agents among states. This enables the impact of uncertainty on preferences be represented and agents to take actions that maximize their expected utility.

In early research on the topic, Von Neumann and Morgenstern (1944) proved a theorem that derives the existence of a utility function from a preference ordering and axioms on such orderings. This demonstrated that a utility function is sufficient to encode preferences over a complicated set of alternatives.

Game theory is divided into two branches, namely cooperative and non-cooperative that differ in how they formalize independence among agents. The terms cooperative and non-cooperative, when used in this context, differ from common vernacular, which would suggest that competition does not occur in the former and cooperation does not occur in the latter, neither of which is the case. Cooperative game theory describes the outcomes that occur when different combinations of agents come together, whereas non-cooperative game theory provides a more detailed model of the options available to individual agents and their individual utility functions. It should be noted, however, that any cooperative game can be converted to a non-cooperative representation. In game theory, the price of anarchy (Koutsoupias & Papadimitriou, 2009; Roughgarden, 2010) can be viewed as a measure of the degree to which competition approximates cooperation in situations whereby agents have access to shared resources. It represents the ratio between the best possible system performance or utility, from the system designer's perspective, that could be obtained if the system designer had complete or full control, as opposed to the performance with a group of strategically motivated autonomous agents (Koutsoupias & Papadimitriou, 2009).

Problems in game theory have different types of solution concepts depending on the type of game being played. A game is said to have a dominant strategy solution if each player has a unique best strategy notwithstanding the strategies adopted by the other players in the game. A dominant strategy does not guarantee that the utilities received by agents will be optimal and, indeed, in the well-known prisoner's dilemma game, this will not be the case. Outside the scope of mechanism design (see Section 3.2), where the goal is to construct games that have dominant strategies that are socially optimal or maximize the revenue for the designer, games that have dominant strategies are rare.

For games that do not have a dominant strategy solution, a more widely applicable solution concept is required. In games where players deterministically choose their moves, a pure strategy Nash Equilibrium represents such a solution concept. A Nash Equilibrium is a stable solution whereby no agent has an incentive to deviate from their current strategy given the strategies of the other agents. Once the players are in Nash Equilibrium, it is in all the players interest to remain in Nash Equilibrium. A problem, however, is that there may not be a unique Nash Equilibrium.

Not all games have pure strategy Nash Equilibria solutions. In games where players can randomize their strategies, however, stable solutions can be obtained in situations where it is assumed that agents are risk neutral and will act in such a manner as to maximize their expected payoff. This type of solution concept is referred to as a mixed strategy Nash Equilibrium. In this case, an agent chooses a probability distribution over their set of possible strategies and independently selects strategies using the probability distribution. Not all games have mixed strategy Nash Equilibria; if there is a finite set of agents and a finite set of strategies, however, the game will have such an equilibrium (Nash, 1951; Nisan *et al*., 2007).

3.2 Mechanism design

Mechanism design is concerned with the problem of formulating the rules of games such that they exhibit desirable outcomes under the assumption that agents are rational and that their behaviour will be in accordance with a specified utility function. One of the key application areas of mechanism design is that of creating games that make minimal assumptions regarding the behaviour models of agents, such as DSIC games. In DSIC games, it is often relatively straightforward for both agents and mechanism designers to reason about what the optimal course of action is and what will occur within the game when compared to other types of games.

Games within mechanism design are described in terms of allocation rules, payment rules, monotone unctions, and implementability. An allocation rule is a function that determines how items are allocated to agents, such as mapping bids within an auction to what agents receive. A payment rule is a function that, given an allocation, determines what agents pay. A function is called monotone if for larger values provided in the domain, larger values are given in the range. For example, in an auction an allocation function would be called monotonic if for larger bids, an agent wins more. For a single parameter environment where the only unknown parameter is that of the agent's private valuations, an allocation rule is said to be implementable if a payment rule for a sealed-bid auction exists such that the sealed-bid auction is DSIC. For the mechanisms discussed in this section, the utility of an agent is defined as a quasilinear function, which is equal to the difference between the value an agent puts on an item minus the price the agent pays for the item if the price is not higher than the valuation. If the price is higher than the valuation, the agent's utility is zero.

A fundamental result in mechanism design is that of Myerson's Lemma (Roughgarden, 2003; Myerson, 2013). Myerson's Lemma states that an allocation rule is implementable if and only if it is a monotone function and, if this is the case, a unique payment rule exists, which is determined by an exact payment formula given in Myerson (2013). Given, from Myerson's Lemma, that implementability and monotonicity are equivalent and that, typically, it is relatively easy to determine if an allocation rule is monotonic, the lemma significantly simplifies the problem of formulating games that are implementable.

The well-known Vickrey second-price sealed-bid auction (Vickrey, 1961) can be viewed as a special case of Myerson's Lemma. In a Vickrey auction, agents submit bids privately to an auctioneer and the price that the winning agent pays is equal to that of the second highest bidder. This provides an incentive for agents to bid in a truthful manner. The reason for this is that if an agent underbids, there is a chance that they will not receive the item if another agent has bid higher. This would be irrational in that the price that the agent would pay for an item, if they had bid truthfully and won, could be less than their valuation in that they only pay that of the second highest bid. If an agent overbids, there is a chance that another agent will make a bid that is higher than the original agent's valuation, but lower than the original agent's bid. In this case, the original agent will pay more for the item than their valuation. Thus, to maximize their expected utility, agents should bid honestly.

The Vickrey auction is useful in situations whereby the objective of the auction is to maximize the social surplus, whereby the winner of the auction is the agent that values the item the most. A modification of the Vickrey auction, using a reserve price, can be used to maximize the expected revenue under assumptions regarding prior distributions of beliefs over agents' valuations, such as that their valuations are drawn from the identical distributions.

Myserson's Lemma is used in the design of DSIC mechanisms in single parameter environments, where agents' private valuations are unknown. For multi-parameter environments, such as combinatorial auctions, whereby agents have different valuations over different items, the Vickrey–Clarke–Groves (VCG) mechanism (Groves, 1973) is used. The main idea in the VCG mechanism is that, when evaluating their reported valuations and the prices potentially paid for items, it forces agents to internalize their externalities or the social cost of their participation in the auction. For instance, if a winning agent did not participate, the utility of other agents would increase in that they would receive items that would have been received by the winning agent. The VCG payment rule quantifies this.

A key problem with using complicated incentive structures in crowdsourcing platforms is that agents are often only presented with the mechanism for a short time period and for a small number of tasks. In conventional platforms, most crowdsourcing tasks have relatively low returns, such as a few euros, and it is for this reason that agents will not spend a significant amount of time considering the implications of a complicated payment scheme. As such, it is expected that the agents often will not fully understand a payment scheme unless it is quite simple, stated clearly and appropriately, or can be intuitively understood. If this is not the case, it cannot be assumed that agents will act precisely as required or behave in a manner that is rationally implied by the rules of the game. Mechanism design aids in addressing this problem by providing a principled approach to the construction of DSIC games that ease the reasoning burden in determining the correct action to take.

4 A taxonomy of game theoretic incentive engineering

To illustrate how the various incentivization mechanisms outlined in this paper are interrelated, this section provides a taxonomy in the form of a directed graph without cycles that represents the hierarchical or inheritance structure and dependencies among the mechanisms and classes of mechanism. As can be seen from Figure 1, there are several approaches to addressing the incentivization problem that draw from several different fields, including behavioural economics, incentive theory, gamification, heuristic approaches, along with game theory. This review is primarily concerned with the use of game theoretic techniques, but a brief overview of the other areas is provided here to put the systems discussed subsequently into a wider research context along with providing arguments for the use of mechanism design and game theory in this type of setting.

4.1 Behavioural economics

A key tenet of game theory is the assumption that agents will act in a rational manner with regard to maximizing their welfare. Behavioural economics takes a different view and, rather than assuming that agents are rational, it makes predictions based on the observed behaviour of people (Kahneman, 2003). It considers the role of cognitive, emotional, psychological, and social factors in the economic

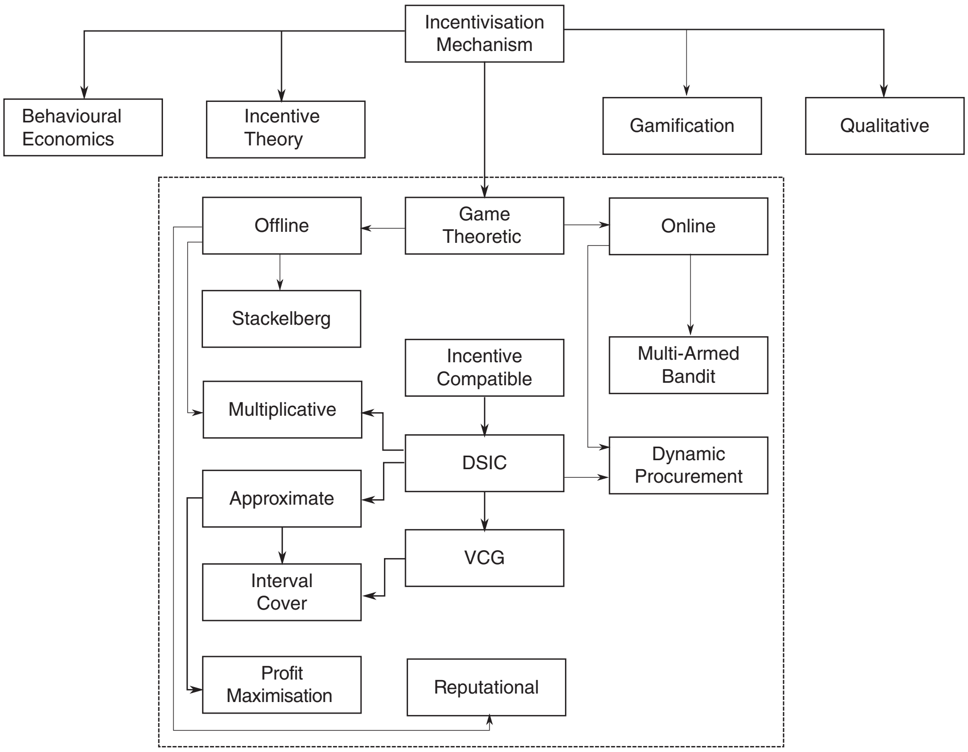

Figure 1 A taxonomy of crowdsourcing incentivization. DSIC = Dominant Strategy Incentive Compatible; VCG = Vickrey–Clarke–Groves

decision-making of individuals and institutions. There are many cognitive biases and psychological traits that inhibit people from making rational decisions and cloud their judgement, such as overconfidence, projection bias, and the effects of limited attention. A core aspect of behavioural economics is that of bounded rationality. An individual's rationality is limited by the tractability of the problem at hand, the cognitive limits of their minds, and the amount of time they have available to make decisions. There are several sources of error regarding human economic decision-making. People often make decisions based on rules of thumb or heuristics rather than using logic. Additionally, people understand and respond to events using mental models that frame the situation based on anecdotes and stereotypes.

Cumulative prospect theory (Tversky & Kahneman, 1992) is an influential theory within the field and represents a model for decision-making under risk and uncertainty. It is based on the observation that people often make decisions using a psychological reference point, such as the *status quo*, and generally care more about avoiding losses than acquiring equivalent gains relative to the reference point. Additionally, people typically overestimate the likelihood of rare events, but underestimate the likelihood of regular events.

Critics of behavioural economics (Berg & Gigerenzer, 2010) argue that it does not make more realistic assumptions than neoclassical economics, as its proponents (Rabin, 2002; Camerer, 2003) claim, and that while it does study what decisions humans make, it does not account for how they make them. The result is that it is even more tied to the way humans think than its neoclassical counterpart. Neoclassical economists have relied on the 'as if' argument, which is to say that although humans do not solve the difficult optimization problems of neoclassical economics when making decisions, they behave as if they do. Behavioural economists point to counter cases from experiments where the predictions of neoclassical economics fail. In offering prospect theory, whereby greater weight is given to losses than gains, behavioural economics provides a better fit for the choices observed in experiments. According to Berg and Gigerenzer (2010), the problem, however, is that prospect theory requires summing up every possible outcome and even more difficult computations to produce a decision than neoclassical optimization. Thus, behavioural economists must also rely on the 'as if' argument.

Mechanism design does not suffer from many of the criticisms directed at game theory in general by behavioural economists, who are mostly concerned with problems that naturally occur in the wild, rather than games that have been specifically designed to have desirable properties. For instance, with game theory, it will often be the case that a game will not have a unique equilibrium, but an array of possible outcomes. Mechanism design enables the creation of games that do not have this property. In the construction of DSIC games, one of the consequences of mechanism design is the easing of the reasoning burden of the participants. A key goal in this area is often the construction of games whereby the optimal course of actions is obvious. Even in situations whereby the optimal course of action is not obvious, if it is not a 'one shot' game, the participants will learn the optimal course of action over time either through direct experience, competition, or through communication via blog posts, etc. As noted in Shah and Zhou (2015), incentive compatibility is relevant in situations whereby people encounter a mechanism over extended periods of time or whereby the financial rewards are large. In such cases, it would be expected that the participants or others, such as bloggers or researchers, would devise strategies to game the mechanism. Behavioural economics could not be used to address this problem.

4.2 Incentive theory of motivation

Incentive theory is a psychological theory of motivation that focuses on the role rewards play in motivating people towards a given course of action. Beyond biological factors, two types of reward are considered in incentive theory, namely intrinsic and extrinsic rewards. Intrinsic rewards arise from internal factors, whereas extrinsic rewards arise from external factors. Incentive theory draws much from the research of the behavioural psychologist B. F. Skinner on operant conditioning (McLeod, 2007). With incentive theory, a reward, which may be either tangible or intangible, is given immediately after the occurrence of an action to increase the likelihood of the action being performed again. As such, it is based on the principle of positive reinforcement. With positive reinforcement, behaviour is altered by a continuously associating a positive meaning to a given course of action over an extended period. For rewards to be motivating, they need to be obtainable. In addition to positive reinforcement, Skinner identified two other responses or operants from the environment that can follow behaviour, namely neutral operants and punishers. Neutral operants neither increase nor decrease the likelihood of the behaviour being performed again, whereas punishers decrease the likelihood and weaken behaviour.

The use of extrinsic rewards can have both positive and negative effects in relation to intrinsic rewards. As noted in Cameron *et al.* (2001), when explicitly tied to performance standards and success, extrinsic rewards have positive effects on intrinsic motivation. In cases whereby extrinsic rewards are loosely tied to behaviour or signify failure, such as when the extrinsic reward received is lower than expected, however, the effects on intrinsic motivation are negative.

Criticisms of incentive theory include that it does not consider the role played by cognitive and inherited factors in relation to learning and behavioural change. It also does not account for social learning whereby humans learn by observation or communication rather than direct personal experience.

In Bosha *et al.* (2017), the use of incentive theory is investigated in bolstering user participation in a participatory crowdsourcing project to collect data related to public safety in East London, South Africa. In certain situations, incentive theory will be useful to induce learning and to bolster participation in crowdsourcing. In many crowdsourcing applications on platforms, such as MTurk, however, the objective is the completion of tasks that aid the requester, not changing the behaviour of users. In a similar manner to behavioural economics (see Section 4.1), incentive theory does not address the adversarial situation whereby participants in a crowdsourcing system adopt strategies to game the platform. Incentive compatibility, from a mechanism design perspective, is required in such situations.

4.3 Gamification

Gamification can be viewed as the application of video game design principles to make work tasks appear like a game so as to increase the motivation for users to engage (Morschheuser *et al.*, 2016). As such, it increases the intrinsic rewards (see Section 4.2) for performing tasks by making task completion more enjoyable. This is important in crowdsourcing applications whereby users may often be participating for free or for very little monetary rewards. In such systems, intrinsic motivations, such as enabling the users to be creative or to gain social recognition, can often outweigh extrinsic factors from a motivational perspective.

Platforms that incorporate gamification features often use design patterns to induce motivations in users regarding feelings of autonomy, mastery, flow, and suspense (Morschheuser *et al*., 2016). Motivational affordances adopted in this space frequently include points, badges, leader boards, avatars, and stories.

In Hamari (2013), the authors investigated the introduction of gamification, and specifically the incorporation of badges, into a utilitarian peer-to-peer trading service known as Sharetribe. The results indicated that the claim of significant overall increases in usage frequency, quality, or social interaction due to gamification could not be supported. Nevertheless, the users that did actively engage with badges were also more likely to engage with the service, leave comments on listings, complete transactions, and list their goods for trade.

For certain types of application, gamification may increase user participation. The goals of gamification and incentive engineering, however, are somewhat different. With gamification the objective is to increase users' intrinsic rewards and engagement, whereas with incentive engineering, it is to ensure that users' incentives are structured to have desirable properties. Indeed, there is no reason both could not be used in the same platform.

4.4 Qualitative approaches

Many problems within the incentivization space are computationally hard and cannot be easily or intuitively modelled using rational utility maximizing agents whereby the utility functions represent all implicit and explicit rewards for a particular course of action. One of the primary goals of mechanism design is to make minimal assumptions regarding the behaviour models of agents, such as that if there is an obvious dominant incentive compatible strategy, the agents will play such a strategy. Such weak behaviour models, however, may not be sufficient to explain the behaviour of agents in other types of scenario whereby certain parameters of the problem are not under the control of the system designer.

Several approaches have been developed that address the incentivization problem using qualitative techniques such as subgraph matching (Ebden *et al*., 2015), content analysis (Shaw *et al*., 2011), volunteering (Mao *et al*., 2013), and SCOUT (Stewart *et al*., 2010). Such systems are not explicitly designed in terms of utility maximizing agents; rather different models are adopted to represent system behaviour. The commonalities and differences of the systems are not expressed in a formal manner, however, and the relationship amongst them is thus difficult to ascertain due to disparity of the techniques used. In situations where the parameters of the problem are under the control of a system designer, the use of mechanism design, in contrast to qualitative approaches, will yield the benefits of DSIC games discussed in Section 3.2.

4.5 Game theoretic incentive engineering

In Sections 5 and 6 several applications of game theory to crowdsourcing are discussed from an offline and online perspective, respectively. The difference between offline and online approaches is related to the adoption of reinforcement learning techniques, which are adopted in the latter, but not the former. As is illustrated Figure 1, the game theoretic incentivization mechanisms discussed in Sections 5 and 6 can be classified in number of ways, such as in relation to whether they are incentive compatible, approximate, or operate in an online manner.

The Stackelberg game (see Section 5.1) and reputational (see Section 5.2) incentivization schemes were not developed using results from mechanism design. In the case of the former, a pre-existing game was used in modelling the desired incentive structures. In the case of the latter, the interaction between requesters and worker agents is captured in the formalism of repeated games.

With mechanism design, it is not assumed that a game pre-exists in the wild or that a given problem should be altered to match a model of a pre-existing game; rather the rules of the game should be engineered such that the expected behaviour of agents will be in accordance with the system designer's objectives. Incentive compatibility is a key concept in mechanism design and refers to situations whereby it is in all agents' interests to provide private information truthfully or act in accordance with their preference. DSIC is a stronger form of the concept and refers to games whereby truth-telling is always in an agent's interest regardless of the actions of other agents. The multiplicative incentivization mechanism (see Section 5.3) is incentive compatible, but it is also DSIC in that the agents within the scenario whereby it is adopted do not interact.

Profit maximization (see Section 5.4.1) and interval cover (see Section 5.4.2) represent two approaches that have been developed using mechanism design. With profit maximization, the incentivization problem is addressed in the context of profit maximization for a mobile sensing scenario. With interval cover, a VCG mechanism is used for the case when agents have identical quality ratings and an approximation algorithm when not.

In terms of online algorithms, multi-armed bandit (MAB) (see Section 6.1) and dynamic procurement auction (see Section 6.2) approaches are considered. Online approaches, which are used in situations whereby the values of courses of action are uncertain, address the trade-off of exploring new courses of action and exploiting pre-existing courses of action that have previously given good results. This is useful in situations whereby the objective is to maximize the utility of the work performed subject to a budget constraint and in which the costs of the workers are unknown.

5 Offline game theoretic algorithms in crowdsourcing

A wide variety of crowdsourcing applications have been developed that incorporate ideas from game theory and mechanism design in the development of incentivization structures for user participation. Several approaches adopt an auction structure. Using mechanism design in the design of auctions enables the development of systems either for profit maximization or for maximizing social surplus. Problems within this area are often computationally hard and several approximate algorithms have been developed that maintain desirable mechanism design properties, but relax the requirement of obtaining a maximal. Other techniques include reputation-based incentivization and multiplicative incentives that incentivize users not to guess answers when presented with alternatives. The systems included in this section represent members of the classes from Figure 1 that do not fall under the online category.

5.1 Platform and user-centric mobile phone sensing

In Yang *et al*. (2012), approaches to incentivizing users to participate in crowdsourcing using smartphone technology are discussed. As noted in Yang *et al*. (2012), smartphones support a wide range of sensing technologies, such as Global Positioning System receivers, microphones, accelerometers, gyroscopes, and cameras, amongst others. These sensors can be used in the creation of novel technologies and in the augmentation of pre-existing services, such as social media services, with information related to the user's or device's context and the environment in which they are embedded. The connectivity of these devices is near pervasive and, as such, smartphone technology has the potential to become the largest sensing platform constructed in history. To enable the creation of a smartphone platform that incentivizes users to allow their mobile device to be used for sensing, two approaches are proposed, namely the platform-centric model and the user-centric model. With the platform-centric model, a reward is provided by the platform, which is shared amongst participating users.

With both approaches, it is assumed that there is a principal agent or platform owner that gains from a discrete set of sensing tasks being performed and a group of agents that have smartphones with the computational, sensing, and communication resources available to perform the tasks, but incur a cost for doing so. With the platform-centric model, the problem is viewed in terms of the goal of the platform owner in maximizing their utility. The platform owner wishes to receive the greatest benefit in terms of the sensing tasks performed whilst minimizing cost. In contrast to this, the smartphone users wish to receive the greatest price (cost to the platform owner) for performing tasks.

The platform-centric model takes the form of a Stackelberg game. In a Stackelberg game or competition, there is a single leader agent that makes the first move and a group of follower agents that move or perform actions subsequently. A Stackelberg game models the situation whereby the leader agent has an advantage of moving first given that they know *ex ante* that the follower agents cannot perform a leader action and that the follower agents will observe their action. In economics, the leader agent is sometimes referred to as the market leader.

For the user-centric model, an auction procedure with the following properties was designed:

- Individual rationality: The utility of the smartphone users is zero if they do not perform sensing tasks. If sensing tasks are performed, users receive a utility of the price paid by the platform minus the user's cost (see Section 3.2). If the user's cost for performing a task is greater than the price paid by the platform, they do not perform the task.
- Profitability: The incentives are structured in such a manner that the value the platform owner receives for the tasks being performed does not go beyond their costs in terms of making payments to users.
- Computational efficiency: Although it is computationally intractable to determine an exact solution to the problem posed, polynomial time approximate algorithms can be used to ensure results with bounded quality.
- Truthfulness: The incentives are structured such that it is not in the agents' interest to deviate from the truthful revelation of their costs or value of the work required to perform the task.

In terms of commercial applications, this research would not be suitable for systems such as MTurk and CrowdFlower (see Section 5.3) in that mobile phone crowdsourcing tasks typically operate over longer durations of time than typical tasks on those platforms. One potential commercial application area, although not specifically intended for such systems, is within Witkeys.

In China, a Witkey is a crowdsourcing site that enables the online purchase of services and information, which is used by users of either a personal or professional nature. Such sites, enable tasks to be posted and users offer a small reward for solutions to the tasks. Other users attempt to solve the tasks and compete to have their solutions selected by the requesters. There are several Witkey sites, including Witkey.com, Taskcn.com, Zhubajie.com, and K68.cn (Yang *et al*., 2008). The founder of Witkey.com, which was the first of such sites, coined the term. A Witkey site plays the role of a third party in a knowledge market, receiving payments from requesters and transferring payment to the providers of completed tasks that have been selected as winning submissions. The tasks posted on Witkey sites usually require a given skill set to complete and a moderate investment of time.

An analysis of the strategic behaviour of users in the Taskcn Witkey site, which is one of the largest, is provided in Yang *et al*. (2008). In Taskcn, market inefficiencies are observed in that with uncoordinated user participation, many users race to complete for the most popular tasks, which leads to a low probability of individual users winning. This is a likely contributing factor as to why most users of the site only make a smaller number of submissions and then become inactive. The users that remain active, however, tend to strategically choose tasks that they have a higher probability of winning and which attract less competition. A small number of users win multiple tasks and increase their submission to win ratio over time. It is unclear, however, as to whether this is because of them learning a better strategy for task selection or whether it due to their increasing reputation rating from past wins and thus having a higher probability of being selected as a winner than other users.

5.2 *Reputation-based incentivization*

In Zhang *et al*. (2012), the authors discuss a reputation-based incentivization scheme, which is based on social norms. The interaction between requesters and worker agents is modelled in the formalism of repeated games. A protocol is proposed that prevents the problem of free riding and incentivizes agents to contribute to task-solving processes. The free riding problem is related to situations whereby agents are paid *ex ante* or prior to completing tasks. In such cases, if correct incentives schemes, such as a reputation-based scheme, are not in place, there could be an incentive for payment to be received, but not for effort to be expended on the task.

The approach presented in Zhang *et al*. (2012) integrates payment and reputation schemes in the design protocols for crowdsourcing. In terms of maximizing the revenue of a crowdsourcing site, there is a trade-off between keeping worker agents incentivized such that they act in accordance with the social norm and lowering the percentage of payments received by workers, and, thus, increasing the revenue of the site. If the payments made to agents are too low, the agents will be less inclined to conform to the social norm and their reputations will decrease, leading to worker agents being prohibited from performing tasks and to a reduction in revenue. The protocol designer, therefore, must achieve a balance between increasing the per transaction revenue for the site and the incentivization of the worker agents in contributing to revenue maximization.

In addition to the commercial platforms discussed in Sections 5.1 and 5.3, a reputation-based incentivization scheme could be useful in the Upwork platform. Upwork, formerly known as Elance oDesk, is a crowdsourcing platform that enables businesses and professionals to connect and collaborate on a freelancing basis. It is the world's largest online platform for outsourced contracts (Ghani *et al*., 2014) and differs from MTurk in several ways. Upwork provides functionality to view workers' profiles, including information related to their identity, feedback, skills, and a rating. It helps to enable workers be chosen by requesters based on selection criteria. Additionally, functionality is provided to clients for interviewing, hiring, and collaborating with freelancers through the platform. An application is provided that tracks time and takes screen shots as freelancers are working. In terms of payments, Upwork takes a percentage of the fees payed to workers.

5.3 Multiplicative incentives

Crowdsourcing is becoming increasingly popular in the production of labelled data for supervised learning within the machine learning community. With supervised machine learning, when learning a classification model, algorithms require a large set of training material that is labelled positively or negatively for a given class or group of classes. In Shah and Zhou (2015), the problem of outsourcing the labelling process to the crowd is considered. The primary problem with crowdsourcing such data is that it will be of low quality/accuracy in that crowd workers are typically not experts and, as such, it can be expensive to obtain good results if appropriate incentives are not put in place.

The approach presented in Shah and Zhou (2015) introduces a requirement, referred to as the no-free-lunch axiom, that is intended to discourage spammers and miscreants. The system works by introducing a gold standard of material, which has a correct, known classification to the system operator, in a random manner to the rest of the material that has not been classified. The users performing the classification are unaware of which material forms part of the gold standard and which is unclassified. The users receive payment for answers they provide in the gold standard. The incentive mechanism takes a multiplicative form. Each answer provided in the gold standard is given a score and the payment given to the user is the product of the scores. For correct classifications, the user's payment is increased by a given ratio; for incorrect answers the user's payment is multiplied by zero, and for skipped answers it is multiplied by one. Thus, if the user provides an incorrect answer on any of the gold standard questions, they will receive a payment of zero. This introduces a truthful incentive to skip questions in situations where the user does not know the answer, rather than to guess.

Potential commercial platforms that could avail of multiplicative updates include MTurk, CrowdFlower, and Witkeys (see Section 5.1). MTurk is the most widely known commercial crowdsourcing platform and is a microtask marketplace for work that requires human intelligence. The platform draws from the fact that there are still many tasks that can be performed by humans more efficiently, or to a higher standard, than by computers. For example, transcribing audio recordings or identifying objects in photographs or videos. Within MTurk there are two types of users, namely requesters and workers (sometimes referred to as Turkers). The microtasks in MTurk are referred to as HITs. The platform charges a fee to requesters on top of the payment to workers. In MTurk, the requesters design and post their tasks and set the price for accepted submissions. The workers decide which tasks they will perform. Subsequently, the requester reviews the work submitted by workers and either accepts or rejects the tasks. The tasks designed by requesters are usually quite simple and could potentially form the elementary components of more complex tasks and, as such, the price for completing a task is usually quite low along with the amount of time it takes to complete.

CrowdFlower is a crowdsourcing somewhat similar to MTurk in that its focus is on relatively simple microtasks. The target market for CrowdFlower is narrower and more focused than that of MTurk in that it primarily targets tasks related to the labelling, enrichment, and cleaning of data for data science and machine learning algorithms. In contrast to MTurk, which offers little in the way of quality control measures, CrowdFlower enables advanced worker targeting and provides management tools, such functionality to monitor the performance of workers on tasks. In terms of the workflow, CrowdFlower uses a similar model to MTurk. Tasks are posted and on a continual basis and contributors discover work through job boards and, subsequently, decide whether to take on the tasks. Once a contributor uploads a completed task, the work is evaluated using a hidden subset of known answers. The contributors' performance on these tasks determines their reputation and the degree to which the system trusts them.

5.4 Approximate algorithms

Often with the design of mechanisms, an important objective is to ensure that the game is DSIC. One reason for this is that it significantly simplifies or eases the burden of reasoning about what moves should be made by players when compared to alternative approaches. Another reason is that the mechanism designer will typically have a good idea of what will happen or what agents will do in such games. In addition to games being DSIC, a common requirement is that the social surplus or revenue be maximized and that the allocation and payment rules be executed in polynomial time. This can cause problems because optimization problems of this type are often NP-hard. As noted in Section 3.2, with social surplus maximization, goods are allocated to the agents that value them the most. This section discusses approximation algorithms that maintain the DSIC requirement, but relax the optimality requirement for maximization. Both the profit maximization and interval cover algorithms discussed here would not be suitable for sort-lived tasks performed on a once-off basis. Thus, in a similar manner to the research discussed in Section 5.1, the algorithms would be most suited (with modifications), from a commercial perspective, to the Witkey platform.

5.4.1 Profit maximization

ProMoT represents a reverse auction for profit maximization in the context of mobile phone crowdsourcing platforms (Shah-Mansouri & Wong, 2015). The allocation rule for profit maximization is proven to be NP-hard through a reduction to the Knapsack problem and, as such, the authors propose an alternative greedy algorithm for the maximization problem along with a payment scheme. In showing that the approximation algorithm maintains the monotonicity of the allocation rule, the authors invoke Myerson's Lemma to ensure that the resulting game is DSIC. Experimental results show that this approach performs better than the approach discussed in Section 5.1. The authors (Shah-Mansouri & Wong, 2015) note that, given that the winner determination problem is NP-hard, the VCG mechanism cannot be used. As the authors use Myerson's lemma and monotonicity to prove the approach is DSIC, the problem must be within a single parameter environment. Thus, although it is claimed the agents make bids on subsets of tasks, two different subsets could not be selected if they contained an overlap. Subsets must represent exclusive alternatives, which are independent coverage tasks. Alternatively, the problem would represent a combinatorial auction and Myerson's lemma could not be used.

5.4.2 Interval cover

For crowdsourcing tasks that have a natural temporal or spatial ordering, the interval cover problem concerns assigning agents to tasks such that the total costs of the task requester are minimized subject to error tolerance constraints for each task (Dayama *et al*., 2015). Allocation problems in mobile pervasive sensing and in choosing suppliers in demand response smart grid scenarios represent two examples whereby determining the allocation of agents to tasks are of this form. For example, in mobile pervasive sensing applications where real-time information is required for air quality or traffic conditions on congested routes in urban areas is required, agents commuting on the routes of interest would provide such information from mobile phone applications and on-board sensors. The costs of the agents, in this case, would include costs related to installing, maintaining, and executing sensors, along with mobile phone applications. In demand response smart grid scenarios, the problem would be related to choosing a mix of suppliers of variable reliability to meet a varying power demand profile given an error tolerance limit regarding the probability of failure or the demand not being met. In this case, suppliers would bid for times to supply power to meet the demand.

A reverse auction approach to addressing the interval cover problem is presented in Dayama *et al.* (2015) where agents submit bids that represent their costs for complete bundles of tasks. In this scenario, multiple tasks must be completed and, unlike typical tasks on MTurk (see Section 5.3), each task involves a significant cost on the part of the worker agents. It is assumed, the requester knows the quality of workers, which determines their probability of success, but not their costs. Tasks have a natural spatial or temporal ordering and agents are, thus, interested in contiguous subsets of tasks, referred to as intervals. Given that workers could have a low-quality rating, multiple workers can be assigned to the same task such that a quality threshold is met.

In the case where agents have different quality ratings—the heterogeneous scenario, the problem of minimizing the cost of the requester is NP-hard. In the homogeneous scenario, where the agents' quality ratings are taken to be the same, cost minimization can be performed in polynomial time. In this case, a VCG mechanism (see Section 3.2) can be used to design an allocation rule that is truthful, individually rational, and maximizes social surplus or allocative efficiency.

In the heterogeneous scenario, however, it is not practical to appeal to VCG as it would involve solving several NP-hard problems. As such, Dayama *et al.* (2015) presents an approximate approach that operates in polynomial time and has a bound on the quality of the solution obtained.

6 Online algorithms

The game theory approaches discussed in Section 5 were offline in the sense that the incentive structure adopted was static and was not altered as the system was in operation. In certain situations, however, it will be beneficial, or necessary in the case of actions that have an uncertain reward, to learn incentive structures at run-time. To achieve this, systems must address the exploration/exploitation trade-off in finding good payment schemes. In this section, scenarios in which the incentive structures vary over time are discussed whereby reinforcement learning algorithms are invoked. Commercial crowdsourcing platforms relevant to this area include MTurk, Witkeys, and CrowdFlower (see Sections 5.1 and 5.3).

6.1 Multi-armed bandits

MAB algorithms explore the trade-off between resource exploration and exploitation. In the context of crowdsourcing, the requester of a set of tasks wishes to maximize their utility in terms of having the tasks performed and receiving results, whilst minimizing their costs. A problem, however, exists in terms of what price should be given to perform tasks in that the cost distributions of the workers will be unknown to the requester. If the price placed on performing a task is too high, the budget of the requester will be exhausted quickly or the price for the completion of all tasks will be high. If the price for performing a task is too low, it will be starved of high-quality workers. With MAB algorithms, the agents repeatedly choose amongst actions that have an uncertain reward. A goal with the development of such algorithms is to asymptotically minimize the regret over time when the actions chosen are compared to what the actions would have been if they had been chosen optimally with hindsight (Blum & Monsour, 2007).

The use of MABs is a general approach to online learning. In Tran-Thanh *et al.* (2014), the problem of assigning tasks to workers with a limited budget is addressed using a MAB in a crowdsourcing context.

6.2 Dynamic procurement auctions

With crowdsourcing tasks, often the requester wishes to develop a payment scheme subject to a budget constraint such that the utility of the tasks performed is maximized, whereas worker agents seek to maximize their individual utility by choosing what tasks to perform and at what price they will participate. In Singla and Krause (2013), the problem of divergent goals between the requester and worker agents is addressed using procurement auctions.

There are two types of models for procurement auctions that could be adopted by the requester; these are referred to as the bidding model and the posted price model. With a bidding model, the workers bid for the price they are willing to perform the crowdsourcing task. With this type of auction, an agent will not bid a price too low in that if it were accepted, they would not be happy with the utility. A price too high, however, will also not be bid in that it would reduce the chance of winning. The model will not always reveal the correct valuation of the agents' costs, however. For example, sometimes the agents will not know their true costs as they could be difficult to determine or the agents will not fully understand the mechanism. Issues, such as whether they trust the requester, could also be factored into the price.

With a posted price model, the workers are sequentially given a price and they either accept or reject the offer. As the workers respond, the offered price is varied, depending on the response given.

The approach adopted in Singla and Krause (2013), which uses a posted price model, is guaranteed to be budget feasible, approximately optimal in terms of the utility obtained by the requester, and truthful regarding the bids made by the workers given limited information or assumptions in relation to their costs. This extends and builds upon research on MABs for online auctions (Babaioff *et al.*, 2015) to the case of procurement auctions under budget constraints.

7 Discussion

Incentive Engineering, in general, is a difficult problem for several reasons. A common assumption in the application of incentive engineering to crowdsourcing is that agents are self-interested. To say that an agent is self-interested is to say that the agent is rational and will act in accordance with its interests or objectives, not necessarily that the agent is selfish or that the agent's interests reflect a desire to increase their own financial benefit to the detriment of others. One of the most significant problems in designing incentives for the crowdsourcing domain is that the system designer does not know *a priori* what the objectives of the users are or what incentivizes them to behave in a given manner. In many cases, however, providing financial rewards can be taken as a proxy to induce behaviour of a given type for a large proportion of the agents participating in crowdsourcing systems.

Nevertheless, for platforms, such as MTurk or CrowdFlower, due to the small sums of money typically on offer and due to the limited amount of time users will spend on tasks, users will not take a significant amount of time in understanding how the financial incentivization mechanisms work, which will be a problem if they are difficult to understand. This should not be viewed as the users acting in an irrational manner, rather that, in such cases, the incentivization mechanism does not account for all costs within the problem, such as the amount of time it takes for its own comprehension by users. Mechanism design represents one approach to addressing problems of this type in that it enables the creation of DSIC games that have obvious dominant strategies. Such games can be more easily explained to the user than non-DSIC games and only depend on weak models of agent behaviour, such as that the users will act in accordance with their interest and play dominant strategies.

There are several problems with game theory for modelling the behaviour of real and computational players. From a computational perspective, although not in the class NP-complete, no known polynomial time algorithms for computing Nash Equilibria in general are available. Given that computers cannot solve such problems in polynomial time, it is unlikely that humans will be able to find exact solutions either in a reasonable amount of time. The 'as if' argument must be adopted in relation to this (see Section 4.1), but it should be noted that the 'as if' argument must also be adopted within behavioural economics whereby even more difficult computations are required (Berg & Gigerenzer, 2010). When games have been engineered using mechanism design such that they have a unique dominant strategy that is always in a user's interest to play, and that has been explained or demonstrated to the user, appropriate strategies will be adopted.

With DSIC games, the computational difficulty of computing Nash Equilibria is avoided. Another problem exists, however, in that optimizing for allocation or payment could be NP-hard. In many cases, however, there are approximation algorithms with guarantees regarding the quality of solution that preserve the DSIC property.

Sometimes, even when using mechanism design, the payment rules for games will be complicated and difficult to understand by the non-expert user. In such cases, approximate approaches which relax the requirement for the game to be DSIC will be required. Approximation, in this case, is required for a different reason to that of approximation to make the maximization of allocation or payment tractable. In this case, rules may need to be simplified to ease in understanding. Consider, for instance, the payment rule used by Google for its highly profitable AdWords sponsored search platform, which uses a generalized second-price auction that is not DSIC and will not be optimal from a theoretical perspective. It is likely that Google were unaware of the DSIC results in this area at the time the platform was developed, and it would be too costly or risky to now change; another possible reason for adopting the generalized second-price auction, however, is that it simplifies how the payment rule can be described to users.

In certain situations, there will be disparities between the behaviour of humans and what is rationally implied by the games they are participating in. The capability of computers to act in an ultimately rational manner in maximizing utility functions, it has been argued (Parkes & Wellman, 2015), make the application of game theory to artificial agents more fruitful or productive than its application to humans in open environments, such as the Internet.

As noted by Shoham (2008), some issues with game theory could be addressed through the development of more pragmatic or more appropriate models of agent behaviour. In game theory, it is assumed that agents have strategies that have minimal structure and which are fully captured in terms of real-valued utility functions. In the real world, and within artificially intelligent systems, it is often useful to adopt a richer vocabulary to describe and reason about agent behaviour. In such systems, some agents can perform actions and others cannot. Also, agents make plans, and have beliefs in relation to their own behaviour and their environment.

A drawback with adopting richer models, however, is that it makes greater assumptions regarding the behaviour of agents, which may not be accurate. Additionally, with non-game theoretic models, certain desirable benefits of mechanism design, such as the creation of DSIC games, cannot be availed of. Notwithstanding this, the use of qualitative approaches, such as discussed in Section 4.4, is often more intuitive for certain types of problem that cannot be easily quantified in terms of costs and benefits.

8 Conclusion

This paper provided a survey of incentive engineering for crowdsourcing systems from a game theoretic perspective. The incentive structures for worker agents and the potential for revenue generation for system operators represent key aspects in determining the effectiveness of such systems. There are several metrics by which the impact of incentivization schemes could be measured, depending on the purpose for which the systems were created. In commercial settings, revenue generation will often be the primary objective. In such cases, it is necessary to ensure that the incentives provided to agents promote behaviour in accordance with systems' goals.

Game theoretic approaches have a rich history and have received significant interest within computer science from an algorithmic design perspective in this last decade. Three key aspects in the design of effective mechanisms are that they are DSIC, they maximize revenue or social surplus, and that they operate in polynomial time. DSIC games ensure that it is always in an agent's interest to be truthful regardless of the behaviour of other agents. Such games are useful in that they ease the reasoning burden of both agents and system operators when compared to alternative games and lead to the development of platforms whereby the outcomes are predictable. In certain cases, however, the form of the games produced will be difficult to understand to non-expert users. In such scenarios, approximate approaches may need to be developed.

In addition to the problem of difficult to understand payment rules, many maximization problems, in terms of either revenue or social surplus, are NP-hard. There are many results within computer science in relation to approximation algorithms for NP-hard problems and, as such, computer science has much to offer this space. Often, the use of such approximation algorithms enables the monotonicity of the allocation rules be maintained and preserves the DSIC structure of the games. There are also situations whereby games will be dynamic and the parameters cannot be hard coded *a priori*. In relation to this, the paper outlined techniques that address the learning and adaptation of games at run-time.

There have been many alternative approaches developed, which are not based on game theory or mechanism design, that provide an intuitive and a less formal approach to addressing incentivization within crowdsourcing. There are problems with using weak agent models and game theory to represent the behaviour of rational agents. For example, for certain problems it may be difficult or unintuitive to specify an agent's motivations entirely in terms of cost and benefits that are encoded within a utility function. In a certain sense, adopting a weak model of agent behaviour can be viewed as both a strength and weakness of mechanism design. It is a strength in that it makes limited assumptions and broadens the scope of application of the mechanisms. It is a weakness, however, in that it limits the expressiveness to which agent behaviour, which may be known for a domain-specific problem, can be represented.

Several commercial platforms exist within the crowdsourcing space. The paper outlined current research from the literature that could be adopted within such systems to potentially increase user participation, reduce the risk of users gaming the system, increase revenue, and assign tasks to users that gain the most from contributing. It is anticipated that the role of incentive engineering will become increasingly important within the crowdsourcing domain with the proliferation of such systems.

Acknowledgements

The authors would like to acknowledge the support of the European Union's Seventh Programme for research, technological development, and demonstration under grant agreement number 308513. Thanks go to the anonymous reviewers for their feedback and comments, which significantly improved the content of the paper.

References

Adriaens, T., Sutton-Croft, M., Owen, K., Brosens, D., van Valkenburg, J., Kilbey, D., Groom, Q., Ehmig, C., Thürkow, F., Van Hende, P. & Schneider, K. 2015. Trying to engage the crowd in recording invasive alien species in Europe: experiences from two smartphone applications in northwest Europe. *Management of Biological Invasions* **6**(2), 215–225.

Babaioff, M., Dughmi, S., Kleinberg, R. & Slivkins, A. 2015. Dynamic pricing with limited supply. *ACM Transactions on Economics and Computation* **3**(1), 4.

Benkler, Y. & Nissenbaum, H. 2006. Commons-based peer production and virtue. *Journal of Political Philosophy* **14**(4), 394–419.

Berg, N. & Gigerenzer, G. 2010. As-if behavioral economics: Neoclassical economics in disguise? *History of Economic Ideas* **18**, 133–165.

Blum, A. & Monsour, Y. 2007. Learning, regret minimization, and equilibria. In *Algorithmic Game Theory*, Nisan, N., Roughgarden, T., Tardos, E. & Vazirani, V. V. (eds). Cambridge University Press, 79–102.

Bosha, E., Cilliers, L. & Flowerday, S. 2017. Incentive theory for a participatory crowdsourcing project in a developing country. *SA Journal of Information Management* **19**(1), 7.

Brabham, D. C. 2013. *Crowdsourcing*. MIT Press.

Camerer, C. F. 2003. Behavioural studies of strategic thinking in games. *Trends in Cognitive Sciences* **7**(5), 225–231.

Cameron, J., Banko, K. M & Pierce, W. D. 2001. Pervasive negative effects of rewards on intrinsic motivation: the myth continues. *The Behavior Analyst* **24**(1), 1.

Cooper, H. M. 1988. Organizing knowledge syntheses: a taxonomy of literature reviews. *Knowledge in Society* **1**(1), 104–126.

David, G., Michael, R., Erwin, F. & Martin, S. 2012. Crowdsourcing information systems: definition, typology and design. In *ICIS 2012 The 33rd International Conference on Information Systems*, Joey F George (ed.). Association for Information Systems/AIS Electronic Library (AISeL), 3562–3572.

Dayama, P., Narayanaswamy, B., Garg, D. & Narahari, Y. 2015. Truthful interval cover mechanisms for crowdsourcing applications. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, 1091–1099*. International Foundation for Autonomous Agents and Multiagent Systems.

Drexler, K. E. & Miller, M. S. 1988. Incentive engineering for computational resource management. *The Ecology of Computation* **2**, 231–266.

Ebden, M., Huynh, D., Moreau, L. & Roberts, S. 2015. Incentive engineering through subgraph matching with application to task allocation. *HAIDM*.

Endriss, U., Kraus, S., Lang, J. & Wooldridge, M. 2011. Designing incentives for boolean games. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, 79–86. International Foundation for Autonomous Agents and Multiagent Systems.

Ghani, E., Kerr, W. R. & Stanton, C. 2014. Diasporas and outsourcing: evidence from odesk and India. *Management Science* **60**(7), 1677–1697.

Ghezzi, A., Gabelloni, D., Martini, A. & Natalicchio, A. 2017. Crowdsourcing: a review and suggestions for future research. *International Journal of Management Reviews* **0**, 1–21.

Groves, T. 1973. Incentives in teams. *Econometrica: Journal of the Econometric Society* **41**, 617–631.

Guo, B., Wang, Z., Yu, Z., Wang, Y., Yen, N. Y., Huang, R. & Zhou, X. 2015. Mobile crowd sensing and computing: the review of an emerging human-powered sensing paradigm. *ACM Computing Surveys (CSUR)* **48**(1), 7.

Hamari, J. 2013. Transforming homo economicus into homo ludens: a field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic Commerce Research and Applications* **12**(4), 236–245.

Higgins, C. I., Williams, J., Leibovici, D. G., Simonis, I., Davis, M. J., Muldoon, C., van Genuchten, P., O'Hare, G. M. P. & Wiemann, S. 2016. Citizen OBservatory WEB (COBWEB): a generic infrastructure platform to facilitate the collection of citizen science data for environmental monitoring. *International Journal of Spatial Data Infrastructures Research* **11**(1), 20–48.

Howe, J. 2006a. Crowdsourcing: a definition. Retrieved 3 June 2006 from http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html.

Howe, J. 2006b. The rise of crowdsourcing. *Wired Magazine* **14**(6), 1–4.

Kahneman, D. 2003. Maps of bounded rationality: psychology for behavioral economics. *The American Economic Review* **93**(5), 1449–1475.

Koutsoupias, E. & Papadimitriou, C. 2009. Worst-case equilibria. *Computer Science Review* **3**(2), 65–69.

Mao, A., Kamar, E., Chen, Y., Horvitz, E., Schwamb, M. E., Lintott, C. J. & Smith, A. M. 2013. Volunteering versus work for pay: incentives and tradeoffs in crowdsourcing. In *First AAAI Conference on Human Computation and Crowdsourcing*.

McLeod, S.A. 2007. BF Skinner: Operant conditioning. Retrieved September 9, 2009.

Morschheuser, B., Hamari, J. & Koivisto, J. 2016. Gamification in crowdsourcing: a review. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 4375–4384. IEEE.

Muller, C.L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A. & Leigh, R.R. 2015. Crowdsourcing for climate and atmospheric sciences: current status and future potential. *International Journal of Climatology* **35**(11), 3185–3203.

Myerson, R. B. 1981. Optimal auction design. *Mathematics of Operations Research* **6**(1), 58–73.

Myerson, R. B. 2013. *Game Theory. Analysis of Conflict*. Harvard University Press.

Nash, J. 1951. Non-cooperative games. *Annals of Mathematics* **54**, 286–295.

Nisan, N., Roughgarden, T., Tardos, E. & Vazirani, V. V. 2007. *Algorithmic Game Theory*, **1**. Cambridge University Press.

O'Grady, M. J., Muldoon, C., Carr, D., Wan, J., Kroon, B. & O'Hare, G. M. P. 2016. Intelligent Sensing for Citizen Science - Challenges and Future Directions. *Mobile Networks and Applications* **21**(2), 375–385.

Parkes, D. C. & Wellman, M. P. 2015. Economic reasoning and artificial intelligence. *Science* **349**(6245), 267–272.

Petersen, K., Vakkalanka, S. & Kuzniarz, L. 2015. Guidelines for conducting systematic mapping studies in software engineering: an update. *Information and Software Technology* **64**, 1–18.

Rabin, M. 2002. A perspective on psychology and economics. *European Economic Review* **46**(4), 657–685.

Roughgarden, T. 2003. The price of anarchy is independent of the network topology. *Journal of Computer and System Sciences* **67**(2), 341–364.

Roughgarden, T. 2010. Algorithmic game theory. *Communications of the ACM* **53**(7), 78–86.

See, L., Fritz, S., Dias, E., Hendriks, E., Mijling, B., Snik, F., Stammes, P., Vescovi, F. D., Zeug, G., Mathieu, P.-P., Desnos, Y.-L. & Rast, M. 2016a. Supporting earth-observation calibration and validation: a new generation of tools for crowdsourcing and citizen science. *IEEE Geoscience and Remote Sensing Magazine* **4**(3), 38–50.

See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., Fritz, S., Kerle, N., Jiang, B., Laakso, M., Liu, H.-Y., Milčinski, G., Nikšič, M., Painho, M., Pődör, A., Olteanu-Raimond, A.-M. & Rutzinger, M. 2016b. Crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information. *ISPRS International Journal of Geo-Information* **5**(5), 55.

See, L., Schepaschenko, D., Lesiv, M., McCallum, I., Fritz, S., Comber, A., Perger, C., Schill, C., Zhao, Y., Maus, V., Siraj, M.A., Albrecht, F., Cipriani, A., Vakolyuk, M., Garcia, A., Rabia, A.H., Singha, K., Marcarini, A.A., Kattenborn, T., Hazarka, R., Schepaschenko, M., van der Velde, M., Kraxner, F. & Obersteiner, M. 2015. Building a hybrid land cover map with crowdsourcing and geographically weighted regression. *ISPRS Journal of Photogrammetry and Remote Sensing* **103**, 48–56.

Senaratne, H., Mobasheri, A., Ali, A. L., Capineri, C. & Haklay, M. 2017. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science* **31**(1), 139–167.

Shaw, A. D., Horton, J. J. & Chen, D. L. 2011. Designing incentives for inexpert human raters. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, 275–284. ACM.

Shah, N. B. & Zhou, D. 2015. Double or nothing: multiplicative incentive mechanisms for crowdsourcing. *Advances in Neural Information Processing Systems* **1**, 1–9.

Shah-Mansouri, H. & Wong, V. W. S. 2015. Profit maximization in mobile crowdsourcing: a truthful auction mechanism. In*2015 IEEE International Conference on Communications (ICC)*, 3216–3221. IEEE.

Shoham, Y. 2008. Computer science and game theory. *Communications of the ACM* **51**(8), 74–79.

Singla, A. & Krause, A. 2013. Truthful incentives in crowdsourcing tasks using regret minimization mechanisms. In *Proceedings of the 22nd International Conference on World Wide Web*, 1167–1178. International World Wide Web Conferences Steering Committee.

Stewart, O., Lubensky, D. & Huerta, J. M. 2010. Crowdsourcing participation inequality: a SCOUT model for the enterprise domain. In Proceedings of the ACM SIGKDD Workshop on Human Computation, 30–33. ACM.

Tran-Thanh, L., Stein, S., Rogers, A. & Jennings, N. R. 2014. Efficient crowdsourcing of unknown experts using bounded multi-armed bandits. *Artificial Intelligence* **214**, 89–111.

Tversky, A. & Kahneman, D. 1992. Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and uncertainty* **5**(4), 297–323.

Vickrey, W. 1961. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance* **16**(1), 8–37.

Von Neumann, J. & Morgenstern, O. 1944. Theorem 3.1.18, Theory of games and economic behavior, Princeton University Press.

Wang, L., Zhang, D., Wang, Y., Chen, C., Han, X. & M'hamed, A. 2016. Sparse mobile crowdsensing: challenges and opportunities. *IEEE Communications Magazine* **54**(7), 161–167.

Yang, D., Xue, G., Fang, X. & Tang, J. 2012. Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, 173–184. ACM.

Yang, J., Adamic, L. A. & Ackerman, M. S. 2008. Crowdsourcing and knowledge sharing: strategic user behavior on taskcn. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, 246–255. ACM.

Zhang, Y. & Van der Schaar, M. 2012. Reputation-based incentive protocols in crowdsourcing applications. In FOCOM, 2012 Proceedings IEEE, 2140–2148. IEEE.

Zhao, Y. & Han, Q. 2016. Spatial crowdsourcing: current state and future directions. *IEEE Communications Magazine* **54**(7), 102–107.