Mechanism Design for Incentivizing Social Media Contributions

Vivek K. Singh, Ramesh Jain, and Mohan Kankanhalli

Abstract Despite recent advancements in user-driven social media platforms, tools for studying user behavior patterns and motivations remain primitive. We highlight the voluntary nature of user contributions and that users can choose when (and when not) to contribute to the common media pool. A Game theoretic framework is proposed to study the dynamics of social media networks where contribution costs are individual but gains are common. We model users as rational selfish agents, and consider domain attributes like voluntary participation, virtual reward structure, network effect, and public-sharing to model the dynamics of this interaction. The created model describes the most appropriate contribution strategy from each user's perspective and also highlights issues like 'free-rider' problem and individual rationality leading to irrational (i.e. sub-optimal) group behavior. We also consider the perspective of the system designer who is interested in finding the best incentive mechanisms to influence the selfish end-users so that the overall system utility is maximized. We propose and compare multiple mechanisms (based on optimal bonus payment, social incentive leveraging, and second price auction) to study how a system designer can exploit the selfishness of its users, to design incentive mechanisms which improve the overall task-completion probability and system performance, while possibly still benefiting the individual users.

1 Introduction

With the emergence of Web 2.0 and multiple related social media applications (e.g. Flickr, Youtube, Facebook, Wikipedia) research interest has grown in multiple as-

V.K. Singh (✉) · R. Jain
University of California, Irvine, Irvine, USA
e-mail: singhv@uci.edu

R. Jain
e-mail: jain@ics.uci.edu

M. Kankanhalli
National University of Singapore, Singapore, Singapore
e-mail: mohan@comp.nus.edu.sg

pects of social media including data sharing, image tagging, media processing, ontologies, retrieval etc. While these contributions have significantly advanced the state of the art from the technology perspective, not much research attention has been given till now to the *end-user* or *social aspect* of social media research. Despite significant interest in concepts like crowd-sourcing [6], collective intelligence [9], human-computation [25] etc., the tools to undertake *user behavior analysis in social media networks* are still in their infancy and no theoretical frameworks are available to mathematically analyze why and how often do users contribute to such social media?

That notwithstanding, social media networks are becoming increasingly relevant each day. Citizen-journalists are already providing interesting event information and images for common benefit to various news agencies [12]. Similarly, multiple users are already using tools like Google Image Labeler [4] to tag images and contributing content to Wikipedia [26], thus creating a rich collective information mechanism which can provide common benefits to a larger society.

An important point to consider in all these applications is that the user contribution is *totally voluntary*. Further the decision making is completely distributed and there are no means for central coordination or explicit communication between the various participating users. This brings us to the important issue of user *motivation* and that the individual users will contribute to such social media networks only based on their personal utility decisions.

Such a setup, though compelling, leads to multiple conflicting goals. While, the task-completion costs are incurred by the individuals, the benefits are common. Thus while the owners of systems such as New York Times, Google Labeler, or Wikipedia are interested in maximizing the tasks accomplished, individual agents may be interested in maximizing their personal utility gain in such a sensing/content-provision mechanism. A key question which arises in such a scenario is how can an individual user optimally decide his/her contribution strategy i.e. *when (and when not) should he/she undertake the social media task*. A system administrator on the other hands is interested in *finding the optimal incentive mechanisms to influence these selfish end-users so that the overall system utility is maximized*.

Clearly, there are no currently available tools which can answer such questions. The analysis and answer to such questions requires explicit modeling of user behavior as well as considering the specific characteristics of the domain being considered. Hence we propose the use of a game-theoretic framework, which models users as rational (selfish) agents and incorporates the dynamics of social media (e.g. voluntary participation, *virtual* reward structure, social benefits, and public-sharing) to gain some insights/explanations for user behavior patterns and also obtain certain prescriptive guidelines for system designers to motivate their users.

We demonstrate how game-theoretic modeling can be used to answer the above-mentioned questions. We study the user-user interaction and show how a user can find her optimal contribution level. Further, we demonstrate how a system designer can draw insights from such behavior patterns and exploit the selfishness of its users, to design multiple incentive mechanisms which help in improving the overall system performance, while possibly still benefiting the individual users.

2 Related Work and Domains

2.1 Why Are Social Media Networks Different?

Similar problems have been studied in multiple contexts in the past. From an economics perspective, the problem of 'public goods' [14] and optimal taxation [2] is well studied. However, emerging web-based 'societies' are fundamentally different as participation in them is voluntary and as such no taxation can be enforced. Only incentives, (if appropriate) can be given. Further, the reward on these social media sites etc. is typically a 'virtual currency' which has very different dynamics than real money. Such virtual currency (e.g. reputation points, extra bandwidth, virtual weaponry, gadgets etc.) is like 'fairy gold-dust' and typically costs the system designers exponentially less than their *perceived* value from user perspective. While some of this virtual currency is starting to be traded by users for real-world money [23], the marginal cost for system designers to grant such currency remains very low.

The issue of selfishness and contributions to a society also differ from scientists, and open-source software developers, as their 'contributions' are typically in-sync with their full time vocation. Hence they have direct and indirect professional benefits (grants, citations, downloads, jobs) from demonstrating their skill level. While a small percentage of contributors in media networks (e.g. directors on Youtube, or photographers on Flickr) might generate some career benefits from their contributions, these benefits remain atypical. More frequently the contributions to social media sites like Google Image Labeler, Photo-synth [11], Wikipedia do not involve rights and recognition as is common in open-source or scientific communities.

Social media network paradigms also differ from P2P [3, 17] or network routing disciplines as you cannot really calculate or regulate the usage characteristics. The produced content is truly a *public good* [14] for everybody like 'sunshine'.

Similarly, typical social media networks are also different from Mechanical Turk because the Turk users are not working to create any 'public good'. They are carrying out well defined tasks for an 'employer' who will 'privately' consume the generated media. Social media networks like Google Image Labeler, GalaxyZoo, Photo-synth, Wikipedia, Ushahidi on the other hand focus on creating 'public goods' for an open community.

Thus social media contributions deal with a unique set of parameters involving, voluntary participation, no taxation, real cost, virtual incentives, rare career benefits, and no regulations on usage characteristics. In this work we consider these differences in modeling and studying the media networks. Table 1 provides a summary of the comparison between different related scenarios and how social media contributions differs from each of them.

2.2 Related Work in Social Media

There have been attempts at enhancing user experience, and using that to get social media related tasks undertaken. *Human-computing* work by Von Ahn [25] is an

Table 1 Distinct characteristics of different type of contribution mechanisms

Case	Participation	Taxation	Reward currency	Career benefits	Usage pattern
Social media	*Voluntary*	*Not enforceable*	*Virtual*	*Rare*	*public good*
Physical societies	Mandatory	Enforceable	Real	–	–
Open-source software & scientific contributions	Voluntary	–	–	Yes	Partially regulated
P2P/ networking	Voluntary	–	Virtual	–	Monitored & regulated
Mechanical Turk	Voluntary	–	Real	Rare	Private

excellent example of this. Similarly 'Cognitive Surplus' [20] work has been arguing a case for providing mechanisms for users to contribute for common good.

Works like [10], study user behavior patterns in terms of the way they interact with social media sites (Youtube). They classify the users into different categories based on their access patterns, comment frequency, subscriptions etc. These are very important studies from the perspective of understanding general user behavior patterns. We maintain our focus in this paper though on user behavior from a *motivation* perspective and on *incentivizing* user contributions for better system performance.

Other works study the motivation of contributors on MovieLens, Wikipedia etc. from a sociological or psychological perspective [5, 19, 22]. Schroer et al. [19] discuss the intrinsic and extrinsic motivations for German Wikipedia contributors. Works like [5, 22] have highlighted how different forms of motivation (e.g. locked-out tools/ features in Slash-dot, extra weapons in World-of-Warcraft forum and comparative reputation/status within community etc.) can be used to motivate users. Similarly, Nov et al. [13] study how factors like 'user tenure', 'structural embeddedness', and 'motivation of self-development' affect user contributions on Flickr. However, all these studies are 'qualitative' and aimed at identifying what *type* of motivations work well in on-line communities. We, however, aim to provide 'quantitative' mechanisms to find out *how many* such extra features, weapons, or *how much* (e.g. extra bandwidth, reputation points etc.) are most suitable for different scenarios.

This book chapter builds upon a workshop paper [21], and reflects the advancements in our thinking process. Specifically, this chapter discusses multiple (instead of one 'paying the bonus' option discussed earlier) mechanism design options (e.g. social incentives, second price auction) available to designers. Further, this version relaxes the constraint of 'individually rational' contributions, and discusses how system designers can handle the cases wherein it is individually irrational for each user to contribute.

Mechanism design (i.e. defining rules of a game to achieve certain outcomes) is an area of growing importance (including recent Nobel prizes) in economics. It is also slowly making inroads into on-line communities (e.g. for creating optimal reputation feedback mechanisms in eBay like auction scenarios [1]).

3 Background: Game Theory

Game theory is a branch of applied mathematics that is used in the social sciences (most notably economics), biology, engineering, political science, international relations and is becoming increasingly relevant in computer science. It is used to mathematically capture behavior in strategic situations, in which an individual's utility from choices undertaken *depends on the choices made by others*. Thus it is often used for studying optimality and stable points in multi-agent problems, as opposed to conventional operation research or calculus based approaches which are well-suited for single agent optimization.

Here, we provide a quick refresher for some of the game-theoretic terms and concepts as relevant to this work.

1. *Game:* A game refers to any situation wherein multiple (2 or more) agents are making strategy decisions, and the chosen strategy effects the utility obtained by that agent *as well as* the other agents involved. Games can be *zero-sum*, wherein one agent's loss is considered other agent's gain e.g. war-like situations, or *non-zero-sum*, where one agent's loss does not necessarily means others gain. In our formalism the users (among themselves in Sect. 4) and the users and system designers (in Sect. 5) are both playing non-zero-sum games.
2. *Nash equilibrium:* Nash equilibrium is a solution concept in game theory which defines a point where each agent knows other agent's strategy options, and from which no agent has anything to gain by changing only her own strategy *unilaterally*. It is useful in giving guaranteed utility bounds to users, as once they choose their strategy based on the Nash equilibrium, it is in other agent's benefit to respond with their Nash equilibrium response. Any other response can only decrease their utility. Please note that multiple Nash equilibria may exist in the same game and agents can move or converge bilaterally to another Nash equilibrium (especially in repeated games).
3. *Mixed Strategy Nash equilibrium (MSNE):* A mixed strategy Nash equilibrium is a probabilistic variant of Nash equilibrium wherein the agents do not fix themselves to a single strategy but rather decide on the appropriate *mixture of strategies* which guarantees that no agent can gain anything by unilaterally diverting from it. Pure Nash equilibriums can be considered to be boundary cases of mixed strategy Nash equilibria.
4. *Mechanism design:* Mechanism design is the process of defining the rules of the game so as to lead it to a certain desired outcomes.

4 A Game-Theoretic Framework: The User Viewpoint

In this section we model user-user interaction patterns, and describe how rational (selfish) users may make optimal contribution decisions.

4.1 Problem Motivation

To motivate and ground the problem to a real life scenario from start, let us consider a citizen-journalism task (T) where a 'suspicious bag' left unattended at a train station can be reported by any of the N persons (agents) walking past. The gain (G) is common but the cost incurred (c_i) is individual. The problem from an agent(i)'s perspective is to find the percentage of times which he/she should report the bag himself/herself.

4.2 Problem Formulation

Let there be N social agents which can undertake a common task T. Let the cost for an agent i to undertake task T be c_i. While the costs are individual, the gains incurred (G) are taken to be common i.e. everybody gains equally. Each user can incur gain G, from a task if either it completes the task, or it does not, but somebody else completes it. Hence, the net utility of each agent EU_i is a function of its cost, gain, its chosen strategy in terms of how often to undertake the task (P_i), as well as the probability of task being undertaken by (at least one of the) *other* agents ($P_{\text{Do}}^{\text{All}-\{i\}}$). Hence, the problem from each agent's perspective is to find its best response strategy, (i.e. probability of doing task) which maximizes its net utility.

$$\underset{P_i \in [0,1]}{\text{argmax}} \, EU_i = f\left(c_i, G, P_i, P_{\text{Do}}^{\text{All}-\{i\}}\right). \tag{1}$$

The provided formulation incorporates two important characteristics of social media networks. Voluntary participation characteristic is innately represented in the problem statement, and the property of common gain once the task is completed is also made explicit.

   In the formulation presented here, we assume that the users are selfish agents, that there is no collusion or agreement between them, and that the cost incurred is positive and less than gain $0 < c \le G$ (we relax this constrain later). Also, in this formulation we use the term 'social media task', generically to include all relevant scenarios like those involving photo/video sharing, image tagging, commenting, content provision, linking, or content flagging. Lastly, we use 'agent' as a neutral term to represent any contributor who can undertake the relevant task.

4.3 Approach

We employ a game-theoretic framework to solve Eq. (1). To solve this problem, we start with a simple two person non-zero-sum game to study the interaction between two homogeneous agents and then iteratively add more complexities.

Table 2 Game between two users: matrix showing net utility for each user, under different strategies adopted by each player

Agent1 \ Other agent	Do	Don't
Do	$G-c, G-c$	$G-c, G$
Don't	$G, G-c$	$0, 0$

Table 3 Game between user (agent$_i$) and others: matrix showing net utility for agent$_i$, under different strategies adopted

Agent$_i$ \ Other agents	Do	Don't
Do	$G-c_i$	$G-c_i$
Don't	G	0

4.3.1 Two Agents, One Social Media Task

As shown in Table 2, for agent 1 there are two possible strategy options. He/she can either 'do' the task or 'don't' do it. If he/she chooses to do the task and the other agent also undertakes the task, then both will get the gain G but also incur the cost c. Agent 1's best case scenario is when he/she does not undertake the task but the other agent does it. Hence agent 1 will receive gain G without any cost while other agent will incur it. However, if both agents choose the 'don't' strategy, then there will be no gain achieved as the task is not undertaken.

The solution concept used in such settings where other agent's decisions affect your utilities is Nash equilibrium. The Nash equilibrium occurs when none of the agents can *unilaterally move to a better rewarding state* [16]. In this particular scenario, both the agents have a choice to either use the strategy 'Do' or 'Don't' and there exist three Nash equilibria. Two pure Nash equilibria exist at states [Do, Don't] and [Don't, Do] for agents 1 and 2, respectively. However, they favor one agent or the other and are unlikely to be maintained in long term. Hence, a mixed strategy solution seems a stable long term solution. The mixed strategy Nash equilibrium guarantees that:

$$P_i^* \in [0, 1], P_i \neq P_i^*: \quad EU_i\left(P_i^*, P_{\text{Do}}^{\text{All}-\{i\}^*}\right) \geq EU_i\left(P_i, P_{\text{Do}}^{\text{All}-\{i\}^*}\right), \quad (2)$$

where P_i^* is the optimal strategy for agent i, and $P_{\text{Do}}^{\text{All}-\{i\}^*}$ represents the cumulative effect of best possible strategy choices made by other agent(s).

A mixed strategy Nash equilibrium (MSNE) can be computed based on the condition of *choice indifference* i.e. when the agents do not gain (or lose) by changing their strategies [16]. This makes for a good equilibrium point because if the agents prefer one choice, obviously they shall go for the better choice. Thus the equilibrium needs to take place at a point where both agents do not stand to gain (or lose) any value by strategy selection.

In the given scenario (Table 2), if we equate the two options for (say) player 1. We get

$$p \cdot (G-c) + (1-p) \cdot (G-c) = p \cdot (G) + (1-p) \cdot 0, \quad (3)$$

where p is the probability of agent 2 choosing strategy 1 i.e. Do. This gives the value of p:

$$p = \frac{G - c}{G}. \tag{4}$$

For heterogeneous agents the formulation changes to

$$p_2 \cdot (G - c_1) + (1 - p_2) \cdot (G - c_1) = p_2 \cdot (G) + (1 - p_2) \cdot 0, \tag{5}$$

where p_2 is the probability of agent 2 choosing strategy 1 i.e. Do. This gives the value of p_2 as:

$$p_2 = \frac{G - c_1}{G} \tag{6}$$

and similarly p_1 is

$$p_1 = \frac{G - c_2}{G}. \tag{7}$$

If each agent chooses to undertake the task with just p_1 (resp. p_2) probability, s/he will get the same net utility as doing the task *always* by him/her self.

4.3.2 *N* Agents, One Social Media Task

For the *heterogeneous*, *N* agent case let us look again at Table 2. For a Nash Equilibrium to exist the two strategy options for agent i must provide same net utility. Thus:

$$(G - c_i) \cdot \left(P_{\text{Do}}^{\text{All}-\{i\}} \right) + (G - c_i) \cdot (1 - P_{\text{Don't}}^{\text{All}-\{i\}}) = P_{\text{Do}}^{\text{All}-\{i\}} \cdot G + 0, \tag{8}$$

where $P_{\text{Do}}^{\text{All}-\{i\}}$ is the probability of the task being 'done' by at least one of the $N - 1$ agents left after removing the ith agent from the set of 'all' agents.

Using the above equation, the equilibrium probability of the task being 'not done' by any of the other agents can be calculated as

$$P_{\text{Don't}}^{\text{All}-\{i\}} = \frac{c_i}{G}. \tag{9}$$

Similar equations can be formulated for all values of i.

$$P_{\text{Don't}}^{\text{All}-\{1\}} = \frac{c_1}{G},$$

$$P_{\text{Don't}}^{\text{All}-\{2\}} = \frac{c_2}{G},$$

$$\vdots$$

$$P_{\text{Don't}}^{\text{All}-\{N\}} = \frac{c_N}{G}. \tag{10}$$

Combining (multiplying) all of these equations gives

$$\{\overline{P_1} \cdot \overline{P_2} \cdots \overline{P_N}\}^{N-1} = \prod_{i=1}^{N} \frac{c_i}{G}, \tag{11}$$

where $\overline{P_i}$ is the probability of the task not being done by agent i.

Thus,

$$P_{\text{Don't}}^{\text{All}} = \sqrt[N-1]{\prod_{i=1}^{N} \frac{c_i}{G}} \tag{12}$$

or:

$$P_{\text{Do}}^{\text{All}} = 1 - \sqrt[N-1]{\prod_{i=1}^{N} \frac{c_i}{G}}, \tag{13}$$

$$P_{\text{Don't}}^{\text{All}} = P_{\text{Don't}}^{\text{All}-\{i\}} \cdot \overline{P_i} \tag{14}$$

and solving for $\overline{P_i}$ using values from Eq. (9) and Eq. (12) gives us the optimal contribution strategy, i.e. equilibrium probability for the agent i to 'not' undertake the task, as follows:

$$\overline{P_i} = \sqrt[N-1]{\prod_{i=1}^{N} \frac{c_i}{G}} \times \frac{G}{c_i}. \tag{15}$$

Note that the feasibility condition of a solution in which agent i must participate in the task is the condition that

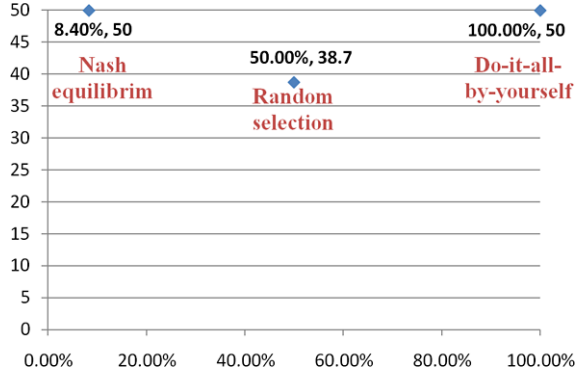$$\frac{c_i}{G} \geq \sqrt[N-1]{\prod_{i=1}^{N} \frac{c_i}{G}} \tag{16}$$

and that for homogeneous case (if applicable) the above equation reduces to

$$\overline{P_i} = \left(\frac{c}{G}\right)^{\frac{1}{N-1}}. \tag{17}$$

4.4 First Insights

To illustrate the basic concepts let us consider how the model and the derived solution works for a simple scenario of the unattended bag reporting where we have three agents each with a cost of 50, 60 and 70, respectively, while the common gain from reporting is 100. Thus using the parameter values ($c_1 = 50$, $c_2 = 60$, $c_3 = 70$,

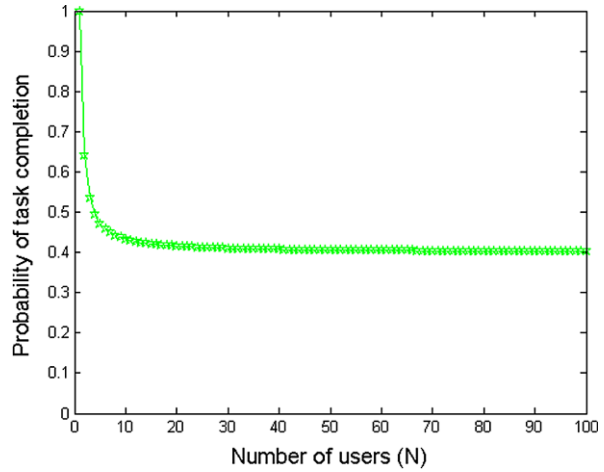**Fig. 1** Effect of different task-completion strategies upon agent 1's utility

$G = 100$) in Eq. (15), we get the equilibrium probabilities for each agent to undertake the task to be 0.084, 0.346 and 0.445, respectively. The utility for each agent was found to decrease if it moved away from the equilibrium point. It was also interesting to note that in spite of doing the task much lesser times, the agents obtained the same utility (values 50, 40, 30 resp.) as they would have achieved by doing the task *always* by themselves. This is shown in Fig. 1, for agent 1, who at MSNE, made a net gain of 50, by contributing just 8.4% times. Not changing the other agents response, this utility is the same as that obtainable by a naive 'do-it-all-by-yourself' approach which involves 100% contribution rate, and is more than that obtainable by random selection which involves 50% contribution rate.

We next proceeded to model the dynamics of a game scenario wherein a large number of selfish users are considering a common task ($c = 50$, $G = 100$, $N \in [1, 100]$). Upon varying the gain and costs we found the expected results of each user's utility increasing if the gain was high and decreasing if the cost was high. Since each user was selfishly guarding his/her incentives the utility of each agent was unaffected by the change in N.

However, the change in N had a dramatic impact on the percentage of times the common task was completed. As can be seen from Fig. 2, the task-completion probability decreases (from 1.0 to 0.5) as N increases. This was an interesting observation as intuitively one thinks that the probability of task getting done should increase with N, because if we multiply individual probabilities for everyone not doing the task, the overall probability of all not doing the task should be extremely low.

However, as the game-theoretic model makes explicit, the knowledge about large N, makes each agent adjust its task-completion probabilities in such a way which guards its individual utilities but can bring down the overall task completion. It reminds us of the apathy which can exist in large groups of selfish individuals undertaking common tasks. In hind-sight it also resonates well with how the 'free-rider' [8] problem is non-existent in single person teams. In n-person teams (especially non-coordinated teams), there is always a finite possibility of everybody deciding to free-ride on a particular task.

Fig. 2 Effect of large number of users on task-completion probability

4.5 Individual vs. Group Rationality

Our discussion so far has focused on the scenario where it was *rational* for an individual agent to undertake the task. This translated into a constraint that $c_i < G$ i.e. each agent should have some utility incentive for undertaking the task. We now relax this constraint i.e. also consider scenarios where an individual agent acting on its own would have *no incentive* for completing the task. This effect can be seen from Table 3, where $c_i > G$ setting means that *Do* strategy can only yield a negative utility. The *Don't* strategy on the other guarantees a zero or higher utility. Hence, in game-theoretic terms, the *Don't* strategy strictly dominates the *Do* strategy. This means that in a society of N agents, it is individually rational for each agent to not undertake the task. Unfortunately this also implies that the task will *never* get done in a such a setting.

Obviously though, from the perspective of the entire society, everybody will benefit if one person does indeed undertake the task. That one person would incur a loss but, $N - 1$ people will gain from it. *Hence what is rational to an individual is not the same as what is rational for the society.* The onus thus lies on the society and the system designer to device ways to encourage that one person to undertake the task.

To summarize, the model studied in Sects. 4.1 through 4.5, makes explicit two problems in a large setup of rational contributors. One, the probability of task completion decreases with larger N, and two, the tasks which are rational at group level, might never get done if they are irrational from individual agent perspective.

Both of these are alarming issues requiring corrective actions from the system designer.

5 The Designer's Viewpoint

While the issues mentioned in the previous section are alarming, the system designer has the advantage that it can actually define the 'rules of the game', which will ben-

efit the system performance. In fact, it can use the knowledge of how a rational user would behave in such systems, to devise effective mechanisms which will steer the system response in the desired direction. In this section, we study different mechanisms available for the system designer to do this, and their relative advantages and disadvantages.

5.1 Problem Formulation

Let us consider a case where the system designer has an interest in maximizing the probability of task completion. Let her benefit from each task completed be G_s. Hence, without any mechanism design the net utility obtained by the system designer is ($P_{\text{NoBonus}}^{\text{All}} \cdot G_s$), where $P_{\text{NoBonus}}^{\text{All}}$ is as found in previous section (Eq. (13)). However, using mechanism design, the system designer can enhance this performance.

Let us assume that the system designer is open to granting an extra benefit b to each user completing the task, so as to influence the users' (selfish) decision process of choosing how often to undertake the task. Clearly, providing the extra bonus b does entail some additional cost on the system.

Thus the overall utility problem for system designer is

$$\underbrace{\text{argmax}}_{C_b \in [0,\infty]} EU_s = \left(P_{\text{Bonus}}^{\text{All}}\right) \cdot G_s - \sum_{i=1}^{N} P_{i,\text{Bonus}} \cdot C_b, \tag{18}$$

where $P_{\text{Bonus}}^{\text{All}}$ is the overall probability of the task being undertaken by any of the agents after the bonus. P_{Bonus}^{i} is the probability of the task being undertaken by the ith agent after the bonus. C_b is the cost incurred by the system designer in granting the bonus.

Thus the system designer's problem is to design the best mechanism for granting the bonus b with system cost of C_b, such that overall system utility EU_s is maximized. EU_s will be maximized when a large increase is observed in probability of task completion due to the extra bonus, but the bonus cost is still low. Such a maxima will clearly involve trade-offs in terms of the bonus b's *amount* as well as *how it is granted*. We study three different mechanisms to support this.

5.2 Mechanism 1: Paying the Extra Benefit

The first option that the system designer can consider is paying the extra benefit b herself, if it provides overall system benefit. Note, however, that as per the dynamics of social media systems, such benefits are quite likely to be 'virtual' (e.g.

granting 'additional bandwidth', 'enhanced weaponry', 'titles/badges' or 'memorabilia'), and cost the system designer exponentially less than their *perceived* benefit by the user.[1]

For the current discussion, let us consider system cost for providing this additional bonus to be $C_b = \alpha \cdot b^{\frac{1}{\beta}}$, where α, β can be chosen based on the domain.

To quantify these parameters let us make simplifying assumptions of homogeneous costs and $C_b = b^{\frac{1}{2}}$ (i.e. $\alpha = 1$ and $\beta = 2$).

Based on extensions of the ideas already discussed under user-user interaction in Sect. 4.3, the various parameters involved in Eq. (18) can be computed as follows.

Extension of Eq. (15) defines the probability of task completion (with bonus) for agent i.

$$\overline{P_{i,\text{bonus}}} = \sqrt[N-1]{\prod_{i=1}^{N} \frac{c_i - b}{G}} \times \frac{G}{c_i - b} \tag{19}$$

which for the homogeneous case translates to

$$P_{i,\text{Bonus}} = 1 - \left(\frac{c - b}{G} \right)^{\frac{1}{N-1}}. \tag{20}$$

Similarly, extension of Eq. (13) can be used to compute the probability of overall task completion after bonus incentive.

$$P_{\text{Bonus}}^{\text{All}} = 1 - \left(\frac{c - b}{G} \right)^{\frac{N}{N-1}}. \tag{21}$$

Note that now we have the values for all the parametric components of Eq. (18), and the value of optimal bonus which maximizes EU_s can be computed using the standard calculus maximization methods or by choosing maxima as obtained by numerical methods.

5.3 Mechanism 2: Social Incentives

The second option which the system designer can consider is exploiting the 'social' aspect of social media networks, and realize that the friendships, contacts and mutual role-play is an important incentive to the users. To study this aspect, let us consider the common gain being discussed so far to be G_1. Let the social gain (G_2) for each user be dependent on the size of her social network:

$$G_{2,i} = k_1 \cdot S_i. \tag{22}$$

---

[1] The use of perceived changes in games to try and influence agent interaction is well studied under hyper-game theory [24].

Table 4 Game between user (agent$_i$) and others in case of social incentives: matrix showing net utility for agent$_i$, under different strategies adopted

Agent$_i$ Other agents	Do	Don't
Do	$G_1 + G_{2,i} - c_i$	$G_1 + G_{2,i} - c_i$
Don't	G_1	0

Over large user set, lets define the size of user's social network to be dependent on the size of the network as follows: $S_i = k_2 \cdot N^{\frac{1}{\beta}}$, where α, β can be chosen based on the domain. For the ease of presentation, let us assume that $S_i = k_2 \cdot \sqrt{N}$. Hence we can rewrite Eq. (22) as follows:

$$G_{2,i} = k_1 \cdot k_2 \cdot \sqrt{N}. \tag{23}$$

As shown in Table 4, and by extension of Eq. (9), we get

$$P_{\text{Don't,social}}^{\text{All}-\{i\}} = \frac{c_i - G_{2,i}}{G_1} \tag{24}$$

or if we define $\phi = c_i - G_{2,i}$:

$$P_{\text{Don't,social}}^{\text{All}-\{i\}} = \frac{\phi_i}{G_1} \tag{25}$$

which gives

$$P_{\text{Social}}^{\text{All}} = 1 - \left(\frac{\phi}{G_1}\right)^{\frac{N}{N-1}} \tag{26}$$

and

$$P_{i,\text{social}} = 1 - \left(\frac{\phi_i}{G}\right)^{\frac{1}{N-1}}. \tag{27}$$

This additional social benefit G_2 can be compared to the bonus (b) granted by the administrator in previous section. In fact, using Eq. (23), we can compute the size of network N at which the socially derived incentive will become equal to the system administrator granted bonus (b):

$$N = \left(\frac{b}{k_1 \cdot k_2}\right)^2. \tag{28}$$

The *'tipping point'* i.e. the size of the network beyond which, network will generate enough benefit to effect the task-completion probability and consequently enhance user/system utilities occurs at a specific point when $b = G - c$ i.e.

$$N_1 = \left(\frac{G - c}{k_1 \cdot k_2}\right)^2. \tag{29}$$

Table 5 Game between user (agent$_i$) and others: matrix showing net utility for agent$_i$, under different strategies adopted

Agent$_i$ / Other agents	Do	Don't
Do	$G - c_i + b^*$	$G - c_i + b^*$
Don't	G	0

There also exists a second threshold N_2 beyond which the bonus exceeds the cost ($b > c$) incurred and all the tasks get completed and system designer's utility reaches its maximum possible value.

$$N_2 = \left(\frac{c}{k_1 \cdot k_2} \right)^2. \tag{30}$$

In handling such large networks, the system designer's responsibility is limited to supporting the infra-structure for larger network of users, which is typically much smaller than the benefits accrued. Thus this option might perform better than that of explicitly providing benefits (as in Sect. 5.2) for networks with large user base.

5.4 Mechanism 3: Second Price Auction Based Mechanism

The 'second price auction' inspired mechanism involves the system designer announcing a bonus value which is 'slightly less' than the second lowest cost user ($b^* = c_2 - \epsilon$). As shown in Table 5, this would ensure that the 'Do' strategy dominates for exactly one user (the one with the lowest cost). This mechanism works on the assumption that the users realize (or are made aware with a one-time broadcast e.g. at networking joining) that there will be always one user whose cost will be lower than the bonus offered. Thus at 'run-time', the agents do not need any additional information, collaboration, or collusion, and are practically assured of task completion by some other agent *if* their cost is greater than the incentive offered.

If their cost is indeed lesser than the bonus then the user *knows* that she should actually undertake the task, both for everybody's benefit (group rationality) as well as individual rationality ('Do' strategy dominates). In a way this strategy works on the basis on (implicitly) assigning the responsibility to one agent (rather than distributing it between everybody), and rewarding her appropriately.

This mechanism (where applicable) ensures task completion with a (theoretical) probability of 1. The overall system performance would touch theoretical bounds, as it is (almost) as efficient as best possible. Further it incentivizes the lowest cost user, by guaranteeing it more than its fair share i.e. the 'Shapley value' of cooperating with such a coalition.

5.4.1 Shapley Value

The Shapley value in Game theory, captures the 'fair' estimate of an individual agent i's contribution to the overall coalition. It is computed as a function of marginal

utility delivered by agent i, averaging over all the different sequences according to which the most profitable set S could be built up from the empty set. Mathematically, it is defined as

$$\phi(i) = \sum_{S \subseteq N - \{i\}} \frac{|S|!(N - |S| - 1)!}{N!} \left(v\big(S \cup \{i\}\big) - v(S) \right), \tag{31}$$

where S is the set of agents deciding to join the 'coalition', N is the size of the overall network, $v(S)$ is the value attainable by the set S.

For the formulation being considered here, $v(S \cup \{i\})$ i.e. the value of the Set including the lowest cost agent i is

$$v\big(S \cup \{i\}\big) = G_s - c_1. \tag{32}$$

The value of the set without the lowest cost agent is the case where the second lowest cost agent undertakes the task i.e.

$$v(S) = G_s - c_2. \tag{33}$$

All the other agents present do not affect this computation, rather they simply gain from whichever agent completes the task. Hence, the critical point is the case with size $n = S - 1$, with the lowest cost agent deciding to join (or not to join) the set. Using Eq. (31), thus the Shapley value of the agent with lowest cost is $c_2 - c_1$, which will be greater than zero. Hence, if the system designer pays the lowest cost agent its cost, plus the additional value which it brings to the coalition, it should have no utilitarian reasons to defect. This in turn assures the system designer of a (near[2]) optimum system performance.

5.5 Case Study

To study the impact of the proposed mechanisms we undertook experimental case studies. We considered a scenario with system gains as $G = 100$ and $G_s = 50$ and varied the N and costs to study different scenarios. The observed effects of each of the three mechanisms designed, on the performance of the system gave interesting insights.

---

[2]Strictly speaking, system designer needs to provide the lowest cost agent with a bonus that is just a fraction above its cost (i.e. $b = c_1 + \epsilon$). This will make the 'Do' strategy dominate for the user. However, in a cooperative setting, it is often considered better to grant a 'fair share' of the additional benefit the agent brings to the system by participating. Further still, the extra bonus serves as an implicit signaling mechanism to ensure overall system gains.

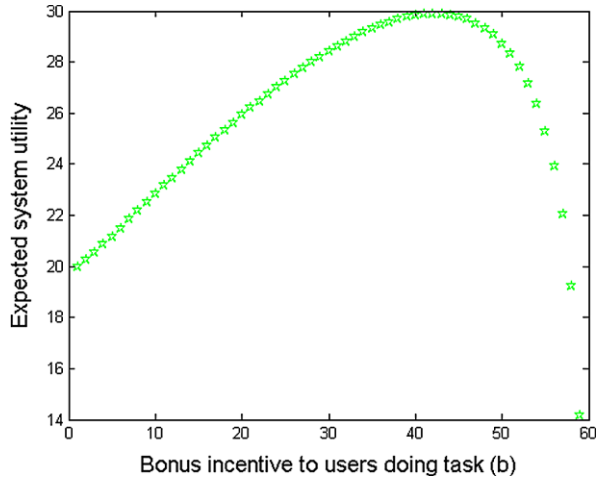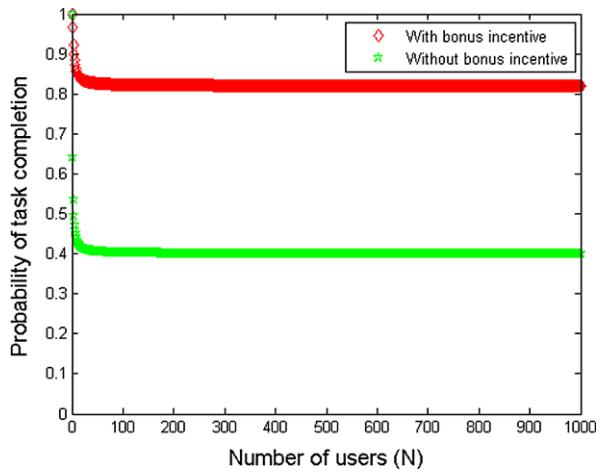**Fig. 3** Mechanism 1: System utility vs. bonus incentive

Fig. 4 (Color online) Mechanism 1: Effect of bonus on task-completion probability

5.5.1 Mechanism 1: Paying the Extra Benefit

We varied the bonus provided to the users in the range $b \in [0, 50]$ (see Fig. 3), and found the system's utility is maximized at $b = 33$. Needless to say, similar values can be obtained for other scenarios too by simple parameter changes.

The net system utility for this task was found to be 31.32, which is higher than the value of 25, as obtained without mechanism design.

As can be seen from Fig. 4, the probability of the task being undertaken by at least one of the N users, increases to around 0.83 (red colored plot) with the use of this additional incentive value. This is as opposed to the 0.50 probability without the incentive mechanism (green plot, also see Fig. 2). The value of the net utility for the homogeneous players was also found to increase to 83 (from 50.0 without bonus).
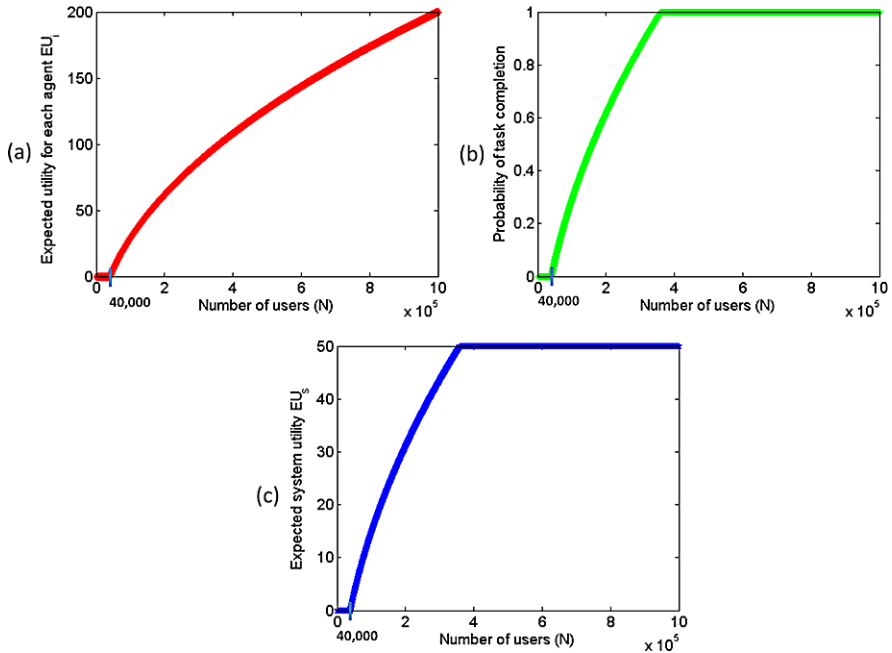
Fig. 5 Mechanism 2: social incentives. Impact of the size of network on the (**a**) profit for each additional user, (**b**) system designer's utility, and (**c**) overall probability of task completion
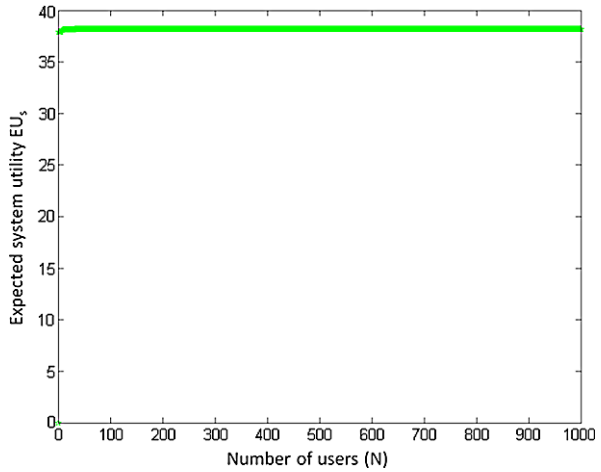
Hence this mechanism was indeed useful in finding an optimal level of incentive level which maximizes the system's net utility and increases the probability of task completion while also enhancing the net utility of each user.

5.5.2 Mechanism 2: Social Incentives

Mechanism 2 'Social incentives' also had a very interesting impact on the system and user performance. For example, under individually irrational (i.e. $c_i > G$) settings the task was completed 0% of times without any mechanism design. However, the social incentives created by the large size of network was indeed found to be effective in providing the social incentive for users to undertake the task. The *tipping point* value of the network size for the settings of $c = 150$, $G = 100$, $N \in [2, 1Million]$ was found to be at $N_1 = 40,000$. This clearly corroborates Eq. (29) which estimates this value.

As shown in Fig. 5(a), each additional agent joining the network beyond this point obtained more returns. we consider this '*tipping point*' to be important because each new user now has increasing benefits to join the network. This can explain the snow-ball effect noticed in social media networks, where the size of the network grows exponentially beyond a certain critical point. This also corroborates well with the notion of 'tipping point' as postulated and observed in both offline [18] and online [15] networks.

Fig. 6 System utility for
Mechanism 3: Second price
auction based mechanism at
different network sizes

The increased number of users, and utility, in turn benefits the system utility and
probability of task completion too. As shown in Fig. 5(b), the utility for the system
designed also tipped at the value of $N_1 = 40,000$ and continued until $N_2 = 360,000$,
causing it to jump to its maximum value of 50. Similarly, the probability of task
completion also kept increasing beyond N_1 and reached its maximum possible value
of 1.0 at N_2.

5.5.3 Mechanism 3: Second Price Auction Based Mechanism

The second price auction mechanism also showed some interesting effects. For ex-
ample, under a large heterogeneous user base where it is individually irrational for
users to contribute ($c_i \in [150 \pm 10]$, $G = 100$, $N \in [2, 1000]$), the base case scenario
resulted in 0% task completion and zero system utility. However, paying the lowest
cost user a perceived bonus of 140.07, resulted in task-completion probability of
1.0, and system utility of 38.17. This is shown in Fig. 6 where we also notice that
system utility (and task-completion probability) were independent of the number of
users on the network. This is quite different from the reduced probability of task
completion for larger network sizes, as was earlier seen in Fig. 2.

5.6 Comparison of Different Approaches

To understand the impact of different mechanisms on the system designer's utility,
we undertook a computational experiment. We studied the best utility attainable
by the system designer under different situations of the application domain. The
variables used for different situations were as follows:

$G_s = 50$ (System designer's gain from task completion)
$G = 100$ (Each user's gain from task completion)

Table 6 Effect of different mechanisms on system performance

S. no	Settings			Mechanism used			
	Homogeneous agents	Very Large N	Individually Rational	Base case	Paying the extra benefit	Social incentives	2nd price auction
1	Yes	Yes	Yes	25.00	31.32	50.00	−15.10
2	Yes	Yes	No	00.00	22.00	50.00	−62.80
3	Yes	No	Yes	25.02	31.33	50.00	−14.89
4	Yes	No	No	00.00	22.01	00.00	−62.40
5	No	Yes	Yes	25.17	−22.53	50.00	43.68
6	No	Yes	No	00.00	−22.53	23.79	38.17
7	No	No	Yes	25.12	−22.80	37.10	43.67
8	No	No	No	00.00	−22.53	00.00	38.17

The cost for each agent (c_i) was 50 for the individually rational situation (i.e. $G > c_i$, hence each agent can have some reasons to undertake the task) and 150 for the individually irrational situation (i.e. $G < c_i$, thus each agent individually has no motivation to undertake the task).

Homogeneous agents all had exactly the above-mentioned costs, while the *Heterogeneous* agent's cost varied randomly in the range ±10 from the mean value.

The *very large N* setting was studied on the agent population of 1,000,000 (1 million) while *non very large N* was studied on a population of 1,000.

As shown in Table 6, there is no *one-fits-all* mechanism available to enhance the system designer performance. However, it is interesting to note that in each situation, one or more mechanisms designed outperform the base case scenario (of not using any mechanism to influence the system performance). Broadly speaking, the base case scenario is unable to generate any system utility under the cases where it is individually irrational for agents to undertake the task. Mechanism 1 (i.e. paying the extra benefit) outperforms the base case in homogeneous settings but does not work well under heterogeneous user setting. The second price auction (Mechanism 3) typically outperforms all other mechanisms in the heterogeneous agent setting. The only exception is the *very large N* scenario, where Mechanism 2 (social incentives) works well. The second price auction mechanism is not suited for homogeneous agents case, and leveraging the social benefits is often a good choice there. In fact, the social incentive mechanism works well irrespective of the homogeneous or heterogeneous nature of agents, if the network size is very large.

While summarizing the mechanisms, this table clearly demonstrates that one or more mechanisms always outperform the base case scenario. It also provides prescriptive guidelines for the system designer to choose the most appropriate mechanism design based on the settings at hand. For example, a system designer may choose to Mechanism 1 and pay the bonus herself initially, but switch to Mechanism 2 i.e. social incentives beyond the 'tipping point'. In fact a modeling approach like ours allows the system designer to estimate (and cater for) this value much before when the tipping point actually occurs in the network.

Note also that the discussed formulation and mechanisms, consider the distinct characteristics of social media systems as we consider voluntary participation, common gain amongst all users, virtual reward currency which typically costs exponentially less to the system, and social benefits which members bring to each other.

6 Discussion and Future Work

We realize that the framework discussed works on a strong rationality or selfishness assumption. In near future, we plan to extend the work to consider a 'bounded rationality' [7] model for humans. Also, our current model considers only explicitly quantifiable incentives. We realize that other than very few scenarios (e.g. second-life monetization [23]), explicit quantification of gains and costs is still difficult. As what happens with all nascent fields, no numerical data are readily available on costs and gains for social media contributions (e.g. Wiki content provision, or Youtube video sharing). While we had to study the current framework using numerical case studies, we want to undertake more work to better quantify such costs and gains.

Our current model considers tasks that only need contribution from one user and other contributions are redundant. We are working on extending the approach to consider cases which require k different contributions or viewpoints (e.g. minimum two images required for stereoscopy or minimum five spam flags for post removal etc.). We are also considering a graded utility model wherein the value of each successive contribution gets lower but is still finite. Lastly, it would also be relevant to consider scenarios involving j tasks, each needing k contributions amongst the N users.

We also intend to broaden the motivation factors to considered to include intrinsic motivation factors and concepts like Maslow's hierarchy and using them appropriately in the future models. Lastly, the enhancements obtained due to mechanism design were gained (amongst other reasons like optimality based incentive levels), because of the setting that benefits granted were 'virtual', while the costs and gains were 'real-world'. However, we feel this is indeed true in many social media environments like citizen-journalism, image-labeling, Wiki-contributions etc.

While we admit, that our modeling is by no means perfect, this is meant to be a first step in drawing research interest toward this area. The value of this paper lies in providing food-for-thought to social media designers and developers charged with creating crowd-sourcing, media applications that require individual contributions to enhance the overall value of the application and its content.

7 Conclusions

In this work we have proposed a game-theoretic framework for studying user behavior and motivation patterns in social media networks. We have modeled users

as rational selfish agents, and considered domain attributes like voluntary participation, *virtual* reward structure, social incentives, and public-sharing to model the dynamics of this interaction. We first studied the aspects of user-user interaction and used that to find the most appropriate contribution strategy from each user's perspective. The model created showed how the probability of task completion may decrease with large *N* and made explicit the concepts like free-rider problem. It also explained how individually irrational tasks may never get done in a collective network. We next studied the dynamics of system-user interaction, and showed how a system designer can design different incentive mechanisms which can help in improving the overall system performance under different settings, while possibly still benefiting the individual users.

References

1. Dellarocas, C.: Reputation mechanism design in online trading environments with pure moral hazard. Inf. Syst. Res. **16**(2), 209–230 (2005)
2. Diamond, P.A., Mirrlees, J.A.: Optimal taxation and public production II: Tax rules. Am. Econ. Rev. **61**(3), 261–278 (1971)
3. Friedman, E., Halpern, J., Kash, I.: Efficiency and Nash equilibria in a scrip system for p2p networks. In: Proceedings of the 7th ACM Conference on Electronic Commerce. ACM, New York (2006)
4. Google Image Labeler: http://images.google.com/imagelabeler/, Last accessed: 07/06/2009
5. Harper, F.M., Li, S.X., Chen, Y., Konstan, J.A.: Social comparisons to motivate contributions to an online community. Persuasive Technology, 4744/2007, 2007
6. Howe, J.: The rise of crowdsourcing. Wired Mag. **14**(6), 1–4 (2006)
7. Kahneman, D.: Maps of bounded rationality: Psychology for behavioral economics. Am. Econ. Rev. **93**(5), 1449–1475 (2003)
8. Kim, O., Walker, M.: The free rider problem: Experimental evidence. Public Choice **43**(1), 3–24 (1984)
9. Lévy, P.: Collective Intelligence: Mankind's Emerging World in Cyberspace. Perseus Books, Cambridge (1997)
10. Maia, M., Almeida, J., Almeida, V.: Identifying user behavior in online social networks. In: SocialNets '08: Proceedings of the 1st Workshop on Social Network Systems, pp. 1–6 (2008)
11. Microsoft photo-synth project: http://photosynth.net/, Last accessed: 07/06/2009
12. New York Times launches 'citizen journalism' sites: http://washingtonsquarepark.wordpress.com/2009/03/03/new-york-times-launches-citizen-journalism-sites/, Last accessed: 07/06/2009
13. Nov, O., Naaman, M., Ye, C.: Motivational, structural, and tenure factors that impact online community photo sharing. In: Proceedings of the Third International AAAI Conference on Weblogs and Social Media (ICWSM 2009). Retrieved November, vol. 13, p. 2009 (2009)
14. Olson, M.: The Logic of Collective Action: Public Goods and the Theory of Groups. Harvard Univ. Press, Cambridge (1971)
15. Onnela, J.-P., Reed-Tsochas, F.: Spontaneous emergence of social influence in online systems. Proc. Natl. Acad. Sci. (2010)
16. Osborne, M.J.: An Introduction to Game Theory. Oxford University Press, Oxford (2003)
17. Park, J., van der Schaar, M.: A game theoretic analysis of incentives in content production and sharing over peer-to-peer networks. In: IEEE Thematic Meetings on Signal Processing: Signal and Information Processing for Social Networks, 2009
18. Schelling, T.C.: Micromotives and Macrobehavior. Norton, New York (1978)

19. Schroer, J., Hertel, G.: Voluntary engagement in an open web-based encyclopedia: Wikipedians and why they do it. Media Psychol. **12**(1), 96–120 (2009)
20. Shirky, C.: Cognitive Surplus: Creativity and Generosity in a Connected Age. Penguin, Baltimore (2010)
21. Singh, V., Jain, R., Kankanhalli, M.: Motivating contributors in social media networks. In: Proceedings of the first SIGMM Workshop on Social Media, pp. 11–18. ACM, New York (2009)
22. Sun, L., Vassileva, J.: Social visualization encouraging participation in online communities. In: Groupware: Design, Implementation, and Use, 2006
23. Terdiman, D.: The Entrepreneur's Guide to Second Life: Making Money in the Metaverse. Sybex, Berkeley (2007)
24. Vane, R.: Advances in hypergame theory. In: Workshop on Game Theoretic and Decision Theoretic Agents—Conference on Autonomous Agents and Multi-Agent Systems, 2006
25. von Ahn, L.: Games with a purpose. Computer **39**(6), 92–94 (2006)
26. Wikipedia: http://www.wikipedia.org