contributed articles

Praise, pay, and promote crowd-member workers to elicit desired behavioral responses and performance levels.

BY OGNJEN SCEKIC, HONG-LINH TRUONG, AND SCHAHRAM DUSTDAR

Incentives and Rewarding in Social Computing

INCENTIVES AND REWARDS help align the interests of employees and organizations. They first appeared with the division of labor and have since followed the increasing complexity of human labor and organizations. As a single incentive measure always targets a specific behavior and sometimes additionally induces unwanted responses from workers, multiple incentives are usually combined to counteract the dysfunctional behavior and produce desired results. Numerous studies have shown the effectiveness[20] of different incentive mechanisms and their selective and motivational effects.[14] Their importance is reflected in the fact that most big and mid-size companies employ some kind of incentive measures.

Expansion of social computing[18] will include not only better exploitation of crowdsourcing[5] but also

solutions that extend traditional business processes (see Figure 1); increasing research interest seems to confirm the trend.[3] Several frameworks aiming to support such new collaboration models are being developed (such as socially enhanced computing[6,7]). These new forms of social computing are intended to support greater task complexity, more intelligent task division, complex organizational and managerial structures for virtual teams, and virtual "careers." With envisioned changes, incentives will also gain importance and complexity to address workers' dysfunctional behavior. This new emphasis calls for automated ways of handling incentives and rewards. However, the social computing market is dominated by flat and short-lived organizational structures, employing a limited number of simple incentive mechanisms. That is why we view the state of the social computing market as an opportunity to add novel ways of handling incentives and rewards.

Here, we analyze incentive mechanisms and suggest how they can be used for next-generation social computing. We start with a classification of incentive mechanisms in the literature and in traditional business organizations, then identify elements that can be used as building blocks for any composite incentive mechanism and show the same elements are also used in social computing, even though the resulting schemes lack the complexity

» **key insights**

■ **Existing social computing platforms lack the ability to formulate, compose, and automatically deploy appropriate incentive mechanisms needed for complex agent collaborations.**

■ **Analyzing incentive mechanisms in traditional companies and in social computing platforms reveals how incentive mechanisms consist of simpler reusable elements that can be formally modeled.**

■ **Formal modeling of incentive mechanisms allows composition, optimization, and deployment of portable and dynamically adaptable incentive schemes for social computing environments.**

contributed articles

Figure 1. Social computing is evolving from social networks and crowdsourcing to include structured crowd organizations able to solve complex tasks.
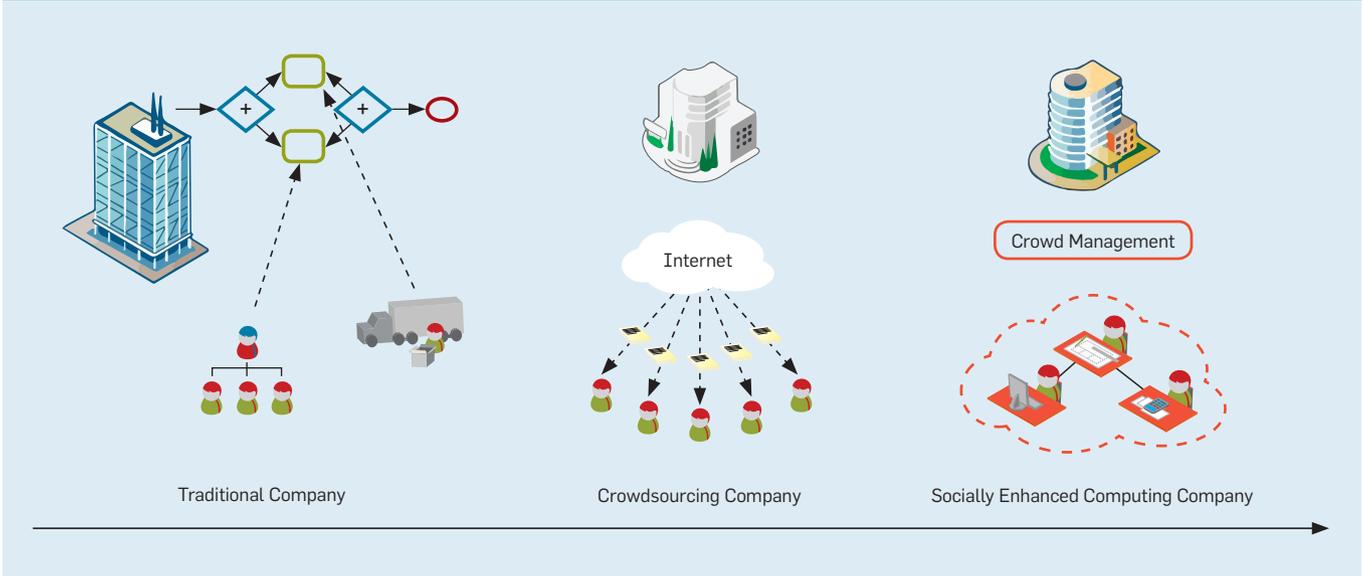
needed to support advanced business processes; we conclude with our vision for future developments.

Related Work

In economics, incentives are predominantly investigated within the models set out in the Principal-Agent Theory,[13,20] introducing the role of a principal that corresponds to owners or managers who delegate tasks to a number of agents corresponding to employees (workers) under their supervision. The principal offers the agents an incentive to disclose part of their personal performance information (signal) to devise an appropriate contract.

Only a few articles in the computer science literature have addressed incentives and rewards, usually within specific application contexts (such as social networks and wiki systems,[9,32] peer-to-peer networks,[23] reputation systems,[22] and human micro-task platforms[16,17,28,29]). Much recent research aims to find suitable wage models for crowdsourcing.[11] However, to the best of our knowledge, the topic has not been comprehensively addressed.

Incentive Mechanisms

The incentive mechanisms we cover here involve most known classes of incentives used in different types of organizations: companies, not-for-profit (voluntary), engineering/design, and crowdsourcing. Different organizations employ different (combinations of) incentive mechanisms to stimulate specific responses from agents:

Pay per performance (PPP). The guiding principle says all agents are to be compensated proportionally to their contribution. Labor types where quantitative evaluation can be applied are particularly suitable. In practice, it shows significant, verifiable productivity improvements—25% to 40%—when targeting simple, repetitive production tasks, both in traditional companies[15] and in human intelligence tasks on Amazon's Mechanical Turk platform.[17] Other studies, as cited by Prendergast,[20] conclude that approximately 30% to 50% of productivity gains is due to the filtering and attraction of better workers, due to the selective effect of the incentive. This important finding explains why greater profit can be achieved even with relatively limited incentives. PPP is not suited for large, distributed, team-dependent tasks, where measuring individual contributions is inherently difficult. However, it is frequently used to complement other incentive mechanisms.

Quota systems and discretionary bonuses. With this mechanism, the principal sets a number of performance-metrics thresholds. When agents reach a threshold they are given a one-off bonus. Quota systems evaluate whether a performance signal surpasses a threshold at predefined points in time (such as annual bonuses). On the other hand,

discretionary bonuses may be paid whenever an agent achieves a performance level for the first time (such as a preset number of customers).

Two phenomena[20] typically accompany this mechanism:

▸ The effort level always drops off following an evaluation if the agent views the time until the next evaluation as too long; and

▸ When the performance level is close to an award-winning quota, motivation is significantly greater.

Appropriate evaluation intervals and quotas must be set in such a way that they are achievable with a reasonable amount of additional effort, though not too easily. The two parameters are highly context-dependent, so can be determined only after observing historical records of employee behavior in a particular setup. Ideally, these parameters are dynamically adjustable.

Deferred compensation. This mechanism is similar to a quota system, in that an evaluation is made at predefined points in time. The subtle but important difference is that deferred compensation takes into account three points in time: t_0, t_1, t_2. At t_0 an agent is promised a reward for successfully passing a deferred evaluation at t_2. The evaluation takes into account the period of time $[t_1, t_2]$, not just the current state at t_2. In case $t_1 = t_0$ the evaluation covers the entire interval.

Deferred compensation is typically used for incentivizing agents working

on complex, long-lasting tasks. The advantage is it allows more objective assessment of an agent's performance at a particular time. Agents are also given enough time $[t_0,t_1]$ to adapt to the new conditions, then prove the quality of their work over some period of time $[t_1,t_2]$. The disadvantage of this mechanism is it is not always applicable, since agents are not always able to wait for a significant portion of their compensation. A common example of this mechanism is the "referral bonus," or a reward for employees who recommend or attract new employees or partners to the company.

Relative evaluation. Although this mechanism can involve many variations the common principle is that an entity is evaluated with respect to other entities within a specified group. The entity can be a human, a movie, or a product. The relative evaluation is used mainly for two reasons:

▸ By restricting the evaluation to a closed group of individuals, it removes the need to set explicit, absolute performance targets in conditions where the targets are not easily set due to the dynamic and unpredictable nature of the environment; and

▸ It has been empirically proved that

people respond positively to competition and comparison with others (such as in Tran et al.[30]).

Promotion. Empirical studies (such as Van Herpen et al.[31]) confirm that the prospect of a promotion increases motivation. A promotion is the result of competition for a limited number of predefined prizes. Promotion schemes are usually treated under the tournament theory,[14] though there are other models, too. The prize is a better position in an organization's hierarchy, bringing higher pay, more decision-making power, and greater respect and esteem. Promotions include basic

Table 1. Adoption of incentive mechanisms in different business environments: + = low, ++ = medium, +++ = high; application considerations (right).

| Mechanism | Usage Environments | | | Application Considerations | | | |
| | Traditional Company | | | | | | |
	SME	Large Enterprise	Social Computing	Positive Application Conditions	Negative Application Conditions	Advantages	Disadvantages
Pay Per Performance	++	+++	+++	quantitative evaluation possible	large, distributed, team-dependent tasks; measurement inaccuracy; when favoring quality over quantity	fairness; effort continuity	oversimplification; decreased solidarity among workers
Quota/ Discretionary Bonus	+	+++	+	recurrent evaluation intervals	constant level of effort needed	allows peaks/ intervals of increased performance	effort drops after evaluation
Deferred Compensation	+	+++	+	complex, risky, long-lasting tasks	subjective evaluation; short consideration interval	better assessment of achievements; paying only after successful completion	workers must accept risk and wait for compensation
Relative Evaluation	+	++	+++	cheap group-evaluation method available	subjective evaluation	no absolute performance targets; eliminates subjectivity	decreases solidarity; can discourage beginners
Promotion	++	+++	+	need to elicit loyalty and sustained effort; when subjective evaluation is unavoidable	flat hierarchical structure	forces positive selection; eliminates centrality bias	decreases solidarity
Team-based Compensation	+	++	+	complex, cooperative tasks; inability to measure individual contributions	when retaining the best individuals is priority	increases cooperation and solidarity	disfavors best individuals
Psychological	+	+	++	stimulate competition; stimulate personal satisfaction	when cooperation must be favored	cheap implementation	limited effect on best and worst workers (anchoring effect)

ideas from relative evaluation and quota systems. They eliminate centrality bias and enforce positive selection. The drawback is that by valuing individual success, agents can be de-motivated from helping each other and engaging in collaborations. They often incorporate subjective evaluation methods, though other evaluation methods are also possible in rare instances.

Team-based compensation. This mechanism is used when the contributions of individual agents in a team environment are not easily identified. With it, the entire team is evaluated and rewarded, with the reward split among team members. The reward can be split equally or by differentiating individual efforts within the team. The latter is a hybrid approach combining a team-based incentive, together with an incentive mechanism targeted at individuals, to eliminate dysfunctional behavior. Some studies (such as Pearsall et al.[19]) show this approach is indeed more effective than pure team-based compensation. One way to avoid having to decide on the amount of compensation is to tie it to the principal's profit, and is called "profit sharing." Team-based compensation is also susceptible to different dysfunctional behavioral responses. Underperforming agents effectively hide within the group, while the performance of the better-performing agents is diluted. Moreover, teams often exhibit the free-rider phenomenon,[12] where individuals waste more resources (such as time, materials, and equipment) than they would if individual expenses were measured. Minimizing these negative effects is the primary challenge when applying this mechanism.

Psychological incentive mechanisms. Psychological incentives are the most elusive, making them difficult to define and classify, since they often complement other mechanisms and can be described only in terms of psychological actions. A psychological incentive must relate to human emotions and be advertised by the principal and be perceived by the agent. The agent's perception of the incentive affects its effectiveness. As this perception is context-dependent, choosing an adequate way of presenting the incentive is not trivial; for example, choosing and promoting an "employee of the month" is

The effort level always drops following an evaluation if the agent views the time until the next evaluation as too long.
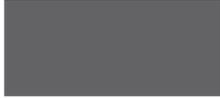
effective in societies where the sense of common good is highly valued. In more individually oriented environments competition drives performance. A principal may choose to exploit this fact by sharing comparisons with the agents. Acting on human fear is a tactic commonly (mis)used (such as through the threat of dismissal or downgrading). Psychological incentives have long been used in video games, as well as in more serious games, to elicit player dedication and motivation. Such techniques (including gamification[4]) are also used to make boring tasks (such as product reviews and customer feedback) feel more interesting and appealing (see Table 1).

Analyzing Incentive Mechanisms
No previous work has analyzed incentives past the granularity of incentive mechanisms, preventing (development of) generic handling of incentives in information systems. Our goal is to identify finer-grain elements that can be modeled individually and used in information systems to compose and encode the described incentive mechanisms (see Figure 2). Such a conceptual model would allow specification, execution, monitoring, and adaptation of various rewarding and incentive mechanisms for virtual teams of humans.

Each incentive mechanism described earlier can be modeled using three incentive elements:

Evaluation method. Provides input on agent performance to be evaluated in the logical context defined in the incentive condition;

Incentive condition. Contains the business logic for certain rewarding actions; and

Rewarding action. Is meant to influence future behavior of agents.

Though we describe these elements informally here, their true power lies in the possibility of being formally modeled. An evaluation method can ultimately be abstracted to an evaluation function, incentive condition to a logical formula, and rewarding action to a function, structural transformation, or external event. These abstractions allow us to formally encode each incentive mechanism and thus program many real-world reward strategies for crowds of agents working on tasks ranging from simple image tagging to

Table 2. Application and composability considerations for evaluation methods.

Evaluation Methods		Advantages	Disadvantages	Active Human Participation	Issues	Alleviated By	Solving	Typical Use
		Application Considerations			**Composability**			
Individual	Quantitative	fairness, simplicity, low cost	measurement inaccuracy	no	multitasking	peer evaluation; indirect evaluation; subjective evaluation	issues due to subjectivity	pay per performance; quota systems; promotion; deferred compensation
	Subjective	simplicity, low cost	subjectivity; inability to assess different aspects of contribution	yes	centrality bias; leniency bias; deliberate low-scoring; embellishment; rent-seeking activities	incentivizing decision maker to make honest decisions (such as through peer evaluation)	multitasking	relative evaluation; promotion
Group	Peer	fairness; low cost in social computing environment	active participation required	yes	preferential attachment; coordinated dysfunctional behavior of voters	incentivizing peers (such as also by peer evaluation)	multitasking; issues due to subjectivity	relative evaluation; team-based compensation; psychological
	Indirect	accounts for complex relations among agents and their artifacts	evaluation-algorithm cost of development and maintenance	no	depends on algorithm used; fitting data to the algorithm	peer voting; better implementation of algorithm	issues due to subjectivity; peer-evaluation issues	relative evaluation; psychological; pay per performance

modular software development.

Individual Evaluation Methods

Quantitative evaluation. Quantitative evaluation represents the rating of individuals based on the measurable properties of their contribution. Some labor types are suitable for precisely measuring an agent's individual contributions, in which case the agent can be evaluated on number of units processed, but apart from the most primitive types of labor, evaluating an agent's performance requires evaluating different aspects of performance, or measurable signals, the most common being productivity, effort, and product quality. Different measures are usually taken into consideration with different weights, depending on their importance and measurement accuracy.

Quantitative evaluation is attractive because it does not require human participation and can be implemented entirely in software. Associated problems are measurement inaccuracy and difficulty choosing proper signals and weights. An additional problem is a phenomenon called multitasking, which, in spite of its counterintuitive name, refers to agents putting most

of their effort into tasks subject to incentives while neglecting other tasks, subsequently damaging overall performance.[10]

Subjective evaluation. When important aspects of human work are understandable and valuable to humans only, we need to substitute an objective measurement with a human (subjective) assessment of work

quality. In this case a human acts as a mapping function that quantifies human-oriented work by combining all undefinable signals into one subjective assessment signal. Even though subjective evaluation is implemented simply and cheaply, it is also inherently imprecise and prone to dysfunctional behavioral responses. Phenomena observed in practice[20] include:

Figure 2. Incentive strategies consist of smaller, easily modeled components.
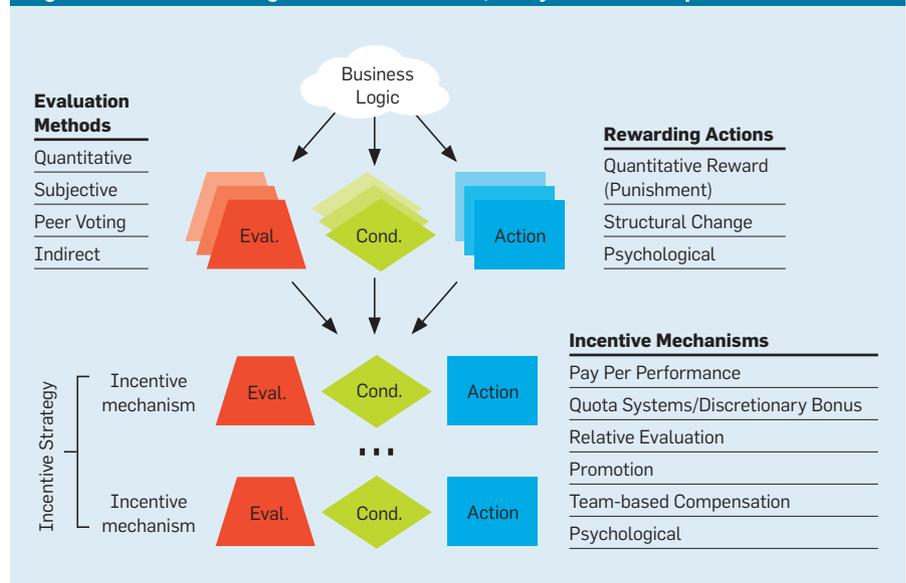
Centrality bias. Ratings concentrated around some average value, so not enough differentiating of "good" and "bad" workers;

Leniency bias. Discomfort rating "bad" workers with low marks; and

Rent-seeking activities. Actions taken by employees with the goal of increasing the chances of getting a better rating from a manager, often including personal favors or unethical behavior.

Group Evaluation Methods

Peer evaluation (peer voting). Peer evaluation is an expression of collective intelligence where members of a group evaluate the quality of other members. In the ideal case, the aggregated, subjective scores represent a fair, objective assessment. The method alleviates centrality and leniency bias since votes are better distributed, the aggregated scores cannot be subjectively influenced, and activities targeting a single voter's interests are eliminated. Engaging a large number of professional peers to evaluate different aspects of performance leaves fewer options for multitasking.

This method also suffers from a number of weaknesses; for example, in small interconnected groups voters may be unjust or lenient for personal reasons. They may also feel uncomfortable and exhibit dysfunctional behavior if the person being judged knows their identity. Therefore, anonymity is often a favorable quality. Another way of fighting dysfunctional behavior is to make voters subject to incentives; votes are compared, and those that stand out discarded. At the same time, each agent's voting history is monitored to prevent consistent unfair voting.

When the community consists of a relatively small group of voted persons and a considerably larger group of voters and both groups are stable over time, this method is particularly favorable. In such cases, voters have a good overview of much of the voted group. Since the relationship voter-to-voted is unidirectional and probably stable over time, voters do not have an interest in exhibiting dysfunctional behavior, a pattern common on the Internet today.

The method works as long as the size of the voted group is small. As the voted group increases, voters are unable to acquire all the new facts needed to pass fair judgment. They then opt to rate better those persons or artifacts they know or feel have good reputations (see Price[21]), a phenomenon known as "preferential attachment," or colloquially "the rich get richer." It can be seen on news sites that attract large numbers of user comments. Newly arriving readers usually tend to read and vote only the most popular comments, leaving many interesting comments unvoted.

In non-Internet-based businesses, cost is the major obstacle to applying this method, in terms of both time and money. Moreover, it is technically challenging, if not impossible, to apply it often enough and with appropriate voting groups. However, the use of information systems, the Internet, and social networks now makes possible a drastic decrease in application costs. A number of implementations exist on the Internet (such as Facebook's Like button, binary voting, star voting, and polls), but lacking is a unified model able to express their different flavors and specify the voters and voted groups.

Indirect evaluation. Since human performance is often difficult to define and measure, evaluating humans is commonly based on properties and relations among the artifacts they produce. As the artifacts are always produced for consumption by others, determining quality is ultimately left to the community. Artifacts are connected through various relations (such as contains, refers-to, and subclass-of) among themselves, as well as with users (such as author, owner, and consumer). The method of mapping properties and relations of artifacts to scores is nontrivial. An algorithm (such as Google's PageRank) tracks relations and past interactions of agents or their artifacts with the artifact being evaluated and calculates the score. A tailor-made algorithm must usually be developed or an existing one adapted to fit a particular environment. The major difference from peer evaluation is the agent does not actively evaluate the artifact, and hence the algorithm is not dependent on interacting with the agent.

The method's advantages and drawbacks fully depend on the properties of the applied algorithm. If the algorithm is suitable it exhibits fairness and prevents false results. The cost of the method depends in turn on the cost of developing, implementing, and running the algorithm. A common problem involves users who know how the algorithm works, then try to deceive it by outputting dummy artifacts with the sole purpose of increasing their scores. Detecting and preventing such attempts requires amending the algorithm, further increasing costs; Table 2 lists common application and composability considerations for these evaluation methods, as well as how drawbacks of a particular evaluation method can be alleviated by combining it with other methods.

Table 3. Incentive mechanisms used by social computing companies.

Incentive Strategy	No. of Companies	Percentage
Relative Evaluation	75	54%
Pay Per Performance	46	33%
Psychological	23	16%
Quota/Discretionary Bonus	12	9%
Deferred Compensation	10	7%
Promotion	9	6%
Team-based Compensation	3	2%

Table 4. Number of incentive mechanisms used by social computing companies; a majority of surveyed companies and organizations employ only one mechanism.

No. of Mechanisms	No. of Companies	Percentage
1	116	83%
2	15	11%
3	6	4%
≥4	3	2%

Rewarding Actions

Agents' future behavior can be influenced through rewarding actions:

Rewards. Rewards can be modeled as quantitative changes in parameters associated with an agent; for example, a parameter can be the wage amount, which can be incremented by a bonus or decreased by a penalty;

Structural changes. Structural changes are an empirically proven[31] motivator. A structural change does not strictly imply position advancement/downgrading in traditional tree-like management structures but does include belonging to different teams at different times or collaborating with different people; for example, working on a team with a distinguished individual can diversify an agent's experience and boost the agent's career. One way to model structural changes is through graph rewriting;[2] and

Psychological actions. Though all incentive actions have a psychological effect, psychological actions are only those in which an agent is influenced solely by information; for example, we may decide to show agents only the results of a couple of better-ranking agents rather than the full rankings. This way, the agents will not know their position in the rankings and can be beneficial in two ways: prevent the anchoring effect[17] for agents in the top portion of the rankings and prevent discouragement of agents in the lower portion. Psychological actions do not include explicit parameter or position change, but the diversity of presentation options means defining a unified model for describing different psychological actions is an open challenge. Effects of these actions are difficult to measure precisely, but apart from empirical evidence (such as Frey and Jegen[8]), their broad adoption on the Internet today is a clear indication of their effectiveness.

Incentive Conditions

Incentive conditions state precisely how, when, and where to apply rewarding actions, with each action consisting of at most three components, or subconditions:

Parameter. Expresses a subcondition in the form of a logical formula over a specified number of parameters describing an agent;
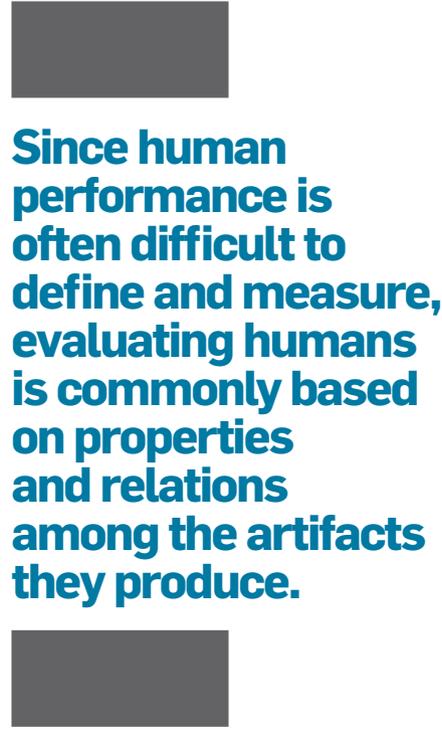
> Since human performance is often difficult to define and measure, evaluating humans is commonly based on properties and relations among the artifacts they produce.

Time. Helps formulate a condition over an agent's past behavior; and

Structure. Filters out agents based on their relationships and can be used to select members of a team or all collaborators of a particular agent.

Using these components at the same time helps make it possible to specify a complex condition (such as "target the subordinates of a specific manager, who over the past year achieved a score higher than 60% in at least 10 months").

Incentive conditions are part of the business logic, and as such are stipulated by HR managers. However, a small company can take advantage of good practices and employ pre-made incentive models (patterns) adapted to its needs. Feedback information obtained through monitoring execution of rewarding actions can help adapt condition parameters.

In Real-World Social Computing Platforms

In the first half of 2012 we surveyed more than 1,600 Internet companies and organizations worldwide that described themselves through keywords like "social computing" and "crowdsourcing," providing a solid overview of the overall domain of social computing. However, we were interested only in those employing incentive measures. Therefore, we manually investigated their reward/incentive practices (such as types of awards, evaluation methods, rules, and conditions) as stated on company websites, classifying them according to the previously described classifications. Overall, we identified and examined 140 companies and organizations using incentive measures.

Survey results. We found it striking that 59 of the 140 companies (42%) used a simple "contest" business model employing a relative evaluation incentive mechanism in which a creative task is deployed to the crowd. Each crowd member (or entity) then submits a design. The best design in the vast majority of cases is chosen through subjective evaluation (85%). That was expected, since the company buying the design reserves the right to decide the best design. In fact, in many cases, it was the only possible choice. When using peer evaluation, a com-

pany delegates the decision as to the best design to the crowd of peers while taking the risk of producing and selling the design. In some cases (such as a programming contest), the artifacts are evaluated quantitatively through automated testing procedures. Worth noting is that peer or quantitative evaluation produces quantifiable user ratings. In such cases, individuals are better motivated to take part in future contests, even if they feel they cannot win, because they can use their ranking as a personal quality proof when applying for other jobs or even as personal prestige.

Apart from the 59 organizations running contests, relative evaluation is used by another 16 organizations, usually combined with other mechanisms. This makes relative evaluation by far the most widely used incentive mechanism in social computing today (54% of those we surveyed) (see Table 3). This is in contrast with its use in traditional (non-Internet-based) businesses, where it is used considerably less,[1] as implementation costs are much greater.

The other significant group includes companies that pay agents for completing human micro-tasks. We surveyed 46 such companies (33%). Some are general platforms for submitting and managing any kind of human-doable tasks (such as Amazon's Mechanical Turk). Others offer specialized human

services, most commonly writing reviews, locating software bugs, translating, and performing simple, location-based tasks. What all these companies have in common is the PPP mechanism. Quantitative evaluation is the method of choice in most cases (65%) in this group.

Quantitative evaluation sometimes produces binary output (such as when submitting successful/unsuccessful steps to reproduce a bug). The binary output allows expressing two levels of the quality of work, so agents are rewarded on a per-task basis for each successful completion. In this case, the company usually requires no entry tests for joining the contributing crowd. In other cases, establishing work quality is not easy, and the output is proportional to the quantity of finer-grain units performed (such as word count in translation tasks), though agents are usually asked to complete entry tests; the pay rate for subsequent work is determined by the test results. Other evaluation methods include subjective and peer/indirect evaluation, both at 17%. Interesting to note is how rarely peer evaluation is employed for double-checking results, as companies find it cheaper to test contributors once, then trust their skills later on. However, as companies start to offer more complex human tasks, quality assurance becomes im-

perative, so we expect so see a rise in peer and indirect evaluation.

Only three companies combined four or five different mechanisms (see Table 4). The most well known is uTest.com; with a business model requiring a large crowd of dedicated professionals, it is clear why it employs more than just simple PPP.

ScalableWorkforce.com is the only company we studied that advertises the importance of crowd (work-force) management, offering tools for crowd management on Amazon's Mechanical Turk to its clients. The tools allow for tighter agent collaboration (fostering a sense of community among workers), workflow management, performance management, and elementary career building.

Of the 140 organizations we surveyed, 12 (8.5%) rely uniquely on psychological mechanisms to assemble and improve their agent communities. Their common trait is their reliance on the indirect influence of rankings in an agent's (non-virtual) professional life; for example, avvo.com attracts large numbers of lawyers in the U.S. who offer a free response and advice to people visiting the website. Quality and timeliness of professionals' responses affect their reputation rankings on avvo.com, which can be used as an advertisement to attract paying customers to their private practices. Another interesting example involves companies like crowdpark.de and prediculous.com that ask their users to "predict" the future by betting on upcoming events with virtual currency. Users with the best predictions over time earn virtual trophies (badges), the only incentive for participation. Crowdsourced odds are also useful for adjusting odds in conventional betting involving real money.

Table 5. Evaluation methods, excluding companies running creative contests.

Evaluation Method	No. of Companies	Percentage
Quantitative Evaluation	51	63%
Peer Voting + Indirect	35	43%
Subjective Evaluation	14	17%

Table 6. Companies using different evaluation methods (columns) within different incentive mechanisms (rows) as of early 2012; they may not be the primary mechanisms used by these companies.

	Quantitative	Subjective	Peer	Indirect
Pay Per Performance	mturk.com	content.de	crowdflower.com	translationcloud.net
Quota/Discretionary Bonus	gild.com		carnetdemode.fr	
Deferred Compensation	advisemejobs.com	bluepatent.com	crowdcast.com	
Relative Evaluation	netflixprize.com	designcrowd.com	threadless.com	topcoder.com
Promotion	utest.com	scalableworkforce.com	kibin.com	
Psychological Incentives	crowdpark.de	battleofconcepts.nl	avvo.com	
Team-based Compensation		mercmob.com	geniuscrowds.com	

Team-based compensation was used by only three companies we surveyed; for example, mercmob.com encourages formation of virtual human teams for various tasks. Agents express confidence in the successful completion of a task by investing part of a limited number of their "contracts," or a type of local digital currency. When invested, the contracts are tied to the task, motivating the agents who accept the task to do their best to self-organize as a team and attract others to join the effort. If in the end the task is completed successfully each agent gets a monetary reward proportional to the number of invested contracts.

Discretionary bonuses or quota systems are used by 11 companies (8%). However, they are always used in combination with another mechanism, usually PPP (64%), as in traditional companies.

Deferred compensation is used by 7% of the companies, usually as their only incentive mechanism; for example, Bluepatent.com crowdsources the tasks of locating prior art for potential patent submissions. The agents (researchers) are asked to find and submit relevant documents proving the existence of prior art. Deciding on the validity and usefulness of such documents is an intricate task, hence the decision on compensation is delayed until an expert committee decides on it. Advisemejobs.com pays out classical referral bonuses to agents suggesting appropriate job candidates.

Only 7% of our surveyed companies offer career advancement combined with other incentive mechanisms. As the crowd structure is usually plain, career advancement usually means higher status, implying a higher wage. We encountered only two cases where advancement also meant structural change, with an agent taking responsibility for leading or supervising lower-ranked agents.

In traditional companies deciding on a particular employee's promotion is usually a matter of subjective evaluation by the employee's superiors. With the promotion being the most commonly employed traditional incentive, the subjective evaluation is also the most commonly used evaluation method. However, if we remove the companies running creative contests, where the artistic nature of the artifacts forc-

es use of subjective evaluation, we see a reversal of the trend in social computing. Subjective evaluation trails quantitative and peer evaluation (see Table 5), as explained by the fact that information systems enable cheaper measurement of different inputs and setting up peer-voting mechanisms.
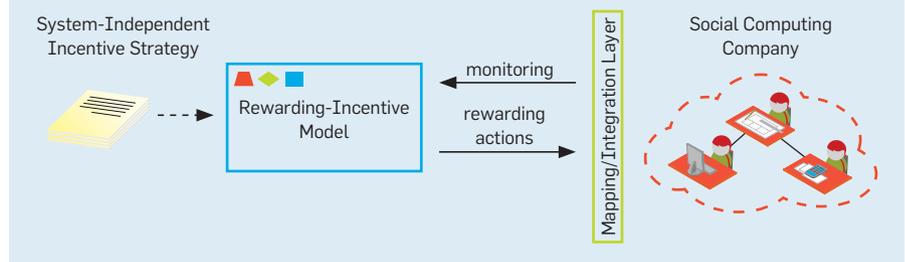
Only a small number of the companies and organizations we surveyed employ a combination of incentive mechanisms. Locationary.com uses agents around the world to expand and maintain a global business directory by adding local business information, employing two basic incentive mechanisms, aided by a number of supporting ones: The first is the so-called conditional PPP; with every new place added and/or corrected, agents win "lottery tickets" that increase the chances of winning a reward in a lottery, though a minimum number of tickets is needed to enter the draw. The second is team-based compensation. Locationary.com shares 50% of the revenues obtained from each directory entry with its agents. Any agent adding new information about a business obtains a number of shares of that directory entry. The reward is then split among agents proportionally to the number of shares they possess. Additionally, each entry in the directory must be approved through votes by trusted agents. Each agent has a trust score calculated by indirect evaluation that accounts for numerous factors like trust scores of voters, number of approved and rejected entries, and freshness of data. Trust influences the number of tickets awarded, thus affecting the odds of winning an award; the actual payout is limited to the agents with a certain level of trust.

Locationary.com uses a combination of PPP and a quota system to motivate overall agent activity. Team-based

compensation is used to incentivize adding high-end clients to the directory first. If an agent is first to add detailed information about, say, a hotel on the Mediterranean Sea, then, in addition to lottery tickets, that agent can expect appreciable income from the hotel's advertising revenue. Adding a local fast-food restaurant could bring the same number of lottery tickets but probably no advertising revenue. Peer voting serves to maintain data accuracy and quality, while indirect evaluation (expressed through trust) identifies and keeps high-quality contributors. In the end, we also see an example of deferred compensation, with money paid to contributors after some length of time but only if at the moment of payout they still have a satisfactory trust level. This example demonstrates how different mechanisms are used to target different agent behaviors and how to compose them to achieve their full effect; Table 6 outlines several companies employing different evaluation methods within a number of incentive mechanisms.

Conclusion

With creativity contests and micro-task platforms dominating the social computing landscape the organizational structure of agents is usually flat or very simple; hierarchies, teams of agents, and structured agent collaborations are rare. In such environments, most social computing companies need to use only one or two simple incentive mechanisms, as in Table 4. Promotion, commonly used in traditional companies, is rarely found in social computing companies. The reason is the short-lived nature of transactions between agents and the social computing companies. For the same reason, team-based compensation is also poorly represented. The

Figure 3. Conceptual scheme of a system able to translate portable incentive strategies into concrete rewarding actions for different social computing platforms.

System-Independent Incentive Strategy

Rewarding-Incentive Model

monitoring

rewarding actions

Mapping/Integration Layer

Social Computing Company

idea of building a "career in the cloud" is a distant dream. On the other hand, most traditional companies combine elaborate mechanisms to elicit particular responses from their agents and retain quality employees.[20] The mechanisms complement one another to cancel out individual drawbacks.

Our survey shows that as the cost of quantitative, peer, and indirect evaluation has decreased, relative evaluation and PPP have become the most popular incentive mechanisms among social computing companies. Subjective evaluation, though well represented, is found largely within companies that base their business models on organizing creativity contests and that psychological incentives and gamification approaches are gaining ground. We expect them to achieve their full potential as amplifiers for other incentive mechanisms.

The expected growth in complexity of business processes and organizational structures due to social computing will require novel automated ways of handling the behavior of agent crowds. That is why it is necessary to develop models of incentive mechanisms and frameworks fitting existing business models and real-world systems (such as workflow, human-provided services,[26] and crowdsourcing), while supporting composability and reusability of incentive mechanisms.

Such systems must be able to monitor crowds of agents and perform runtime adaptation of incentive mechanisms to prevent diverse negative effects (such as free riding, multitasking, biasing, anchoring, and preferential attachment), switching when needed between different evaluation methods, rewarding actions, and incentive conditions at runtime, while minimizing overall costs. This way, particular agent sub-groups can be targeted more efficiently.

Additional benefits include:

Optimization. Historical data can help detect performance bottlenecks, preferred team compositions, and optimal wages. Additionally, we can choose the optimal composition of incentive mechanisms, opening up novel possibilities for achieving indirect automated team adaptability through incentives;

Reusability. For certain business models, proven incentive patterns cut time and cost, with incentive patterns tweaked to fit particular needs based on feedback obtained through monitoring;

Portability. By generalizing and formally modeling incentive mechanisms, we can encode them in a system-independent manner; that way, they become usable on different underlying systems, without having to write more system-specific programming code (see Figure 3); and

Incentives as a service. Managing rewards and incentives can be offered remotely as a Web service.

We are developing an incentive framework supporting these functionalities[24,25] evaluated on social-compute-unit systems[27] for maintaining large-scale distributed cloud-software systems.

Acknowledgment

References
1. Armstrong, M. *Armstrong's Handbook of Reward Management Practice: Improving Performance Through Reward.* Kogan Page Publishers, London, 2010.
2. Baresi, L. and Heckel, R. Tutorial introduction to graph transformation: A software engineering perspective. In *Proceedings of the First International Conference on Graph Transformation* (Barcelona). Springer-Verlag, London, 2002, 402–429.
3. Davis, J.G. From crowdsourcing to crowdservicing. *IEEE Internet Computing 15*, 3 (2011), 92–94.
4. Deterding, S., Sicart, M., Nacke, L., O'Hara, K., and Dixon, D. Gamification: Using game design elements in non-gaming contexts. In *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems* (Vancouver, B.C.). ACM Press, New York, 2011, 2425–2428.
5. Doan, A., Ramakrishnan, R., and Halevy, A.Y. Crowdsourcing systems on the World-Wide Web. *Commun. ACM 54*, 4 (Apr. 2011), 86–96.
6. Dustdar, S., Schall, D., Skopik, F., Juszczyk, L., and Psaier, H. (Eds.) *Socially Enhanced Services Computing: Modern Models and Algorithms for Distributed Systems.* Springer, 2011.
7. Dustdar, S. and Truong, H. L. Virtualizing software and humans for elastic processes in multiple clouds: A service-management perspective. *International Journal of Next-Generation Computing 3*, 2 (2012).
8. Frey, B.S. and Jegen, R. Motivation crowding theory. *Journal of Economic Surveys 15*, 5 (2001), 589–611.
9. Hoisl, B., Aigner, W., and Miksch, S. Social rewarding in wiki systems: Motivating the community. In *Proceedings of the Second International Conference on Online Communities and Social Computing* (Beijing). Springer, 2007, 362–371.
10. Holmstrom, B. and Milgrom, P. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization 7* (1991), 24–52.
11. Horton, J.J. and Chilton, L.B. The labor economics of paid crowdsourcing. In *Proceedings of the 11th ACM Conference on Electronic Commerce* (Cambridge, MA). ACM Press, New York, 2010, 209–218.
12. Kerrin, M. and Oliver, N. Collective and individual improvement activities: The role of reward systems. *Personnel Review 31*, 3 (2002), 320–337.
13. Laffont, J.-J. and Martimort, D. *The Theory of Incentives.* Princeton University Press, Princeton, NJ, 2002.
14. Lazear, E.P. Personnel economics: The economist's view of human resources. *Journal of Economic Perspectives 21*, 4 (2007), 91–114.
15. Lazear, E.P. Performance pay and productivity. *The American Economic Review 90*, 5 (2000), 1346–1361.
16. Little, G., Chilton, L.B., Goldman, M., and Miller, R.C. TurKit: Tools for iterative tasks on Mechanical Turk. In *Proceedings of the ACM SIGKDD Workshop on Human Computation* (Paris). ACM Press, New York, 2009, 29–30.
17. Mason, W. and Watts, D.J. Financial incentives and the performance of crowds. In *Proceedings of the ACM SIGKDD Workshop on Human Computation* (Paris). ACM Press, New York, 2009, 77–85.
18. Parameswaran, M. and Whinston, A.B. Social computing: An overview. *Communications of the Association for Information Systems 19* (2007), 762–780.
19. Pearsall, M.J., Christian, M.S., and Ellis, A.P.J. Motivating interdependent teams: Individual rewards, shared rewards, or something in between? *Journal of Applied Psychology 95*, 1 (2010), 183–191.
20. Prendergast, C. The provision of incentives in firms. *Journal of Economic Literature 37*, 1 (1999), 7–63.
21. Price, D.D.S. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science 27*, 5 (1976), 292–306.
22. Resnick, P., Kuwabara, K., Zeckhauser, R., and Friedman, E. Reputation systems: Facilitating trust in Internet interactions. *Commun. ACM 43*, 12 (Dec. 2000), 45–48.
23. Sato, K., Hashimoto, R., Yoshino, M., Shinkuma, R., and Takahashi, T. Incentive mechanism considering variety of user cost in P2P content sharing. In *Proceedings of the IEEE Global Telecommunications Conference* (New Orleans). IEEE, 2008, 1–5.
24. Scekic, O., Truong H.L., and Dustdar, S. Modeling rewards and incentive mechanisms for social BPM. In *Proceedings of the 10th International Conference on Business Process Management* (Tallinn, Estonia, Sept.). Springer, 2012, 150–155.
25. Scekic, O., Truong, H.L., and Dustdar, S. Programming incentives in information systems. In *Proceedings of the 25th International Conference on Advanced Information Systems Engineering* (Valencia, Spain, June 2013).
26. Schall, D., Dustdar, S., and Blake, M.B. Programming human and software-based Web services. *IEEE Computer 43*, 7 (2010), 82–85.
27. Sengupta, B., Jain, A., Bhattacharya, K., Truong, H. L., and Dustdar, S. Who do you call? Problem resolution through social compute units. In *Proceedings of the 10th International Conference on Service-oriented Computing* (Shanghai). Springer, Berlin, Heidelberg, 2012, 48–62.
28. Shaw, A.D., Horton, J.J., and Chen, D.L. Designing incentives for inexpert human raters. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work* (Hangzhou, China). ACM Press, New York, 2011, 275–284.
29. Tokarchuk, O., Cuel, R., and Zamarian, M. Analyzing crowd labor and designing incentives for humans in the loop. *IEEE Internet Computing 16*, 5 (2012), 45–51.
30. Tran, A. and Zeckhauser, R. *Rank As An Incentive.* HKS Faculty Research Working Paper Series RWP09-019, John F. Kennedy School of Government, Harvard University, Cambridge, MA, 2009.
31. Van Herpen, M., Cools, K., and Van Praag, M. Wage structure and the incentive effects of promotions. *Kyklos 59*, 3 (2006), 441–459.
32. Yogo, K., Shinkuma, R., Konishi, T., Itaya, S., Doi, S., Yamada, K., and Takahashi, T. Incentive-rewarding mechanism to stimulate activities in social networking services. *International Journal of Network Management 22*, 1 (2012), 1–11.

Ognjen Scekic (oscekic@dsg.tuwien.ac.at) is a Ph.D. student in the Vienna PhD School of Informatics and the Distributed Systems Group at the Vienna University of Technology, Vienna, Austria.

Hong-Linh Truong (truong@dsg.tuwien.ac.at) is a senior researcher in the Distributed Systems Group at the Vienna University of Technology, Vienna, Austria.

Schahram Dustdar (dustdar@dsg.tuwien.ac.at) is a full professor and director of the Distributed Systems Group at the Vienna University of Technology, Vienna, Austria.