ECAI 2014
T. Schaub et al. (Eds.)

doi:10.3233/978-1-61499-419-0-591

Information-based Incentivisation when Rewards are Inadequate

Samhar Mahmoud and **Lina Barakat** and **Simon Miles** and **Adel Taweel** and **Brendan Delaney** and **Michael Luck**[1]

Abstract. In many cases, intermediaries play a major role in linking between service providers and their target users. Yet, attracting intermediaries at a marketplace to promote a service to their existing customers can be very challenging, since they are usually very busy and would incur additional cost as a result of such promotion. In response, this paper presents an information-based incentivisation framework, which combines financial rewards with other motivating information, in order to incentivise intermediaries at a marketplace to undertake service promotion. Specifically, the intermediaries are associated with a group of incentivising agents, capable of learning the individual motivational needs of these intermediaries, and accordingly target them with the most effective incentives. The incentivising agents collaborate with each other to gather motivational information, by sharing their observations on intermediaries. The proposed incentivisation approach is evaluated through a corresponding agent-based simulation, and the experimental results obtained demonstrate its effectiveness.

1 INTRODUCTION

In order to be viable, service providers must face the well-known problem of attracting sufficient customers to purchase their products or use their services. Competition can be very high, and choosing where and how to advertise and sell a service can be crucial. In many cases, this may best be achieved through various intermediaries (third parties) in different marketplaces with a ready supply of potential target users. For example, in online advertising, providers could market their products to a large user population via placing advertisements on popular web sites (e.g. yahoo). However, intermediaries within such marketplaces are often heavily engaged in promoting existing services, and it can be very challenging to attract them to devote adequate resource for promoting another service.

A promising way to tackle this challenge is via applying suitable *incentivisation* mechanisms to encourage an intermediary to promote a service. In fact, various economic and psychological studies emphasis the importance of incentives in directing an entity towards a desired behaviour [1, 3]. Such studies reveal that there is typically a mixture of motives driving an entity to undertake a particular task, which could be intrinsic (e.g. out of personal interest in the task) or extrinsic (e.g. to gain financial reward or social approval), and may differ among entities. Targeting such motives with relevant incentives would thus allow pushing the entity's behaviour in the desired direction (e.g. towards service promotion).

The applicability of such incentivisation in computational systems has been investigated by a number of researchers. In particular, most work here has focused on studying the effect of punishment (in the form of fines) as a monetary incentivisation to establish a desired behaviour among a population of self-interested agents in a multi-agent system [6, 8, 11, 13]. However, monetary incentivisation alone could be very costly, not sufficient, or even not effective in certain cases (e.g. when it negatively affects the intrinsic motivation of an entity [3, 4]).

Few researchers have also considered other forms of incentivisation to promote cooperation among agents, including violation alerts [15], deferred reciprocity [12], and reputation [5]. Yet, again these efforts are restricted to specific motivational factors, without the ability to explore other, potentially relevant ones.

In response, this paper proposes a generic computational incentivisation framework, capable of supporting monetary rewards with any number of other relevant incentivisation mechanisms, offered to an entity in the form of informational incentives. This framework allows learning and reasoning about the effect of various incentives on an entity's behaviour (which is not known in advance and may change over time), and reflecting such reasoning on the design of more effective (personalised) future incentives for the entity.

The framework is presented in the context of the problem of motivating an intermediary to promote a service, and a number of specific motivational factors relevant in this regard are provided. Yet, it should be noted that the framework is not restricted to these factors, and can be easily extended to include others. Moreover, although the framework makes no assumptions of any particular behavioural model of the intermediary, an example of such a model is presented and utilised for evaluation purposes.

The paper is organised as follows. Section 2 introduces a motivating scenario, followed by the proposed information-based incentivisation framework in Section 3. A simulation model enabling the framework evaluation and the corresponding experimental results are presented in Section 4 and Section 5, respectively. Finally, Section 6 concludes the paper.

2 MOTIVATIONAL SCENARIO

Consider a scenario from the health domain, where medical researchers aim to recruit sufficient numbers of patients for clinical trials. Such trials are the gold standard by which medical research is evaluated. They are used to study various aspects of medical science, as well as being a vital stage in the deployment of new drug treatments. Currently, however, such trials are frequently unsuccessful at recruiting sufficient patients. Indeed, a review by the UK's Medical Research Council found that only 31% of trials actually recruited to their planned target, with 30–40% of costs arising during the recruit-

[1] King's College London, London, UK, email:firstname.surename@kcl.ac.uk

ment phase alone [10]. In this context, the main challenge for patient recruitment lies in locating and contacting patients in a sufficiently timely manner to allow them to participate in a trial. As a result, the best way of doing so is via the clinic of their general practitioner (GP), who has regular contact with patients seeking treatment. The recruitment process thus often involves a human recruitment agent visiting clinics in an attempt to locate suitable patients (e.g. asking GPs or searching local medical records), yet this creates significant overhead as it is both slow and costly.

In response, several researchers [14, 9] have proposed replacing human recruiters with software agents that reside at local clinics. In this view, each agent would maintain a local repository of information about active clinical trials. Whenever a patient enters a clinic for a consultation, the agent would inspect this patient record in real-time to ascertain if they are eligible for any trials, with instant notifications being presented to a GP during consultation to inform them of the patient's eligibility. This would allow patients to immediately be recruited, negating the need for laborious effort on the part of the patient, clinical researcher or human recruiter.

However, both of the above solutions ignore the problem that some GPs may not be sufficiently motivated to recruit patients. This may be related to the limited time available to them, the lack of interest in (or information about) the trial, or the fact the patients might be resistant to signing up. Consequently, some means of overcoming these factors, and ensuring that GPs are more willing to participate in promoting clinical trials to patients, is needed. Note that, in the scenario above, the medical researcher, the clinical trial, the GP, and the patient, represent the service provider, the service, the intermediary, and the target user, respectively.

3 INFORMATION BASED INCENTIVISATION

Clearly, both monetary reward (the commission received by the intermediary per each attracted target) and cost (the resources consumed by the intermediary to attract a target) have a major impact on the intermediary's decision regarding service promotion. For example, in the scenario of Section 2, such monetary reward and cost correspond to the payment received by the GP per patient recruitment, and the time spent by the GP talking to a patient about the trial, respectively. Yet, other factors, such as service popularity, peer pressure, and intermediary reputation, could also have an important influence on the intermediary's promoting behaviour, by modifying expectations about potential monetary gains (rather than modifying monetary gains themselves), or by triggering desires for altruistic behaviour towards particular direction, especially when compared to such behaviour by others. Unlike monetary reward and cost, which are usually easily accessible to the intermediary (from the agreement with the provider and personal performance), information on such other factors is much more challenging for the intermediary to obtain (especially given its dynamic and distributed nature), and hence could be overlooked or wrongly estimated. Consequently, a strategic service provider should not rely on the intermediary's ability to acquire such motivational information, but rather seek to communicate this information appropriately to the intermediary (in the form of incentives), to ensure its consideration in the intermediary's decision making. Moreover, since different intermediaries may exhibit different sensitivity towards various factors, targeting them with information should be tailored to meet their *individual* motivational needs.

To achieve this, we propose associating each intermediary with a *personalised incentivising agent*, capable of reasoning about the

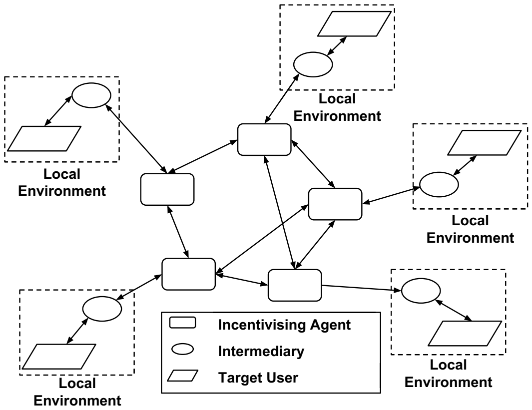


Figure 1: System Architecture

intermediary's service promoting behaviour, learning the intermediary's preferences, and correspondingly targeting the intermediary with the most promising incentivisation information. The architecture enabling such incentivisation capabilities for a single provider is depicted in Figure 1, showing the interactions between the different roles involved: the incentivising agent delegated by the provider, the intermediary (which could be a human or another agent), and the target user.

Since the main focus of this paper is the interaction between the incentivising agent and the intermediary to encourage the latter to promote the service, the target user role is ignored in the rest of the paper. Also, we make the following two assumptions. First, each incentivising agent can interact with (and observe the promoting behaviour of) the corresponding intermediary. Second, incentivising agents are willing to share true observations of their interactions with intermediaries, and these are accessible to all such agents. This is reasonable to assume since all incentivising agents represent the interest of the same provider, and can be established by having all the agents reporting their results to a central agency (the service provider). Note that this ability to benefit from collected aggregated data on a service, allows the incentivising agent to obtain more accurate service-related information (e.g. service popularity), which is not accessible to the intermediary.

Next, we elaborate in more detail on the potential informational factors for intermediary incentivisation, followed by the proposed incentivisation model for each incentivising agent.

3.1 Informational factors

In addition to monetary reward (MonReward), and promoting cost (Cost), other factors may also affect an intermediary's behaviour towards promoting a service, including: service popularity (SerPop), peer pressure (PeerPress), and intermediary reputation (Rep).

Service popularity refers to the acceptance rate of the service amongst target users, e.g. clinical trial popularity is the probability that a patient would agree to be recruited in this trial if advertised. Intuitively, the higher the popularity of a service, the higher the chance of receiving the monetary reward per each promoting attempt, and consequently the more encouraged an intermediary would be to promote this service.

Peer pressure, in this context, refers to the promoting behaviour of an intermediary's relevant peers (e.g. direct competitors), which may naturally affect this intermediary's own behaviour (e.g. driven by competition). For example, in the clinical trial recruitment scenario, it is very likely for a GP to be influenced by those GPs directly known to them or sharing common interests. In the proposed system, such peer relations among intermediaries are reflected via a *topological structure* connecting their respective incentivising agents, where each connection represents a relation between an intermediary and one of its peers (e.g. direct competitors).

Finally, *intermediary reputation* refers to how active a particular intermediary is in promoting a particular service. Clearly, the desire to maintain a high reputation may positively influence the intermediary's inclination to promote the service.

3.2 Incentivisation model

The incentivisation model of a particular incentivising agent (assigned to particular intermediary) with respect to a particular service can be represented as a tuple, $(F, vl^c, utl, dcsn, msg, obs)$, detailed below.

F is the set of candidate informational factors that could affect the tendency of the intermediary to promote the service. That is, $F = \{\text{SerPop}, \text{PeerPress}, \text{Rep}\}$ (note, however, that the model can be easily extended to include other factors).

$vl^c : F \times T \rightarrow \mathbb{R}^+$ is a factor value function, which assigns to each factor $f \in F$, its corresponding value at time step $t \in T$. Factor values could differ over time and per intermediary, and are estimated by the incentivising agent based on observing the intermediary's promoting behaviour and the information shared by other incentivising agents of the same provider, as follows:

$$vl^c(\text{SerPop}, t) = \frac{\text{population-level } \textit{successful} \text{ promotions at time } t}{\text{population-level promotions at time } t}$$

$$vl^c(\text{PeerPress}, t) = \frac{\text{peers promoted the service at time } t}{\text{peers invited to promote the service at time } t}$$

$$vl^c(\text{Rep}, t) = \frac{\text{intermediary-level promotions up to time } t}{\text{intermediary-level promotion opportunities up to time } t}$$

$utl : F \times T \rightarrow \mathbb{R}^+$ is a factor utility function, which assigns to each factor $f \in F$, its corresponding utility at time step t, with respect to incentivising the intermediary towards promoting the service (i.e. a higher utility indicates a more positive influence of the factor on the intermediary's promoting behaviour).

$dcsn : T \rightarrow 2^F$ is a factor selection function, which determines the factors to be provided as incentives to the intermediary at time step t, such that $|dcsn(t)| \leq limit$. Here, $limit$ is the maximum number of incentives allowed at each time step. Such a limit may exist due to many reasons, including communication cost and intermediary constraints (e.g. a GP may simply refuse to be overloaded with incentivisation information). Intuitively, the factor selection decision should be guided by the factor utilities, i.e. $dcsn(t) = fun^d(\{utl(f, t)\}_{f \in F})$.

$msg : T \rightarrow 2^{F \times \mathbb{R}^+}$ is the incentivisation message function, which returns the message to be sent to the intermediary at time step t. That is, $msg(t) = \{(f, vl^c(f, t) \mid f \in dcsn(t)\}$.

Finally, $obs : T \rightarrow \{\text{'y'}, \text{'n'}\}$ is the intermediary behaviour observation function, which records the decision made by the intermediary following receipt of message $msg(t)$. Here, $obs(t) = \text{'y'}$ indicates that the service was promoted, while $obs(t) = \text{'n'}$ indicates that the service was not promoted, by the intermediary (note that t refers to the time of sending the incentivisation message, regardless of the time of decision observation, which may occur later).

The history of past interactions with the intermediary is utilised by the incentivising agent to learn the former's behavioural model, and to update the utilities of the various informational factors, so that more strategic incentivisation decisions, tailored towards the specific needs of the intermediary, are made in the future. In other words, $utl(f, t) = fun^u(f, OBS_t)$, where $OBS_t = \{(msg(i), obs(i))\}_{i=1}^{t-1}$ are the prior interactions with the intermediary up to time t.

Concrete examples of functions fun^u and fun^d are presented in Section 4.

4 SIMULATION MODEL

To study the effectiveness of the proposed information-based incentivisation, we conducted an agent-based simulation incorporating a set of intermediary agents (referred to as $AGENT^{int}$), connected according to a particular interaction topology, and their corresponding incentivising agents (referred to as $AGENT^{inc}$). The simulation proceeds on the basis of rounds (or time steps). In each time step, four different phases can be distinguished: promoting; incentivising agent reconsideration; incentivisation; and intermediary reconsideration. These phases are presented in detail next.

4.1 Promoting phase

In the promoting phase, each intermediary agent, $ag \in AGENT^{int}$, is asked to promote the service. An intermediary's decision regarding whether to do so in the current round, i.e. the current time step $t \in T$, is determined according to its *promoting probability*, $p^{prom}(t)$, which represents the intermediary's tendency to promote the service (higher values correspond to higher likeliness to undertake the promotion). Having agreed to promote the service, the intermediary agent then receives a reward for doing so, and calculates the time required to do so. The decision made by each intermediary is then observed and recorded by its corresponding incentivising agent. Note that, initially, the promoting probabilities for all intermediaries are set to random values, which are updated as the simulation progresses (see Section 4.4).

4.2 Incentivising agent reconsideration phase

In the incentivising agent reconsideration phase, each incentivising agent, $ag \in AGENT^{inc}$, adjusts its respective model in order to reflect the information obtained from the promoting phase. Specifically, based on the observed effect of the previous incentivisation message (i.e. the message sent to the intermediary in the previous round) on the intermediary's behaviour, the utility of each informational factor $f \in F$, is updated according to Equation 1 (which corresponds to function fun^u of Section 3.2), where δ^u represents the learning step.

$$utl(f,t) = \begin{cases} utl(f,t-1) + \delta^u & \text{if } (f \in dcsn(t-1)) \wedge \\ & \quad (obs(t-1) = \text{'y'}) \\ max(utl(f,t-1) - \delta^u, 0) & \text{if } (f \in dcsn(t-1)) \wedge \\ & \quad (obs(t-1) = \text{'n'}) \\ utl(f,t-1) & \text{if } (f \notin dcsn(t-1)) \end{cases} \tag{1}$$

Such utility modification allows the incentivising agent to gradually *learn* the sensitivity of the intermediary towards various informational factors, thus consequently making more effective incentivisation decisions, tailored to the intermediary's specific motivational needs. Note that this learning is excluded from the first round t_0, with all informational factors being initially assigned equal utilities: $\forall f \in F,\ utl(f, t_0) = \mathrm{u}_0$ (where u_0 is a randomly generated value).

4.3 Incentivisation phase

In the incentivisation phase, each incentivising agent, $ag \in AGENT^{inc}$, has the opportunity to motivate its corresponding intermediary to promote the service. For the sake of this simulation, we assume that only one incentivising factor can be sent to the intermediary per round, i.e. $limit = 1$. The decision of which informational factor to select as the incentive in the current round $t \in T$, is based upon a probability distribution P over the factors, $P(f \in dcsn(t)) = p^{sel}(f,t)$, which is determined with respect to their utilities (this distribution facilitates the simulation of function fun^d of Section 3.2). That is, factors f with higher utilities $utl(f,t)$ are assigned higher selection probabilities $p^{sel}(f,t)$, such that $\sum_{f \in F} p^{sel}(f,t) = 1$. A simple example of such a probability function (the one adopted in this simulation) is: $p^{sel}(f_i,t) = \dfrac{utl(f_i,t)}{\sum_{f \in F} utl(f,t)}$.

Note that, the incentivising agent is allowed an *exploration* period $\{t_0, ..., t_{explr}\}$ at the beginning of the simulation, during which all informational factors are always considered equally probable, regardless of their utilities, i.e. $\forall t \in \{t_0, ..., t_{explr}\},\ P(f \in dcsn(t))$ corresponds to a uniform probability distribution.

4.4 Intermediary reconsideration phase

In the intermediary reconsideration phase, each intermediary agent, $ag \in AGENT^{int}$, reconsiders its promoting probability. For this purpose, we assume the following intermediary behavioural model (note, however, that the proposed incentivisation model is independent of any particular behavioural model of the intermediary). Based on the outcome of the intermediary's previous promoting decision (in terms of the money gained and the cost incurred), and the incentivisation message received from the incentivising agent in the current round t, the intermediary re-assesses its current *motivation degree* regarding promoting the service, $mtvdegree(t)$, and correspondingly adjusts its promoting probability for the next round, $p^{prom}(t+1)$, reflecting the cumulative motivation so far, as shown in Equation 2.

$$p^{prom}(t+1) = \begin{cases} min(p^{prom}(t) + \delta^p(t), 1) & \text{if } mtvdegree(t) > 0 \\ max(p^{prom}(t) - \delta^p(t), 0) & \text{if } mtvdegree(t) < 0 \\ p^{prom}(t) & \text{if } mtvdegree(t) = 0 \end{cases} \tag{2}$$

where $\delta^p(t)$ is the learning rate in the current round t, which could be either fixed, or adapted (set proportional) to the current motivation degree $mtvdegree(t)$.

The motivation degree, $mtvdegree(t)$, is estimated according to the domination between the motivating and demotivating aspects, as follows: $mtvdegree(t) =$

$$\sum_{f \in F \cup \{\text{MonReward}\}} (w(f) \times vl^m(f,t)) - vl^m(\text{Cost}, t) \tag{3}$$

where function $w(f) \in [0,1]$ returns the relative importance of factor f for the intermediary, such that $\sum_{f \in F \cup \{\text{MonReward}\}} w(f) = 1$ (these weights are assigned to each intermediary on a random basis at the beginning of the simulation); and function $vl^m(f,t) \in [0,1]$ returns the value currently held by the intermediary for factor f, scaled with respect to the factor's minimum and maximum possible values. Specifically, the values of MonReward and Cost correspond to the outcome of the intermediary's decision in the current round's promoting phase, while those of the informational factors $f \in F$ are either acquired from the current round's incentivisation message $msg(t)$ (if $f \in dcsn(t)$), or set to 0 (if $f \notin dcsn(t)$).

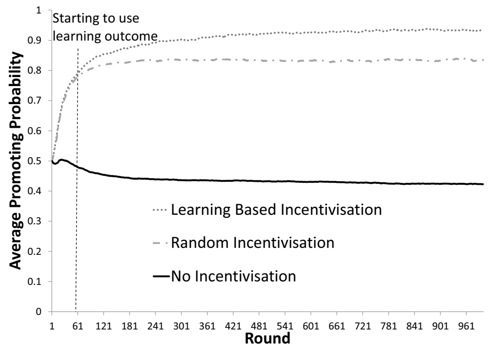


Figure 2: Comparison between incentivisation strategies - population level results

5 EXPERIMENTAL RESULTS

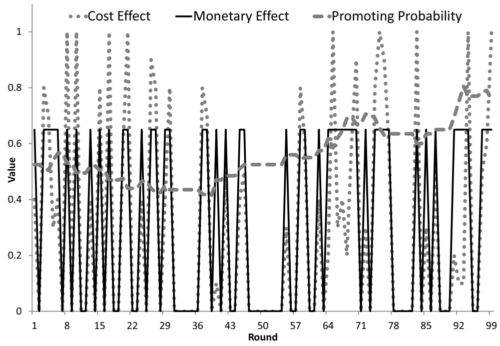
In order to evaluate our approach, we ran several experiments comprising three sets. As a baseline, the first set of experiments involves *no incentivisation*, where incentivising agents provide no incentivisation messages, i.e. $\forall t,\ dcsn(t) = \emptyset$ for all agents. The second set of experiments utilises a *random incentivisation strategy*, where incentivising agents produce incentivisation messages on a random basis, i.e. the *exploration* period t_{explr} covers all the runs. Finally, the last set of experiments utilises the *learning-based incentivisation strategy*, where incentivising agents adopt the utilities learned according to Equation 2 to select the most suitable incentivisation messages (see Sections 4.2 and 4.3).

The results of each experiment set are averaged over 100 simulation runs. Each run involves 1000 intermediaries and corresponding incentivising agents, with a duration of 100,000 rounds, a sufficiently large number to avoid the development of misleading results that could be affected by the length of runs [2, 7]. The exploration period in the learning-based incentivisation strategy lasts until round 60, after which incentivising agents start using the learned outcome. The MonReward is fixed among all intermediaries, while the Cost is generated randomly for each intermediary at each round.

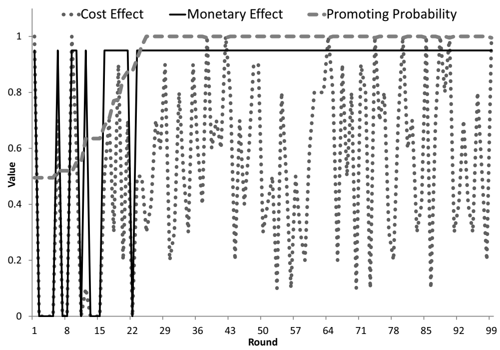
For the interaction topology among intermediaries, we adopt an example suited to the kinds of system we want to model. While lattice topologies are regular structures, as opposed to random topologies, Watts and Strogatz [16] noted that many biological, technological and social networks lie somewhere between the two: neither completely regular nor completely random. They instead proposed *small world networks*, which we adopt in our experiments, as a variation of regular lattices in which agents are connected to a number of hops (on the ring) away, but with some of the connections replaced by connections to other randomly selected nodes in the network, in line with some specific rewiring probability (set to 0.2 in our experiments). This topology fits well with our motivational scenario, where we assume that a GP is more concerned with competition from other GPs in close geographical proximity, and a few that are known to them through other means.

5.1 Population-level results

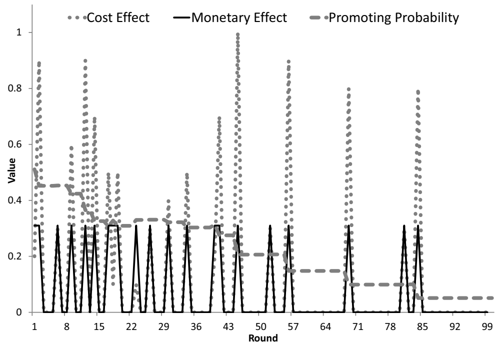
Here, we study the changes in the promoting behaviour of intermediaries over time, from the overall population perspective. In this regard, Figure 2 shows the evolution of the population promoting probability (averaged among all intermediaries), as a result of applying the three different strategies outlined above.



(a) Balanced Weights



(b) High MonReward Weights



(c) Low MonReward Weights

Figure 3: No incentivisation - agent level results

As can be concluded from the figure, runs without incentivisation achieve a midrange average promoting probability p^{prom}, which can be explained by the uniform distribution of preferences towards various motivational factors (in terms of weights $factorweight$) between the intermediaries. In particular, a good proportion of these agents have high MonReward weights, making the MonReward alone sufficient for them to overcome the demotivation incurred by the Cost, and consequently increasing their promoting probabilities p^{prom}. Yet, a larger proportion are still significantly affected by the Cost (due to the absence of other informational incentives), and thus decide to reduce their probabilities p^{prom}, causing the average promoting probability to drop to a low midrange level.

When utilising the random incentivisation strategy, the results indicate that the average promoting probability p^{prom} is pushed to a relatively high level. This is because this strategy manages to support intermediaries with additional motivational factors, occasionally succeeding to target them with the most relevant ones. Finally, as expected, learning-based incentivisation achieves the best performance among all strategies, with an almost immediate effect on the intermediary's promoting behaviour (the average promoting probability p^{prom} starts to increase to a higher level after round 60). More detail regarding the results is provided next.

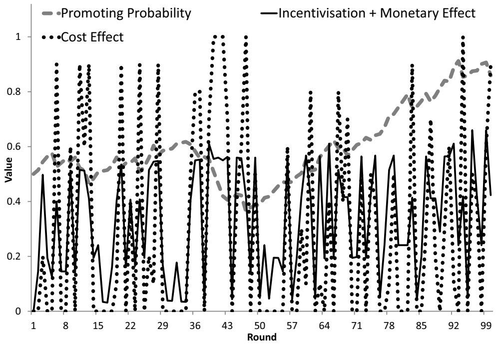
5.2 Agent-level results

In order to provide further explanation and analysis of the results above, we present here a more detailed view on these results with regard to individual agents. Specifically, we study the behavioural evolution of an individual intermediary agent, showing the changes in the promoting probability p^{prom} of this agent over time, along with the motivating and demotivating aspects affecting such probability, i.e. $\sum_{f\in F\cup\{\text{MonReward}\}}(w(f)\times vl^m(f,t))$ and $vl^m(\text{Cost},t)$, respectively (see Equation 3).

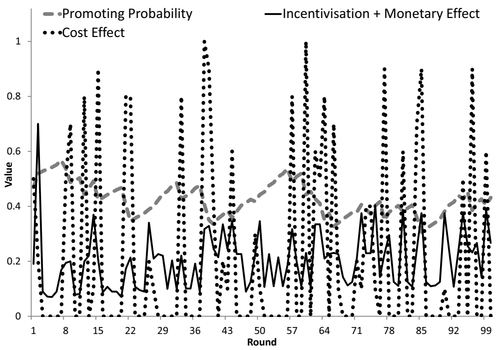
In runs that have no incentivisation, MonReward and Cost are the only contributing factors to p^{prom}, since the intermediary here has no access to other information, i.e. $f\in F$, $vl^m(f,t)=0$. Figure 3 shows the balance between the effect of these two factors (i.e. $w(\text{MonReward})\times vl^m(\text{MonReward},t)$ against $vl^m(\text{Cost},t)$), and p^{prom}, for three types of intermediary agents: an agent with balanced weights for all the positive factors (Figure 3(a)), i.e. $w(\text{MonReward})\approx\frac{1}{4}$; an agent with high MonReward weight (Figure 3(b)), i.e. $w(\text{MonReward})\approx 1$; and an agent with low MonReward weight (Figure 3(c)), i.e. $w(\text{MonReward})\approx 0$. As expected, intermediary agents that favour MonReward (those not much influenced by other informational factors) increase their promoting probabilities to a high level, since the monetary reward is very likely to be sufficient to compensate for the cost incurred. In contrast, agents that do not favour money decrease their promoting probabilities to a very low level, since the reward has little impact on their decisions. Finally, agents with balanced weights maintain midrange promoting probabilities since the MonReward occasionally outweighs the Cost (when the cost is not very high).

In runs that utilise the random incentivisation strategy, in addition to MonReward and Cost, other informational factors are also considered in the intermediary's decision making. Figure 4 shows the results in two cases: an intermediary with balanced weights for all positive factors $F\cup\{\text{MonReward}\}$ (Figure 4(a)), and one with unbalanced weights for these factors (Figure 4(b)). As can be seen from the figure, in the balanced weights case, random incentivisation works well in terms of motivating the intermediary towards promoting the service more often. This is because, the intermediary here has similar interest in all informational factors, and thus incentivisation with any such factor guarantees a positive influence. The same does not hold, however, for the non-balanced weights case, where the random incentivisation strategy achieves much lower performance due to the possibility of targeting intermediaries with irrelevant (uninfluencing) incentives.

This limitation is handled when using the learning-based incentivisation strategy (Figure 5), which is capable of pushing intermediaries of such type (i.e. with biased weights) to increase their promoting probabilities to a high level, by providing them with the most affecting incentivisation information. Note that, for reasons of clarity, only 50 time steps (rounds) are shown in Figure 5, starting from round 251, in order to show the effect after learning has taken place. Also note that the learning-based incentivisation strategy similarly achieves high performance for all the other weight settings (i.e. high MonReward weight, low MonReward weight, and balanced



(a) An agent with balanced weights



(b) An agent with unbalanced weights

Figure 4: Random incentivisation - agent level results

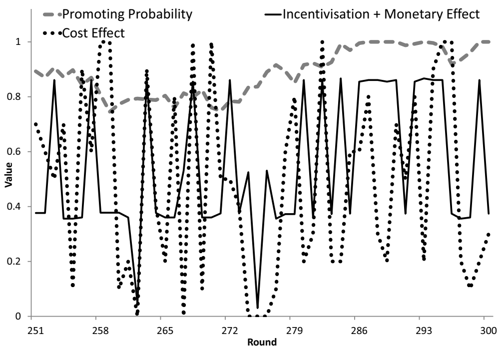


Figure 5: Learning-based incentivisation - agent level results (the unbalanced weights case)

weights), for the same reasons explained above for the other strategies. However, due to space constraints, we omitted these results here, and focused only on the case in which the other strategies fail to perform well.

6 CONCLUSION

The paper presented an information-based incentivisation approach, utilising various motivational information to support monetary rewards. In particular, to encourage an intermediary to promote a service for a provider, the provider utilises a set of incentivising agents, able to learn intermediary-specific preferences towards different informational factors (e.g. service popularity, peer pressure and reputation), and target them respectively with the most promising ones. The results of our experiments, conducted through an agent-based simulation, illustrated that such additional informational incentivisation can indeed be effective, especially in situations where monetary rewards are insufficient or unimportant (where the intermediary values other factors). Although the proposed approach addresses the service promotion problem, we believe that it forms a basis for tackling the more general problem of establishing a desired behaviour in a community. Future work involves investigating negative incentivisation effect, integrating other interaction topologies, and applying the proposed idea in a real-world application [9] (a software agent is currently deployed in over 60 clinics in the UK to help GPs recruiting patients for clinical trials, and we plan to support this software with the incentivisation capabilities outlined in this paper).

REFERENCES

[1] P. B. Clark and J. Q. Wilson, 'Incentive systems: A theory of organizations', *Administrative Science Quarterly*, **6**(2), pp. 129–166, (1961).

[2] J. M. Galan and L. R. Izquierdo, 'Appearances can be deceiving: Lessons learned re-implementing Axelrod's evolutionary approach to norms', *Journal of Artificial Societies and Social Simulation*, **8**(3), (2005).

[3] U. Gneezy, S. Meier, and P. Rey-Biel, 'When and why incentives (don't) work to modify behavior', *Journal of Economic Perspectives*, **25**(4), 191–210, (2011).

[4] U. Gneezy and A. Rustichini, 'Pay enough or don't pay at all', *The Quarterly Journal of Economics*, **115**(3), 791–810, (2000).

[5] M. Heitz, S. König, and T. Eymann, 'Reputation in multi agent systems and the incentives to provide feedback', in *Multiagent System Technologies*, volume 6251 of *Lecture Notes in Computer Science*, 40–51, (2010).

[6] D. Helbing, A. Szolnoki, M. Perc, and G. Szab, 'Punish, but not too hard: how costly punishment spreads in the spatial public goods game', *New Journal of Physics*, **12**(8), 083005, (2010).

[7] S. Mahmoud, N. Griffiths, J. Keppens, and M. Luck, 'An analysis of norm emergence in Axelrod's model', in *NorMAS'10: Proceedings of the Fifth International Workshop on Normative Multi-Agent Systems*. AISB, (2010).

[8] S. Mahmoud, J. Keppens, N. Griffiths, and M. Luck, 'Efficient norm emergence through experiential dynamic punishment', in *Proceedings of the 20th European Conference on Artificial Intelligence*, pp. 576–581. IOS Press, (2012).

[9] S. Mahmoud, G. Tyson, S. Miles, A.Taweel, T. Vanstaa, M. Luck, and B. Delaney, 'Multi-agent system for recruiting patients for clinical trials', in *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*, (2014). to appear.

[10] A. McDonald, R. Knight, M. Campbell, V. Entwistle, A. Grant, J. Cook, D. Elbourne, D. Francis, J. Garcia, I. Roberts, and C. Snowdon, 'What influences recruitment to randomised controlled trials? a review of trials funded by two uk funding agencies', *Trials*, **7**(1), 9, (2006).

[11] N. Nikiforakis, 'Punishment and counter-punishment in public good games: Can we really govern ourselves?', *Journal of Public Economics*, **92**, 91–112, (2008).

[12] M. Rodrigues and M. Luck, 'Cooperative interactions: An exchange values model', in *Coordination, Organizations, Institutions, and Norms in Agent Systems II*, volume 4386 of *Lecture Notes in Computer Science*, 356–371, (2007).

[13] B. T. R. Savarimuthu, M. Purvis, M. Purvis, and S. Cranefield, 'Social norm emergence in virtual agent societies', in *Declarative Agent Languages and Technologies VI*, volume 5397 of *Lecture Notes in Computer Science*, pp. 18–28, (2009).

[14] S. Treweek, E. Pearson, N. Smith, R. Neville, P. Sargeant, B. Boswell, and F. Sullivan, 'Desktop software to identify patients eligible for recruitment into a clinical trial: using SARMA to recruit to the ROAD feasibility trial.', *Informatics in primary care*, **18**(1), 51–58, (2010).

[15] D. Villatoro, G. Andrighetto, J. Sabater-Mir, and R. Conte, 'Dynamic sanctioning for robust and cost-efficient norm compliance', in *Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Spain, 16-22 July 2011*, (2011).

[16] D. J. Watts and S. H. Strogatz, 'Collective dynamics of 'small-world' networks.', *Nature*, **393**(6684), 440–442, (1998).