Incentive-Compatible Mechanisms for Norm Monitoring in Open Multi-Agent Systems (Extended Abstract)*

Natasha Alechina[1], **Joseph Y. Halpern**[2],**Ian A. Kash**[3] and **Brian Logan**[1]
[1] University of Nottingham
[2] Cornell University
[3] Microsoft Research
{nza,bsl}@cs.nott.ac.uk, halpern@cs.cornell.edu, iankash@microsoft.com

Abstract

We consider the problem of detecting norm violations in open multi-agent systems (MAS). In this extended abstract, we outline the approach of [Alechina *et al.*, 2018], and show how, using ideas from *scrip systems*, we can design mechanisms where the agents comprising the MAS are incentivised to monitor the actions of other agents for norm violations.

1 Introduction

Norms have been widely proposed as a means of coordinating and controlling the behaviour of agents in a multi-agent system (MAS). Norms specify the behaviours that agents should follow to achieve the objectives of the MAS. For example, the designer of a system to allow agents to post content (invitations to tender for work, prices of goods or services, etc.) may wish to ensure that the postings are relevant, accurate and up to date.

In a MAS where norms must be enforced, the responsibility for enforcing norms lies with a system component termed the *normative organisation* [Dastani *et al.*, 2009], which continuously monitors the actions of the agents (and perhaps carries out other tasks on behalf of the MAS). If an action (or the state resulting from an action) violates or would violate a norm, the action is either prevented, or the agent that performed the action is penalised (incurs a sanction). The effective monitoring of agent actions is therefore key to enforcing norms in a MAS. However, in large systems with many agents, maintaining a separate component to monitor the actions of the agents may involve significant overhead for the MAS.

In [Alechina *et al.*, 2018], we propose an approach to norm monitoring in open multi-agents systems in which the monitoring of agent actions is performed by the agents comprising the MAS. We term this *decentralised monitoring*. We focus on norms that prohibit certain actions (or the resulting state), for example, posting irrelevant or inaccurate content may be prohibited. The novelty of our approach is that the MAS does not need to bear the cost of paying for monitoring; at the same time we do not need to assume that fines can be levied on the agents who violate the norms and used to pay for monitoring, as done in [Fagundes *et al.*, 2014]. The latter assumption does not hold for many open systems where sanctioned agents can always leave the system and, if needed, rejoin it later under a different identity. Hence, a key issue for our approach is how to incentivise the agents to monitor the actions of other agents. We show how, using ideas from *scrip systems* [Friedman *et al.*, 2006; Kash *et al.*, 2012; 2015], we can design incentive-compatible mechanisms where the agents do the monitoring themselves. We can think of scrip as "virtual money" or "tokens". Performing an action costs a token, and detecting violations is rewarded with tokens. The main difference between our setting and that of [Kash *et al.*, 2015] (hereafter KFH) is that the agents are not always rewarded after they monitor, but only if they discover a violation. This requires a non-trivial adaptation of the techniques developed by KFH.

In this abstract, we present the main ideas from [Alechina *et al.*, 2018]. We consider two settings. In the first, the *inadvertent setting*, actions that violate a norm are assumed to be unintentional: violating a norm does not increase an agent's utility. In the second, the *strategic setting*, actions that violate the norm are intentional: violating the norm increases the agent's utility, and an agent chooses whether to try to violate the norm. We describe a mechanism that achieves *perfect enforcement* in the inadvertent setting; in equilibrium, all actions are monitored and hence there are no violations of the norm. In the strategic setting, we prove that there can be no equilibrium with perfect enforcement. However, the probability of violations can be made arbitrarily small: for all $\epsilon > 0$, we can design a mechanism where, in equilibrium, the probability of violations is at most ϵ.

2 Running Example

We consider a MAS where agents want to post content on the web. There are norms regarding the content that may be posted; for example, copyrighted images should not be posted, and comments should not be abusive or defamatory. We assume that agents may occasionally submit postings that violate the norm. If such content appears on the web, this may

---

[*] This paper is an extended abstract of an article in the Journal of Artificial Intelligence Research [Alechina *et al.*, 2018].

cause considerable harm to the MAS (e.g., it can be fined or sued). (As is standard in the MAS literature, we are assuming that the MAS is an entity independent of the agents that use it, that may incur the computational costs associated with monitoring and can be fined or sued.)

It is therefore in the MAS's interest that submitted postings are checked for compliance with the norm before they appear on the web. We assume that it is possible to check objectively if a particular item of content violates the norm. (For simplicity, we assume that if a posting that violates the norm is checked, the violation will be detected. We can easily modify our approach to handle the case where there is some probability ρ of the violation being caught.) Checking whether a posting is 'bad' (violates the norm) requires some work, and incurs a small utility cost. Although checking requires some resources, we assume that if a violation is found, evidence of the violation can be provided that can be checked in negligible time (so we do not need to deal with disputes about whether content violates the norm). If the content does violate the norm, the posting is discarded and no violation occurs. We assume a basic infrastructure that ensures that content submitted by an agent is signed, and that the digital signatures can be trusted. The signature identifies the agent, and is interpreted as a statement by the agent that the submitted content conforms to the norm. Note, however, that the infrastructure does not itself enforce the norm; it serves only to ensure auditability. We believe that such a separation of concerns is good design: the same basic infrastructure may be used by different systems with different norms.

There is a system-level objective that content conform to the norm, but since the cost to the MAS to check all submitted postings may be prohibitive, we would like to distribute the monitoring of submissions among the agents that use the system. Just as for the MAS, checking a submission incurs a small negative utility for an agent. This means that agents must be appropriately incentivised to monitor. It should be clear that the ideas exemplified by this scenario are applicable far more broadly.

We formalise the submission and monitoring of content for norm violations as a non-cooperative game. This scenario (and the resulting game) is similar to the scenario considered by KFH, but differs in several key respects. In the KFH setting, an agent requests a service, and the problem is to incentivise provision of the service; if the service is not provided, the requesting agent will not be satisfied. Here, it is not necessary that each submission be monitored for the submitting agent to be satisfied. We assume that, if no agent monitors, it is possible for the submitting agent to post and benefit from it; however, a norm violation may be missed. A more significant difference is that, in our setting, a submission may violate the norm. This has no analogue in the setting of KFH, and does complicate matters, as we shall see. Despite this, many of the ideas used by KFH can be used in our setting. In particular, we adopt the idea of using *tokens* as payment for submitting a posting and as a reward for monitoring. In order to submit a posting, an agent must pay one token; finding a bad posting is rewarded by receiving one or more tokens as payment. This encourages agents to volunteer to monitor submissions. The exact mechanisms and amounts are discussed below.

3 Unintentional Violation

In this section, we formalise the "incentivisation game". We show that there exists an equilibrium using threshold strategies if norm violations are unintentional, and that in this equilibrium all violations are detected.

In the inadvertent setting, bad submissions happen with a constant probability b, but agents are unaware that they are violating the norm when they submit something inappropriate. For technical reasons, we assume that b is a rational number (our results hold as long as we use a sufficiently good approximation to the true probability, so this assumption is really without loss of generality). The game in the inadvertent setting is described by the following parameters:

- a finite set of n agents $1, \ldots, n$;
- the time between rounds is $1/n$;
- in each round t an agent is picked at random to submit a posting (we implicitly assume that agents always have something that they want to post);
- probability of a submission being bad: b;
- utility of posting to the agent who posts: 1, (we assume that the utility of posting is independent of whether what is posted violates the norm, since violations are unintentional);
- disutility of monitoring to the agent doing the monitoring: $-\alpha$ (where $0 < \alpha < 1$);
- discount factor: $\delta \in (0, 1)$ (like KFH, we assume that agents discount future payoffs).

The game runs forever. We assume for simplicity that the system is homogeneous: all agents get the same utility for posting (1), the same disutility for monitoring ($-\alpha$), have the same probability of being chosen to submit a posting ($1/n$), and have the same discount factor (δ).[1] At each round an agent gets utility 1 if its submission is posted, utility $-\alpha$ if it monitors and utility 0 otherwise.

To incentivise monitoring, we use tokens as payment for submitting a posting and as a reward for monitoring. An agent must pay one token in order to submit a posting. Agents are rewarded with tokens only if they monitor and detect a bad submission. We argue below that in order for the system to function successfully (agents being able to submit postings, and some agents always available for monitoring), the 'right' amount to pay for finding a bad submission is $1/b$.[2] This means, in expectation, an agent gets one token for finding a bad submission. Thus, the price of submitting a posting is equal to the expected reward for checking a submission.

We need some additional notation to describe what happens:

- $p^t \in \{1, \ldots, n\}$ is the agent chosen to submit a posting in round t;
- $V_i^t \in \{0, 1\}$ denotes whether agent i volunteers to monitor in round t ($V_i^t = 1$). Whether an agent volunteers is

---

[1] Using ideas from [Kash *et al.*, 2012; 2015], we can extend the approach discussed here to deal with different *types* of agents, characterised by different parameters.

[2] We are implicitly assuming that tokens are divisible into units such that it is possible to transfer $1/b$ tokens.

determined by the agent's strategy (which may depend on p^t).

- $v^t \in \{0, \dots, n\} \setminus \{p^t\}$; $v^t = j$ if agent $j \neq p^t$ such that $V_j^t = 1$ is chosen to monitor in round t, and $v^t = 0$ if no one is chosen to monitor in round t;
- $f^t \in \{0, 1\}$; $f^t = 0$ if the content submitted in round t is good, $f^t = 1$ if it is bad.

Given that good and bad postings have the same utility (1), the utility of an agent i in round t is:

$$u_i^t = \begin{cases} 1 & \text{if } i = p^t,\ i \text{ has} \geq 1 \text{ token, and } v^t = 0 \text{ or } f^t = 0; \\ -\alpha & \text{if } v^t = i; \\ 0 & \text{otherwise.} \end{cases}$$

Thus, an agent gets utility 1 in round t if it is chosen to submit a posting ($p^t = i$), it has at least one token, and either the posting is not monitored ($v^t = 0$) or it does not violate the norm ($f^t = 0$). Given the discount factor δ, the total utility U_i for agent i is $\Sigma_{t=0}^{\infty} \delta^{t/n}\ u_i^t$. Note that the number of tokens that an agent has does not affect the agent's utility. However, as an agent requires a token to submit a posting, the number of tokens that an agent has does have an indirect impact on utility; if the agent has no tokens, then it will not be able to submit a posting, and thus will forego the opportunity to get utility 1.

Paying agents $1/b$ for finding a bad posting makes the situation similar to that considered by KFH, where the agent wanting work done pays one token, and the agent doing the work gets one token. However, the fact that in the current setting payment is made only if a submission is found to be bad complicates matters. An expected payment of 1 token is not equivalent to an actual payment of 1 token! To understand the issue here, note that the most obvious way to deal with the payment of tokens is to have the agent who wants to submit a posting pay one token to the normative organisation, and then have the normative organisation pay $1/b$ tokens to the monitor if a violation is detected. But there are problems with this approach. If monitors have a long run of "bad luck" and do not find submissions that violate the norm, agents will have very few tokens on average; on the other hand, if monitors get lucky, and find quite a few submissions that violate the norm, agents will end up with many tokens on average. As pointed out by KFH, having both too few or too many tokens will cause problems. Intuitively, with too many tokens, (almost) everyone will have plenty of tokens, so no one will volunteer to monitor; with too few tokens, it will often be the case that the agent who is chosen to submit a posting will not have a token to pay for the submission. This problem does not occur in the setting of KFH, because there the payment made by the agent requesting a service always matches exactly the payment received by the agent providing the service.

We solve this problem by having the agents rather than the normative organisation perform the role of the "bank". When agent i wants to submit a posting, it pays a randomly chosen agent who has fewer than the maximum number of tokens allowed (see below) 1 token; if an agent j monitors and finds a violation, a randomly chosen agent with at least $1/b$ tokens gives j $1/b$ tokens. This ensures that the number of tokens in 'circulation' remains constant. (Note also that since the agent that pays the monitoring agent is randomly chosen from among those agents with at least $1/b$ tokens, agents may submit a posting as long as they have a single token. If the submitting agent were to pay the monitoring agent when a violation is detected, agents could not submit a posting unless they have $1 + 1/b$ tokens.) It may seem counterintuitive to have a random agent pay the penalty. However, since violations are inadvertent and occur randomly, each agent pays the same amount in penalties in expectation as if we had charged the violator; moreover this approach has the advantage of making the system run more smoothly.

We assume that all agents follow a *threshold strategy* when deciding whether to monitor. There is a fixed threshold k such that agents volunteer iff they have fewer than k tokens. It is easy to see that there is an equilibrium in threshold strategies if everyone uses a threshold of 0. In that case, no one ever volunteers to monitor a submission, so everyone gets to post, without monitoring. Of course, no agent has any incentive to deviate from this strategy. On the other hand, this equilibrium runs counter to the purposes of the MAS. We are thus interested in nontrivial equilibria in threshold strategies, where everyone uses a threshold $k > 0$. Note that the maximum number of tokens any agent may have is then $k + 1/b$. If an agent i has at least one token and is chosen to submit a posting ($i = p^t$), p^t gives a randomly chosen agent with fewer than $k + 1/b$ tokens one more. The subitting agent p^t than asks for volunteers to act as monitor. All agents with fewer than k tokens volunteer. If at least one agent volunteers, one, v^t, is chosen at random to act as monitor. If v^t confirms the submission conforms to the norm, it is posted. If v^t detects a violation of the norm, then the posting is discarded, and a randomly chosen agent with at least $1/b$ tokens gives v^t $1/b$ tokens.

Theorem 1 *For all $\epsilon > 0$, there exist a δ sufficiently close to 1 and an n sufficiently large such that if all n agents have a discount factor $\delta' \geq \delta$, then there exists a k such that the mechanism above with all agents using a threshold of k is an ϵ-Nash equilibrium.*

4 Intentional Violation

In this section, we show that there exists an equilibrium in threshold strategies when norm violations are intentional, and that in this equilibrium most violations are detected.

In the strategic setting, we assume that when an agent is chosen to submit a posting, it can either submit something good (i.e., that does not violate the norm) or something bad. The parameters of the game are the same as in Section 3, except that:

- there is no longer a probability b of a submission being bad (the quality of a submission now becomes a strategic decision); and
- the utility of a bad posting is no longer 1, but $\kappa > 1$ (we must assume $\kappa > 1$ here, otherwise no agent would ever submit anything bad: the utility of doing so is no higher than that of submitting something good, and the violation may be detected).

As before, monitoring agents get paid only if they find a bad submission. With these assumptions, it is not hard to show that there does not exist an equilibrium with perfect enforcement.

Theorem 2 *In the setting of strategic violations, the game has no equilibrium with perfect enforcement.*

Although we cannot achieve perfect enforcement in the strategic setting, we can achieve the next best thing: we can make the probability of a bad posting as low as we want. More precisely, for all $\epsilon, \epsilon' > 0$, there is an ϵ-Nash equilibrium such that the probability of a bad posting is ϵ'.

The idea now is that, with some probability, a submitted posting will not be checked; there will be no attempt to get volunteers to monitor that submission. Let $c^t = 0$ if there is no monitoring in round t; $c^t = 1$ otherwise. The decision regarding whether to monitor is made *after* the agent chosen to submit, p^t, submits their posting (otherwise, p^t will always submit something bad in round t if $c^t = 0$). If $c^t = 0$, then whatever p^t submits is posted in that round, whether it is good or bad. As before, if an agent submits a bad posting and there is monitoring, we assume that the bad posting definitely will be detected and discarded, so the submitting agent gets utility 0 in that round. The utility of agent i in round t now becomes

$$u_i^t = \begin{cases} 1 & \text{if } i = p^t \text{ and } f^t = 0; \\ \kappa & \text{if } i = p^t,\ f^t = 1,\ c^t = 0 \vee (c^t = 1, v^t = 0); \\ -\alpha & \text{if } i = v^t; \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that the normative organisation decides that postings will be monitored with probability $1 - 1/\kappa$. Further suppose that an agent uses a randomised algorithm: with probability β it submits a good posting, and with probability $1 - \beta$ it submits a bad posting. The agent's expected payoff is then $\beta + (1 - \beta)(1/\kappa)\kappa = 1$, independent of β. Thus, we get an equilibrium in the single-shot game if monitoring occurs with probability $1 - 1/\kappa$ and agents submit bad postings with probability β, for all choices of β, provided that there is always guaranteed to be a monitor available. We will show that again there is an equilibrium in threshold policies. As long as there are not too many tokens in the system, there are bound to be some agents with fewer than the threshold number of tokens, so there will be a volunteer.

We assume the designer of the MAS specifies a value β^* (that, intuitively, should be a small 'tolerable' probability of a violation occurring). If a monitor that finds a violation is paid $1/\beta^*$ tokens, then essentially the same type of mechanism as that proposed for the case of unintentional violations will work, provided that we get bad submissions with probability exactly β^*. So, perversely, in this setting, while all strategies are equally good for the submitting agent, the MAS actually wants to *encourage* agents to submit something bad with probability β^*, so that monitors again get an expected payment of 1 token. The way to do this is for the normative organisation to announce that it will track the number of bad submissions, and if the fraction of submissions that have been bad up to round t is β, checks will happen with probability $1 - \beta^*/(\beta\kappa)$. Thus, if $\beta = \beta^*$, then checks happen with probability $1 - 1/\kappa$, and we have an equilibrium. Moreover, the payment ($1/\beta$ tokens) is exactly what is needed to ensure that, in equilibrium, monitoring occurs with probability β^*. For if $\beta < \beta^*$, then the check will happen with probability less than $1 - 1/\kappa$, which means that agents will want to make more bad submissions. On the other hand, if $\beta > \beta^*$, then monitoring will happen with probability greater than $1 - 1/\kappa$, and agents will want to make fewer bad submissions. Thus, in equilibrium, we get bad submissions with probability exactly β^*.

To summarise, we have the following mechanism, given a threshold k. If an agent is chosen to submit, it submits bad content with probability β^* and good content with probability $1 - \beta^*$. After the agent has decided what to submit and made the posting available, the normative organisation decides whether the posting will be monitored. For an initial period (say 1,000 rounds), a submission is monitored with probability $1 - 1/\kappa$; afterwards, if the fraction of submissions that have been discovered to be bad due to monitoring is β and β is more than (say) two standard deviations from β^*, then monitoring occurs with probability $1 - \beta^*/(\beta\kappa)$. If the decision has been made to monitor, and the submitting agent has at least one token (so that a submission can be made), the submitting agent asks for volunteers; all agents with fewer than k tokens volunteer to monitor and one is chosen at random to be the monitor. As in the case of unintentional violations, if at least one agent volunteers, then the submitting agent gives a randomly chosen agent with less than $k + 1/\beta^*$ tokens one more. If the monitor approves the submission, it is posted. If the monitor finds a problem with the submission, then a randomly chosen agent with at least $1/\beta^*$ tokens gives the monitor $1/\beta^*$ tokens and the submission is discarded.

Theorem 3 *For all $\epsilon > 0$, there exist a δ sufficiently close to 1 and an n sufficiently large such that if all n agents use a discount factor $\delta' \geq \delta$, then there exists a k such that the mechanism above with all agents using a threshold of k is an ϵ-Nash equilibrium.*

Note that, in the equilibrium whose existence is stated in Theorem 3, the probability of a bad submission is β^*, as desired.

5 Additional Content

In [Alechina *et al.*, 2018], we provide proofs of the theorems, describe how to make the system more distributed and reduce the role of normative organisation, and show using simulations that our theoretical results, that apply to systems with a large number of agents, hold for multi-agent systems with as few as 1000 agents—the system rapidly converges to the steady-state distribution of scrip tokens necessary to ensure monitoring and then remains close to the steady state. We also compare our work to related work on scrip systems and monitoring in multi-agent systems.

Acknowledgements

Joe Halpern was supported in part by NSF grants IIS-0534064, IIS-0812045, IIS-0911036, and CCF-1214844, and by AFOSR grants FA9550-08-1-0438, FA9550-09-1-0266, and FA9550-12-1-0040, and ARO grant W911NF-09-1-0281.

References

[Alechina *et al.*, 2018] Natasha Alechina, Joseph Y. Halpern, Ian A. Kash, and Brian Logan. Incentive-compatible mechanisms for norm monitoring in open multi-agent systems. *Journal of Artificial Intelligence Research*, 62:421–446, 2018. (To appear).

[Dastani *et al.*, 2009] Mehdi Dastani, Davide Grossi, John-Jules Ch. Meyer, and Nick Tinnemeier. Normative multi-agent programs and their logics. In John-Jules Ch. Meyer and Jan Broersen, editors, *Knowledge Representation for Agents and Multi-Agent Systems*, volume 5605 of *Lecture Notes in Computer Science*, pages 16–31. Springer Berlin Heidelberg, 2009.

[Fagundes *et al.*, 2014] Moser Silva Fagundes, Sascha Ossowski, and Felipe Meneguzzi. Analyzing the tradeoff between efficiency and cost of norm enforcement in stochastic environments. In *21st European Conference on Artificial Intelligence (ECAI 2014)*, pages 1003–1004. IOS Press, 2014.

[Friedman *et al.*, 2006] Eric J. Friedman, Joseph Y. Halpern, and Ian A. Kash. Efficiency and Nash equilibria in a scrip system for P2P networks. In *Proceedings 7th ACM Conference on Electronic Commerce (EC-2006)*, pages 140–149. ACM, 2006.

[Kash *et al.*, 2012] Ian A. Kash, Eric J. Friedman, and Joseph Y. Halpern. Optimizing scrip systems: crashes, altruists, hoarders, sybils and collusion. *Distributed Computing*, 25(5):335–357, 2012.

[Kash *et al.*, 2015] Ian A. Kash, Eric J. Friedman, and Joseph Y. Halpern. An equilibrium analysis of scrip systems. *ACM Transactions on Economics and Computation*, 3(3), 2015. Article No. 13.