Decentralised Norm Monitoring in Open Multi-Agent Systems

(Extended Abstract)

Natasha Alechina
University of Nottingham
nza@cs.nott.ac.uk

Joseph Y. Halpern*
Cornell University
halpern@cs.cornell.edu

Ian A. Kash
Microsoft Research
iankash@microsoft.com

Brian Logan
University of Nottingham
bsl@cs.nott.ac.uk

ABSTRACT

We propose an approach to norm monitoring in open multi-agents systems (MAS) in which monitoring is performed by the agents comprising the MAS. Using ideas from *scrip systems*, we show how to design mechanisms where agents are incentivised to monitor the actions of other agents for norm violations. Crtically, the cost of providing incentives is not borne by the MAS. Instead, monitoring incentives come from (scrip) fees for accessing the services provided by the MAS. In some cases perfect monitoring (and hence enforcement) can be achieved: no norms will be violated in equilibrium. In other cases, we show that, while it is impossible to achieve perfect enforcement, we can get arbitrarily close; we can make the probability of a norm violation in equilibrium arbitrarily small.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multi-agent systems*

1. INCENTIVISING MONITORING

We propose an approach to norm monitoring in open multi-agents systems (MAS) in which the monitoring of agent actions is performed by the agents comprising the MAS. We focus on norms which prohibit certain actions (or the state resulting from an action). The novelty of our approach is that the MAS does not bear the cost of monitoring; nor does it levy fines on violating agents to pay for monitoring as in, for example, [1]. (Levying fines is not possible in many open systems as the agents can always leave the system and rejoin later under a different identity.) Instead, monitoring incentives come from (scrip) fees [2] for accessing the services provided by the MAS.

As a simple example, consider a MAS where agents want to post content on the web. There are norms regarding what 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 posts that violate the norm. If such content appears on the web, the MAS loses significant utility (e.g., it can be fined or sued). It is therefore in the MAS's interest that submitted posts 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 post 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 post 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 post is discarded and no violation occurs. We would like to distribute the monitoring of posts among the agents that use the system. Just as for the MAS, monitoring incurs a small negative utility for an agent. This means that agents must be appropriately incentivised to monitor. The scenario we describe applies to a wide range of normative multi-agent systems. The only requirement is that norms apply to single actions, actions can be checked in advance, and their execution prevented if they are found to violate the norm. Examples other than web postings include bids for jobs and bookings of facilities.

We formalise the posting and monitoring of content for norm violations as a non-cooperative game. This scenario (and the resulting game) is similar to that considered by Friedman et al. [2], but differs in several key respects. As in [2], we adopt the idea of using *tokens* as payment for posting and as a reward for monitoring.

2. UNINTENTIONAL VIOLATION

We first consider a scenario in which bad posts are unintentional, and happen with a constant probability b. In this scenario, agents are unaware that they are violating the norm when they post something inappropriate. The game in the inadvertent scenario is described by the following parameters:

- a finite set of n agents $1, \dots, n$;
- the time between rounds is $1/n$;
- at each round t an agent is picked at random to submit a post (we implicitly assume that agents always have something that they want to post);
- probability of a post being bad: b;

*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.

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

- utility of posting (to the agent doing the posting): 1
- disutility of monitoring (to the agent doing the monitoring): $-\alpha$ (where $0 < \alpha < 1$);
- discount rate: $\delta \in (0, 1)$.

The game runs forever. As is standard in the literature, we assume that agents discount future payoffs (this allows us to compute the total utility derived by an agent in the infinite game).

We need some additional notation to describe what happens:

- $p^t \in \{1, \dots, n\}$ is the agent chosen to submit a post at t;
- $v^t \in \{0, \dots, n\} \setminus \{p^t\}$; $v^t = j$ if agent $j \neq p^t$ is chosen to monitor at t, and $v^t = 0$ if no one is chosen to monitor; at t;
- $f^t \in \{0, 1\}$; $f^t = 0$ if the content posted at t is good, $f^t = 1$ if it is bad.

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

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

Given the discount factor δ, the total utility U_i for agent i is $\Sigma_{t=0}^{\infty} \delta^{t/n}\, u_i^t$.

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

Paying agents $1/b$ for finding a bad posting makes the situation similar to that in [2], where the agent wanting work done pays one token, and the agent doing the work gets one token. However, the fact that in our setting payment is made only if a problem is found complicates matters. Specifically, the agents cannot simply pay the MAS to post, or the MAS pay the agents for finding a bad post. If monitors have a long run of "bad luck" and do not find postings that violate the norm, there will be very few tokens left in the system; on the other hand, if monitors get lucky, and find quite a few postings that violate the norm, the MAS will end up pumping quite a few tokens into the system. Having both too few or too many tokens in the system causes problems [2]. Intuitively, with too many tokens in the system, (almost) all agents will have plenty of tokens, so no one will volunteer to monitor; with too few tokens in the system, it will often be the case that the agent who wants to post will not have a token to pay for it.

To ensure that the number of tokens in "circulation" remains constant, the agents rather than the MAS perform the role of the "bank". 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. If an agent i has at least one token and is chosen to submit a post ($i = p^t$), p^t gives a randomly chosen agent with fewer than $k + 1/b$ tokens one more. The posting 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 that the post conforms to the norm, it is posted. If v^t detects a violation of the norm, then the post 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.*

Proofs of this and all other results, as well as more details and intuition, can be found in the full version, available at arxiv.org/.../.. Note that in the equilibrium the existence of which is stated in Theorem 1, we get perfect enforcement; all bad posts will be detected.

3. STRATEGIC VIOLATION

We now consider the scenario in which violations are strategic. When an agent is chosen to submit a post, 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 2, except that there is no longer a probability b of a posting being bad (the quality of a posting 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 post anything bad: the utility of doing so is no higher than that of posting something good, and the violation may be detected.) As before, monitoring agents get paid only if they find a bad post.

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, there can be no equilibrium with perfect enforcement.*

Although we cannot achieve perfect enforcement in the strategic setting, 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 post is ϵ'.

We have the following mechanism, given a threshold k. If an agent is chosen to post, it submits bad content with probability β^* and good content with probability $1 - \beta^*$. After the agent has decided what to post and made the posting available, the MAS decides whether the posting will be monitored. For an initial period (say 1,000 rounds), a posting is monitored with probability $1-1/\kappa$; afterwards, if the fraction of postings 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 posting agent has at least one token (so that a post can be made), the posting agent asks for volunteers and all agents with fewer than k tokens volunteer to monitor and one is chosen to be the monitor. As in the case of unintentional violations, if at least one agent volunteers, then the posting agent gives a randomly chosen agent with less than $k + 1/\beta^*$ tokens one more. If the monitor approves the posting, it is posted. If the monitor finds a problem with the posting, then a randomly chosen agent with at least $1/\beta^*$ tokens gives the monitor $1/\beta^*$ tokens.

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 posting is β^*, as desired.

REFERENCES

[1] M. S. Fagundes, S. Ossowski, and F. 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. 2014.

[2] E. J. Friedman, J. Y. Halpern, and I. 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. 2006.