

Using Incentives to Obtain Truthful Information

Boi Faltings

Artificial Intelligence Laboratory (LIA)
Swiss Federal Institute of Technology (EPFL)
IN-Ecublens, CH-1015 Ecublens, Switzerland
boi.faltings@epfl.ch

Abstract. There are many scenarios where we would like agents to report their observations or expertise in a truthful way. Game-theoretic principles can be used to provide incentives to do so. I survey several approaches to eliciting truthful information, in particular scoring rules, peer prediction methods and opinion polls, and discuss possible applications.

1 INTRODUCTION

The internet has opened many new possibilities for gathering information from large numbers of individual agents. For example, people rate services in reputation forums, they annotate maps with location information, and they answer questions in online forums. In the future, software agents will control networks of sensors and report measurements such as air quality, radio spectrum, or traffic congestion.

An implicit assumption is that agents will make their best effort to report such information truthfully. However, when they are self-interested, this can not always be assumed. For example, in online reputation forums, leaving a rating is a time-consuming operation and most users will not do this unless they have a motive. Thus, one can often observe skewed distributions of ratings that indicate that most reviews were left by users who either loved or hated the item they rated ([Hu et al., 2006]). It is not clear whether ranking items by taking averages of such reviews is very helpful. Similar, sensors may save energy by providing inaccurate measurements or no measurements at all, or they may be manipulated to provide skewed reports that are beneficial to the interests of their owner.

To obtain better quality information, it is important to reward agents who contribute ratings and thus increase participation of agents even without ulterior motives. Such reward schemes could be useful both as incentives to human agents as well as for software agents operating sensors: rewards could finance the operation of the sensors and direct their deployment towards the most useful measurements ([Aberer et al., 2010]).

Furthermore, it is possible to scale the rewards so that they specifically reward truthful reporting, and can even counter exterior incentives to report false information. These mechanisms are based on scoring rules that reward correct prediction of a future outcome once that outcome becomes known. In peer prediction methods, these rules are extended to situations where the true outcome never becomes known. Instead, they take the predictions of other agents as the ground truth to compare to. This makes truthfulness an equilibrium, i.e. the best response strategy when all other agents are also

truthful. Finally, I show how to design mechanisms that achieve this independently of agent beliefs and are thus easier to apply in practice, for example for encouraging truthfulness in opinion polls.

2 Truthful Reporting Through Scoring Rules

In many cases, agents are asked to provide information about an outcome that will eventually become known with certainty. For example, experts may predict the weather, the future of the economy, or the completion date of a project. When this is the case, incentives for reporting this information truthfully can be provided through *proper scoring rules* ([Savage, 1971]). Agents provide information in the form of a probability distribution on different possible outcomes. Once the true outcome becomes known, they get paid a reward that depends on how well their prediction matched the observed outcome. This reward is computed by a scoring rule that takes the report and the true outcome as arguments. A scoring rule is called *proper* if it provides the highest expected reward exactly when the agent reports its probability distribution truthfully.

Assume that the task is to predict which of k outcomes o_1, \dots, o_k will actually occur, and that an expert agent has a probability distribution $\underline{p} = (p(o_1), \dots, p(o_k))$ for the true outcome. The agent reports this distribution as $\underline{q} = (q_1, \dots, q_k)$. We would like to provide incentives so that it is optimal to report $\underline{q} = \underline{p}$.

This can be provided for example using the *quadratic scoring rule*:

$$\text{pay}(o_t, \underline{q}) = a + b \left(2q_t - \sum_{j=1}^k q_j^2 \right)$$

where o_t is the outcome that actually occurred and a is a non-negative and b a positive constant. It is straightforward to show that this scoring rule is proper in that the expected payment:

$$\begin{aligned} E[\text{pay}](\underline{q}) &= \sum_{i=1}^k p(o_i) \text{pay}(o_i, \underline{q}) \\ &= a + b \left[2 \sum_{j=1}^k p(o_j) q_j - \left(\sum_{j=1}^k p(o_j) \right) \left(\sum_{j=1}^k q_j^2 \right) \right] \\ &= a + b(2\underline{p} \cdot \underline{q} - |\underline{q}|) \end{aligned}$$

is maximized by maximizing $\underline{p} \cdot \underline{q}$, which is the case exactly when the vectors \underline{p} and \underline{q} are aligned. Thus, reporting truthfully is a dominant strategy for agents.

As an example, consider predicting whether the next day's weather will be good (g) or bad (b) as a vector of two probabilities (p(g), p(b)). Let the scoring rule be

$$\text{pay}(o_t, \underline{q}) = 1 + \left(2q_t - \sum_{j=1}^k q_j^2 \right)$$

An expert's true belief could be that the weather will be good with probability 0.8, and bad with probability 0.2. Now consider the expected payoff for reporting this distribution truthfully. If the weather turns out to be good, the expert receives a payment of

$pay(g, (0.8, 0.2)) = 1 + 2 \cdot 0.8 - 0.68 = 1.92$; if it turns out to be bad, the payment is $pay(b, (0.8, 0.2)) = 1 + 2 \cdot 0.2 - 0.68 = 0.72$. Thus, the expected payoff for truthfully reporting the probability distribution is:

$$0.8pay(g, (0.8, 0.2)) + 0.2pay(b, (0.8, 0.2)) = 1.68$$

Now consider a false report, for example $(0.5, 0.5)$. Now the reward in case of good and bad weather is identical and equal to $pay(g/b, (0.5, 0.5)) = 1 + 2 \cdot 0.5 - 0.5 = 1.5$, and thus the expected payment is also equal to 1.5. This is significantly less than what is expected for truthful reporting.

There are other proper scoring rules, such as the logarithmic scoring rule:

$$pay(o_t, \underline{q}) = a + b \ln q_t$$

where o_t is the outcome that actually occurred and a is a non-negative and b a positive constant. These may lead to lower expected payments or wider margins for truth-telling, but can have other drawbacks. For example, with the logarithmic scoring rule payments can become negative.

Proper scoring rules can also be constructed for eliciting averages and other properties of distributions. Recently, [Lambert and Shoham, 2009] have characterized the questions to which truthful answers can be elicited using scoring rules.

3 The Peer Prediction Method

Proper scoring rules can be applied whenever the ground truth that is being observed can eventually be verified. However, there are many cases where this condition is not satisfied. Consider for example ratings reported for products and services on the internet: it is not possible to independently verify whether these ratings were given truthfully. Similarly, measurements taken by sensors would often not be verifiable by other means. A similar situation exists when reporting opinions about hypothetical scenarios, such as what would happen if interest rates were raised by different degrees: since only one of these scenarios will actually be implemented, predictions about the others cannot be verified.

However, in such cases it is still possible to make truthful reporting an equilibrium strategy for agents by applying a proper scoring rule based on the prediction of another agent, called a reference report. Provided the other agent made a truthful prediction and both have the same knowledge and observing the same signals, truthful reporting is the best response. Thus, for a population of agents with the same knowledge, reporting truthfully is a Nash equilibrium. This is called the *peer prediction method* in ([Miller et al., 2005]).

As an example, consider reporting the quality of service received by a plumber. Two agents A and B both report on the quality of service they received. The key idea is that the quality of service A received will influence its expectation of the quality that B received: if A observed good service, then its belief for the probability $p(g|g)$ that B also received good service is higher than the value $p(g|b)$ if A received bad service. Assume for this example that $p(g|g) = 0.8$ and $p(g|b) = 0.4$.

Now we apply the same scoring rule mechanism we mentioned earlier, but consider B's report the ground truth. If A observed good service, its probability distribution for B's report is $(p(g|g), p(b|g)) = (0.8, 0.2)$, and just like in the weather prediction example it's expected reward for the scoring rule:

$$pay(o_t, \underline{q}) = 1 + \left(2q_t - \sum_{j=1}^k q_j^2 \right)$$

is 1.68, provided that A's probability distribution for B's experience is indeed (0.8, 0.2).

If A did not experience good service, it would expect B's observation to follow a different probability distribution, in this case (0.4, 0.6). If it nevertheless reports good service, the expected reward is only $0.4 \cdot 1.92 + 0.6 \cdot 0.72 = 1.2$. On the other hand, when A truthfully reports bad service, the mechanism treats this as a prediction of the probability distribution (0.4, 0.6) for B's experience. The payments for truthfully reporting bad service are calculated using the probabilities (0.4, 0.6) and would lead to a higher expected reward for truthful reporting of $0.4 \cdot (1 + 2 \cdot 0.4 - 0.52) + 0.6 \cdot (1 + 2 \cdot 0.6 - 0.52) = 1.52$.

Note that, contrary to the weather prediction, we are not asking A to report this probability distribution, but only whether it received good or bad service. Thus, the designer of the reward scheme needs to know how an observation influences A's beliefs about the observations of another agent B with reasonable precision in order to compute the payments. It can in part be deduced from the general expectations of the quality of service, but also involves an assumption of how the individual agents would update their beliefs in response to a positive or negative experience.

Furthermore, the original peer prediction method suffers from the weakness that truthful reporting is not the only equilibrium strategy: any strategy where agents all report the same is also a Nash equilibrium. In fact, since actual observations or predictions are likely to be noisy, the highest-paying equilibrium is always one where agents always report the same, independently of their true knowledge!

This problem can be overcome by constructing scoring rules that refer not to one, but several reference reports. [Jurca and Faltings, 2007a, Jurca and Faltings, 2009] show that when at least 3 reference reports are used, truthful reporting can be made the highest-paying Nash equilibrium. Furthermore, they show that truthful reporting can be made the *only* Nash equilibrium and thus completely eliminate the problem of collusive reporting strategies.

It has recently been shown that peer prediction methods can be generalized to scenarios where agents report not on identical events, but events that are merely correlated ([Witkowski, 2009]). This makes it applicable for example to measurements in sensor networks, where different sensors measure quantities that are correlated by not equal.

4 Opinion polls

A major weakness of the peer prediction method is that it requires all participating agents to share the same probability distribution of the reported events. If this is not the

case, proper scoring rules can still be designed, but the rewards that must be paid to agents quickly become very large ([Jurca and Faltings, 2007b]).

To counter this effect, it is possible to design peer prediction schemes as opinion polls that publish the current results of the poll. Agents whose probability distribution is sufficiently close to this published one will have truthful reporting as their best strategy, while agents that consider the public distribution as grossly wrong may instead be merely helpful by making reports that will drive the public poll closer to what they consider to be the true distribution.

Such a mechanism was first shown in ([Jurca and Faltings, 2008]) for aggregating opinions about a hidden signal that could be either good (g) or bad (b). At time t , the published polls shows the average fraction R_t of good reports. An agent A_i has its own probability distribution $p_i(r|s)$ that characterizes the conditional probability distribution of a reference report r given its own observation s of the signal, where the reference report is filed by another agent that observes the same signal and the same public poll. The mechanism compares the report s filed by agent A_i to a reference report r filed by another agent B, and rewards A_i if the two reports match:

- for matching a good report, the reward is $c(1 - R_t)$.
- for matching a bad report, the reward is cR_t .

where c is a positive constant to scale the average reward, for example to ensure that it compensates for the effort required to file it.

To analyze the incentives for agent A_i , we distinguish three cases:

- a) A_i considers the current poll result reasonable, characterized by the fact that $p_i(g|b) < R_t < p_i(g|g)$.
- b) A_i considers the poll result unreasonably high, characterized by the fact that $R_t \geq p_i(g|g)$, which means that no matter what A_i observes, it would always expect other agents to observe a bad signal with a higher probability than the current poll result.
- c) A_i considers the poll result unreasonably low, characterized by the fact that $p_i(g|b) \geq R_t$, symmetrically on the other side.

In the case where the poll result is reasonable, the agent is best off reporting truthfully. Consider the case where it observes a good signal, then the expected rewards are:

- for reporting good (truthful):

$$p_i(g|g)c(1 - R_t) > R_t c(1 - R_t)$$

- for reporting bad (non truthful):

$$p_i(b|g)cR_t = (1 - p(g|g))cR_t < (1 - R_t)cR_t$$

Thus, the expected reward for reporting truthfully is strictly greater than the expected reward for a non-truthful report. A symmetric analysis can be made for the case of a bad observation.

As an example, consider that agents A and B both hire a plumber that according to the public reputation scheme provides good service 90% of the time, based on 10

previous reports. Suppose that A sees the plumber at work and he does a good job. Then A might consider that the current poll value is accurate or slightly too low and report good service, expecting a payment of $10/9$ with a probability of higher than 0.9, so above 1 in expectation.

However, if the agent considers the poll unreasonably high, its best strategy is to report bad, independently of its own observation. While this behavior is not truthful, it can be considered *helpful* in that the agent drives the outcome of the poll closer to its own opinion. For example, suppose A observes the plumber at work and realizes that he is completely incompetent, but still by chance receives good service. Now, A might have a private probability that B will receive bad service that is much higher than the 10% that would be expected from the poll, let's say 50%. Now A would be better off reporting poor service, as its expected reward would be $1/0.1 = 10$ with probability 0.5, which is much higher than his expectation in case of truthful reporting. However, the report could still be considered helpful in that it drives the value of the opinion poll towards A's true opinion.

The advantage of this mechanism is that agents can have different and unknown prior distributions for the signal, whereas scoring rules require this distribution to be known to the mechanism designer.

5 Applications

The techniques reported here have numerous applications. The most obvious ones are forums such as reputation and review forums. Leaving such feedback is cumbersome and thus often done by agents who have ulterior motives and thus do not leave honest reports. Here it would be useful to reward raters for their effort, and it would be even better to scale these payments to encourage honest feedback.

Another range of applications is in ensuring quality of crowdsourcing. For example, consider an image labeling task as in the ESP game proposed in [von Ahn and Dabbish, 2004]: two people are independently asked to give keywords that describe the content of an image. They get a reward when they provide matching keywords. This game has the flaw that people will tend to use very common words, and so these have to be explicitly excluded. A more general strategy based on the opinion poll mechanism given above would be to scale the rewards according to the frequency of the matching word: a less common word would fetch a higher reward. One can imagine many other applications in crowdsourcing where rewards depend on the agreement with other worker's results.

Further applications can be found in sensor networks. The peer prediction method can be generalized to settings where agents do not measure exactly the same signal. It is sufficient that measurements are correlated in a known way ([Witkowski, 2009]). Thus, one can design a reward scheme that rewards truthful operation of a network of sensors that sense related values, for example air pollution ([Aberer et al., 2010]). This could be applied in sensor networks, in particular when sensors are operated by different entities who might save cost by inaccurate measurements, or even maliciously want to manipulate measurements.

Services such as internet access, cloud computing, or wireless communications require monitoring of the quality of service. This would be most easily done by the cus-

tomers themselves, but the difficulty is that they often have an incentive to misreport since they stand to gain refunds or other claims if service is deemed to be insufficient. Somewhat surprisingly, it turns out that incentive mechanisms are entirely sufficient to solve this problem, as shown in [Jurca et al., 2007]. Provided that the entire user population is sufficiently large, it would take a significant coalition of users to shift the average reported quality enough to obtain a refund for poor service. However, as long as such a large coalition has not formed, a reward scheme based on peer prediction is sufficient to punish each individual user for deviating from truthful reporting, and can be realized at low cost. Thus, a lying coalition would have to be created in a coordinated fashion, and such coordinated action would be detectable by other means. This opens another wide range of applications.

6 Conclusions

The internet has enabled wide distribution of user-contributed content whose correctness cannot be verified. Much of this content is reported by agents with ulterior motives and may often not reflect the truth. I have discussed ways of providing incentives to agents to provide such content truthfully. I believe that such mechanisms are of fundamental importance for the future use of reputation forums, sensor nets and crowdsourcing applications on the internet. They also have other applications in multi-agent systems, such as service monitoring.

While work so far has shown an interesting range of mechanisms to encourage truthful reporting, many open questions remain. The biggest issue is clearly the dependence on knowledge of prior probability distributions that are not always available. The opinion poll mechanism we described is a first step but still has to be generalized to elicit more complex information than just binary signals. Also, as it stands it has little protection against collusive behavior.

Another issue is how to provide rewards. Paying monetary rewards is often not practical, and one needs to experiment with other forms of rewards, such as reputation or privileges that will be valued in similar ways as money.

Acknowledgements

I thank Radu Jurca who has worked with me on this topic for many years, and Karl Aberer for fruitful discussions on reputation and community sensing.

This work has been supported in part by Opensense project (839-401) in the Nanotera.ch program.

References

- [Aberer et al., 2010] Aberer, K., Sathe, S., Chakraborty, D., Martinoli, A., Barrenetxea, G., Faltings, B., and Thiele, L. (2010). Opensense: Open community driven sensing of environment. In *ACM SIGSPATIAL International Workshop on GeoStreaming (IWGS)*.

- [Hu et al., 2006] Hu, N., Pavlou, P. A., and Zhang, J. (2006). Can online reviews reveal a product's true quality?: empirical findings and analytical modeling of online word-of-mouth communication. In Feigenbaum, J., Chuang, J. C.-I., and Pennock, D. M., editors, *ACM Conference on Electronic Commerce*, pages 324–330. ACM.
- [Jurca et al., 2007] Jurca, R., Binder, W., and Faltings, B. (2007). Reliable qos monitoring based on client feedback. In *Proceedings of the 16th International World Wide Web Conference (WWW07)*, pages 1003–1011, Banff, Canada.
- [Jurca and Faltings, 2007a] Jurca, R. and Faltings, B. (2007a). Collusion resistant, incentive compatible feedback payments. In *Proceedings of the ACM Conference on Electronic Commerce (EC'07)*, pages 200–209.
- [Jurca and Faltings, 2007b] Jurca, R. and Faltings, B. (2007b). Robust incentive-compatible feedback payments. In Fasli, M. and Shehory, O., editors, *Agent-Mediated Electronic Commerce*, volume LNAI 4452, pages 204–218. Springer-Verlag, Berlin Heidelberg.
- [Jurca and Faltings, 2008] Jurca, R. and Faltings, B. (2008). Incentives for expressing opinions in online polls. In *Proceedings of the 2008 ACM Conference on Electronic Commerce*, pages 119–128. ACM.
- [Jurca and Faltings, 2009] Jurca, R. and Faltings, B. (2009). Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research (JAIR)*, 34:209–253.
- [Lambert and Shoham, 2009] Lambert, N. S. and Shoham, Y. (2009). Eliciting truthful answers to multiple-choice questions. In Chuang, J., Fortnow, L., and Pu, P., editors, *ACM Conference on Electronic Commerce*, pages 109–118. ACM.
- [Miller et al., 2005] Miller, N., Resnick, P., and Zeckhauser, R. (2005). Eliciting honest feedback: The peer prediction method. *Management Science*, 51(9):1359–1373.
- [Savage, 1971] Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):pp. 783–801.
- [von Ahn and Dabbish, 2004] von Ahn, L. and Dabbish, L. (2004). Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '04*, pages 319–326, New York, NY, USA. ACM.
- [Witkowski, 2009] Witkowski, J. (2009). Eliciting honest reputation feedback in a markov setting. In Boutilier, C., editor, *IJCAI*, pages 330–335.