Distributed Reputation Systems Using Blockchain Records

Kate Sills
Twitter: @kate_sills, katelynsills@gmail.com

> *"One of the major benefits of the Internet is that it enables potentially beneficial interactions, both commercial and noncommercial, between people, organizations, or computers that do not share any other common context." (Nisan 2007, 677)*

The Problem: Opportunism Among Strangers

One often proposed use case for smart contracts is that they can be used to create "smart bonds"—loans where enforcement is automated and guaranteed. As Gideon Greenspan points out, this use case is currently impossible: "If the funds used for coupon payments are controlled by the bond's smart contract, then those payments can indeed be guaranteed. But this also means those funds cannot be used by the bond issuer for anything else. And if those funds aren't under the control of the smart contract, then there is no way in which payment can be guaranteed. In other words, a smart bond is either pointless for the issuer, or pointless for the investor" (Greenspan 2016).

This paper proposes that a distributed reputation system is the solution to problems of opportunism where smart contracts alone are insufficient. Decentralization will be essential, as centralized reputation systems have significant opportunity for coercion and abuse as exemplified by China's social reputation system. (China's systems have been described as "basically a big data gamified version of the Communist Party's surveillance methods [Botsman 2017]).

In *Order without Law: How Neighbors Settle Disputes*, Yale Law professor Robert C. Ellickson observes how cattle ranchers in rural California deal with issues of broken fences and wayward cattle. Rather than following legal rules, their behavior is governed by norms about what it means to be a good neighbor. Ellickson's hypothesis is that welfare-maximizing norms, such as those used by the ranchers, can emerge in a close-knit community, which he defines as a "social network whose members have credible and reciprocal prospects for the application of power against one another and a good supply of information on past and present internal events." He states that close-knit communities do not necessarily need to be small—they only need certain characteristics: high quality of information, reciprocal power, and ease of enforcement. Furthermore, the community must have foundational "protection of a member's bodily integrity and also guarantee that a member has a way of acquiring and holding personal property" (Ellickson 1991).

A blockchain provides for all of these requirements, sometimes incidentally. For instance, information about previous transactions is publicly available (a rarity in real life), bodily protection is ensured incidentally due to the virtual nature of the internet, and public key encryption and the security of the ledger ensure that an entity can acquire and hold personal property.

Measuring Success

The success of a reputation system can be defined as welfare maximizing, that is, minimizing “the members’ objective sum of (1) transaction costs and (2) deadweight losses arising from failures to exploit potential gains from trade” (Ellickson 1991, 184). While theoretically useful, this definition unfortunately attempts to objectively assess something which is naturally subjective. Therefore, the main measure of an operational reputation system should be whether people voluntarily decide to use it. If so, it would follow that value has been added, either by lowering transaction costs or enabling trade that hadn’t been possible before. Additionally, a backwards-looking measure of success might be the predictive power of the model—did an agent defect when cooperation was predicted, and vice versa.

Characteristics of Potential Solutions

While a full solution to this problem is still unknown, I propose that solutions are likely to have the following characteristics:

1. **Reputational information should be a second-party good.**

By this, I mean that a reputation should not be owned by the first-party (the entity themselves) or a third-party (such as the government.) A reputation depends on the outcomes of previous transactions, information should only be shared with the consent of at least one person involved. When Alice wants to make a promise with Bob, she should be able to independently ask the other parties with previous transactions with Bob for information, even *against Bob’s wishes*. This is contrary to a self-sovereign view of reputation in which Bob has the final say over what information can be released. If Bob was able to pick and choose which parts of his history to reveal, a problem of adverse selection would arise in which Bob would be able to hide his history and disguise himself among newcomers with little or no history.

2. **Interpretation of the outcome as cooperation or defection could be baked into smart contracts.**

In order to make a reputational system possible, transactions need to be graded as positive or negative experiences. This can be achieved through reviews like those on Yelp, but smart contracts allow us to more clearly define bad behavior—was money sent to this contract by this time or not? Did the oracle note receipt of a real-world good or not? Smart contracts allow the parties themselves to easily turn qualitative assessments into quantitative ones. Indeed, they could even include a function specifically for external observers to call, which outputs a score of how each party performed in the contract.

3. **Because the system is decentralized, the user can pick and choose which signals to listen to in determining who to trust.**

Unlike the Chinese government's reputation systems, in a distributed second-party reputation system, users are able to choose what matters to them. Instead of valuing whether someone is properly submissive to authoritarians, for example, a user could instead value altruism, timeliness, or any number of features. This ability to select the basis of your own reputation system can be seen as a demand for diverse reputational information, whereas the fact that smart contracts can define good performance themselves, as observed earlier, can be seen as supply. Different communities could define certain desired values and ways of measuring. Thus, there's no significant need for agreement on values between communities if there is no interaction between them, which lessens the amount of infighting compared to a centralized solution.

4. **The system must be sybil-proof.**

The addition of new sock-puppet accounts controlled by the same entity should not affect the outcome of the reputation system. Potentially, "sybil-proofing" could be a distinct step from the reputation algorithm, and can even done on a centralized basis.

Next Steps

In a chapter on "Manipulation-Resistant Reputation Systems" in *Algorithmic Game Theory* (Nisan 2007), Eric Friedman, Paul Resnick, and Rahul Sami outline a number of centralized reputation systems such as PageRank and Pathrank, and very briefly describe decentralized solutions. The chapter also provides references of previous work on reputation from both economic and computer science backgrounds. Next steps would include evaluating these solutions to see whether they fit the characteristics outlined above.

Bibliography

Botsman, Rachel. "Big data meets Big Brother as China moves to rate its citizens." WIRED. November 28, 2017. Accessed February 05, 2018. https://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion.

Ellickson, Robert C. *Order Without Law: How Neighbors Settle Disputes.* Cambridge, MA: Harvard University Press, 1991.

Greenspan, Gideon. "Why Many Smart Contract Use Cases Are Simply Impossible." CoinDesk. April 18, 2016. Accessed February 12, 2018. https://www.coindesk.com/three-smart-contract-misconceptions/.

Nisan, Noam, Tim Roughgarden, Éva Tardos, Vijay V. Vazirani, and Christos H. Papadimitriou. *Algorithmic Game Theory.* New York: Cambridge University Press, 2008.