

DRAFT

Reputational Systems

One of the most valuable aspects of the Internet is the potential to reach a huge audience with minimal infrastructure investment. However, in many cases this scalability comes with an asymmetry: although any processes that can be automated scale easily, any concomitant human involvement must scale at the same rate and thus becomes very expensive, very quickly. It's easy to forget that originally Yahoo was purely a human-edited directory to the Internet. Obviously that approach didn't last long.

The scalability imperative has led to a quest...really many quests...to automate processes. Thus search engines have quickly eclipsed edited directories. But a lot of processes requiring human judgment and experience, often involving finding the wheat among the vast chaff of the Internet, have proven and will continue to prove stubbornly resistant to algorithmic replacement, and humans are still needed. However, there is an economical solution to this dilemma found when one realizes that scale is two-sided: an increase in demand usually means an increase in audience, and audiences are of course composed of...humans.

When computers cannot replace people, the Internet can still help by harnessing the power of mass participation. Peer-to-peer (P2P) technology was once defined as “an aggregation of underutilized resources at the network edge”. Although this phrase naturally brings to mind computer hardware, the largest “underutilized resource at the network edge” is not composed of machines, but of the people sitting at them. Thus the smartest companies are harnessing the efforts of their audiences, rather than just computers and employees, to automate their processes.

Here are some examples:

At Amazon.com anyone can submit book reviews, and regular contributors rise up the ranks of “top reviewers” based on the feedback of readers who, in turn, learn to use the reviewer's status as an indication of trustworthiness.

At eBay both parties in a transaction are encouraged to rate each other, and flawless eBay ratings are so valuable that criminals try to hijack the accounts of honest members in order to pose as honest merchants.

At Slashdot.org readers both submit and discuss news items and then actively rate each others contributions; not only do the highest rated comments bubble to the top, but their authors also accumulate “karma points” which subsequently increase the weight of both their comments and votes.

In each of these examples the participants perform valuable functions that could not be done by software algorithms, and all three are archetypes of what we refer to as a “re-

DRAFT

putational system” because the quality of the participants’ efforts, as rated by their peers, determines their reputation, which in turn affects the weight of their contribution. The magic of this approach is the **built-in incentive of reputation itself**. That is, not only is reputation a useful mechanism for filtering content, but status and peer recognition alone are often sufficient rewards to induce participation. So the approach not only solves the problem of subjective evaluation but is also its own self-reinforcing motivator. Automation, indeed.

Also note that reputational systems don’t need to be obvious or overt. Despite the examples used in this paper, reputation doesn’t have to be something is consciously seen or used by human participants. Google is one simple example: page rank (a kind of reputation) is determined by links, so when someone links to a page they are “voting” for it, albeit not intentionally. In fact, systems like this that use participants’ normal actions to determine reputation are probably more reliable than those that rely on explicit voting.

Characteristics of a Reputational System

Each of the previous examples exhibits three characteristics that are necessary (and perhaps defining) of a reputational system:

1. **Participation**: members of the community are active contributors of the core, valued content of the system.
2. **Peer Evaluation**: participants rate the contributions of others, and these ratings are recorded, aggregated, and displayed as the participant’s “reputation”.
3. **Persistent Identity**: participants maintain consistent (although perhaps pseudonymous) identities.

Beyond being mere requirements or prerequisites, the presence of these three characteristics may lead (inevitably?) to the natural emergence of a reputational system. Let’s look at each of the characteristics individually:

Participation

Beyond just giving ratings, participation must involve actual contribution. This can take many forms, such as the book reviews, auction participation, and comments/submissions in the above three examples. Note that at Slashdot and eBay these contributions essentially *are* the site, whereas at Amazon the reviews are only supporting the primary content of the site, the books. Amazon could exist without user participation; eBay and Slashdot could not.

In both cases, however, users are not just evaluators but active contributors; they are making their efforts visible on the Internet and thus opening themselves up to both criticism and praise by others.

DRAFT

Also note that not all users are required to be active participants. Probably only a fraction of Amazon users rate reviews, and an even smaller fraction write them, but everybody can read them.

Peer Evaluation

In the absence of peer evaluation reputations could (and probably would) still exist within a community, for example in the world of Usenet bulletin boards where some users are recognized as authoritative, but in such cases a member only knows about individuals with whom he or she has interacted in the past, and newcomers arrive with no conception of the local pecking order. Thus the size of a community is constrained by the inability of humans to track an unlimited number of relationships.

In a community with a formalized reputation system, however, any member can immediately know another member's reputation. Prior interactions are unnecessary, and thus the scale of the community effectively unconstrained.

Generally the actual evaluations are of contributions, not of the individuals themselves. The reputation emerges from the quality of an individual's participation. You are the sum of your actions, at least in the eyes of your peers.

Persistent Identity

An essential component of a reputational system is that individuals are identifiable across multiple interactions. Pseudonymity (in which a participant's true identity is hidden) is fine, as long as the pseudonym remains constant, but true anonymity must be minimized. Anonymous contributions to Slashdot are allowed, for example, but all are tagged with the username "Anonymous Coward" as an indication of the community's general sentiment toward the practice.

Most successful implementations encourage long-term consistent use of identities by making reputation something that has been earned over time, making participants loathe to start over from scratch. Reputation thus becomes a valued asset, and reinforces good behavior.

Implementation

Unfortunately there is no generic platform for reputational systems; each of the examples mentioned here were custom built. There are many community and bulletin board software packages available, both commercially and open source, but none of them are focused on reputation. A general-purpose reputation engine would lend itself well to the open source approach (i.e. broadly applicable but non-differentiating layer) but until then any implementations will have to be home grown.

DRAFT

The first question for any potential implementor, of course, is whether a reputational system is appropriate. The general answer is that any situation where a community's feedback could help solve an automation challenge is a candidate, but here are the specific criteria:

Large Audience: "Large" is, of course, a subjective term, but the group must be big enough that if only a fraction participate there are still enough people to average out anomalous results, vandalism, and gaming attempts.

Repeated Interactions: For reputations to work, people must have more than one interaction. In other words, you need to have a "community" so that people will care about what other people think. Sites with a high rate of unique visitors will have a hard time building reputational systems, even if daily traffic is high. So, for example, it might be hard to build community among buyers and sellers of homes since most people engage in those transactions relatively infrequently, but a community of real estate brokers, who deal with real estate on a daily basis, might be more successful.

Signal-to-Noise Challenge: What reputational systems do best is find the good amidst the mediocre, particularly when the difference is a matter of judgment or human expertise. So systems that contain or generate a lot of information of which only a small portion is truly valuable are candidates for this approach.

As mentioned previously, in many cases status and peer recognition alone are enough of an incentive to drive participation. An early reputational system (pre-Internet, in fact) is a copier repair database built by Xerox that tracked the usefulness of its contents. What might have been resented extra work became instead an opportunity for recognition, and some technicians effectively competed to contribute the best data. Again, success required enthusiasm from only some of the participants.

Status can be instantiated to some extent by giving high-reputation members extra rights and privileges. Basic administrator rights (e.g. the ability to alter/delete contributions of junior members) both reinforces high status and further reduces the workload of paid employees.

In many situations, however, status in itself is either insufficient incentive and must be augmented with something more tangible such as prizes or, in cases like eBay, the nature of the system itself attaches monetary value to high reputation. The network of inter-linked blogs (called the "Blogosphere") is an emergent reputational system in which popularity drives monetization through advertising. Google works in a similar way: a page's rank is determined largely by the numbers of links to it, and again there is a significant monetary advantage to having a high Google rating.

The more tangible the rewards of high reputation, the more effort will be put into gaming the system. Google fights a constant war against gamers who try to outwit Google's bots. In addition to dealing with outright identity theft, eBay also wrestles with

DRAFT

gaming attempts. There is no magic bullet to dealing with this problem, but implementors of reputational systems need to recognize that as rewards become more tangible gaming attempts become more determined and require greater vigilance.

Of course, fraud detection is really just a matter of accuracy in how we compute reputation, which turns out to be hard even in the absence of fraud. But this first tentative step into the formulae of reputation immediately raises questions: should the votes of high-reputation individuals carry more weight? Should more recent votes count more than older votes? Should voting itself be rewarded? How about frequency of participation/contribution? Should overly positive or negative voters be normalized to the mean?

An even bigger question is transparency. Since reward only reinforces behavior if two are connected, do we openly describe our algorithms, or do we keep them opaque as obstacles to fraud? Amazon and eBay are again good case studies, for at first glance they seem to simply count up votes. But there is probably something more sophisticated going on under the hood. The right answer is a balance: make it obvious what kinds of behaviors are rewarded, but keep the detailed inner workings secret.