HOW GOOGLE GOOGLES: REPUTATION ON THE INTERNET

Andrew Clausen [1]

[1] Department of Computer Science & Software Engineering,
The University of Melbourne

EXTENDED TUTORIAL ABSTRACT

The anonymity possible on the internet poses new challenges for trust and reputation in online social interaction and commerce. However, the masses of information available on the internet such as hyperlinks, emails and recommendations present a huge opportunity for obtaining reputation information and enforcing accountability. Therefore, the rules of the game of trust, fraud and deception have changed in the online world, and it's not clear where the cards will fall.

If fraud and deception are to be resisted, then it is crucial that we are able to utilize this mass of information. This will be difficult for several reasons. Firstly, the sheer size of the information makes it too much for a human to process alone. Secondly, much of the information will be unreliable, as attackers attempt to manipulate those using it.

Google's PageRank algorithm assigns a reputation score to every web page it can find on the internet, based on the hyperlink and domain name structure. This algorithm has been quite successful at picking the reputable rather than "spam" websites for users searching the internet. We will discuss in detail how and why it works.

eBay has to solve a much more difficult problem. The payoff for deception in online commerce is much greater than for spam websites. Also, all parties to a transaction must trust each other. eBay derives its security from the credit card system. eBay requires that all recommendations must be associated with a transaction. Since it requires some work to obtain a credit card (even by identity theft), this is reasonably effective at preventing deception.

In this tutorial there will be detailed discussion about how algorithms and the communities in which they are applied interact. Attacks such as Sybil attack, identity theft, bribes and tit-for-tat counter-complaints pan out in these different online environments. Finally, there will be discussion about traditional trust and reputation, such as secured loans, credit rating agencies and small worlds theory.

There will be some mathematical content, but no mathematics background is required.