

# Cover Page

**Title of submission:** Avoiding the Prisoner's Dilemma of the Web

**Category of submission:** Conceptual Sketch

**Name and full contact address (surface, fax, email) of the individual responsible for submitting and receiving inquiries about the submission:** Contact Pete Mortensen, Jump Associates LLC, 101 S. Ellsworth Ave. Ste. 600, San Mateo, 94401, USA, +1 650.373.7240, peterm@jumpassociates.com

---

# Avoiding the Prisoner's Dilemma of the Web

**Peter Mortensen**

Communications Lead  
Jump Associates LLC  
101 S. Ellsworth Ave. Ste. 600  
San Mateo, CA 94401  
peterm@jumpassociates.com

**Conrad Wai**

Project Lead  
Jump Associates LLC  
101 S. Ellsworth Ave. Ste. 600  
San Mateo, CA 94401  
conrad@jumpassociates.com

**Abstract**

Everyone talks about increasing safety, security and privacy on the web. But in spite of decades of work to achieve these ends, people still find it hard to know which individuals they meet on the Internet they can trust.

Worse still, many sites, including classified ad giant Craigslist, only function when both parties act honestly, an outcome Game Theory has shown to be least likely in such a situation. If the web is to reach its maximum utility, this Prisoner's Dilemma of the Web must be resolved.

In this paper, we detail the causes – and existing remedies for – distrust between individuals on the web. By outlining seven strategies for fostering peer-to-peer trust, we set up a scenario that combines a few of these strategies with Craigslist to create a theoretical model for breaking the web's Prisoner's Dilemma.

**Keywords**

Trust, security, individuals, transactions, strategy, prisoner's dilemma, utility

**Project/problem statement**

Everyone talks about increasing safety, security, and privacy on the web. From day one, talk about the World Wide Web has centered on how to make it safer. Concerns about identity theft, incidental disclosure of private information and computer viruses and hacking

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Copyright © 2007 AIGA | The professional association for design.

dominate this discourse. Over time, things have changed for the better. Entering a credit card number on the site of an online-only merchant like Amazon has transformed from a terrifying fringe activity into a dominant way to do routine holiday shopping. Information sources like the New York Times online carry as much credibility as their offline counterparts. But despite decades of brand-building, as well as research into antiviral software, secure forms of communication, increased penalties for cyber-crimes and other efforts, the web still feels utterly wild in one particular area: After all these years, people still find it hard to know which online individuals they can trust.

**Wikipedia** is a website that attempts to aggregate all relevant knowledge in the entire world through the contributions of thousands of individual users. As users challenge existing articles, they are revised to reflect other opinions.

**Craigslist** is a website comprising dozens of locally focused web pages that provide free classified advertising for everything from apartment rentals to used car sales. It functions much like classified ad departments in newspapers of yore.

This adds up to real trouble for sites like Craigslist and Wikipedia, which depend on people behaving well to function. After all, if people suddenly decided that answering a Craigslist ad in person was an invitation to get robbed, then the company stands to lose much of its momentum. Until Wikipedia can verify all its information with expert references, the invaluable Internet encyclopedia will be eyed with suspicion in academic circles. When no one's identity is a given, the benefits of certain kinds of crowd-sourced or crowd-driven sites are severely restricted. And no amount of research into more powerful encryption or proposals for stiffer penalties for cyber-fraud will ever make it easier to believe that a stranger is who she says she is. In this paper, we will highlight why individuals are much more difficult to trust than institutions, define a set of possible strategies for forging trust between individuals and finally sketch a scenario that upgrades an existing classified ad website.

## Background

This project was conducted by Peter Mortensen, first author, and Conrad Wai, second author and illustrator.

From May 31 to July 13, we set out to explore ways of fostering openness and trust on the web. Most methods for creating trust on the web have focused on validation through external authorities. The creation of security certificate authorities in the 1990s sparked the e-commerce revolution – and the first dotcom bubble. The rapid rise of bookseller Amazon.com, computer-maker Dell and thousands of also-rans grew directly out of the invention of secure transactions. Over time, most people with Internet access have come to view credit card purchases over the network as routine. Along with this technical solution, researchers at Stanford [1] have identified 10 strategies for building an institution's online credibility, emphasizing demonstrations of verifiable information, including specific contact details, third-party references and others.

Perhaps ironically, many of the same strategies that best build group credibility online have either neutral or negative effects on individual credibility. For example, if someone started a web site with the URL Stevejobs.com, filled it with absolutely true information sourced from credible off-site sources and then provided the address and phone number for Apple headquarters, no one would believe it to be the personal site of the company's CEO, Steve Jobs. The reason for this is that it is much easier to pose as someone else on the Internet than it is in person. Because everyone knows this, people eye anyone with suspicion who claims to use their real name and publish full contact information online.

## Challenge

Any person hoping to transform a virtual relationship into a visceral one faces a Prisoner's Dilemma [2]. In this classic game theory scenario, two suspects are arrested by police for the same crime, but investigators have insufficient evidence for a conviction of more weight than a light sentence of six month for each. Both get offered the same deal: If one will testify and the other remains silent, the state's witness goes free while the quiet suspect would get 10 years in prison. If both betray each other, however, each gets five years in prison. The vast majority of the time, both prisoners will choose to betray his counterpart, because testifying leads to a better outcome for him regardless of whether the other prisoner testifies or stays silent. The win/win situation almost never arises.

This has tremendous implications for interactions between individual strangers on the Internet, where actual identities play a critical role. That's because such relationships can only function properly when both parties decide to pursue the win/win scenario in the Prisoner's Dilemma and provide authentic information about themselves – by far the least likely outcome in the game. Craigslist.org provides a wonderful forum to buy, sell and rent everything you need.

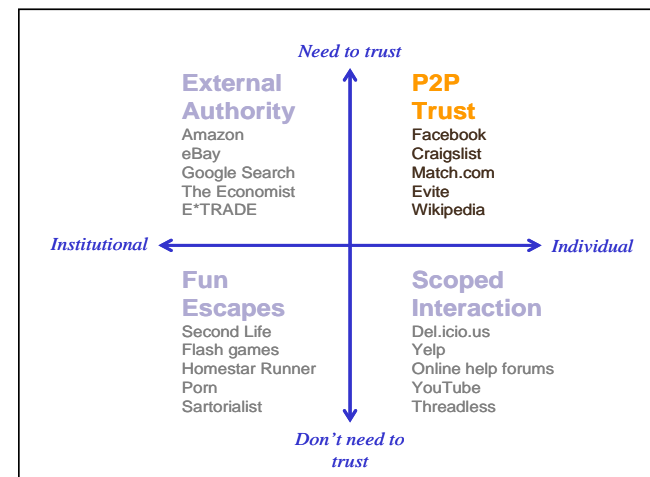
But the system breaks down if either the buyer or seller lies about themselves. That happened in 2006, when an armed robber in Silicon Valley, California [3], arranged to sell a laptop to three different buyers, only to incapacitate them with a taser and rob their homes. The same summer saw many violent robberies associated with Craigslist across the country. How could this have been prevented? Though Craigslist requires people to exchange contact information to make deals

in person, nothing stops unethical people from providing false information to exploit the trusting – while the site assumes no liability.

## Solution

A. Process:

The web doesn't need to engender much trust to be useful. Fun escapes and pointed information/ content fulfill legitimate needs. But the creation of institutional trust has allowed the web's utility to grow by leaps and bounds, opening up e-commerce and whole new levels of information dissemination. By sorting web sites based on the importance of trust in institutions and individuals, the authors uncovered an opportunity for growth in sites that foster individual trust (Fig. 1). Connecting with other people on a deep level will bring us closer to fulfilling on the web's promise as a social medium. But how do we achieve Individual Trust?



**Figure 1.** Authors sorted a variety of Internet sites based on the necessity of trust to their function.

## B. Solution Details:

Through analyzing the spectrum of websites hoping to foster commercial activity or communication between individuals over the Internet, recurrent patterns have emerged. The following seven strategies comprise the current strategies for building personal trust on the web.

**Match.com** is a dating website for singles on the Internet.

**MySpace** is a wildly popular social networking site.

**LinkedIn** is a networking site for professionals, who can recommend job candidates and reconnect with colleagues.

**Facebook** is a social network site begun at Harvard University that originally only allowed college students to join.

**eBay** is the most popular online auctions site on the Internet.

**1.** In a normal conversation, people learn to trust each other as they learn new, more intimate information over the course of time through a process of **Mutual Incremental Disclosure**. The same holds true on the Internet. For example, at Match.com, people scan widely for people they might be interested in. Through a series of “winks” and then e-mails, users can hold more extended conversations until each is ready to meet with one another via phone call or home address exchange. Many long-running message board communities rely on similar practices to initiate members. Current implementations of this strategy do not require people to provide accurate information.

**2.** Many sites and individuals verify who they are by moving information from unverified communication sources to known direct communication touchpoints, a **Second Channel**. On a personal level, it could mean using a web-only message system (such as MySpace messaging) to request that an e-mail be sent from an address only the real individual could have access to. When the recipient replies from that second address, the loop is relatively secure to verify. This second, privileged channel proves that the user is honest.

**3.** Many sites where people provide each other with recommendations or services provide venues for

feedback that allow them to keep each other honest, **Individual Endorsement**. At Linked In, for example, people ask their friends and professional networks who they should hire for particular positions. Since an individual stakes her reputation on the referral, they have a strong incentive to only promote people they can really vouch for and truly believe to be competent.

**4.** Many individuals develop their credibility on the Internet by affiliating themselves with larger organizations known to be trustworthy. This **Institutional Reputation** helps, for example, journalists that blog on their publication's websites audiences quickly both because of the exposure and because readers can tell right off the bat that they're reading the work of the same reporter whose columns they enjoy in the paper. Interestingly, Facebook used to follow this method closely, requiring all its users to provide an official university e-mail to register, thereby tying a user name to an existing, officially verified identity. Since the site opened up to the public, this relative assurance has been diminished.

**5.** In many cases, endorsements of individuals come neither from institutions nor from previously trusted acquaintances. Instead, a large number of people with limited knowledge of a person or his work provide aggregated evaluations that can help new users make their own decisions, also known as **Wisdom of Crowds** [4]. eBay and Wikipedia both rely on this tactic. For eBay, a high positive feedback rating makes both buyers and sellers more reputable, while Wikipedia presumes that a large enough number of people reviewing articles will create something factual.

**Flickr** is a site where people can share their photos with either the public or a small group of private users.

6. It's easier to recognize friends from ordinary life online when they populate their profiles with recent photos, providing **Visual Verification**. Combining a user name suspected of being that of a friend's with a set of non-commercial photos that depicts her goes a long way to prove an identity. Flickr, Facebook and MySpace all benefit from this strategy.

7. Many websites connecting individuals function on the operating principle that most people are basically decent, a **Faith in Humanity** system. Craigslist can't work at all unless both buyer and seller provide authentic information and supply goods that are up to snuff. Unfortunately, this faith is easy to undermine, as shown in the Craigslist robber case of 2006.

C. Results:

While multiple websites already rely on trust between individuals to support their businesses, no one has yet created an individual trust solution that is as robustly foolproof as secure transactions are for institutions. We believe that the answer to these challenges exists within the seven strategies outlined previously. Not in any one in particular, but in mashing them up together to create personal profiles people can believe in.

To a certain extent, these tactical mash-ups are already happening. Facebook, at the time it launched, had a fairly rich solution that relied on Institutional Reputation (a college e-mail address), Visual Verification (profile photos and personal galleries) and Individual Endorsement (only friends can view profiles or see other users' friend lists). As a result, the site grew a loyal following and avoided the case cultivated by similar sites such as social network MySpace.

One possible mash-up that might prove powerful in the very near future is what might be called **Secure Disclosure**, a combination of Institutional Reputation, Second Channel and Mutual Incremental Disclosure. The process blends the sensed intimacy of learning the details of a person's life with the verifiability of a business online. The following illustrative scenario describes how it might work on Craigslist, a site this is particularly vulnerable if a large number of people intent on theft began to congregate there.

*Craigslist Plus offers the piece of mind that comes with knowing exactly who you're buying from and selling to. A 27-year-old lawyer, Tony, decides to register for the service, concerned as he is about his personal safety. He heads to the Craigslist Plus site, enters a major credit card number and his social security number. His credit card company links the address and phone numbers it has on file for him to his new account. This information, however, never passes into the hands of Craigslist itself. Tony's account has no associated profile for other users to read, but he is granted immediate access to the Craigslist Plus network. He finds a couch he'd like to buy, a great deal.*

*Initial contact between buyer and seller occurs exactly as it does now, as each person asks questions via anonymized e-mail to determine if they want to do business. However, the answers to these questions can be verified as never before. Rather than sending an e-mail claiming he lives in the Noe Valley neighborhood in San Francisco, Tony clicks the NEIGHBORHOOD button in his Craigslist Plus dashboard. This action triggers a notification, the couch's seller if she would like to exchange the location of her neighborhood in return for Tony's. When the seller, a 36-year-old software*

*developer in Palo Alto named Frances, clicks her NEIGHBORHOOD button, both users simultaneously receive the information, which is transmitted directly from credit card company records to secure Craigslist pages, eliminating opportunities for tampering or deception.*

*This process continues, with each requesting and exchanging more information simultaneously. Eventually, each user clicks the button to set up a meeting, arranging for pick up at Frances' home. Tony and Frances exchange addresses and phone numbers – again, from credit card records, not personal entry – and a smooth couch transaction occurs the following Saturday.*

This scenario is simply meant to illustrate the potential power granted to individual users by allowing them to disclose information about themselves that has already been vetted by credible institutions. By taking manual entry out of the equation, efforts to prove identity are much more effective than any number of measures initiated independently. It's quite easy to imagine a system standardized across the web that would allow people to transmit their verified identities on multiple networks, continually pulling information from the trusted third-party. This system would be no more inherently secure than online transactions are now, but merely by bringing individuals up to this same level would be a step in the right direction for all of us. For the future, we need to test and implement additional mash-ups of individual trust strategies.

## References

- [1] Kopytoff, Verne. Craigslist users hit by robbers. San Francisco Chronicle (San Francisco, July 2006)
- [2] Poundstone, W. Prisoner's Dilemma. Doubleday (New York, 1992)
- [3] Fogg, B.J. Stanford Guidelines for Web Credibility. A Research Summary from the Stanford Persuasive Tech Lab. Stanford University. (Palo Alto, May 2002) [Webcredibility.org/guidelines](http://webcredibility.org/guidelines)
- [4] Surowiecki, James. The Wisdom of Crowds. Anchor Books (New York, 2005)

## Acknowledgements

Thanks to Fake Leander, Fake Steve, Callously and Evil Doug Coupland. You're an inspiration to us all.