Pseudonym Parties: An Offline Foundation for Online Accountability (PRELIMINARY DRAFT)

Bryan Ford
Massachusetts Institute of Technology

March 27, 2007

Abstract

Many unsolved Internet security vulnerabilities reduce to a lack of *user accountability*: any user who misbehaves—e.g., by spamming from a free E-mail account or stuffing an online ballot box—can simply open other anonymous accounts or connect from other IP addresses. The obvious solution of requiring all users to identify and authenticate themselves to online services, through a universal public-key infrastructure (PKI) for example, is inconvenient and impractical to deploy universally, and raises serious privacy concerns. Ensuring accountability does not in general require *identifying* users, however: it only requires enforcing a principle of *one person, one persona* for a given online service. This paper proposes *pseudonym parties*, a decentralized scheme that combines technical tools (pseudonymous online accounts) with in-person social occasions (parties) to provide online accountability while preserving the ability of users to participate anonymously in online services. This approach is fully decentralized, can be deployed incrementally at minimal cost, and may even be fun to participate in.

1 Introduction

Today's online ecosphere continually suffers from its inability to tell who is a genuine, unique human and who is not. Because open-access messaging systems cannot isolate or authenticate the *human* source of messages for the purpose of suppressing abuses, spam has already relegated USENET to historical obscurity [15], threatens the usability of E-mail [18], and is even advancing on popular Voice-over-IP systems.[1] The automated "Turing tests" many web sites now employ [17] lock out visually impaired users [4, 11] and are vulnerable to attack using artificial intelligence [3] or social engineering [5]. Wikipedia progressively tightens its editing rules to combat the rising tide of anonymous vandalism [10, 8, 16]. Voting-based participatory systems such as Slashdot operate reliably only to the extent that nobody cares about the results of votes enough to bother opening many accounts and stuffing the ballot boxes. Banning detected abusers by IP address frequently prevents access by other legitimate users on the same ISP [9], and many attacks come from compromized zombie machines not under the control of their owners [7]. Forcing all users to register with real identities that can be securely authenticated in some way (e.g., by entering a credit card number) is inconvenient to users, costly to service providers, and raises important privacy concerns.

What most online services and Internet-based communities need, however, is not to determine exactly *who* a given user is in the real world, but merely to ensure that each real, human participant can obtain only one account or persona at once on a particular online service. If services could reliably enforce a *one person, one persona* rule when appropriate without having to obtain and verify personal information, then online personas would no longer be disposable and thus would provide a substantial degree of accountability—without compromising the user's privacy. Online services could temporarily or permanently revoke the access rights of abusers, such as E-mail spammers or Wikipedia vandals, without affecting innocent users or permitting the same abuser to reappear immediately under a different name. Voting-based systems for peer review or online democratic deliberation could protect voting anonymity while preventing ballot box stuffing.

Pseudonym parties provide an alternative method of enforcing the *one person, one persona* rule, taking advantage of the fact that real humans can only be in one place

[1] Try googling for '`skype spam`'.

at a time. On a specific day every year, participating organizations or ad hoc groups of people host parties in their local areas at which they pass out a set of pseudonymous online credentials to anyone who shows up in person. The physical presence requirement, combined with suitable hand-out procedures, ensures that each user can obtain only one such set of credentials per year. Given this set of credentials, a user can create any number of pseudonymous identities at participating online services—but only one identity per service. Using these pseudonymous credentials to obtain accounts on online services, instead of registering anonymously, could offer users with immediate benefits such as special voting privileges and exemption from IP-based blacklists, waiting periods, or other cumbersome verification procedures, in addition to the larger collective benefits to Internet communities.

This paper is organized as follows. Section 2 outlines previously proposed approaches to user accountability in more detail. Section 3 then presents and discusses the proposed scheme, deliberately focusing more on the scheme's social aspects than on its technical details in order to promote discussion. Section 4 briefly points out some important concerns related to deploying and implementing the scheme, and Section 5 concludes.

2 The Human Recognition Problem

The abuse of an online system by creating many illegitimate virtual personas has become known in the peer-to-peer networking community as a *sybil attack* [6], after a famous case of multiple personality disorder [12]. Variants of the sybil attack currently plague many online services, however, not just peer-to-peer networks.

2.1 "Authenticating" on IP Addresses

Although E-mail spam is so pervasive and problematic that it is usually treated separately from other Internet maladies, spam is just a specific variety of sybil attack. Any particular identifiable source of spam, once detected, can be fairly easily blacklisted: the problem is that a single spammer can co-opt thousands of compromised hosts with frequently-changing dynamic IP addresses to create the illusion that the spam is coming from an endless variety of fake (or forged) E-mail identities. The botnet is the spammer's tool to circumvent the weak authentication the Internet provides in the form of source IP addresses. If E-mail messages arriving at a mail server somehow came reliably tagged with an opaque token that was guaranteed to have a one-to-one association with the human user that caused that E-mail to be sent—even if that token provided no information whatsoever about that user that could in any way be used to trace him—countering a detected spammer would be a simple, one-time matter of blacklisting that user's token. To counter most forms of abuse it is not necessary to be able to hunt down the abuser in person; it is merely necessary to be able to associate his future actions reliably with his past actions. But IP addresses reliably offer no such association.

The futility of holding users to account based on their IP address is further illustrated by the frequency of online ballot-stuffing attacks. In November 1999 students at MIT and CMU deployed automatic voting scripts to rig a Slashdot poll about the best CS grad school, then in 2006 MIT students similarly rigged a Doonesbury poll. One of the techniques used was to set up a machine that simply roams among unused IP addresses on the vast Class A network (2^{24} addresses) that MIT was allocated in the Internet's early days, casting one vote from each IP address. A user of America Online or another of the large ISPs that place their customers behind NATs or web proxies [2], on the other hand, may not be so lucky as to be able to cast even one vote, if another user of the same ISP has already cast a vote via the same NAT or proxy. In effect, how many votes you have depends not on how many people you are but on how early your organization joined the Internet.

2.2 Identity-based Authentication

Many proposed solutions to this user accountability problem revolve around authenticating a user's identity in some way before granting (full) access to an online service. Most E-commerce sites already require users to identify themselves by entering personal information and a credit card or even a bank account number. Single sign-on initiatives such as Microsoft Passport, the Liberty Allianc, and OpenID try to centralize a user's personal information at a single "identity provider", which various online services contact to authenticate the user. Pervasively forcing users to reveal their true identities unnecessarily in order to use a service or participate in an online community creates grave privacy concerns, however: anonymous communication and social participation is widely viewed as a crucial tool to safeguard basic democratic values such as privacy and freedom of speech [14], and the rights of minorites [13].

Some E-commerce sites support emerging anonymous payment services such as paysafecard.com, Xrost, and Ukash. Although anonymous payment services may help protect a user's privacy while actually engaging in an on-

line financial transaction, they do not help E-commerce sites that need other forms of user accountability *before* the transaction—such as when offering a promotional deal like "get the first 3 downloads free," without effectively offering unlimited free downloads to any user who persistently signs up for successive free accounts. Some sites require the user to enter a verifiable credit card number merely for identification purposes without charging the card, but this approach once again requires the user to divulge her real identity.

2.3 Slowing Down Attacks

Another commonly proposed defense against the sybil attack is to increase the cost of creating new identities. Such defenses only work to the extent that abusers are actually limited by these artifically imposed costs, however. Online "Turing tests" such as CAPTCHAs [17], even when they work [3], may prevent fully-automated attacks, but cannot protect against a determined user who is simply willing to sit and solve the puzzles repeatedly—in order to stuff an online ballot with tens or hundreds of votes instead of just one, for example. Computational puzzles [1] similarly only slow down attacks by some factor and are easily countered by an abuser with a large botnet.

Clever heuristic algorithms have been proposed that evaluate the shape of a social network graph to limit the extent of the damage any one particular sybil attack may cause [19]. While such an algorithm might be effective against an isolated abuser's attempt to create thousands of fake identities, however, it would probably not be effective against many abusers who create only a few fake identities each, or against coordinated groups of abusers. In an online community that uses voting for decision making or peer review, *everyone* has the incentive to cheat by creating fake identities, and a heuristic algorithm is unlikely to catch a cheater who does not cheat too much more than everyone else cheats. In any case, some attacks require only two online identities: in one abuse known on Wikipedia as *sock puppetry*, a user creates one fake account with which to vandalize random pages, then subsequently uses his *real* account to revert those edits in order to bolster the real account's online "reputation."

3 Pseudonym Parties

While distinguishing one human from another may be fundamentally difficult for computers, humans manage this task all the time when we meet in person. A relatively informal and inexpensive bit of organizational infrastructure in the *offline* world might provide exactly the foundation we need for anonymous but accountable online participation, allowing users to create multiple privacy-preserving online personas while protecting online services against sybil attacks.

Suppose that on some particular day of every year - let's call it *Pseudonym Day* for the moment - groups of people interested in having private but accountable online personas gather somewhere in their nearby physical neighborhood to throw a party. Any group of people may organize such a *pseudonym party*, provided the group follows certain procedures standardized by some broader distributed network of pseudonym party organizers and opens itself to oversight by organizers of other parties to ensure that the required procedures are followed correctly. Outside of this standardized framework, each pseudonym party's organizers are free to run their party as they see fit: e.g., as a festive social occasion in which participants are invited to bring food and drinks to share (hence the term "party"), as a conference or workshop in which to discuss online social issues and the like, or as a purely functional affair minimizing cost and volunteer effort.

Regardless of form, at the heart of each pseudonym party is a procedure in which each participant receives login information granting access to one, and only one, anonymous *pseudonym account* on a designated web site run by the party's organizers or an affiliated organization. Pseudonym accounts store no personal information, and no one needs to show identification or meet any requirements regarding age, citizenship, home location, or other personal characteristics, in order to obtain an account. A New Yorker visiting Paris on Pseudonym Day can just show up at any pseudonym party in Paris to get his yearly account, and since the basic procedure is standardized, he shouldn't even need to know French. The only requirement is to be alive and able to show up and follow the required procedure.

As with elections in countries that have no formal voter registration, each participant gets an indelible ink mark in some obvious place as they are given their pseudonym account, preventing them from obtaining several accounts on the same day (e.g., at different parties in the same area). Lest we worry that people would be reluctant to submit themselves to being marked in this fashion, we observe that bars and nightclubs everywhere—not to mention amusement parks—tend to do exactly the same thing to everyone who enters, and few apparently complain.

3.1 Creating Online Personas

A pseudonym account is not affiliated with any particular online service, but acts as a front-end through which users

create or renew accounts on participating online services. A user might login to his pseudonym account and enter `wikipedia.org`, for example, to create a personal Wikipedia account for himself and automatically log him into it. The pseudonym account server is responsible for enforcing the *one person, one persona* rule: if the user enters `wikipedia.org` again in his pseudonym account, he simply finds himself in his existing Wikipedia account. Since each person can obtain one new pseudonym account per year, online services may expire accounts created this way after a year to prevent users from gradually accumulating many accounts, and may offer users a way to transition smoothly from one year's account to the next, but such policies are specific to the online service.

A pseudonym account holder may create one personal account on each of any number of distinct online services in this way. To protect the user's privacy, the pseudonym account server ensures that online service providers cannot tell which two accounts on different services correspond to the same pseudonym account, and hence the same user, unless the user explicitly gives them personal information establishing such a link. If a user uses his pseudonym account to create both a professional profile for himself on `LinkedIn.com` and a steamy personal profile on `AdultFriendFinder`, for example, no one can tell that the two profiles represent the same person even if the two web sites collude or are hacked—unless, of course, the user gives away the connection, by posting the same photo in both profiles for example. Even then, the user could deny the connection, claiming that someone had simply downloaded his photo from LinkedIn and used it to fabricate an embarrassing profile on AdultFriendFinder, and there would be no way to prove the connection existed short of compromising the pseudonym account server. In effect, the pseudonym account server expressly defends the user's right to exhibit a controlled form of "multiple personality disorder" across different services in order to protect his privacy.

Although we wish to avoid specifying too many technical details at this point in the interest of focusing the discussion for now on higher-level social and usability issues, one specific way the pseudonymous authentication mechanism might operate is as follows. When the user directs the pseudonym account server to log him into a particular online service, the pseudonym server computes and sends to the online service a keyed secure hash of the online service's own public host name (e.g., '`wikipedia.org`'), using as the key a per-user secret that the pseudonym server keeps closely guarded and never reveals to any online service or any user. The online service uses the resulting hash as a token that pseudonymously represents the user's identity, confident that the resulting token will always be the same for a given user and online service hostname, although one user's tokens for different online services appear unrelated because they are hashes of a different hostname. Straightforward extensions to this mechanism could allow online services to differentiate identity contexts at other granularities: e.g., Wikipedia might allow its users to edit different Wiki pages under different identities, while still ensuring that users only adopt one identity *on each page*.

3.2 Coexistence with Other Mechanisms

Pseudonym accounts need not replace existing account creation and login mechanisms, but could serve as a "premium" mechanism offering rewards to users who make the effort to help organize or at least show up at a pseudonym party once a year. E-mail authenticated via a pseudonym account might be exempted from heuristic spam filters, for example, preventing the loss of legitimate E-mail due to false positives. Wikipedia might still allow users to create traditional accounts, but could exempt pseudonym account users from IP address bans or from waiting periods imposed between account creation and editing. (An abusive E-mail or Wikipedia user logged in via a pseudonym account can still be banned by his pseudonym account's pseudonymous identity, of course, and will be unable to create a fresh pseudonym account until the next Pseudonym Day.) Online services that support peer review by voting might require a pseudonym account in order to cast votes, while allowing less sensitive forms of access via a traditional anonymous account or no account.

3.3 Trust and Security

Unlike identity-based single sign-on schemes, users need not trust their pseudonym account provider with their personal information, because they never had to provide any personal information to obtain the account. Users may need to trust the pseudonym party organizers to protect their privacy in the account assignment process: e.g., not to try to identify and keep tabs on who was assigned which account. Users also must trust the party's designated account server to protect the relationship between different online service accounts the user obtains via the same pseudonym account. Each pseudonym party's organizers may run their own server or work with a support organization they consider trustworthy; each user in turn has a free choice of which pseudonym party in their area to attend, and thus which group of organizers to trust.

Online service providers must similarly trust pseudonym party organizers and their designated account servers to enforce the one person, one persona principle. Each service's administrators ultimately choose which pseudonym account servers to trust, but the network of pseudonym party organizers may also need to establish a system of mutual oversight and operational peer review to monitor each party's health and trustworthiness. Initially, each pseudonym party might simply stake its claim to legitimacy on the public, online or offline reputations of its organizers.

4 Deployment

Unlike identity-based single sign-on services or traditional public-key infrastructure (PKI), pseudonym account services do not need to be widely deployed "all at once" before they become useful at all.

Non-profit organizations and special-interest groups that operate primarily within a local geographic region, for example, might initially both run online services of interest to the local public and organize pseudonym parties to provide pseudonymous credentials for accessing their own online services, protecting their own online community forums from abusers both geographically local and remote. Ad hoc groups and organizations might in this way start with a purely local focus and gradually expand the useful geographical scope of the pseudonymous credentials they hand out by federating with other similarly developing groups and organizations in other geographic areas. Ideally a pseudonym account obtained on Pseudonym Day anywhere in the world should eventually be usable to create accountable pseudonymous identities on online services anywhere else in the world, but this long-term ideal need not be achieved all at once.

Popular web sites that represent global participatory communities operating using deliberative democratic procedures, such as Wikipedia and Slashdot, are particularly sensitive to sybil attacks in the form of ballot stuffing or sock puppetry, but these same communities also tend to have many users who are concerned with preserving privacy and the ability to participate anonymously. Since pseudonym parties currently appear to be the only proposed solution that can address both strong accountability and privacy at the same time, these online services could benefit greatly from such a scheme, and might therefore represent a likely context for initial experimentation with and deployment of pseudonym parties.

4.1 Other Implementation Issues

There are of course many additional issues and details to work out in the implementation of such a scheme, whose solutions are left for now as topics for discussion. Here are a few:

- Is there a safe way to give new users "first-time" pseudonym accounts immediately when they learn about the system, without forcing them to wait up to nearly a year until the next Pseudonym Day?
- Should there be "backup" mechanisms to obtain pseudonym accounts in case a person is sick or otherwise immobile on Pseudonym Day, or is at a location where there is not yet any organized pseudonym party?
- Might pseudonym parties be allowed to give users the choice of showing ID and attaching personal information to their pseudonym account, so that they could use the same account for both anonymous and identity-based single sign-on if they wish to?
- What specific software do pseudonym account servers and participating online services need, and what is the protocol by which they interact? Could existing identity-based single sign-on infrastructure be reused and adapted to this purpose?
- To what extent, if any, should pseudonym parties and affiliated organizations be allowed or encouraged to build ties or accept the support of governments or corporations?

5 Conclusion

Combating the sybil attacks at the heart of many online problems such as spam, wiki vandalism, and online ballot box stuffing, need not and should not force us to give up our privacy. Pseudonym parties would protect users' ability to maintain multiple disconnected, potentially anonymous online personas, while ensuring accountability and allowing online services to enforce the democratic "one person, one vote" principle when appropriate.

References

[1] Adam Back. Hashcash — a denial of service counter-measure, August 2002. `http://www.cypherspace.org/adam/hashcash/`.

[2] Martin Casado and Michael J. Freedman. Peering through the shroud: The effect of edge opacity on IP-based client identification. In *4th NSDI*, Cambridge, MA, April 2007.

[3] Kumar Chellapilla, Kevin Larson, Patrice Simard, and Mary Czerwinski. Computers beat humans at single character recognition in reading based human interaction proofs (HIPs). In *2nd Conference on Email and Anti-Spam*, July 2005.

[4] Curtis Chong. Graphical verification: Another accessibility challenge. *The Braille Monitor*, November 2003.

[5] Cory Doctorow. Solving and creating CAPTCHAs with free porn. *Boing Boing*, January 2004.

[6] John R. Douceur. The sybil attack. In *1st International Workshop on Peer-to-Peer Systems*, March 2002.

[7] Joris Evers. ISPs versus the zombies. *CNET News*, July 2005.

[8] Katie Hafner. Growing Wikipedia refines its 'anyone can edit' policy. *New York Times*, June 2006.

[9] Adam Kalsey. Why IP banning is useless, February 2004. `http://kalsey.com/2004/02/why_ip_banning_is_useless`.

[10] Will Knight. Wikipedia tightens editorial rules after complaint. *New Scientist*, December 2005.

[11] Matt May. Inaccessibility of CAPTCHA: alternatives to visual turing tests on the web, November 2005. W3C Working Group Note 23.

[12] Flora Rheta Schreiber. *Sybil: the true story of a woman possessed by sixteen separate personalities*. Warner Books, 1973.

[13] Edward Stein. Queers anonymous: Lesbians, gay men, free speech, and cyberspace. *Harvard Civil Rights-Civil Liberties Law Review*, 38(1), 2003.

[14] Al Teich, Mark S. Frankel, Rob Kling, and Ya ching Lee. Anonymous communication policies for the Internet: Results and recommendations of the aaas conference. *Information Society*, May 1999.

[15] Brad Templeton. I remember USENET. *O'Reilly Network*, December 2001.

[16] Bill Thompson. Not as wiki as it used to be. *BBC News*, August 2006.

[17] Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford. CAPTCHA: using hard AI problems for security. In *Eurocrypt*, 2003.

[18] Paul Wouters. Personal spam statistics 1997-2004, January 2005. `http://www.xtdnet.nl/paul/spam/`.

[19] Haifeng Yu, Michael Kaminsky, Phillip B. Gibbons, and Abraham Flaxman. SybilGuard: defending against sybil attacks via social networks. *SIGCOMM Computer Communications Review*, 36(4):267–278, 2006.