Leveraging Social Networks for Online Identity Verification

AssertID / Joon Nak Choi (Chief Scientist) and Kevin Trilli (Founder/CEO)
http://www.assertid.com

Keywords
identity verification, age verification, identity oracle, social network, COPPA.

Functional Goals
- Limit harmful contact between adults and minors
- Limit/prevent minors from accessing inappropriate content on the Internet
- Prevent harassment, unwanted solicitation, and bullying of minors on the Internet

PROBLEM INTRODUCTION

Online communities face increasing pressure to *limit harmful contact between adults and minors*. The National Center for Missing and Exploited Children found that 13% of *all* children online received sexual solicitations. In a one-third of such cases, the solicitor attempted to meet the child offline. Furthermore, 4% of children online were asked for nude pictures of themselves [1]. Recognizing this problem, forty-nine of the fifty US State Attorney Generals have signed an agreement with online social networks to work together to find a solution for protecting children online. Also, the Federal Trade Commission is monitoring the situation as an extension of its existing role preventing children from accessing adult-oriented websites.

AssertID limits harmful contact between adults and minors by functioning as an *identity oracle*. According to Bob Blakley of Burton Research, an identity oracle *collects* and *warehouses* private information about online individuals. Identity oracles can use this data to answer questions about an individual--without actually disclosing the private data [2]. Consider a *hypothetical* example. A user (OpenID: Kid123) applies to join Club Penguin. To check if Kid123 is really a minor, Club Penguin queries AssertID about Kid123's age. AssertID possesses private information about Kid123--that he is Joe Kidd (age 12) living at 123 Main Street in Boston. AssertID releases Kid123's age to Club Penguin while keeping his other attributes private.

It is important to note that AssertID *verifies* its data. Most online communities today trust their users to tell the truth about themselves--i.e. *self-assert* accurate data about themselves. Yet, many users self-assert false information. Notably, sexual predators often pretend to be minors to gain their intended victims' confidence. To limit such misrepresentations, AssertID using the following logic:

- In the absence of age verification, any user can lie about his (or her) age. Thus, users' *self-asserted* ages cannot be assumed as accurate.
- A typical user is connected with people online who know him offline--real-world friends and colleagues. These people know *something* about his real age.
- If such people *verify* that the user is telling the truth about his age (*vouching* for the user), even outsiders (i.e. strangers) can have greater confidence in the user's self-asserted age.
- Users verified by many other users can be trusted more than users verified by few other users.
- Users verified by other users who *themselves* have been verified can be trusted to an even greater extent; they are verified by others known to be trustworthy.

Age is only one of several socio-demographic attributes verifiable through this logic. Gender, marital status and geographic location can be verified in much the same way.

AssertID provides an easy-to-interpret *score* representing the likelihood that a user is self-asserting his *actual* age. These scores are computed by a patent-pending algorithm based on social network analysis. AssertID enhances the identity oracle concept, by providing not only users' self-asserted ages, but also its *degree of confidence* in this data.

The same approach has at least two other applications. It can *limit/prevent minors from accessing inappropriate web content*. When an online user applies to enter an adult-only website, the site would query AssertID about the user's age. If AssertID is reasonably sure that the user is 18 or over (21 in some jurisdictions), the site grants user access. AssertID can also reduce *online harassment and bullying* using a similar approach. Cyber-bullies gain much of their power by misrepresenting themselves online. If online communities validate users' self-asserted attributes (i.e. age, gender, etc.) using AssertID, bullies will find it much more difficult to misrepresent themselves.

PROPOSED SOLUTION

AssertID has two main components. AssertID's front end functions as an *identity oracle*, collecting and warehousing self-asserted private information about its users. Behind the scenes, AssertID functions as a *web of trust* validating self-asserted information about its users.

AssertID as an Identity Oracle

- Queries users to self-assert socio-demographic attributes, including name, age, gender, marital status, address and picture. This process incorporates trusted registration processes and data sources when available, to more powerfully verify identity attributes.
- Queries users for their online identities, including social network profiles, email addresses, and OpenIDs.

- Binds this data to a unique identifier (primary email) and warehouses this data.
- Users control third-party access to their private data, granting different permission levels for different entities (e.g. online communities, individuals, etc.)
- When queried on the age associated with an online identity, AssertID provides the age, provided that the user has granted permission for this information to be shared with the party asking the question. Note that age is only one of several socio-demographic attributes that can potentially be provided by AssertID.

This is a straightforward implementation of the identity oracle concept described by Blakley [4] and others [5]. However, AssertID provides an additional piece of information making socio-demographic data much more useful--a quantitative score indicating its confidence that the self-asserted age is actually true.

AssertID as a Web of Trust

AssertID obtains such a quantitative score in two stages: (1) building a *web of trust* amongst its users and (2) computing users' embeddedness within the web of trust.

AssertID builds its *web of trust* by asking users to validate each other's self-asserted attributes. Each validation has two consequences. On one hand, the validated person's attributes become more reliable. However, a validation also has second-order effects--positive externalities. Users do not *directly* assess each other's trustworthiness--but end up doing this *indirectly*. Consider user *A*, who validates another user *B*'s *attributes*. By doing so, *A* is indicating his belief that *B* is telling the truth. This says something about *B*'s trustworthiness as a *user*. Thus, *attribute* validations are a proxy for *user* validations.

As users validate each other's attributes, they build a network of implicit user-level validations. Users who are more "entangled" in this network can be trusted more than their less-entangled peers because they have been verified by many users--who themselves have been verified by still other users. This builds on sociological research finding that: (1) human beings are "embedded" (i.e. entangled) in webs of social relationships; (2) the *way* they are entangled (i.e. embedded) affects their behaviors; (3) with greater embeddedness in a social network, people are less likely deceive and/or cheat other members of that network [6]. The final point speaks to AssertID's objective; greater *embeddedness* indicates greater *trustworthiness*.

A metaphor is useful for conceptualizing embeddedness. Imagine a butterfly caught in a spiderweb. How can the butterfly be freed from this web? Two possibilities exist. The butterfly can be freed by cutting a small number of web strands directly adjoining it. Conversely, the butterfly can be freed by cutting a much larger number of strands that are further away. A user's embeddedness in the web of trust can be seen in the same way. Compared with their less-embedded peers, highly-embedded users have two characteristics: they are verified (trusted) by a greater number of other users who in turn are each verified (trusted) by a greater number of still other users.

AssertID uses a quantitative measure that captures both aspects of this metaphor. In a social network, *Bonacich centrality* (henceforth *b-centrality*) measures a person's *number of connections*, with each connection *weighted* by the value of its *own* connections. In other words, a focal individual gains greater b-centrality by connecting with other well-connected individuals. In mathematical terms, the b-centrality of individual *i* in a social network (graph), who is connected with *j* neighbors, is calculated by:

$$c_i = \Sigma_j(\beta c_j)$$

where c_i is the centrality of point *i* and c_j is the centrality of point *j*.β indicates how much c_j should contribute towards c_i. $\beta = 1$ indicates that the full value of c_j is added to c_i; in contrast, $\beta = 0$ indicates that the c_j does add to c_i at all [7].

AssertID uses a proprietary variant of b-centrality to measure users' embeddedness in its web of trust. Even in its raw form, b-centrality is an excellent indicator of user trustworthiness--users with high b-centrality scores can be considered more likely to self-assert true (actual) attributes than users with low b-centrality. AssertID uses proprietary modifications to this measure to further enhance trust in the system. The resulting metric--the *AssertID score*--shows AssertID's *degree of confidence* that one particular attribute (e.g. age) of one particular user is actually true.

AssertID computes this score for all of its users using a proprietary algorithm. Batch-computed several times a day, this algorithm updates user scores to incorporate their most recent verifications. *This algorithm runs in linear time.* Thus, it is scalable to largest of social networks.

ISafe Partnership

A *web of trust* can grow organically via viral propagation. However, this web can grow faster when seeded with initial members—who can be highly trusted by the system and can bring in their friends. AssertID has joined ISafe, a non-profit group focused on educating children how to use the Internet safely, to design and implement such an initiative to seed its web of trust.

Note that a vicious cycle has prevented identity and age verification for minors. Unlike US-based adults, who can be verified by credit bureaus and public records databases, no such data exists on minors. This has discouraged firms and entrepreneurs from attempting to validate minors' identities, which in turn has hampered efforts to collect identity data on minors in a secure manner. To break this cycle, AssertID must create a database of minors by registering trusted users into an authoritative database.

One potential solution is to require every user to go through an in-depth identity confirmation using physical documents (e.g. birth certificates). However, this creates tremendous friction for end users, reducing online communities' ability to recruit new users. Therefore, any successful solution provider must minimize this friction.

AssertID's solution is to partner with ISafe, which currently reaches 6 million children through its current programs. AssertID and ISafe are developing a process to bring these children into a web of trust, by incorporating an AssertID "safe-surfing" credential into ISafe programs. The process will leverage a beta test conducted by personnel currently at ISafe and AssertID, to safely register children and provide them with an online credential identifying them by age and gender [3]. The process verifies that an adult (claiming to be the child's parent) has given the child permission to participate. It *also* binds the child's profile to a specific *school*. This provides a second level of security in addition to parental validation; teachers and school administrators also verify a child's identity. Over time, this process will reduce the number of registrations that will be required at a given site as users will be able present pre-existing credentials created in through process, versus forcing the end site to register 100% of its users.

Current and Future Implementations

The AssertID beta has been implemented as a Facebook widget, and verifies First Name and Last Name. It will soon incorporate other attributes like photo, and eventually age range (e.g., Minor – Less than 18, or Adult: At least 18). Note that AssertID has two important synergies with social networks like Facebook. On one hand, AssertID imports users' self-asserted attributes directly from their Facebook profiles, eliminating AssertID's need to query users for this information. More importantly the verification procedure strongly resembles the "add a friend" process that Facebook users are already familiar with. On the other hand, Facebook gains a built-in identity oracle, reducing users' ability to deceive others regarding their age, marital status, etc. Overall, AssertID generalizes an identity credential created within Facebook (and eventually, other online social networks) so that any 3rd party can trust it in a standardized fashion.

This implementation functions as follows:

- When an individual signs up for a Facebook account, Facebook asks him about key socio-demographic data, including date of birth.
- Such socio-demographic information is displayed on that user's Facebook profile per the user's privacy settings--he can determine which information to release to his friends and the general public.
- AssertID provides an application to be embedded on this page, which obtains his self-asserted attributes from the site in accordance with Facebook's terms of service.
- AssertID's application also asks users to send verification requests to Facebook friends. Such friends can verify the requester by adding the widget.
- The self-asserted attributes and verification data is sent to an AssertID server, which computes its degree of certainty regarding the attributes.
- AssertID posts its attribute scores on the widget embedded on the user's profile, which is available to anyone viewing the profile.
- More importantly, AssertID provides an external directory system which enables users to leverage this identity in other communities.

How would users and websites actually use these scores? *Users* and *websites* can have a better idea whom they are dealing with online. Consider three hypothetical situations:

- A 15-year old girl has met a fellow Miley Cyrus fan online. After exchanging messages, her new friend wants to meet in the physical world. Knowing that AssertID is highly confident that this new friend is 14, the girl feels more secure about setting up a meeting. Had AssertID been less certain about her friend's age, the girl could have asked for additional proof that her friend was actually who she purported to be.
- A 13-year old boy has been harassed online by local classmates posing as geographically-distant persons. Seeing that AssertID has low confidence in the harassers' self-asserted geographic locations, the boy can more easily figure out who is harassing him.
- A website can use AssertID-validated ages--ages listed on Facebook profiles that exceed a certain confidence score--to grant or deny access to age-limited content.

In addition to the current Facebook widget, AssertID can be embedded into other social networks (e.g. MySpace) or work on a standalone basis across several networks.

Strengths and Weaknesses

What differentiates AssertID from competing approaches to age verification is its universal applicability. AssertID does not solely rely on public records databases that possess data on US-based adults but no one else. Thus, it can be used to verify identity and age for anyone with an online social network profile, anywhere in the world, of any age.

AssertID has a second major strength: it is a user-controlled identity system. This is better-aligned with broader privacy trends in the Internet space. People online are fatigued by providing personal information (e.g. financial records) for identification purposes. In contrast, AssertID is built upon a system that people are *already* comfortable with: their social network profiles.

AssertID has two weaknesses. First, it involves a (positive) change in behavior for users: users have to manage their own identity and verify others within their network. Even when embedded in a social network, AssertID requires users to ask friends to verify them--and for these friends to make these verifications. Note that AssertID can reduce

these costs with greater integration with the online social networking sites. Second, note that AssertID assigns a *probability* that someone is self-asserting the correct age--it does not assign *absolute certainty* to a given profile's self-asserted age. Thus, a very small but consequential number of "failures" may occur. Finally, AssertID is an early stage entity built on a revolutionary concept. This concept will need real-world testing and tuning.

Reliance on law and policy

Although AssertID is not dependent on legislation and/or policies mandating age verification in online communities, potential legislation provides a focus for a first application of a general identity verification solution. Therefore, as AssertID develops its identity verification platform, it will actively promote its perspectives to the public sector as a possible solution to this problem. Note, however, that AssertID can potentially provide any type of identity verification, including photo verification. Thus, AssertID will eventually be deployed for a wide range of uses.

Testing and validation

At submission, AssertID's alpha product has just launched. Thus, AssertID is generating empirical data to test its algorithm. This will provide the first quantitative tuning results for its algorithm and process.

The child registration process was tested in an ISAFE beta in 2004. There is low incremental risk associated with relaunching that process, as process simplifications have been made using the lessons learned from that beta test.

EXPERTISE AND COMPANY OVERVIEW

The core AssertID team has expertise in Internet security, social network analysis and software development:

- **Kevin Trilli (Founder and CEO)**: Kevin has 15 years of operating experience across product management, strategy marketing, engineering and operations. Prior to AssertID, Kevin spent 8 years at VeriSign, where he served as Director of Product Management across several lines of business and managed over a dozen products in his tenure, including VeriSign's flagship Web Site Security (SSL) business and identity proofing products. Kevin was the primary product manager at the time of the ISAFE project and worked intimately with ISAFE to develop and market the solution. Kevin holds a Masters of Science in Management from Stanford University and a B.S. in Chemical Engineering from the University of Illinois at Urbana-Champaign.
- **Joon Nak Choi (Chief Scientist)**: Joon Nak (JC) is a PhD Candidate in Sociology at Stanford University. His research focuses on trust within social networks, analyzing using quantitative and computational techniques. JC's responsibility is the establishment of the principles by which the AssertID will function to ensure trust and integrity among the AssertID network. JC has received numerous academic honors, including the Stanford Graduate Fellowship and a Honorable Mention in the National Science Foundation Graduate Fellowship program. JC holds a MA in Sociology from Stanford and an AB in Economics, International Relations and Urban Studies from Brown University.
- **Murali Vivekanandan (Lead Architect/VP Eng)**: Murali has fifteen years of experience designing and developing object-oriented, distributed, and Internet applications on Windows NT and UNIX platforms. Prior to AssertID, Murali was Lead Architect and Co-Founder of Radaptive, a comprehensive end-to-end solution to manage complex IT operations. Murali spent several years as a Lead Architect at Cisco, working on the Cisco.com website which served as a repository for 172 different applications accounting for $2 billion of annual sales. Murali holds a BS Physics from Madras University, and a MS Computer Science from Pune University.
- **Advisors**: Nico Popp (VP of Innovation at VeriSign), Stewart Bonn (Serial Entrepreneur and Investor)

Company

AssertID is an early stage technology company actively raising funds. Having filed intellectual property and built a prototype earlier this year, the team is actively looking to secure its first professional round of funding to productize its offerings and bring them to market.

BUSINESS MODEL OVERVIEW

AssertID faces three primary costs: (1) related to registering new users into the system via the ISafe partnership; (2) 3rd party data used to verify users as needed and (3) the cost of maintaining and delivering the data as a service to meet the needs of the application.

AssertID will offer its data in a traditional licensing model, charged on a per-lookup (user) basis much like credit and other data bureaus. Project integration costs using AssertID's API will be charged on a per-project basis.

AssertID will provide aggressive pricing for its initial customers and provide open considerations for non-profits.

MORE INFORMATION

AssertID.com provides the latest information on the beta.

CONTACT INFORMATION

Kevin Trilli, CEO, AssertID, Inc., 1341 Bay Street, San Francisco, CA, 94123, kevin@assertid.com, 415-734-8514

Joon Nak Choi, Chief Scientist, AssertID, Inc., 1341 Bay Street, San Francisco, CA, 94123, jnchoi@stanford.edu.

For queries on ISAFE: Teri Schroeder, CEO, ISAFE.org, tschroeder@isafe.org, (760) 603-7911 Ext. 12.

CERTIFICATION

I certify that I have read and agree to the terms of the Internet Safety Technical Task Force Intellectual Property Policy. [Signed: Joon Nak Choi]

REFERENCES

1. National Assessment Center. At Risk Online: National Assessment of Youth on the Internet and the Effectiveness of i-SAFE Internet Safety Education (2006). Available at media@isafe.org.
2. Blakley, B. Identity and Community in Human Society. Available at http://podcast.burtongroup.com/ip/2006/06/identity_and_co.html.
3. ISafe Incorporated. i-STIK Product Literature (2004). Available at http://www.isafe.org/imgs/pdf/istik.pdf.
4. Ibid.
5. Kaushik, N. Revisiting the Identity Oracle Concept. Available at http://blogs.oracle.com/talkingidentity/2007/10/revisiting_the_identity_oracle.html
6. Granovetter, M. Economic Action and Social Structure: The Problem of Embeddedness. American Journal of Sociology 91,3 (1985), 481-510.
7. Bonacich, P. Power and Centrality: A Family of Measures. American Journal of Sociology 92 (1987), 1170-82.