*An extended abstract of this paper appeared in the Proceedings of IEEE SPSN, Boston, MS, USA, October 2011.*

# Recognizing Your Digital Friends

Patrik Bichsel*, Jan Camenisch* and Mario Verdicchio†
* IBM Research – Zurich, Switzerland, *Email: {pbi, jca}@zurich.ibm.com*
† Università degli Studi di Bergamo, Italy, *Email: mario.verdicchio@unibg.it*

***Abstract*—Personal relationships are more and more managed over digital communication media, and electronic social networks in particular. Digital identity, conceived as a way to characterize and recognize persons on the Internet, has thus taken center stage, although this concept still remains vague in many of its aspects. This work aims at shedding some light on this topic, by sketching a basic conceptual framework, analyzing the issues for Internet users, and proposing possible solutions that promote a better use of digital identity.**

***Keywords*-Computer security, communication system security, computer network management, identity management systems, social network services, authentication.**

## I. Introduction

The Internet has radically changed the way people interact in recent years. The Web 2.0 wave, in particular, has widened the focus of the users' interest to include not only content, but also the people who have generated it. The huge success of electronic social networks has undoubtedly provided a boost for persons to take center stage on several websites.

Because of the differences between the physical world in which persons exist and interact and the electronic realm of the Internet through which their data is exchanged, several new challenges are posed. This work focuses on the gap between what traditionally characterizes a person and what is made available in digital form. In particular, we focus on how we can recognize a person when we interact with them over the Internet.

In the physical world, a person has some characteristics (e.g., name, hair color and length, facial features) that enable others to identify her, that constitute her *identity*. A question, then, rises on the role of these characteristics in the definition of an identity, and it is legitimate to wonder whether the same concepts and mechanisms work on the Internet. We will not get into long-going philosophical debates on essential characteristics of entities [9], or attempt to classify them based on whether they change over time [13]. We think that such distinctions are not very significant in the usual practice of identity management: in our view, a characteristic $c$ can enable us to identify a person $P$, as long as it has not changed since the encounter when we registered $c$ as belonging to $P$, and $c$ is uniquely characterizing $P$ in the multitude of other people from which we are singling out $P$.

We take inspiration from how identities of persons are managed in the real world, to build a model of *digital identity* on the Internet, and see whether the criteria we use offline to recognize people are supported. Moreover, we tackle the new issues rising from the lack of physical interaction in the context of the Internet.

## II. Basic Concepts

Let us now present the fundamental concepts we are going to use in our proposal. We focus on *persons* that exist in the real world, and whose identities are well defined in the traditional sense: all the characteristics that are normally associated to them are there, and can be checked in the usual way. An example of such a person is a US citizen with a driver's license. Making a comprehensive list of all the characteristics of persons that allow us to identify them is beyond our scope. We focus instead on the restricted set of such characteristics that can be translated in digital form and that a person is willing to divulge on the Internet.

Given a person $P$, we call *digital identity* (DiD) of $P$ the set of all data that has been published by $P$ (or by an authorized person) on the Internet, and that is the digital counterpart of what people in the physical world would normally use to identify and describe $P$. For instance, the DiD of a person can be comprised of a Facebook page, a blog on Tumblr, a Twitter page, or any other data created and managed by the person herself, including her email correspondence. A DiD can be a very complex and dynamic set of information that is hardly ever processed all at once. More often, people view only a small fraction of it, in the form of a social network profile, for instance. We call *facet* a subset of a DiD which is presented in a unitary way. The information provided by a facet, possibly in the form of text, pictures, or multimedia files, can be considered as the digital counterpart of what is presented when people meet in the physical world.

## III. Digital Identity Issues

Several issues rise in a context where, for the lack of physical contact, persons present themselves through the facets of their DiDs.

### A. Security

Like every other type of information that is transmitted through the Internet, a digital identity must deal with the problem of security. We can distinguish two types of possible attacks from malicious users, based on *alteration* and *duplication* of facets of a DiD, respectively. Alteration attacks target an already existing facet of a DiD to change it or add new content. For example, it still happens often, especially with small family-run hotels, that we are required to fax our credit card data to make a reservation. If the webpage of the hotel's owner has been altered to show a different fax number, it is easy to imagine the consequences. Duplication attacks aim at creating a new facet designed to look like it is part of the DiD of a person $P$. The most common example of duplication attack on the Internet are phishing websites, but the problem affects social networks as well. For instance, in front of a page representing $P$, users are naturally led to think that $P$ has published the displayed data, and manages the page. Such supposition is true in most cases, but it cannot be taken for granted, especially if one considers that digital content can be easily copied and reused. A Facebook page presenting itself as the official fan club of a pop star could trick a considerable number of people into giving out their email address with the pretense of a competition.

The facets in these attacks, whether the result of an alteration or created from scratch, are not part of the DiD of the person $P$ they are referring to, because they have not been published by $P$ or an authorized person. In other words, these facets lack the property we call *authenticity*. To support a correct use of digital identities on the Internet, users need instruments that guarantee the authenticity of the facets they are viewing. That said, a multitude of facets, all showing the same picture, do not automatically imply an attack. This is a way for a person to show that these facets are part of her DiD, i.e., a way to support her *recognizability* among users. However, there exists no universally accepted specification on how to represent the fact that different facets are all part of the same DiD, as the websites that host them do not share a uniform data structure for their users. Authenticity implies recognizability, i.e., if we are guaranteed about the entities behind the facets, then we know which facets belong together. We need to understand whether the means to check authenticity can also be used to support recognizability without the burden of too restrictive standards for websites.

### B. Privacy

It has been said that the best way to keep a secret is to never have it. In this context, we may say that the best way to avoid *privacy* issues is to never sign up for anything on the Internet. Still, many voices want to be heard without revealing whom they belong to. The issue here is the exact opposite of what we presented before: some, or possibly all, of the content published on the Internet by a person $P$ may not be supposed to be ascribed to $P$; in other words, sometimes there is the need for avoiding *linkability* between different facets of a DiD. Such need can rise in many different contexts: we are not only thinking about a controversial political blog in countries with controlled media, but also much more mundane cases like a teacher who manages a comic book discussion forum and does not wish to be recognized by his students. This may simply look like a call for *anonymity*, but in the context of digital identity, users often have more complex needs. Complete anonymity, in fact, would not serve the purposes of the above-mentioned blogger, for instance: the blog's existence itself relies on the connection between all the entries, which is normally given by the URL at which they are published. Should the blog be transferred to another address because of technical or safety reasons, how could the readers recognize it when it is back online at a different site? We are looking for solutions to support *pseudonymity*, a way to tackle the trade-off between having an easily recognizable DiD and keeping the details on the person behind it private.

## IV. Dealing with the Issues

Here follow the solutions we propose to tackle the above-mentioned problems with security and privacy of digital identities.

### A. Secure DiDs

In the physical world, we recognize people mainly by means of their physical attributes. Let us first check the authenticity of a facet by comparing the features it shows with the attributes of the person it refers to.

*Authentic attributes.* We can identify three different categories of attributes that assert authenticity in the digital domain. Firstly, in case a facet offers the possibility to directly transfer physical attributes, authentication closely resembles the recognition process in the real world. An example is provided by the "hangouts" of Google+, where users can start a video chat. Secondly, even when physical attributes are not shown, certain features, that are tightly bound to a person and hard to copy, can be transmitted over the Internet. For instance, we may recognize a person based on her writing style, humor, or quirks that we perceive during a chat session, or through an email. Finally, if a user has already interacted with a person $P$, e.g., met her in person or called her on the phone, all attributes that are coherent with the information exchanged during the interaction and published in a facet of $P$'s DiD, increase the confidence in the authenticity of the facet. For example, if $P$ mentions her vacations in Rome on the phone with $Q$, $Q$ gains confidence that the Flickr account with the Colosseum pictures belongs to the DiD of $P$.

*Certified attributes.* Another way to approach the problem is to rely on certificates (e.g., X.509 [7], U-Prove [10], idemix [12]), with which $P$ can add certified attributes to her facets. Let us assume $P$ has a credential from her government that certifies her name, first name, and birth date among other attributes. When registering at a host (e.g., Facebook), $P$ can provide the certified attributes instead of simply inserting them into a Web form. The host would provide a mechanism to distinguish certified from non-certified attributes and show which entity provided the certification. Consequently, a user visiting the facet can verify the set of certified attributes and decide how confident she is in the fact that it authentically represents $P$. However, this approach imposes several requirements. First of all, the hosts would need to adapt the registration process and incorporate mechanisms for showing the certification of attributes. Moreover, the requirement of possessing a certificate is today only practical for entities such as companies or larger organizations. As governments (e.g., Belgium, Germany) start distributing electronic identity (eID) cards, certified attributes may become available for the general public. Finally, users need to have trust in the certificate issuers of their digital friends as well as in the host of the facet. Note that while the increase in trustworthiness of attributes seems to be coupled with a loss of privacy of a user w.r.t. the host of a facet, the use of privacy-friendly authentication techniques [8], [4] can eliminate this issue.

*Community-certified attributes.* Previously, we have focused on mechanisms that are based on a single user verifying the authenticity of a facet. However, after a user has assessed the authenticity of an attribute she could share her findings, e.g., by assigning a confidence rating. Such rating or recommendation requires the user to authenticate in order for other users to trust in the rating. Consequently, we may view such rating as a certification provided by a community of users. An approach in bootstrapping trust in the authenticity of attributes has been proposed in [3], where it is used to initiate a public key infrastructure (PKI). Differently from such proposal, we assume that users trust the host, thus, we can use a mechanism that does not rely on cryptography. Still, as with the externally certified attributes, this approach requires the host to offer a system where users can rate attributes and such ratings are properly displayed.

Let us now focus on mechanisms that support recognizability of facets, that is, help show that they belong to one DiD.

*Unique reference.* A straightforward solution to link several facets is to publish them endowed with a unique reference. This reference is supposed to work as an identifier, showing viewers uniquely of which DiD they are observing one facet, assuming that the reference works across multiple domains of the Internet. A public cryptographic key or a uniform resource identifier (URI) are possible ways to achieve such result. Let us consider the following example. If user $U$ visits some blog on Blogger and sees a comment by John Smith, she should be able to recognize whether he is the John Smith that $U$ knows from Facebook. A unique reference can be established in accordance with the trust model we rely on. If there are trusted hosts, then identity providers have the possibility to endow a DiD with a unique reference using technology like OpenID [11]. When relying on such hosts, we are assuming that the username of the DiD on the trusted host is unique. For the reference to be fully recognizable, it should explicitly include the trusted host's name, but this is not part of the current practice of many websites, so that Web users see that a John.Smith entered a comment in a blog, but there is no way to automatically establish that it is John.Smith@facebook.com, i.e., the John Smith $U$ already knows. Publishing a unique reference requires to adapt the current practice of how facets of DiDs are handled. If hosts of the different facets agreed on a mechanism supporting such solution, this would allow for an automated detection of several facets. Such agreement, although very desirable, looks unlikely. Instead, alternative solutions have emerged. A dedicated service has been introduced that provides a unique reference to all social network activites of an entity called about.me[1]. Another, more simple practice is to publish as an information item within one facet a link to another facet (e.g., Facebook users often post a link to their Flickr account).

*Corresponding attributes.* The same information published under different facets seems to imply a link among such facets, although the simplicity of copying digital information makes it easy to create a facet that is seemingly equal to another one. It is then important to remark that relying on the equality of general attributes (e.g., the same name in several facets) or on similar information (e.g., different facets stating that they are leaving for vacation) is not *per se* a guarantee. However, the correspondence between a new Skype status message "Vacations in Rome. Yeah!" and new Colosseum pictures on a Flickr page can increase a viewer's confidence in the fact that those facets belong to the same DiD. The measure of such confidence increase should depend on how easily such evidence may be fabricated.

### B. Private DiDs

The lack of physical contact may look like a disadvantage when it comes to identity management as it induces the need for verification of the authenticity that determines whether a DiD actually represents the implied person. However, this very lack of a physical touch can introduce new and interesting types of interaction. Unless we are in specific lawful contexts requiring a user to release her attributes according to some real-world definition, there is no limit in the choice of her published characteristics. In such situations users are free to create DiDs that present a meaningful coherence and make them look like they represent an existing entity, without actually corresponding to any real person. This is the case, for instance, of the above-mentioned blogger who wants to protect her real identity, but still wants to be represented on the Internet with a DiD. Such DiD thus works as a pseudonym for a person. One can also give up any pretense of realism, and create DiDs with such unrealistic features that it becomes obvious that they are fictional.

Pseudonymous DiDs are also affected by recognizability issues. For instance, the DiD of a blogger may also have a social network page. Users should be able to recognize that such page belongs to the DiD that also has a blog, while the actual person's privacy is still guaranteed. The mechanisms described in Section IV-A for recognizing that several facets belong to the same DiD work also for pseudonymous DiDs, although we are in a different context, in which the link between a DiD and the person behind it needs to be kept private. To achieve this goal, users are given two possibilities, according to the trust relations they have with the entity hosting their facets. Let $P$ be a person whose DiD $\mathcal{P}$ has a facet on host $H$. When $P$ trusts $H$ not to leak any information, her real identity can be considered protected, and all $P$ needs to do to manage her DiD is to authenticate to $H$. This is usually done by means of a username/password pair. When such trust is missing, $P$ needs to rely on an authentication mechanism,

[1] https://about.me/

that protects her identity also against $H$. A possible solution is offered by anonymous credential systems [4]. They prescribe the use of certified attributes, that could maintain the level of assurance $H$ needs, while at the same time allowing $P$ to remain pseudonymous.

Another cryptographic primitive that supports the management of pseudonymous DiDs is *verifiable encryption* [5]. It prescribes that, when entity $S$ communicates an attribute type and its value to entity $R$, $R$ receives the information encrypted in such a way that it can be decrypted only by a designated mediator, but the attribute type can be nevertheless verified by $R$. $S$ and $R$ agree on the terms under which the mediator is supposed to decrypt the encrypted attribute value. Verifiable encryption enables pseudonymous DiDs to be passed on. For instance, the above-mentioned blogger, whom we call $A$, can be substituted by a new author $B$, without anyone else knowing about the change, as follows. We assume that $A$ verifiably encrypts her public key and publishes it with each post. The host of the blog, then, checks that the public key verifiably encrypted with the previous post matches the public key used to sign the current post, to be assured that the post was submitted by the legitimate author. All $A$ needs to do to pass the authorship to $B$ is to verifiably encrypt $B$'s public key in her last post.

## V. Related Work

Researchers from several fields have investigated deeply on the analogies and the differences in the concepts of identity in the physical and in the digital world.

Allison et al. provide an overview of the concept from several different perspectives: legal (authorship and ownership issues), philosophical (logical relations among digital objects), and historical (chronological models and records of the evolution of digital identities) [1]. Cameron attempts to provide a more unified definition of the concept, with a synthesis of all its aspects into a list of “laws of identity” [6].

Other efforts point at singling out the available technologies to implement the principles that are traditionally attached to digital identity. Windley, for instance, considers the support of digital identity fundamental for businesses on the Internet to succeed, and provides several pointers to existing proposals and standards [13].

When it comes to standard proposals, two main research guidelines can be found in the literature. Low-level computational instruments keep on being elaborated in the context of cryptographic research, to expand the boundaries of what can be provided to users in terms of security and privacy. For instance, the endeavors of Lysyanskaya et al. aim at handling pseudonyms or anonymous access [8]. On a higher level, in the context of distributed system research, standards are proposed to support the expression of identity attributes for authentication and access control purposes, like in OpenID [11], and more and more of these works, see for instance Ardagna et al. [2], consider privacy issues as fundamental.

## VI. Conclusion and Future Work

Internet users deal with digital identities in a similar way to how people deal with each other's identity in the real world. Nevertheless, the lack of the physical dimension leads to a bigger freedom and anonymity, which allows for new types of identity to rise in the digital context of the Internet. People look for, and find each other based on the attributes that they exchange through their digital counterparts. This work aimed at shedding light on the basic concepts related to digital identities, and proposed solutions based on existing technologies to support recognition of people over the Internet, with an eye on both security against attacks, and privacy for users who intend to stay anonymous.

Our next steps on this research path will deal with digital identities of organizations, which have the peculiarity of either being managed by more than one person at the same time, or by different people throughout their life cycle. We consider this topic particularly interesting, because it calls for a compromise in the trade-off between anonymity of the users on the Internet and the accountability of their actions within their organization.

## VII. Acknowledgements

This work was partially supported by the EC within the 7th Framework Programme, under grant agreement 257129 (PoSecCo), and by the MIUR in the framework of the PRIN project Gatecom.

## References

[1] A. Allison, J. Currall, M. Moss, and S. Stuart. Digital identity matters. *Journal of the American Society for Information Science and Technology*, 56(4):364–372, 2004.

[2] C. A. Ardagna, S. De Capitani di Vimercati, S. Paraboschi, E. Pedrini, and P. Samarati. An XACML-based privacy-centered access control system. *WISG '09: Proceedings of the first ACM workshop on Information security governance*, p.49–58, New York, NY, USA, 2009.

[3] P. Bichsel, S. Müller, F.-S. Preiss, D. Sommer, and M. Verdicchio. Security and trust through electronic social network-based interactions. *Workshop on Security and Privacy in Online Social Networking (SPOSN09)*, v.4, p.1002–1007. Aug. 2009.

[4] S. Brands. *Rethinking Public Key Infrastructures and Digital Certificates: Building in Privacy*. MIT Press, 2000.

[5] J. Camenisch and V. Shoup. Practical verifiable encryption and decryption of discrete logarithms. http://eprint.iacr.org/2002/161, 2002.

[6] K. Cameron. The laws of identity. http://www.identityblog.com/?page_id=354, May 2005.

[7] D. Cooper, S. Santesson, S. Farrell, S. Boeyen, R. Housley, and W. Polk. Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. RFC 5280 (Proposed Standard), May 2008.

[8] A. Lysyanskaya, R. L. Rivest, and A. Sahai. Pseudonym systems. *Proceedings of SAC 1999, volume 1758 of LNCS*, p.184–199. 1999.

[9] G. Matthews. Aristotelian essentialism. *Philosophy and Phenomenological Research*, 50:251–262, 1990.

[10] C. Paquin. U-Prove cryptographic specification V1.1. Technical report, Microsoft Corporation, February 2011.

[11] D. Recordon and D. Reed. OpenID 2.0: a platform for user-centric identity management. *DIM '06: Proceedings of the second ACM workshop on Digital identity management*, p.11–16, New York, NY, USA, 2006.

[12] Security Team, IBM Research Zurich. Specification of the Identity Mixer cryptographic library. IBM Research Report RZ 3730, IBM Research Division, Apr. 2010.

[13] P. Windley. *Digital Identity*. O'Reilly Media, Inc., 2005.