

# Utopia

## Providing Trusted Social Network Relationships within an Un-trusted Environment

William Gauvin, Benyuan Liu, Xinwen Fu, and Jie Wang

Department of Computer Science,  
University of Massachusetts Lowell  
{wgauvin, bliu, xinwenfu, wang}@cs.uml.edu

**Abstract.** This paper introduces an unobtrusive method and distributed solution set to aid users of *on-line social networking* sites, by creating a *trusted environment* in which every member has the ability to identify each other within their private social network by name, gender, age, location, and the specific usage patterns adopted by the group. Utopia protects members by understanding how the social network is created and the specific aspects of the group that make it unique and identifiable. The main focus of Utopia is the protection of the group, and their privacy within a social network from predators and *spammers* that characteristically do not fit within the well defined usage boundaries of the social network as a whole. The solution set provides defensive, as well as offensive tools to identify these threats. Once identified, client desktop tools are used to prevent these predators from further interaction within the group. In addition, offensive tools are used to determine the origin of the *predator* to allow actions to be taken by automated tools and law enforcement to alleviate the threat.

**Keywords:** On-line Social Networks, Trusted Environment, Predators, Privacy, Cloud Service.

## 1 Introduction

Successful social networking websites, such as Facebook, MySpace, and LinkedIn, have experienced an explosion in growth. This growth can be attributed to the ease in which members can find each other and share common interest. The user can typically post photos, send messages, comment on friends profiles, join user groups, and generally interact and build online communities of users who share common interests. The amount and types of information that can be shared in these social networking environments is vast, for example, favorite quotes, music and videos are used to introduce the user to the world. The users network can grow over time as the user connects to more and more users and share more and more information.

Social networking sites unfortunately provide an anonymous avenue for those that seek to prey on the young and naive. There are numerous examples of

unscrupulous activity; examples are, cyberbullying, cyberstalking[1][2][3][4][6][7] and underage solicitation for sex [5]. The National Center for Missing and Exploited Children has identified the following problems inherent within on-line socialization; *Child Pornography, Enticement of Children for Sexual Acts Sex Tourism Involving Children, Extrafamilial Child Sexual Molestation, Unsolicited Obscene Material Sent to a Child, Misleading Domain Names and Misleading Words or Digital Images on the Internet*. Many of these categories can directly or indirectly be linked to social networking activity and MySpace specifically[8]. The concern to protect individuals has led to specific agreements between MySpace and government agencies, whose job is to serve and protect the users of MySpace. One of the more detailed works in this area is the “Joint Statement on Key Principles of Social Networking Sites Safety”[9], originating from the Attorney General Martha Coakleys office in Massachusetts and including 49 other Attorney Generals, as well as MySpace. This document provides guidelines and measurements that must be taken to “Provide children with a safer social networking experience”, specifically directed towards the “operators of social networking sites”[9]. This paper also advocates the use of online safety tools, as well as design and functionality changes that are geared towards the protection of children. The overall emphasis of the paper is the concept of providing on-line identity authentication tools to include age verification.

Utopia is a systematic approach to address the problems identified above with social networks and provides a means to influence the usage patterns of the individuals within a group. Utopia focuses on the interaction between the owner of a profile and those that publish to that owner’s wall, i.e., their friends. Using heuristics gathered from the conception of MySpace to the present; it ascertains local and global usage patterns, and uses these predictable patterns to identify anomalous activity. In addition, these patterns can be applied as behavioral templates to identify both good and bad usage traits. Further characterization is achieved using social *honey-pots*, which emulate an individual buddy of a social network and monitor the activity of the group, with respect to spam or *splogs*. Utopia also addresses privacy concerns by providing a method to encrypt blogs, such that only the members of a group may view the contents, even within a public environment. The goal of Utopia is to facilitate, using unobtrusive techniques, the means to evaluate social network groups of an individual, and ascertain specific characteristics that are deemed unhealthy with respect to the general use of the social networking environment[10]. Once identified, this undesired content is eliminated from the users presence, and is no longer a factor in the overall social network experience.

## 2 Utopian Approach

The primary goal of Utopia is to provide a means for individual users of on-line social networking sites and law enforcement to work independently and unobtrusively. The user’s MySpace experience must not be negatively impacted by the tools required to protect such individuals and law enforcement must be

provided the information required to actively police the social network domain. User level transparency is a major goal of Utopia, it is important to protect, but not interfere with the social network experience. Utopia may be deployed in two environments, they are:

1. A standalone environment
2. A fully distributed environment

In a standalone environment, members of a social network create “Trust-Groups”. These groups form ties with each other and share common knowledge about each member of the group. They use a rating system, based on content analysis, that is shared and updated by all members to identify good and bad usage patterns. The ranking system provides a means to associate reputation with a specific publisher, using a global view of the user, as indicated by the behavior and tendencies of the publisher for all members, not just a single individual profile.

The second scenario is a fully distributed environment, in which Trust-Groups are created within the Utopia Social Network *Cloud Service*. The Cloud Service provides a dynamically scalable and virtualized social network environment, which uses software as a service technology to facilitate the assimilation of social network activity into temporal usage patterns which are projected to managed clients. In this environment, members create private social networking groups, as in the standalone environment, but group membership and the usage information is managed and distributed by the cloud service. In addition, the cloud service shares global usage pattern templates with the individual social network group to provide a stronger capability for rating and evaluation.

Figure 1 is an example of both the standalone and distributed environments. In the standalone environment, the group “Family” has ties to each member within an internal home network. They form a private trust group named “Family” and do not participate in the distributed environment offered by the Utopia cloud service. It is important to note that membership is not restricted to computers within the internal network as there are external “Family” members.

The distributed solution set is demonstrated by the group “Team”. This group uses the Utopia cloud service to form the group. The membership information is stored using the cloud service and members are added and removed by updates to the Utopia cloud database as required. The database is protected such that only members of a group may obtain the group ranking, usage templates and other characteristics specific to the individual group. Membership within the distributed environment have the added benefit of global usage patterns, policies and tools which may be inherited by the private groups.

The high-level design of Utopia is a set of components which reside on the clients and a cloud service component. The standalone components implement a peer-to-peer network in which clients may distribute membership and ranking information. The implementation of the standalone environment is a sub-set of the distributed cloud service. Users of the tool that are concerned about privacy may restrict the level of information that is reported and prevent unwanted data leakage out side their sphere of influence. Others may choose to

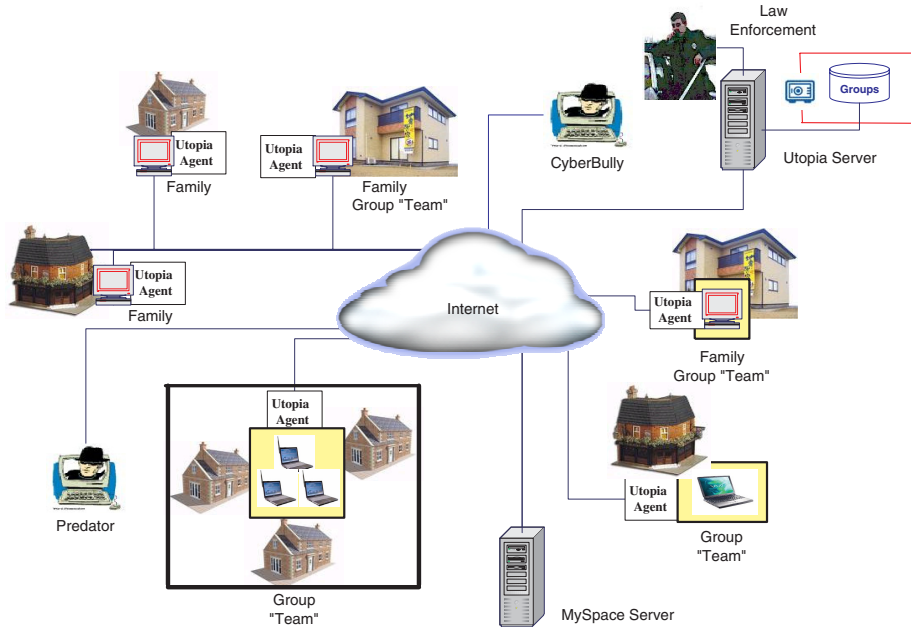


Fig. 1. Utopia High-Level Overview

be highly involved with information sharing of the distributed model and host the group templates in the cloud server; as well as receive updates on predator list and known usage patterns. In addition, both scenarios may report events of questionable origin or material to the cloud service to aid in the collection, identification and assimilation process.

## 2.1 Utopia Standalone Environment

The strength of Utopia resides in the flexibility of its design to be molded into a solution which meets the specific requirements of the members of the group; with out over burdening the system with unwanted features. In the standalone environment, Utopia provides the means to protect users of social network sites by creating a virtual group, for which group membership is consensual. Members are added to groups using an acceptance model. Each member is analyzed to determine the level of trust and usage reputation. A member's reputation is derived by examining the content being posted by the individual, as well as the hygiene of that content. Members that continually post content associated with splogs (Spam Blogs) and questionable content have a lower rating than those who continually adhere to specific guidelines defined for the group for content publishing. For underage users, the profile guidelines are defined by a guardian.

The goal of the rating system is to teach members the acceptable usage patterns for a specific social network group; a browser plug-in is used to provide "pop-ups"

with feedback to the individual to train them on correct usage patterns and suggest better means for providing information. Pop-ups act as a warning mechanism when sensitive information such as phone numbers, addresses or other person information is being published. The overall goal is to train members on the correct methods to use when active within a social networking environment

Once a member’s rating goes below a defined watermark, the content posted by that individual and the access to that person’s wall is restricted. The Utopia agent consists of a rendering engine, which removes the content of the restricted member by filtering and reformatting the response from the social network server and rendering only “clean” content to the users browser. The mechanism provides a means to protect the user when policies and practices are lacking within the social network provider services.

The general design of the Utopia standalone client services (as well as cloud services) are outlined in figure 2. In this figure, the standalone environment is depicted as Client 1 to Client n with a peer-to-peer communication path (lower orange arrow not within the Internet cloud).

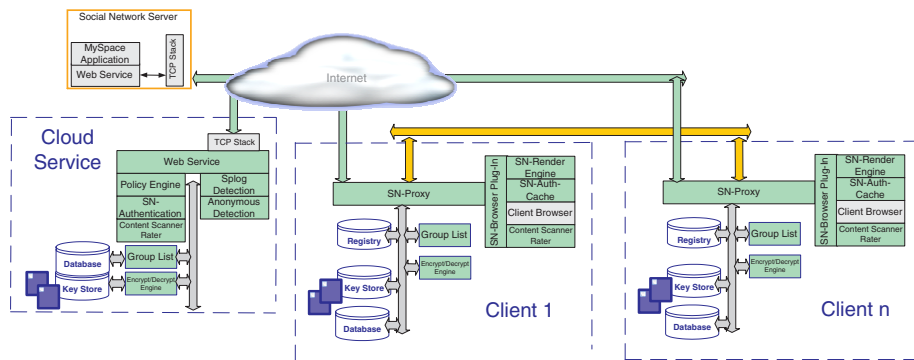


Fig. 2. Utopia Design

The client components consist of the following:

1. A Social Network proxy, which is used to provide a peer-to-peer network for the management of all group associations. The proxy provides an encrypted/authenticated communication channel to distribute rated content to all members within a specific group.
2. A Rendering engine, which modifies or deletes questionable content or total blogs of users which are not members of the group or whose reputation is lower then the weighted value required for viewing. The rendering engine also provides “pop-ups” to suggest better usage patterns to train the user of the tool.
3. An authentication/caching component, which is used by the SN-Proxy to aid in the establishment of communication channels for all groups defined and provide the last measurements, using the cache when clients are not active.

4. A content scanner and rater, which facilitates the means to weigh publishers reputations based on content posted for specific users of the social network group. Included in the content scanner is a honeypot to aid in the local evaluation of splogs and predator friend request.
5. A privacy feature, integrated into the scanner/rendering components.

The result of the standalone solution is a tightly coupled distributed peer-to-peer network which allows all members of the group to participate in globally ranking and propagating reputation of all members of the group. The ranking mechanism is directly reflected in the rendering of the content for each member of the group. In addition, the usage patterns defined for the group are provisioned by a teaching mechanism, which is used to train the individual user of correct on-line behavior, which carries over to all other aspects of their on-line activity.

Utopia accomplishes privacy by using steganographic images posted/hosted on the social network site on a user's profile and the browser plug-in previously defined as a client component in figure 2 that is used to identify these images and translate them to the desired content, rendered by the viewer's browser. A user friendly hosting tool is used to create an image which will be used to embed additional link references and/or text. The image may contain links to other images, or additional content which is only posted if the viewer either is a member of a specific group or knows the password to unlock the image. General viewers of the public profile will only see the image desired by the publisher, this could be the popular "No Image Available" image. No other text or images need be available for the posted blog. A viewer, using the helper plug-in would be evaluated for membership within a group when the image is being pulled down and before rendering; the content for multiple groups may be embedded in the image. Depending on membership, the specific group content would be used to render the content on the viewers browser. The content can include reference links to other photos that the publisher would like to make available as well as text. All would be formatted to be correctly rendered, as if by the hosting site.

Using this method, a person could post an image personalized for their parents, friends and the general public. Figure 3 depicts the use of the Utopia privacy technique for groups *Family*, *Friend* and *Team*. In this figure, an opaque image is rendered on the MySpace server, embedded with private content based on group membership. Non-members only have access to the opaque image. Members of the group *Family* are rendered a picture of the family, with the context string "This is the latest family photo". Members of group *Friend* are rendered a picture of the "friends" and the context string "Remember the dance?". Membership of the group *Team* are rendered a picture of the team with the context string "The rally was great". Membership in multiple groups results in the rendering of multiple individual blogs representing each group.

## 2.2 Utopia Distributed Environment

The distributed environment uses a globally available server, which contains usage templates and offensive tools used for the detection and assimilation of

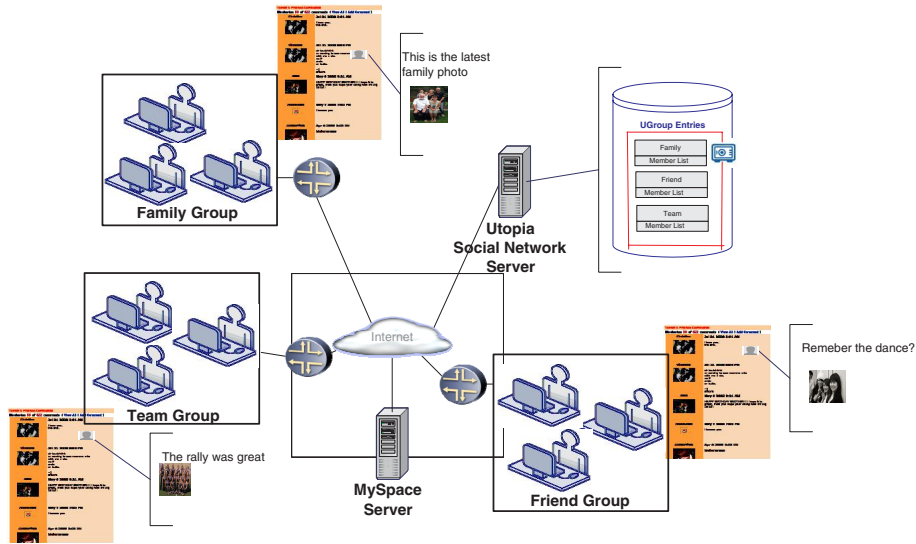


Fig. 3. Utopia Privacy Feature

information to participating clients. It offers a strong solution set to combat cyber-bullying, cyber-stalking and predator identification.

In the distributed environment, the cloud service replaces the localized database in the standalone client solution, as shown in figure 2. Group members use the communication channel to the cloud service (depicted as the green upper arrow in the Internet cloud) as the mechanism to receive reputation and usage templates. Clients obtain the usage templates and policies for the specific groups they authenticate into from the cloud service. Client content is transparently altered and rendered based on the policy and results of the reputation database and usage templates.

Clients may be members of many groups; the ranking of which can vary dependent on the policy defined within the client for each specific group. Some variance is allowed between policies, based on trust and the hygiene of the publishers. This allows a flexible and dynamic means for applying policy to “well-behaved” environments that stray slightly from the desired behavior. Client “pop-ups” are used to suggest better usage patterns when this occurs; and may suggest alternative means for rendering the content. It also provides warnings when known anomalies have been detected, such as splogs or predator friend request and outlines key aspects of the content as reference points for future individual observation for detection.

The overall design of both the client and cloud service within the Utopia system is defined in figure 2. The client design does not change drastically from the standalone environment. This provides an easy migration path from standalone to the distributed environment, as well as the converse. As observed above, the client component may use a standalone environment using peer-to-peer techniques or use the cloud service in a client/server environment. In the distributed

client/server environment, the client is presented with a rich set of usage templates and a predator database used for ranking member content. This solution provides the strongest protection overall. The major features of the cloud service are the following:

1. A web service to provide client/server communication for group members.
2. A distributed policy engine which contains definitions for client groups.
3. An authentication engine, used during client logon to the social network, the authentication process facilitates the means of mapping clients to their pre-defined groups.
4. A content scanner/rater, used to scan groups defined within the cloud service using the global database and usage patterns. The results of the ranking and pattern templates are distributed to their corresponding client members.
5. A splog detection engine; used to identify spam within a social network and provide the means to filter and remove such content. The detection mechanism uses a honeypot that acts as a member of the social network to receive splogs and sexual solicitations. This information is shared by the rating services when evaluating groups.
6. An anonymous network detection engine, used to determine when unsolicited communication traffic is generated from such a domain. This information is used in the global rating system. The source of the engine takes advantage of the new features added by the social network provider, such as the source address of the image posted to the users profile.
7. A database that manages the user groups, privacy and link service.
8. Encryption/Decryption provisioning; to provide secure channels between the cloud service and its client members.

The virtual service manages groups and the validation of groups using the same means defined above for Trust and Privacy. For privacy, the cloud service provides a virtual link feature, which is used within the steganographic images. These links point to virtual links within the cloud service, and gives the user the ability to change the reference of the virtual links of photos, which is generally desired by many. The rating system includes a crawler, which evaluates various groups to determine hygiene based on publisher identities and content. It facilitates the ability to detect predators that have infiltrated the social network group, spam that taints the content within the group; and behavior recognition software that determines patterns which are disallowed within the group. Its important to note, that clients police themselves through the client side software defined in the standalone section of this document. This information is propagated to the cloud service to trigger cloud service analysis to provide a more detailed evaluation when anomalies exist. Not displayed within the diagram in Figure 2 is the relationship between the rating system and the interaction with the profiles of Registered Sex Offenders identified through the Sentinel SAFE technology, of which Utopia uses to provide a stronger detection mechanisms.

The Utopia web service offers a management console used to configure policies, the notification of events, and generate reports. It is used to give a complete

view of activity and provides the means to report scrupulous activity to law enforcement.

### 3 Total Solution

Utopia provides the means for individual groups to define the policies and usage patterns comfortable for group members. It allows guardians to interject policy based on well defined distributed usage scenarios using unobtrusive methods. Utopia trains members on the correct usage patterns and uses embedded plug-ins to enlighten the users. Privacy considerations are managed by secure content, for which only group members have the access to “unlock”. Utopia provisions counter measures to prevent circumvention. These measures prevent users, specifically under-age children, from creating accounts that do not truly represent themselves. Utopia is a total solution that provides the desired level of trust. Member requirements may consists of a private solution, for which the peer-to-peer strategy works well; or, they may desire a rich set of features and participate in “global knowledge” sharing, for which the distributed environment facilitates. Most important, it provides a bridge between information gathered by law enforcement and global techniques and propagates this information to specific individuals where it can best be used to protect and serve its intended audience.

### References

1. Cyberbullying (2009), <http://www.cyberbullying.org/>
2. Hinduja, S., Patchin, J.W.: *Bullying Beyond the Schoolyard Preventing and Responding to Cyberbullying*. Corwin Press (2009)
3. Patchin, J.W., Hinduja, S.: Bullies move beyond the schoolyard: A preliminary look at cyberbullying. *Youth Violence and Juvenile Justice* 4(2), 148–169 (2006)
4. Hinduja, S., Patchin, J.W.: Cyberbullying: An Exploratory Analysis of Factors Related to Offending and Victimization. *Deviant Behavior* 29(2), 129–156 (2008)
5. Hinduja, S., Patchin, J.W.: Offline Consequences of Online Victimization: School Violence and Delinquency. *Journal of School Violence* 6(3), 89–112 (2007)
6. Reuters, Fatal MySpace internet hoax mother is charged (2008), <http://www.news.com.au/heraldsun/story/0,21985,23711115-663,00.html>
7. Lindsay, S.: Boy who posed with guns convicted, *Rocky Mountain News* (2006), [http://www.rockymountainnews.com/drmn/local/article/0,1299,DRMN\\_15\\_4595681,00.html](http://www.rockymountainnews.com/drmn/local/article/0,1299,DRMN_15_4595681,00.html)
8. MySpace exposes sex predators (5/22/2007), <http://www.news.com.au/heraldsun/story/0,21985,21775032-11869,00.html>
9. Commonwealth of Massachusetts, Joint Statement on Key Principles of Social Networking Sites Safety (2008), [http://www.mass.gov/Cago/docs/press/2008\\_01\\_14\\_myspace\\_agreement\\_attachment1.pdf](http://www.mass.gov/Cago/docs/press/2008_01_14_myspace_agreement_attachment1.pdf)
10. MySpace, ParentCare (2009), <http://www.myspace.com/parentcare>