Social Honeypots: Making Friends With A Spammer Near You

Steve Webb
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332
webb@cc.gatech.edu

James Caverlee
Dept. of Computer Science
Texas A&M University
College Station, TX 77843
caverlee@cs.tamu.edu

Calton Pu
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332
calton@cc.gatech.edu

Abstract

Social networking communities have become an important communications platform, but the popularity of these communities has also made them targets for a new breed of social spammers. Unfortunately, little is known about these social spammers, their level of sophistication, or their strategies and tactics. Thus, in this paper, we provide the first characterization of social spammers and their behaviors. Concretely, we make two contributions: (1) we introduce *social honeypots* for tracking and monitoring social spam, and (2) we report the results of an analysis performed on spam data that was harvested by our social honeypots. Based on our analysis, we find that the behaviors of social spammers exhibit recognizable temporal and geographic patterns and that social spam content contains various distinguishing characteristics. These results are quite promising and suggest that our analysis techniques may be used to automatically identify social spam.

1 Introduction

Over the past few years, social networking communities have experienced unprecedented growth, both in terms of size and popularity. In fact, of the top-20 most visited World Wide Web destinations, six are now social networks, which is five more than the list from only three years ago [2]. This flood of activity is remaking the Web into a "social Web" where users and their communities are the centers for online growth, commerce, and information sharing. Unfortunately, the rapid growth of these communities has made them prime targets for attacks by malicious individuals. Most notably, these communities are being bombarded by *social spam* [14, 24].

Some of the spam in social networking communities is quite familiar. For example, message spam within a community is similar in form and function to email spam on the wider Internet, and comment spam on social networking profiles manifests itself in a similar fashion to blog spam. Defenses against these familiar forms of spam can be easily adapted to target their social networking analogs [14]. However, other forms of social spam are new and have risen out of the very fabric of these communities. One of the most important examples of this new generation of spam is deceptive spam profiles, which attempt to manipulate user behavior. These deceptive profiles are inserted into the social network by spammers in an effort to prey on innocent community users and to pollute these communities. Although fake profiles (or fakesters) have been a "fun" part of online social networks from their earliest days [22], growing evidence suggests that spammers are deploying deceptive profiles in increasing numbers and with more intent to do harm. For example, deceptive profiles can be used to drive legitimate users to Web spam pages, to distribute malware, and to disrupt the quality of community-based knowledge by spreading disinformation [24].

Understanding different types of social spam and deception is the first step towards countering these vulnerabilities. Hence, in this paper, we propose a novel technique for harvesting deceptive spam profiles from social networking communities using *social honeypots*. Then, we provide a characterization of the spam profiles that we collected with our social honeypots. To the best of our knowledge, this is the first use of honeypots in the social networking environment as well as the first characterization of deceptive spam profiles.

Our social honeypots draw inspiration from security researchers who have used honeypots to observe and analyze malicious activity. Specifically, honeypots have already been used to characterize malicious hacker activity [21], to generate intrusion detection signatures [16], and to observe email address harvesters [19]. In our current research, we create honeypot profiles within a community to attract spammer activity so that we can identify and analyze the characteristics of social spam profiles. Concretely, we constructed 51 honeypot profiles and associated them with distinct geographic locations in MySpace, the largest and most active social networking community. After creating our social honeypots, we deployed them and collected all of the traffic they received (via friend requests). Based on a four month evaluation period from October 1, 2007 to February 1, 2008, we have conducted a sweeping characterization of the harvested spam profiles from our social honeypots. A few of the most interesting findings from this analysis are:

- The spamming behaviors of spam profiles follow distinct temporal patterns.
- The most popular spamming targets are Midwestern states, and the most popular location for spam profiles is California.
- The geographic locations of spam profiles almost never overlap with the locations of their targets.
- 57.2% of the spam profiles obtain their "About me" content from another profile.
- Many of the spam profiles exhibit distinct demographic characteristics (e.g., age, relationship status, etc.).
- Spam profiles use thousands of URLs and various redirection techniques to funnel users to a hand full of destination Web pages.

The rest of the paper is organized as follows. Section 2 summarizes related work. Section 3 provides background information about social networking communities and describes the social spam that is currently plaguing these communities. In Section 4, we present our methodology for creating social honeypots and collecting deceptive spam profiles. In Section 5, we report the results of an analysis we performed on the spam profiles that we collected in these honeypots. Section 6 concludes the paper.

2 Related Work

Due to the explosive growth and popularity of social networking communities, a great deal of research has been done to study various aspects of these communities. Specifically, these studies have focused on usage patterns [7, 11], information revelation patterns [7, 15], and social implications [8, 9] of the most popular communities. Work has also been done to characterize the growth of these communities [17] and to predict new friendships [18] and group formations [3].

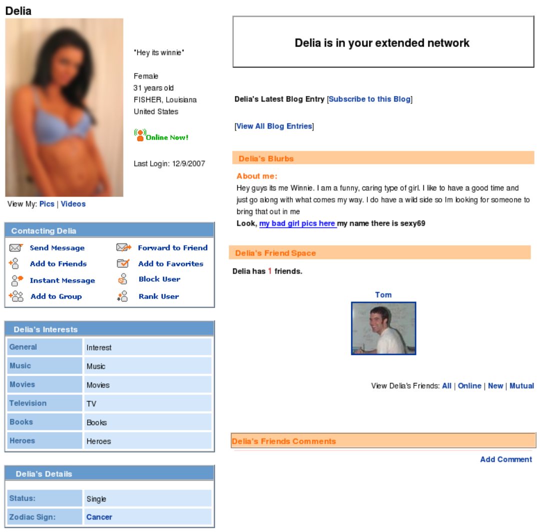
Recently, researchers have also begun investigating the darker side of these communities. For example, numerous studies have explored the privacy threats associated with public information revelation in the communities [1, 3, 4, 12]. Aside from privacy risks, researchers have also identified attacks that are directed at these communities (e.g., social spam) [14]. In our previous work [24], we showed that social networking communities are susceptible to two broad classes of attacks: traditional attacks that have been adapted to these communities (e.g., malware propagation) and new attacks that have emerged from within the communities (e.g., deceptive spam profiles).

Unfortunately, very little work has been done to address the emerging security threats in social networking communities. Heymann et al. [14] presented a framework for addressing these threats, and Zinman and Donath [25] attempted to use machine learning techniques to classify profiles. In our previous research [6], we proposed the SocialTrust framework to provide tamper-resilient trust establishment in these communities. However, the research community desperately needs real-world examples and characterizations of malicious activity to inspire new solutions. Thus, to help address this problem, we present a novel technique for collecting deceptive spam profiles in social networking communities that relies on social honeypot profiles. Additionally, we provide the first characterization of deceptive spam profiles in an effort to stimulate research progress.

3 Social Spam

Social networking communities, such as MySpace, provide an online platform for people to manage existing relationships, form new ones, and participate in various social interactions. To facilitate these interactions, a user's online presence in the community is represented by a *profile*, which is a user-controlled Web page that contains a picture of the user and various pieces of personal information. Additionally, a user's profile also contains a list of links to the profiles of that user's friends. Each of these friend links is bidirectional and established only after the user has received and accepted a *friend request* from another user.

Aside from friend requests, MySpace also provides numerous communication facilities that enable users to communicate with each other within the community (e.g., messaging, commenting, and blogging systems). Unfortunately, spammers have already begun exploiting these systems by propagating spam (e.g., message spam, comment spam, etc.) through them [24]. Even



Delia

"Hey its winnie"

Female
31 years old
FISHER, Louisiana
United States

Online Now!

Last Login: 12/9/2007

View My: Pics | Videos

Contacting Delia

Send Message
Forward to Friend
Add to Friends
Add to Favorites
Instant Message
Block User
Add to Group
Rank User

Delia's Interests

General	Interest
Music	Music
Movies	Movies
Television	TV
Books	Books
Heroes	Heroes

Delia's Details

Status:	Single
Zodiac Sign:	Cancer

Delia is in your extended network

Delia's Latest Blog Entry [Subscribe to this Blog]

[View All Blog Entries]

Delia's Blurbs

About me:

Hey guys its me Winnie. I am a funny, caring type of girl. I like to have a good time and just go along with what comes my way. I do have a wild side so Im looking for someone to bring that out in me

Look, my bad girl pics here my name there is sexy69

Delia's Friend Space

Delia has 1 friends.

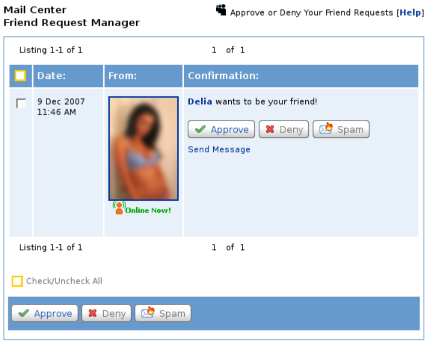
Tom

View Delia's Friends: All | Online | New | Mutual

Delia's Friends Comments

Add Comment

Figure 1: An example of a deceptive spam profile.



Mail Center
Friend Request Manager

Approve or Deny Your Friend Requests [Help]

Listing 1-1 of 1 1 of 1

□	Date:	From:	Confirmation:
□	9 Dec 2007 11:46 AM	Online Now!	Delia wants to be your friend! ✓ Approve ✗ Deny Spam Send Message

Listing 1-1 of 1 1 of 1

□ Check/Uncheck All

✓ Approve ✗ Deny Spam

Figure 2: An example of a spam friend request.

more troubling is the fact that spammers are now polluting the communities with deceptive spam profiles that aim to manipulate legitimate users.

An example of a MySpace spam profile[1] is shown in Figure 1. As the figure shows, spam profiles contain a wealth of information and various deceptive properties. Most notably, these profiles typically use a provocative image of a woman to entice users to view them. Then, once the profiles have attracted visitors, they direct those visitors to perform an action of some sort (e.g., visiting a Web page outside of the community) by using a seductive story in their "About me" sections. For example, the profile in Figure 1 provides a link to another Web page and promises that "bad girl pics" will be found there.

After spammers have constructed their deceptive profiles, they must attract visitors. To generate this traffic for their profiles, spammers typically employ two strategies. First, spammers keep their profiles logged in to MySpace for long periods of time. This strategy generates attention because many of the MySpace searching mechanisms give preferential treatment to profiles that are currently logged in to the system. Consequently, when users are browsing through profiles, the spam profiles will be prominently displayed. The second strategy is much more aggressive and involves sending friend requests to MySpace users. Figure 2 shows an example friend request that corresponds to the profile shown in Figure 1. Unlike the first strategy, which passively relies on users to visit the spam profiles, this strategy actively contacts users and deceives them into believing the profile's creator wants to befriend them.

4 Social Honeypots

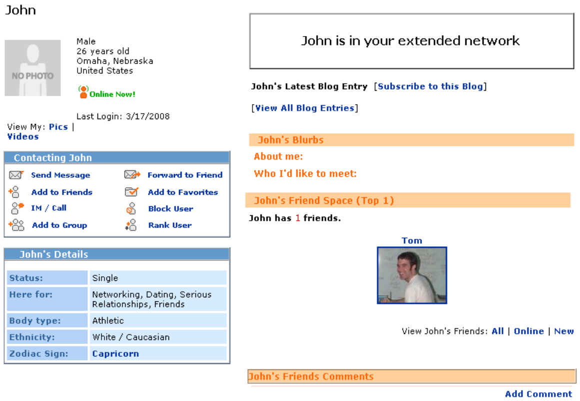
For years, researchers have been deploying honeypots to capture examples of nefarious activities [16, 19, 21]. In this paper, we utilize honeypots to collect deceptive spam profiles in social networking communities. Specifically, we created 51 MySpace profiles to serve as our social honeypots. To observe any geographic artifacts of spamming behavior, each of these profiles was given a specific geographic location (i.e., one honeypot was assigned to each of the U.S. states and Washington, D.C.). With the exception of the D.C. honeypot, each profile's city was chosen based on the most populated city in a given state. For example, Atlanta has the largest population in Georgia, and as a result, it was the city used for the Georgia honeypot. We used this strategy because we assumed that spammers would target larger cities due to their larger populations of potential victims.

All 51 of our honeypot profiles are identical except for their geographic information (see Figure 3 for an example). Each profile has the same name, gender, and birthday. Additionally, all of the demographic information was chosen to make the profiles appear attractive to spammers. Specifically, all of the profiles share the same relationship status (single), body type (athletic), and ethnicity (White / Caucasian). These demographic characteristics were also among the most popular in our previous large-scale characterization of MySpace profiles [7].

To collect timely information and increase the likelihood of being targeted by spammers, we created custom MySpace bots to ensure that all of our profiles are logged in to MySpace 24 hours a day, 7 days a week [2].

[1] All of the provocative images in the paper have been blurred so as not to offend anyone.

[2] We experienced a few short outages (on the order of hours) due to MySpace system updates, which forced us to



John

Male
26 years old
Omaha, Nebraska
United States

Online Now!

Last Login: 3/17/2008

View My: Pics | Videos

Contacting John

Send Message	Forward to Friend
Add to Friends	Add to Favorites
IM / Call	Block User
Add to Group	Rank User

John's Details

Status:	Single
Here for:	Networking, Dating, Serious Relationships, Friends
Body type:	Athletic
Ethnicity:	White / Caucasian
Zodiac Sign:	Capricorn

John is in your extended network

John's Latest Blog Entry [Subscribe to this Blog]

[View All Blog Entries]

John's Blurbs

About me:

Who I'd like to meet:

John's Friend Space (Top 1)

John has 1 friends.

Tom

View John's Friends: All | Online | New

John's Friends Comments

Add Comment

Figure 3: An example of a social honeypot.

In addition to keeping our honeypot profiles logged in to the community, our bots also monitor any spamming activity that is directed at the profiles. Specifically, the bots are constantly checking the profiles for newly received friend requests. To avoid burdening MySpace with excessive traffic (and to avoid being labeled as a spam bot), each of our bots follows a polling policy that employs random sleep timers and an exponential backoff algorithm, which fluctuates sleep times based on the current amount of spamming activity (with a minimum and maximum sleep time of five minutes and one hour, respectively). Therefore, when a honeypot profile is receiving spam, its corresponding bot polls MySpace more aggressively than when the profile is not receiving spam.

After one of our honeypot profiles receives a new friend request, the bot responsible for that profile performs various tasks. First, the bot downloads the spam profile that sent the friend request[3], storing a copy of the profile along with a honeypot-specific identifier and a timestamp that corresponds to the time when the friend request was sent to the honeypot profile. Then, after storing a local copy of the profile, the bot rejects the friend request. We decided to reject the friend requests for two reasons. First, we wanted to identify spam profiles that are repeat offenders (i.e., they continuously send friend requests until they are accepted). Second, we did not want our honeypot profiles to be mistaken as spam profiles. If we blindly accept all of the spam friend requests, our honeypot profiles will appear to be helping the spam profiles in a manner similar to a Web spam page that participates in a link exchange or link farm [13]. Thus, to avoid suspicion by MySpace, our honeypot profiles conservatively reject the friend requests that they receive.

Many of the spam profiles contain links in their “About me” sections that direct users to Web pages outside of the social networking community. We wanted to study the characteristics of these Web pages; hence, in addition to storing the spam profiles, our bots also crawl the pages that are being advertised by these profiles. Specifically, after one of our bots stores a local copy of a profile, the bot parses the profile's “About me” section and extracts its URLs. Then, the bot crawls the Web pages corresponding to those URLs, storing them along with their associated spam profile.

Not surprisingly, almost all of the URLs that are advertised by spam profiles are entrances to sophisticated redirection chains. To identify the final destinations in these chains, our bots follow every redirect. First, the bots attempt to access each of the URLs being advertised by a spam profile. If a bot encounters HTTP redirects (i.e., 3xx HTTP status codes), the bot follows them until it accesses a URL that does not return a redirect. Then, the corresponding Web page is stored and parsed for HTML/javascript redirection techniques using the redirection detection algorithm we presented in our previous research [23]. If our algorithm extracts redirection URLs, our bots attempt to access them. Once again, if the bots encounter HTTP redirects, they follow the redirects until they find URLs that do not return redirects. Finally, the corresponding Web pages are stored. After completing this process, we are left with a collection of final destination pages and the intermediary pages that were crawled along the way.

5 Social Honeypot Data Analysis

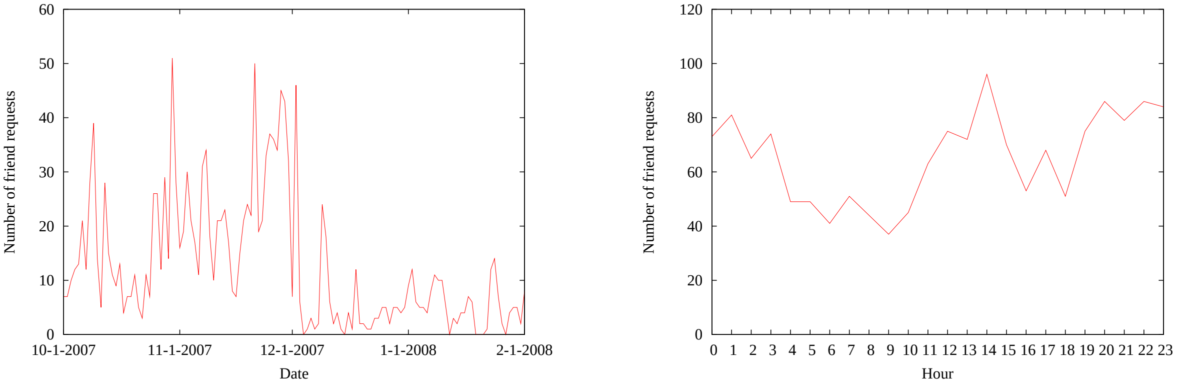
In this section, we investigate various characteristics of the 1,570 friend requests (and corresponding spam profiles) that we collected in our social honeypots during a four month evaluation period from October 1, 2007 to February 1, 2008. First, we characterize the temporal distribution of the spam friend requests that our honeypots received. Then, we analyze the geographic properties of social spam. Next, we investigate duplication in spam profiles and identify five popular groups of spam profiles. After our duplication analysis, we identify interesting demographic characteristics of spam profiles. Finally, we analyze the Web pages that are advertised by spam profiles.

5.1 Temporal Distribution of Spam Friend Requests

Since all of our honeypot profiles were constantly logged into MySpace during our four month evaluation period, we were able to observe various temporal patterns for spamming activity. In Figure 4(a), we

slightly modify our bots.

[3] Band profiles also send unsolicited friend requests; however, our bots simply reject these requests without processing them.



(a) Monthly distribution

(b) Hourly distribution

Figure 4: Temporal distributions of the spam friend requests received by our social honeypots.

present the number of friend requests that our honeypots received on each day of this four month period. This figure is interesting for a few reasons. First, we observe three peaks in spamming activity that occur around holidays. Specifically, our honeypots received the most friend requests the day before, the day of, and the day after Columbus Day (79), Halloween (95), and Thanksgiving (90). One possible explanation for these spikes is that legitimate users might spend more time online during these holiday periods, giving spammers a larger audience for their deceptive profiles.

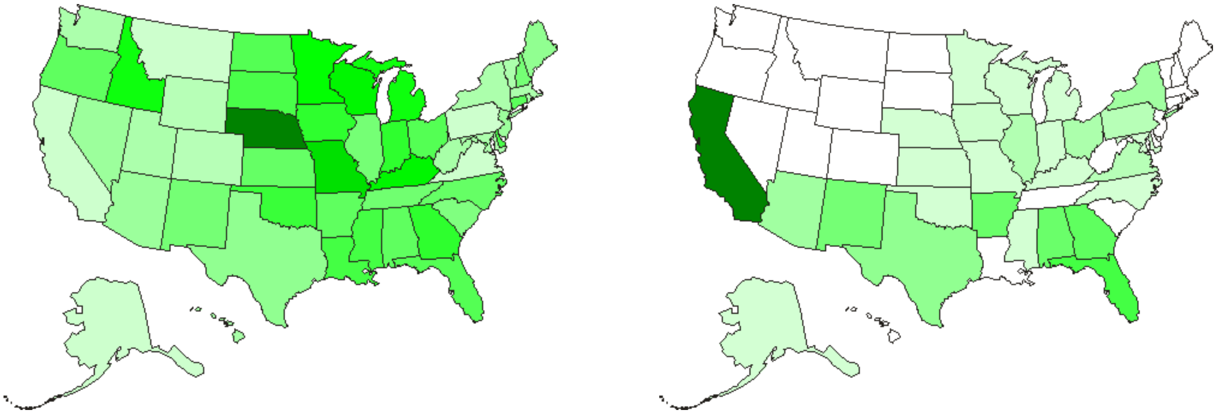
Another intriguing observation from Figure 4(a) is that our honeypots began receiving significantly fewer friend requests after December 2. In fact, of the 1,570 friend requests that our honeypots received, only 299 (19.0%) of them were received after this date. We are still investigating the reasons behind this reduced activity, but one hypothesis is that spammers realized the underlying purpose of our honeypot profiles. Since all of our honeypots reject friend requests after the corresponding spam profiles have been stored, spammers should eventually recognize that each of the honeypots represents a wasted friend request. As we explained in Section 4, we decided to reject spam friend requests because we wanted to avoid having our honeypots labeled as spam by MySpace. As part of our ongoing research, we are revisiting this decision to investigate whether it affects the spamming activity we observe.

To analyze finer-grained temporal patterns, Figure 4(b) shows the hourly distribution of the friend requests that our honeypots received[4]. As the figure shows, for every hour of the day, our honeypots received at least 35 friend requests from spammers. Additionally, distinct hourly patterns emerge from the figure. Most notably, spamming activity is at its peak around 2pm and from 10pm to 1am, and it is at its lowest levels between 4am and 9am. These patterns are particularly interesting because they mirror previous results about the communication patterns of legitimate users in social networking communities [11]. The similarities between legitimate and spam activity patterns are somewhat intuitive for at least two reasons. First, spammers want to be active when their targets are active because they want to increase the chances of successfully deceiving those users. Second, by blending their traffic in with legitimate traffic, spammers reduce the risk of being identified by the operators of these communities.

5.2 Geographic Distribution of Spam Friend Requests

Since each of our honeypot profiles claims to be in a unique geographic location (i.e., one honeypot is in each of the fifty U.S. states and Washington, D.C.), we are able to analyze the geographic properties of spamming behavior. Figure 5(a) shows a color-coded map of the United States, which represents the relative popularity of our geographically dispersed honeypots. States with darker shades of green represent honeypots that received more friend requests than the honeypots in states with lighter shades of green. As the figure shows, a large fraction of the spamming activity was

[4] All of the times are normalized based on the time zone that corresponds to the honeypot profile's location.



(a) Target distribution

(b) Originator distribution

Figure 5: Geographic distributions of spam profiles and their targets.

directed at the Midwestern states. In fact, the five most popular targets were the honeypots in Omaha, Nebraska (80 friend requests), Kansas City, Missouri (58 friend requests), Milwaukee, Wisconsin (56 friend requests), Louisville, Kentucky (56 friend requests), and Minneapolis, Minnesota (53 friend requests). In our previous research [7], we found that MySpace users from Midwestern states began using MySpace considerably later than users from Western states because MySpace was founded in California. As a result, one explanation for why Midwestern users are frequently targeted by spammers is that those users might be less MySpace-savvy and thus a more attractive target for deception by spammers. This hypothesis is also supported by the fact that our Los Angeles, California honeypot received fewer friend requests (10) than any of the other honeypots.

Figure 5(b) shows another color-coded map of the United States. However, unlike Figure 5(a), this figure shows the relative popularity of various states as locations for spam profiles. States with darker shades of green have more spam profiles affiliated with them than states with lighter shades of green. As the figure shows, the most popular locations for the spam profiles were California and Southeastern states. Specifically, the five most popular states for spam profiles were California (186 spam profiles), Florida (92 spam profiles), Georgia (78 spam profiles), Arkansas (74 spam profiles), and Alabama (73 spam profiles).

After investigating the most popular originating and target locations for spam profiles, an obvious question is how much overlap (if any) exists for those locations. Concretely, we wanted to know how often the declared location of a spam profile matches the declared location of a targeted profile. Our original hypothesis was that we would identify a significant number of matches because we believed victims might be hesitant to accept a friend request from someone outside of their city or state. However, we were surprised to find that 1,534 (97.7%) of the friend requests were from spam profiles that reported a location that did not match the city or state associated with the honeypot profile that received them.

One explanation for this disconnect between the locations of spam profiles and their victims is that a clear tension exists between increasing the deceptive properties of a spam profile and making the profile broadly applicable to a large number of potential victims. Obviously, spammers would prefer to create personalized spam profiles for every potential victim because that would greatly increase the likelihood of a successful deception. However, it is costly to create personalized profiles for every potential victim, and as a result, spammers focus on casting as wide a net as possible.

5.3 Spam Profile Duplication

While investigating the geographic distribution of the friend requests that our honeypot profiles received, we noticed that many of our honeypots received a friend request from the same spam profile. In fact, 65 spam profiles sent a friend request to more than one of our honeypots, generating a total of 148 friend requests. 40 (78.4%) of our honeypots received at least one of these duplicate friend requests, and the honeypots that received the most friend requests (i.e., the Omaha, Nebraska honeypot and the Kansas City, Missouri honeypot) also received the most duplicates (11 duplicates and 8 duplicates, respectively). Surprisingly, none of our honeypots received more than one friend request from a given spam profile (i.e., none of the spam profiles were repeat offenders). Thus, after one of our honeypots rejected a spam profile's friend request, that profile was intelligent enough not to send the honeypot another friend request.

After we identified the existence of duplicate friend requests, we wanted to determine how much lag time (if any) exists between the first arrival of a friend request and the arrivals of its duplicates. To quantify the delays between duplicate friend requests, we created 65 time series – one for each set of the duplicate friend requests. Then, for each time series, we computed the size of the time window that includes the first and last point. Based on this analysis, we found that 63 (96.9%) of the time windows close in less than 4 minutes, and 53 (81.5%) of the time windows close in less than a minute. Therefore, when these spam profiles sent friend requests, they sent a large number of them in a short period of time (i.e., they were not particularly stealthy).

Once we determined the number of unique profiles in our collection (1,487), we wanted to know how many of those profiles possess content that is a duplicate (or a near-duplicate) of another profile's content. In our previous work [23], we found that only one-third of Web spam pages are unique, and we wanted to determine if the same level of duplication exists among spam profiles. To quantify the amount of content duplication in our collection of 1,487 unique spam profiles, we used the shingling algorithm from our previous work [23] on all of their HTML content to construct equivalence classes of duplicate and near-duplicate profiles.

First, we preprocessed each profile by replacing its HTML tags with white space and tokenizing it into a collection of words (where a word is defined as an uninterrupted series of alphanumeric characters). Then, for every profile, we created a fingerprint for each of its n words using a Rabin fingerprinting function [20] (with a degree 64 primitive polynomial p_A). Once we had the n word fingerprints, we combined them into 5-word phrases. The collection of word fingerprints was treated like a circle (i.e., the first fingerprint follows the last fingerprint) so that every fingerprint started a phrase, and as a result, we obtained n 5-word phrases. Next, we generated n phrase fingerprints for the n 5-word phrases using a Rabin fingerprinting function (with a degree 64 primitive polynomial p_B). After we obtained the n phrase fingerprints, we applied 84 unique Rabin fingerprinting functions (with degree 64 primitive polynomials p_1, ..., p_{84}) to each of the n phrase fingerprints. For every one of the 84 functions, we stored the smallest of the n fingerprints, and once this process was complete, each spam profile was reduced to 84 fingerprints, which are referred to as that profile's *shingles*. Once all of the profiles were converted to a collection of 84 shingles, we clustered the profiles into equivalence classes (i.e., clusters of duplicate or near-duplicate profiles). Two profiles were considered duplicates if all of their shingles matched, and they were near-duplicates if their shingles agreed in two out of the six possible non-overlapping collections of 14 shingles. For a more detailed description of this shingling algorithm, please consult [5, 10].

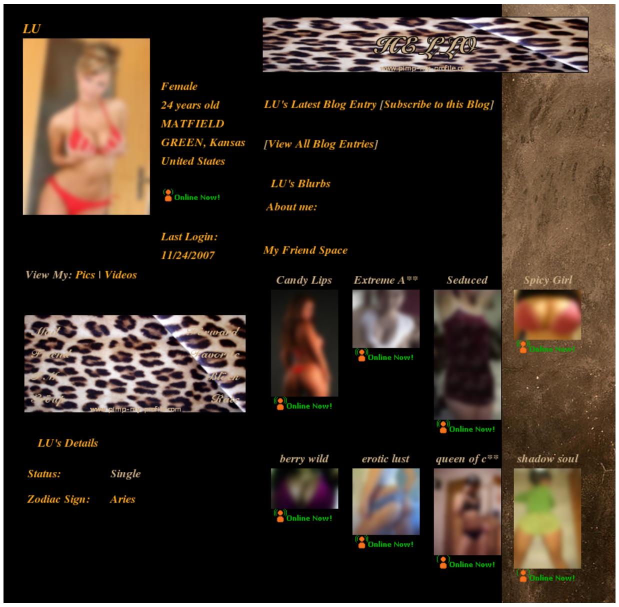
After this clustering was complete, we were left with 1,261 unique clusters of duplicate and near-duplicate profiles. Thus, only 226 (15.2%) of the profiles have the same (or nearly the same) HTML content as one of the remaining 1,261 profiles. This level of duplication is significantly less than what we observed with Web spam pages; however, we do not believe this is an accurate measure of spam profile duplication. Since most of a spam profile's deceptive text is found in the "About me" section, a more reasonable metric for profile duplication is actually "About me" duplication. Hence, in addition to running our shingling algorithm over all of the HTML content in a profile, we also extracted the "About me" content and built equivalence classes using that data. This "About me" clustering generated 637 unique clusters, which means 850 (57.2%) of the profiles have the same (or nearly the same) "About me" content as one of the remaining 637 profiles.

Based on the results of our content duplication analysis, we can conclude that duplication among spam profiles is on par with duplication among Web spam pages. 15.2% of the spam profiles obtain all of their HTML content from another profile, and 57.2% of the spam profiles obtain their "About me" content from another profile. This observation is quite encouraging because it implies that the problem of identifying all spam profiles can actually be reduced to the problem of identifying a much smaller set of unique profiles.

5.4 Spam Profile Examples

After we completed our content duplication analysis, we manually investigated the profiles in our various clusterings. Based on this investigation, we discovered that most of our spam profiles fall into one of five categories:

- **Click Traps**: Each profile contains a background image that is also a link to another Web page. If users click anywhere on the profile, they are immediately directed to the link's corresponding Web site. One of the most popular (and most deceptive) examples displays a fake list of friends, which is actually a collection of provocative images that direct users to a nefarious Web page (see Figure 6 for an example).



LU

Female
24 years old
MATFIELD
GREEN, Kansas
United States

Online Now!

Last Login:
11/24/2007

View My: Pics | Videos

LU's Details

Status: Single

Zodiac Sign: Aries

LU's Latest Blog Entry [Subscribe to this Blog]

[View All Blog Entries]

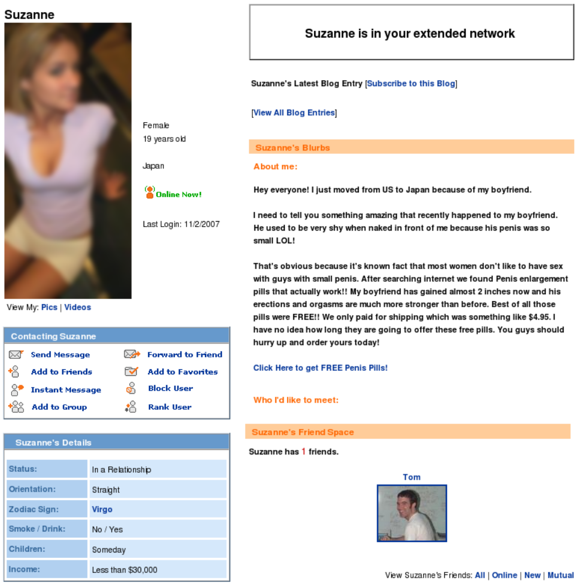
LU's Blurbs

About me:

My Friend Space

Candy Lips
Online Now!
Extreme A**
Online Now!
Seduced
Online Now!
Spicy Girl
Online Now!
berry wild
Online Now!
erotic lust
Online Now!
queen of c**
Online Now!
shadow soul
Online Now!

Figure 6: An example of a Click Trap.



Suzanne

Female
19 years old

Japan

Online Now!

Last Login: 11/2/2007

View My: Pics | Videos

Contacting Suzanne

Send Message	Forward to Friend
Add to Friends	Add to Favorites
Instant Message	Block User
Add to Group	Rank User

Suzanne's Details

Status:	In a Relationship
Orientation:	Straight
Zodiac Sign:	Virgo
Smoke / Drink:	No / Yes
Children:	Someday
Income:	Less than $30,000

Suzanne is in your extended network

Suzanne's Latest Blog Entry [Subscribe to this Blog]

[View All Blog Entries]

Suzanne's Blurbs

About me:

Hey everyone! I just moved from US to Japan because of my boyfriend.

I need to tell you something amazing that recently happened to my boyfriend. He used to be very shy when naked in front of me because his penis was so small LOL!

That's obvious because it's known fact that most women don't like to have sex with guys with small penis. After searching internet we found Penis enlargement pills that actually work!! My boyfriend has gained almost 2 inches now and his erections and orgasms are much more stronger than before. Best of all those pills were FREE!! We only paid for shipping which was something like $4.95. I have no idea how long they are going to offer these free pills. You guys should hurry up and order yours today!

Click Here to get FREE Penis Pills!

Who I'd like to meet:

Suzanne's Friend Space

Suzanne has 1 friends.

Tom

View Suzanne's Friends: All | Online | New | Mutual

Figure 7: An example of a Japanese Pill Pusher.

- **Friend Infiltrators**: These profiles do not have any overtly deceptive elements (aside from their images – and even those are innocuous in some cases). The purpose of the profiles is to befriend as many users as possible so that they can infiltrate the users' circles of friends and bypass any communication restrictions imposed on non-friends. Once a user accepts a friend request from one of these profiles, the profile begins spamming that user through every available communication system (e.g., message spam, comment spam, etc.).

- **Pornographic Storytellers**: Each of these profiles has an "About me" section that consists of randomized pornographic stories, which are bookended by links that lead to pornographic Web pages. The anchor text used in these profiles is extremely similar, even though the rest of the "About me" text is almost completely randomized.

- **Japanese Pill Pushers**: These profiles contain a sales pitch for male enhancement pills in their "About me" sections. According to the pitch, the attractive woman pictured in the profile has a boyfriend that purchased these pills at an incredible discount, and if you act now, you can do the same. An example is shown in Figure 7.

- **Winnies**: All of these profiles have the same headline: "Hey its winnie." However, despite this headline, none of the profiles are actually named "Winnie." In addition to a shared headline, each of the profiles also includes a link to a Web page where users can see the pictured female's pornographic pictures. An example of one of these profiles was shown in Section 3 (Figure 1).

5.5 Spam Profile Demographics

In our content duplication analysis, we analyzed the HTML and "About me" sections of spam profiles in a general sense. To observe more specific features of these profiles, we investigated demographic characteristics of the 1,487 spam profiles that we captured in our honeypots. These characteristics include traditional demographics (e.g., age, gender, etc.) as well as profile-specific features (e.g., number of friends, headlines, etc.).

Our first observation from this demographic analysis is that many of the spam profiles share various demographic characteristics. Specifically, all of the profiles are female and between the ages of 17 and 34 (85.9% of the profiles state an age between 21 and 27). Additionally, 1,476 (99.3%) of the profiles report that they are single. None of these characteristics are particularly surprising because they all reinforce the deceptive nature of these profiles. Specifically, these demographic features make each profile appear as though it was created by a young, "available" woman.

Our second observation is that many spam profiles include additional personal information to enhance their deceptive properties. The profiles that are most adept at leveraging personal information to their advantage are the Japanese Pill Pushers. These profiles are the only ones that claim to be in a relationship, but this

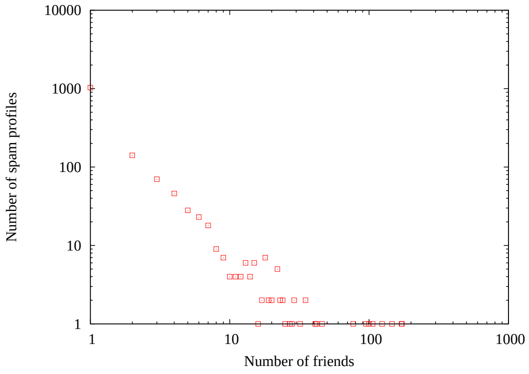


Figure 8: Distribution of the number friends associated with spam profiles.

relationship status is warranted because the profiles mention a boyfriend in their "About me" sections. Additionally, these profiles list "less than $30,000" as their annual income. This is the lowest allowable option on MySpace, and as a result, this annual income value makes it seem like the profile's creator is not particularly wealthy, which reinforces the affordability of the male enhancement pills that these profiles are advertising.

Our final observation is that many of the spam profiles successfully befriended legitimate users. Figure 8 shows the distribution of friends associated with each of the spam profiles. It is important to note that this distribution is skewed towards the low end of the spectrum because our bots visited and stored the spam profiles very quickly (and potentially before any other users had a chance to accept the spam friend requests). However, despite this fact, 455 (30.6%) of the profiles had more than one friend when our bots collected them. Thus, almost a third of the spam profiles were already attracting victims when our bots visited them.

5.6 Advertised Web Pages

Since the purpose of spam profiles is to deceive users into performing an action (e.g., visiting a Web page), most of the profiles contain links to Web pages outside of the community. Specifically, 1,245 (83.7%) of the profiles contain at least one link in their "About me" sections. The remaining 242 profiles are all examples of Friend Infiltrators, and as a result, they postpone their promotional activities until after they have befriended users.

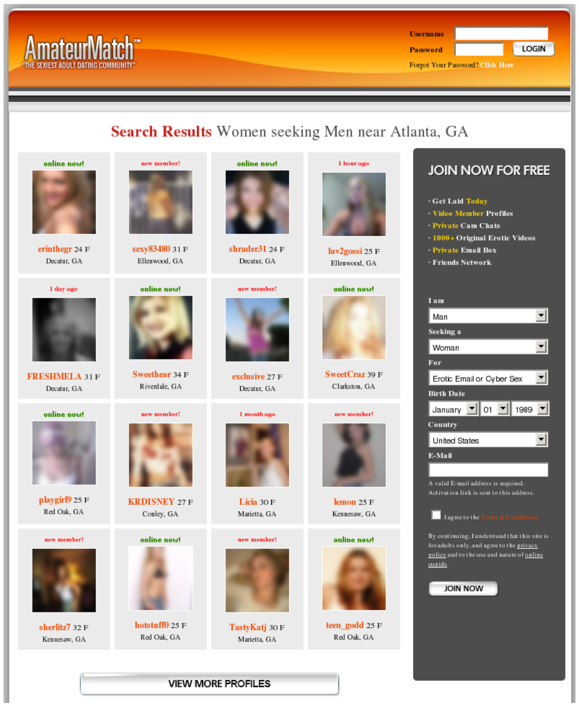
From the 1,245 profiles that contain links, our bots were able to extract and successfully access 1,048 *profile URLs*. Of these 1,048 profile URLs, only 482 (46.0%) of them were unique, which means more than half of the URLs that appear in spam profiles are duplicates. When our bots attempted to crawl the profile URLs, 339 (32.3%) of them returned a total of 657 HTTP redirects. After following these HTTP redirect chains, our original 482 unique profile URLs funneled our bots to only 148 unique destination URLs. Therefore, of the 1,048 Web pages that our bots ultimately obtained with the profile URLs, 900 (85.9%) of them have duplicate URLs.

To investigate this duplication even further, we performed a shingling analysis on the HTML content of these 1,048 Web pages. Based on this analysis, we discovered only 6 unique clusters of duplicate and near-duplicate Web pages. Thus, 1,042 (99.4%) of the Web pages contain content that was duplicated from the other 6 Web pages. Three of these clusters, which account for 93.3% of the pages, contain pages that act as intermediary redirection pages (i.e., the pages immediately redirect users using HTML/javascript redirection techniques). Two of the clusters, which account for 6.6% of the pages, contain pornographic Web pages, and the last cluster contains a single Web page, which executes a phishing attack against MySpace.

Since 93.3% of the pages employ redirection techniques, we parsed those pages for HTML/javascript redirects using our redirection detection algorithm from previous research [23]. Based on this redirection analysis, we identified redirects that use HTML meta refresh tags, javascript location variable assignments, and HTML iframe tags. In total, our algorithm identified 1,307 *redirection URLs*. However, of those 1,307 URLs, only 136 (10.4%) of them were unique; hence, over 90% of the redirection URLs are duplicates.

When our bots crawled these redirection URLs, 959 (73.4%) of them returned a total of 1,288 HTTP redirects. After following these HTTP redirect chains, our bots were eventually funneled to only 15 unique URLs. Thus, of the 1,307 Web pages that our bots crawled using the redirection URLs, only 15 (1.1%) of them have unique URLs. Even more striking is the fact that only 5 domain names are used in those 15 unique URLs, and of those 5 domain names, `fling.com` and `amateurmatch.com` Web pages account for 975 (74.6%) of the Web pages. An example of one of the `amateurmatch.com` Web pages is shown in Figure 9.

Based on the results of our Web page analysis, we can conclude that all of the URLs that are advertised in spam profiles point to an extremely small number of destination pages. Specifically, 1,048 profile URLs funneled our bots to only 6 destinations, and 1,307 redirection URLs funneled our bots to only 5 destina-



Search Results Women seeking Men near Atlanta, GA

online now! erinthegr 24 F Decatur, GA	new member! sexy83480 31 F Ellenwood, GA	online now! shrader31 24 F Decatur, GA	1 hour ago luv2gossi 25 F Ellenwood, GA
1 day ago FRESHMELA 31 F Decatur, GA	online now! Sweethear 34 F Riverdale, GA	new member! exclusive 27 F Decatur, GA	online now! SweetCraz 39 F Clarkston, GA
online now! playgirl9 25 F Red Oak, GA	new member! KRDISNEY 27 F Conley, GA	1 month ago Licia 30 F Marietta, GA	new member! lemon 25 F Kennesaw, GA
new member! sherlitz7 32 F Kennesaw, GA	online now! hotstuff0 25 F Red Oak, GA	new member! TastyKatj 30 F Marietta, GA	online now! teen_godd 25 F Red Oak, GA

I am: Man
Seeking a: Woman
For: Erotic Email or Cyber Sex
Birth Date: January 01 1989
Country: United States
E-Mail:

A valid E-mail address is required. Activation link is sent to this address.

□ I agree to the Terms & Conditions

By continuing, I understand that this site is for adults only, and agree to the privacy policy and to the use and nature of online cupids.

Figure 9: An example of a Web page that is advertised by a spam profile.

tions. This observation is quite valuable because it significantly reduces the problem of identifying the Web pages that are advertised by spam profiles. Instead of dealing with 2,355 URLs, we must simply identify 11 destinations.

6 Conclusions

In this paper, we presented a novel technique for automatically collecting deceptive spam profiles in social networking communities. Specifically, our approach deploys honeypot profiles and collects all of the spam profiles associated with the spam friend requests that they receive. We also provided the first characterization of deceptive spam profiles using the data that we collected in our 51 social honeypots over a four month evaluation period. This characterization covered various topics including temporal and geographic distributions of spamming activity, content duplication, an analysis of profile demographics, and an evaluation of the Web pages that are advertised by spam profiles. In our ongoing work, we are using our analysis results to automatically identify social spam.

References

[1] A. Acquisti and R. Gross. Imagined communities: Awareness, information sharing, and privacy on the facebook. In *Proc. of PET '06*, 2006.

[2] Alexa. Alexa top 500 sites. `http://www.alexa.com/site/ds/top_sites?ts_mode=global`, 2008.

[3] L. Backstrom, C. Dwork, and J. Kleinberg. Wherefore art thou r3579x?: Anonymized social networks, hidden patterns, and structural steganography. In *Proc. of WWW '07*, 2007.

[4] D. Boyd. Social network sites: Public, private, or what? *The Knowledge Tree: An e-Journal of Learning Innovation*, 2007.

[5] A. Broder et al. Syntactic clustering of the web. In *Proc. of WWW '97*, 1997.

[6] J. Caverlee, L. Liu, and S. Webb. Socialtrust: Tamper-resilient trust establishment in online communities. In *Proc. of JCDL '08*, 2008.

[7] J. Caverlee and S. Webb. A large-scale study of myspace: Observations and implications for online social networks. In *Proc. of ICWSM '08*, 2008.

[8] J. Donath and D. Boyd. Public displays of connection. *BT Technology Journal*, 22(4), 2004.

[9] N. Ellison, C. Steinfield, and C. Lampe. Spatially bounded online social networks and social capital: The role of facebook. In *Proc. of ICA '06*, 2006.

[10] D. Fetterly, M. Manasse, and M. Najork. Detecting phrase-level duplication on the world wide web. In *Proc. of SIGIR '05*, 2005.

[11] S. A. Golder, D. Wilkinson, and B. A. Huberman. Rhythms of social interaction: Messaging within a massive online network. In *Proc. of CT '07*, 2007.

[12] R. Gross and A. Acquisti. Information revelation and privacy in online social networks. In *Proc. of WPES '05*, 2005.

[13] Z. Gyöngyi and H. Garcia-Molina. Link spam alliances. In *Proc. of VLDB '05*, 2005.

[14] P. Heymann, G. Koutrika, and H. Garcia-Molina. Fighting spam on social web sites: A survey of approaches and future challenges. *IEEE Internet Computing*, 11(6), 2007.

[15] S. Hinduja and J. W. Patchin. Personal information of adolescents on the internet: A quantitative content analysis of myspace. *Journal of Adolescence*, 31(1), 2008.

[16] C. Kreibich and J. Crowcroft. Honeycomb: Creating intrusion detection signatures using honeypots. *ACM SIGCOMM Computer Communication Review*, 34(1), 2004.

[17] R. Kumar, J. Novak, and A. Tomkins. Structure and evolution of online social networks. In *Proc. of KDD '06*, 2006.

[18] C. Lampe, N. Ellison, and C Steinfield. A familiar face(book): Profile elements as signals in an online social network. In *Proc. of CHI '07*, 2007.

[19] M. Prince et al. Understanding how spammers steal your e-mail address: An analysis of the first six months of data from project honey pot. In *Proc. of CEAS '05*, 2005.

[20] M. Rabin. Fingerprinting by random polynomials. Technical Report TR-15-81, Center for Research in Computing Technology, Harvard University, 1981.

[21] L. Spitzner. The honeynet project: Trapping the hackers. *IEEE Security & Privacy*, 1(2), 2003.

[22] R. Wade. Fakesters: On myspace, you can be friends with burger king. this is social networking? `http://www.technologyreview.com/Infotech/17713/`.

[23] S. Webb, J. Caverlee, and C. Pu. Characterizing web spam using content and http session analysis. In *Proc. of CEAS '07*, 2007.

[24] S. Webb, J. Caverlee, and C. Pu. Granular computing system vulnerabilities: Exploring the dark side of social networking communities. *Encyclopedia of Complexity and System Science (to appear)*, 2008.

[25] A. Zinman and J. Donath. Is britney spears spam? In *Proc. of CEAS '07*, 2007.