[原文刊載於 RUN!PC 旗標資訊月刊 2008 年 2 月號]

無名小站中真實姓名洩露的危機

林英發 陳寬達（陳昇瑋）

中央研究院資訊科學研究所

作者簡介

林英發

中央研究院資訊科學研究所研究助理，研究方向為 Web 2.0 隱私及安全議題。

陳寬達（陳昇瑋）

中央研究院資訊科學研究所助研究員，研究方向為網路量測、網路安全及線上遊戲。網址: http://www.iis.sinica.edu.tw/~swc

摘要

近年來，社群網路服務（Social Networking Services, 簡稱 SNSs）非常流行。為了參與線上的社交活動，使用者通常會把他們的個人資料、相片及影像等資訊上傳到網路。這可能會在無意間洩露了個人隱私 - 因為我們不知道除了自己的朋友外，還有誰會看到我們的資料。即使有些使用者不在社群網路上公開自己的個人資訊，有心人士還是有可能從他們的交友資訊或朋友對他們的描述中推測出使用者的性別、興趣、職業甚至姓名等等。

在本篇論文中，我們測量及分析在社群網路服務中使用者真實姓名被他人洩露的可能性。我們從國內最龐大的社群網路服務 - 無名小站[1] - 收集使用者資料並進行分析，

[1] http://www.wretch.cc

發現無名小站所提供的「好友描述」功能可能導致使用者真實姓名在非自願的情況之下被洩露。在我們蒐集到的樣本中，有 30% 使用者的全名以及 72% 使用者的名字（不包括姓）可以從本篇論文所提出的演算法推測而得，我們同時在本篇論文中討論防治此問題的幾種可行方案。

關鍵詞：網路測量、線上社群網路、網路隱私、資訊洩露

1. 引言

社群網路服務（Social Networking Services, 簡稱 SNSs）例如Myspace[2]、Facebook[3]、Flickr[4]、Orkut[5]以及Yahoo! 360[6]等，被歸納為Web 2.0 服務[7]，它們提供平台讓使用者在網路上介紹自己、認識朋友以及跟其他使用者互動。這些服務近年來非常流行並吸引大量的使用者參與；但，隨著社群網路服務迅速成長 [15]，大量的個人資料也史無前例地被使用者上傳到網際網路。基於網際網路的開放性，已上傳的個人資料很可能就變成了全球分享的資料。因此，社群網路服務使用者的隱私問題如今已成為學術界關注的議題 [7]。

在網路上公開的個人資料很可能會被駭客或惡意使用者利用，例如使用在垃圾郵件（spamming）、網路釣魚（phishing）及惡意追蹤（stalking）等。一般來說，若一封郵件包含接收者的真實姓名，往往能提高它的可信程度，同時較不會遭垃圾郵件過濾器刪除。當使用者將他們的個人資料公開到網路上後，惡意攻擊者就可能會從社群網路服務中查詢到使用者的姓名、生日及職業等資料，再把這些個人資料應用在寄給受害者的詐騙信件或網頁裡。因為包含個人資料的詐騙信件更容易取得受害者的信任 [21]，加上使用者傾向在社群網路上公開個人資料，社群網路服務的普及可能會使網路詐騙的情況更加嚴重。

在本篇論文中，我們研究社群網路服務中非自願姓名洩露（involuntary name leakage）的程度及成因。我們以國內用戶數最多的社群網路服務—無名小站作為研究案例。雖然無名小站的使用者大部分都是匿名參與，我們的分析卻發現，在無名小站上許多使用者的真實姓名可以從朋友之間的相互描述推測而得。儘管使用者沒有在個人資料公開其真實姓名，有心人士還是可能從其朋友對於他（她）的描述中推測而得。

---

2 http://www.myspace.com

3 http://www.facebook.com

4 http://flickr.com

5 http://www.orkut.com

6 http://360.yahoo.com

7 http://en.wikipedia.org/wiki/Web 2

為了量化非自願姓名洩露的程度，我們由無名小站下載 70 餘萬使用者的個人資料與交友狀況為樣本，並提出一套演算法，根據使用者的朋友對他們的描述來推測其真實姓名。**我們的結果顯示，在蒐集的樣本中，至少 78% 的使用者遭受非自願姓名洩露的風險。**其中，38%使用者的全名（連名帶姓）遭受洩露可能，其餘的 62% 使用者的名字（不含姓）亦可能遭受洩露。我們同時分析非自願姓名洩露的原因，發現許多使用者習慣在網路上以真實姓名稱呼他們的朋友。針對這項問題，我們討論三個可能的解決方案，包含提供個人隱私性設定、提供瀏覽限制範圍設定以及取得使用者的確認。

接下來的部分，在第二節我們討論社群網路服務的相關研究，在第三節中我們說明資料收集的方法以及簡單分析蒐集到的資料。在第四節我們提出推測使用者真實姓名的方法以及分析結果；在第五節中，我們討論姓名洩露的成因及其可能引發的資安風險。最後，我們在第六節作出結論。

2. 相關研究

社群網路服務近年的快速成長吸引了不少的學者以及媒體的注意 [9]，而社群網路服務也提供前所未有的機會讓研究者大規模地分析線上社群網路（online social network）的性質及其圖形結構（graph structure）。Ahn 等研究者分析線上社群網路的結構，並找出真實社群網路（off-line social network）與線上社群網路之間的一致性 [2]。Alan 等研究者對數個線上社群網路在固定時間點的圖型結構進行大規模分析，證明線上社群網路擁有真實社群網路的圖形特性 [16]。Kumar 等研究者分析兩個線上社群網路的結構演化，發現了線上社群網路的高連結核心（highly-connective core）以及星形結構（star structure）[15]。OMurchu 等學者對於流行的線上社群及商業社群作出評價、分析及比較 [17]。

使用者在社群網路上的互動及相互關係亦吸引不少學者的關注。Boyd 從人因（human factor）角度分析社群網路服務，發現社群網路服務的使用者彼此不需要認識或信任就可以互相設為朋友 [3]。此外，研究者也發現線上社群網路中使用者之間的連結關係比在真實社群網路中來得弱 [9]。

Gross 等學者從 Facebook.com 網站收集 4,000 個學生的個人資料，並分析線上社群網路中個人隱私洩露的危險性。他們發現網路使用者普遍 「慷慨地」提供個人資訊，只有極少數的使用者更改系統所提供的資料保密性設定 [9]。

3. 資料收集

3.1 無名小站

無名小站（www.wretch.cc）成立於 1999 年，現在由雅虎台灣所經營。它吸引了

約 392 萬的使用者，是國內用戶數最多的社群網路服務[8]。無名小站提供整合式的相簿、個人部落格、BBS系統（bulletin board system）、影音串流及留言板等服務。它提供的基本服務是免費的，但使用者可以選擇成為付費的「金卡VIP會員」或「銀卡VIP會員」以獲得更大的儲存空間以及更多的功能。免費使用者的好友名單上限人數為 50 人，而付費的使用者的好友人數上限分別是 100 人（銀卡）及 500 人（金卡)[9]。

在無名小站中，使用者的朋友關係有兩種，分別是「加我好友」(incoming friends, 或連入朋友）及「我的好友」(outgoing friends, 或連出朋友)。若 A 設 B 為好友，則 B 為 A 的「我的好友」，而 A 為 B 的「加我好友」。在無名小站中，使用者設定其他使用者為朋友不需要經過該使用者的同意，亦即，A 設 B 為好友不需要經過 B 的同意。

3.2 資料採集

我們撰寫一支網頁自動分析 / 下載程式（crawler）來抓取無名小站使用者的「名片」網頁（個人資訊）以及「好友」網頁（交友資訊），然後我們透過分析網頁 HTML 原始碼的方式來取得所需資訊。由於使用者個人資訊中大部分資料欄位都無格式限定（i.e., free-formed），我們利用關鍵字及範圍限定 (limited range) 來驗證它們的值。例如在驗證「體重」欄位時，我們會檢查欄位資料是否為數字並且介於 10 到 300 之間。對於無法驗證或未通過格式檢查的欄位資料我們都視它們為空值(等於使用者沒有填)。

我們首先從無名小站的主頁取得 10 個熱門的使用者帳號，從這些帳號開始，透過使用者之間的好友連結取得無名小站的社群網路。下載使用者資料時，程式也會從他們的好友名單取得其他使用者的帳號，如果發現未曾下載的帳號，程式就會把它們儲存到等候列表 (waiting list) 中。我們的程式會從等候列表隨機取得帳號，將之由列表中刪除，進行下一步的資料擷取，並持續資料擷取的動作直到等候列表清空為止。

表 1 資料概況

無名小站使用者樣本資料	
使用者數	766,972 (20%)
有效使用者數	592,548 (15%)
總連結數	7,619,212
平均用戶連結數	11.5

從 2007 年 9 月到 11 月，我們共下載了 766,972 份使用者資料，約無名小站總用戶量的 20%。因為其中有不少使用者只註冊帳戶卻沒有實際參與社交活動，我們只選擇其中的「有效使用者」，即同時最少有 1 個連入朋友和 1 個連出朋友的使用者，進行分析。在我們的樣本中，有效使用者的總數為 592,548，約無名小站使用者數的 15%，

[8] http://www.wretch.cc/ad/
[9] http://www.wretch.cc/blog/WretchFAQ&article id=6614002

共包含 7,619,212 個朋友連結。表 1 是無名小站樣本的資料概況，每個使用者的平均朋友數目為 11.5。

圖 1 使用者性別及年齡分佈

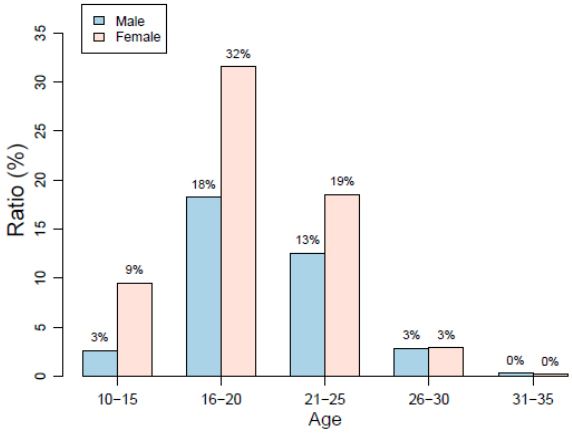
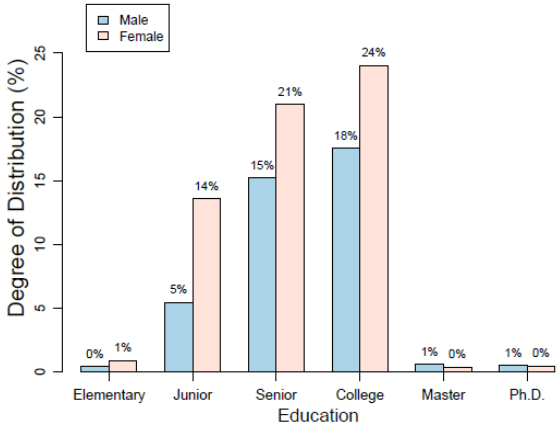


圖 2 使用者性別及學歷分佈



3.3 使用者人口統計

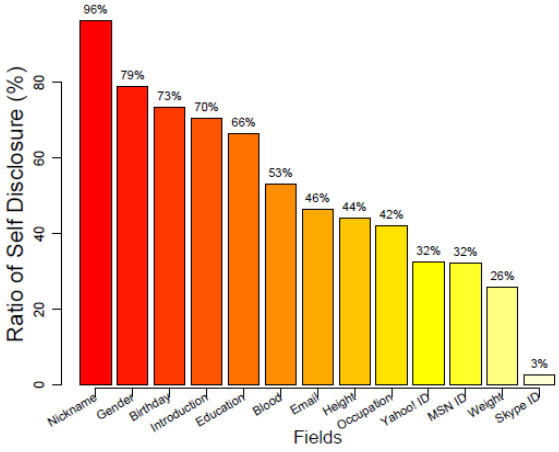
在採集到的無名小站資料中，有 28% 的使用者是男性，51% 的使用者是女性，剩餘的 21%使用者沒有公開性別。從圖 1 中我們可以看到大部分的使用者（82%）年齡分佈在 16 至 26 歲之間。圖 2 是使用者的性別與學歷（樣本的 34%），使用者公開的學歷大部分為高中職到大學。此外，93% 有公開其職業的使用者（佔樣本的 42%），宣稱他們的職業為學生。

3.3.1 自我資訊透露

自我透露（self-disclosure）是「告訴別人一些他們之前不知道的訊息，而這些訊息則成為了彼此共享的消息」[13]。它的功能為加強人與人之間的了解、建立信任、強化人與人之間的關係以及促成朋友或愛情的關係 [11]。圖 3 呈現無名小站樣本中使用者自我資訊透露的狀況，幾乎每位使用者都會填寫暱稱（96%），性別（79%）跟生日

（73%）也非常普遍的被公開。在其他生理相關的資訊中，有不少使用者公開血型跟身高（分別為 53%與 44%），體重卻是最少使用者公開的資訊（26%）。

圖 3 使用者資訊透露情況



3.3.2 自我透露程度

我們定義自我透露程度（degree of self-disclosure, 簡稱 DSD）來量化使用者的自我資訊透露。DSD 的定義如下（對於每個使用者）：

$$DSD = \sum_{i=1}^{n} F_i \times W_i$$

其中 n 為個人資料欄位的總數，F_i 是一個布林值（boolean），記錄使用者的每一個欄位 i 是否公開。而 W_i 則是每一個欄位 i 的權重（weight），它的計算方法如下：

$$W_i = 1 - \frac{R_i}{\sum_{i=1}^{n} R_i}$$

R_i 為樣本中每個欄位 i 的平均自我資訊透露比例（average ratio of disclosure），它的計算方法是把整個樣本中，公開欄位 i 的使用者數目除以總使用者數目。例如，若 1,000 個使用者中，有 100 個填寫「體重」欄位，則體重欄位的 R_w 為 100 / 1000 = 0.1。

在 DSD 的計算中，我們排除無法確認有效性的欄位，包含「暱稱」及「自我介紹」。此外由於不見得所有使用者都擁有「MSN 帳號」、「Yahoo!帳號」以及「Skype 帳號」等資訊，這些欄位也不列入 DSD 的計算。

圖 4 使用者自我資訊透露程度的年齡分佈

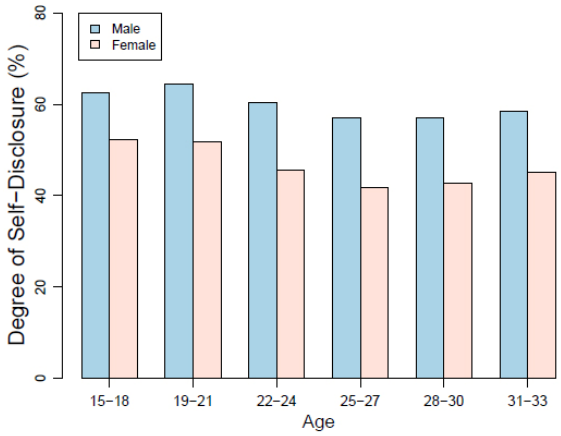
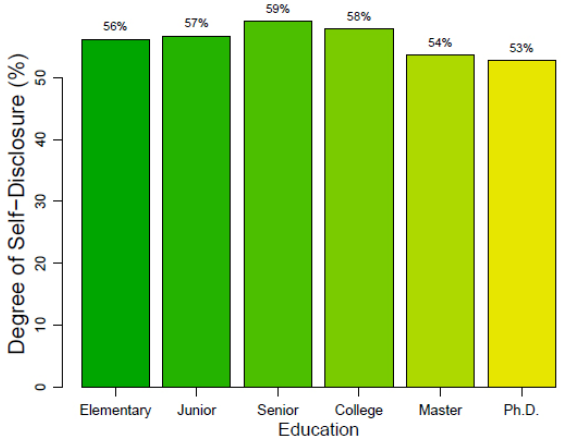


圖 5 使用者自我資訊透露程度的學歷分佈



3.3.3 自我透露程度分析

從圖 4 中我們可以看到使用者的自我透露程度 (DSD) 與其年齡以及性別的關係。使用者的年齡越大，自我透露的程度則會下降。普遍來說，男性使用者的自我透露程度比女性使用者的高，表示在無名小站中男性使用者與別人建立關係的動機比女性使用者的強。DSD 在年齡為 27 歲以後的使用者群有稍為上昇的趨勢，這可能是因為這些使用者已進入社會，相較於年紀較輕的使用者有更強的動機去建立社交連結 [12]。

圖 5 描述 DSD 與使用者學歷的關係。自我透露程度較高的使用者，其學歷主要為高中到大學，而無名小站的使用者大部分為高中生及大學生，由此可知無名小站樣本中大部分使用者的自我透露程度都相對偏高。

4. 名字洩露分析

4.1 好友描述

無名小站提供一個「好友暱稱」的功能供使用者描述他們的好友。使用者通常會在

描述中使用其好友的真實全名或名字。雖然在無名小站中無法直接查看所有好友（連入好友）對於某位使用者的描述（只能看到某一位使用者對他所有連出好友的描述），但我們依然可以手動或自動收集某位使用者的好友們對他（她）的描述。如此一來，即使使用者沒有公開自己的個人資訊，第三者也可能從其他使用者對他的描述中推測其真實姓名。

但是，無名小站中的使用者如何知道好友的真實姓名呢？我們列舉兩個可能：首先，在線上社群網路之中，「朋友」的關係大致可以分成三種，它們分別是真實社會相識（real-world acquaintances）、網路認識（online acquaintances）以及商業聯繫（business contacts）[16]。其中真實社會朋友是最有可能知道他人真實姓名的一種。第二個可能是使用者已公佈其真實姓名在個人資料中（例如暱稱或個人介紹），在這種情況下使用者名字洩露的情況只能歸類為自我資訊透露，而非非自願姓名洩露。我們將會在 4.5.3 節中討論這個可能性。

4.2 研究方法

我們從無名小站樣本中取得每個使用者的連入朋友以及他們對於該使用者的描述，接著藉由這些描述推測使用者的真實姓名。首先，我們把「候選字串」(有可能是名字的字串）從好友描述中擷取出來：我們以常用的中文連接詞把描述斷開，得到一或多個子字串。由於中文的姓名一般來說是三個字，其中姓氏是一個中文字，名字是兩個中文字，所以我們只分析長度為兩或三個字的字串：長度為三個字的字串我們歸類為「全名候選字串」(real name candidate)，而長度為兩個字的字串我們則歸類為「名字候選字串」(first name candidate)。此外，我們也計算每個候選字串在所有好友對該使用者的描述中的重複次數（有多少個描述中包含這個候選字串），作為推測名字時使用的參考資料。

表 2 提供朋友描述及候選字串的概況。每個使用者平均有 7 個連入朋友，代表每個使用者平均最多會有 7 個人給予描述。資料顯示使用者平均獲得 6.8 個描述，表示幾乎每個使用者都會為他們的連出好友給予描述。每個使用者平均大概有 4 個候選字串，他（她）的全名或部分姓名可能分佈在這些候選字串之中。

表 2 朋友描述與候選字串概況

朋友描述與候選字串	
平均連入朋友數	7.10
平均朋友描述數	6.81 (96%)
平均有共同候選字串的朋友描述數	3.46 (49%)
平均不重複候選字串數	3.81

我們使用下列的方法來推測使用者的真實姓名，這些方法主要基於三個方向，它們

分別是：

1. 分析在不同描述中重覆出現的候選字串；
2. 將候選字串與常用姓名名單比對；
3. 將候選字串與常用詞比對，排除可能為常用詞的候選字串。

因為我們只推測最常見的三個字姓名（姓氏為一個字，名字為兩個字)，所以實際可推測到姓名的比例應該會更高。

4.2.1 推測使用者全名

我們使用五種方法來推測使用者的全名（三個字，姓氏 + 名字)。除了第五個方法使用到第四個方法的結果之外，每個方法都獨立進行。

一. **常用姓氏：**如果任何一個全名候選字串的第一個字為常用姓氏 [22]，我們則推測這個全名候選字串為使用者全名。這個方法中我們只會考慮重複次數至少為 1 （即最少在兩個描述中都出現）的全名候選字串。

二. **全名與名字：**我們檢查是否有任何全名候選字串（三個字）的後兩個字重覆出現在名字候選字串（兩個字）之中。因為中文的名字一般是在全名的第二到第三個字，所以如果符合這個條件，我們則推測這個全名候選字串是使用者全名。

三. **常用全名名單：**我們把全名候選字串比對到常用全名名單（包含 1994-2007 大學聯考榜單，共 574,010 個全名）[24]，如果候選字串跟名單裏任何一個名字完全相同，我們就推測這個全名候選字串為使用者的全名。

四. **暱稱與全名：**我們發現使用者普遍會使用暱稱相關詞加上真實全名的部分組成暱稱。例如陳大強的暱稱有可能是：「老陳」、「大強兄」、「強哥」等等。在這個方法中我們利用暱稱相關詞 (人工準備，共 38 個相關詞) 找出可能是暱稱的全名候選字串。如果全名候選字串包含暱稱相關詞，我們就認為這個全名候選字串是暱稱；接著從候選字串中去掉暱稱相關詞，把剩下餘的字串（可能是使用者名字的片段）與其他非暱稱的全名候選字串比對，若比對成功，我們就推測這個全名候選字串為使用者全名。

五. **移除常用詞語：**此方法接續方法四得到的非暱稱全名候選字串，但只選用重複次數最少為 1 的非暱稱全名候選字串。我們把候選字串與常用詞語列表比對（共 100,511 個常用詞）[22]。如果全名候選字串包含常用詞（例如候選字串為「台北車站」而常用詞為「車站」）或常用詞包含全名候選字串（例如候選字串為「乘法」而常用詞為「乘法表」），我們就認為這個候選字串為常用詞而將它排除。最後剩下的全名候選字串中我們以出現次數最多，且

次數大於 1 的全名候選字串（如果存在的話）為全名。

4.2.2 推測使用者名字

推測使用者名字（兩個字，不包括姓氏）的方法同 4.2.1 所列的方法三到五，並改用字串長度為兩個字的名字候選字串。在第三個方法中我們改用常用名字名單（結合 [23] 跟 [24]的名字部分，去掉重複名字，共 208,581 個名字）來比對。此外，在推測使用者名字的各個方法中我們只考慮重複次數最少為 1 的名字候選字串。

表 3 推測到真實姓名的比例（綜合）

名稱的種類	推測到的比例
暱稱	60%
全名	30%
名字	72%
全名或名字	78%

表 4 以不同方法推測到真實全名的比例

方法名稱	推測到的比例
常用姓氏	3%
全名與名字	11%
常用全名名單	9%
暱稱與全名	2%
移除常用詞語	27%

表 5 以不同方法推測到真實名字的比例

方法名稱	推測到的比例
常用名字名單	57%
暱稱與名字	14%
移除常用詞語	69%

4.3 推測名字結果

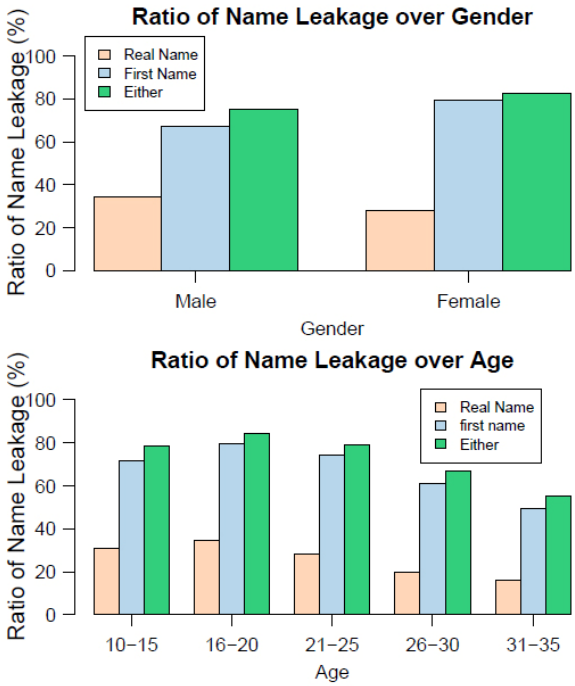
表 3 是推測使用者姓名的整體結果，我們發現 72% 使用者的真實全名以及 30% 使用者的真實名字能夠成功被推測；這表示平均每 10 個使用者中，有 3 個跟 7 個使用者，分別遭受真實名子及真實全名被洩露的風險。共有 78% 的使用者遭受真實姓名(全或或名字）洩露的風險。而表 4 及表 5 顯示每個不同方法的推測結果。

為了驗證推測到的姓名是否有效，我們隨機選擇了 1,000 個推測到全名或名字的使用者，以人工的方式檢查。我們發現約有 74% 的使用者（738 個）其由自動推測的姓名可判定為真實姓名，而剩餘的使用者其誤判原因主要是使用者的暱稱被誤認為姓名。

4.4 姓名洩露比例

我們定義姓名洩露比例（ratio of name leakage）為樣本中姓名可被推測的使用者數目除以樣本大小。圖 6 的上圖顯示男性全名的姓名洩露比例比女性高，而女性名字洩露比例比男性高。這表示在網路上，使用者較常連名帶姓地描述男性，並較常以名字描述女性。圖 6 的下圖顯示姓名洩露比例隨著使用者的年齡下降。我們認為其主因是年齡越大的使用者朋友數目也較少，因此姓名被洩露的機會也較低。

圖 6　姓名洩露比例的性別與年齡分佈



4.5 原因與風險分析

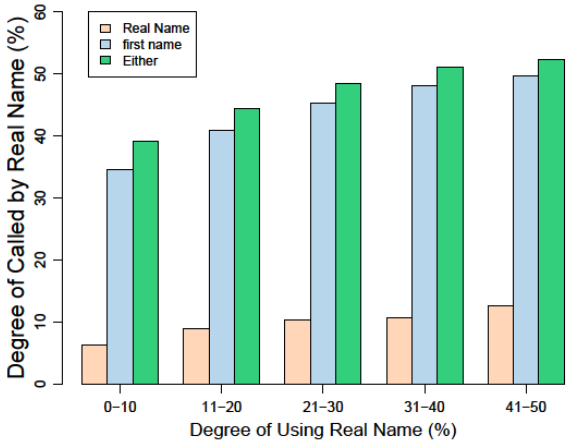
4.5.1　以真實姓名稱呼朋友傾向

我們使用 DUR（degree of using real name, 簡稱）來量化使用者以真實姓名稱呼朋友的傾向。DUR 的定義為一個使用者對於其朋友的描述中，有多少比例包含其朋友的真實姓名。我們使用 DUR 分析姓名洩露的原因。

4.5.2　被朋友以真實姓名稱呼的比率

我們使用 DCR(degree of called by real name)來量化使用者被其朋友以真實姓名稱呼的比率。它的定義為使用者的連入好友為其撰寫的描述中，有多少比例包含該使用者的真實姓名。我們使用 DCR 分析姓名洩露的結果。

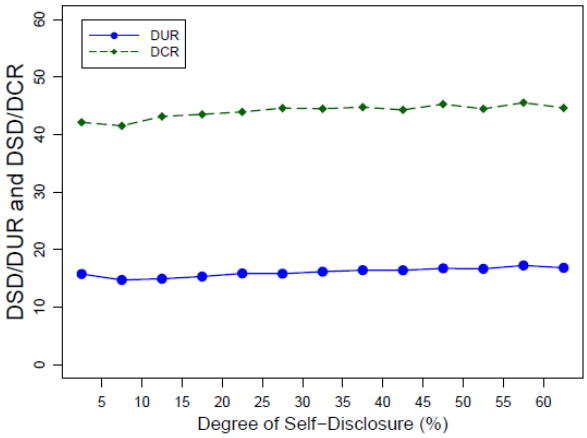
圖 7　DUR 與 DCR 的關係



4.5.3　DUR 與 DCR 的關係

從圖 7 中，我們可看出 DUR 與 DCR 有極強的正相關。被朋友以真實姓名稱呼的比率會隨著以真實姓名稱呼朋友傾向上升而增加。我們推測，此現象是因為不少使用者與其朋友在網路社交服務中互為好友（即彼此同時為連入和連出朋友），並以真實姓名互稱。圖 8 中，我們比較 DUR, DCR 與 DSD 的關係。圖中顯示自我資訊洩露程度跟洩露它人姓名的傾向並沒有明顯的關係。

圖 8　DUR, DCR 與 DSD 的關係



我們考慮到姓名洩露的主因可能是使用者本身已把他們的真實名字透露在個人資料之中，在某些社群網路服務，例如 Facebook.com，使用者通常以直接公開直實姓[9]。在這種情況下，姓名洩露只能算是自我資訊透露的一種。我們檢查從無名小站樣本中推測到的使用者真實姓名，有多少比例出現在其個人資料的暱稱及自我介紹欄位，發現只有不到 0.1% 的使用者主動透露自己的真實姓名。因此，我們推斷無名小站上以好友描述所導致的姓名洩露絕大部分為非自願性，而非使用者的自我資訊透露。

5. 討論

在本節中，我們討論姓名洩露的危險性以及建議的防治方法。

5.1 姓名洩露可能造成的攻擊

5.1.1 以朋友名義發出的垃圾郵件

垃圾郵件是近年很嚴重的問題，它其實是一個商業活動：垃圾郵件發送者（spammer）利用大量的電子郵件地址名單，發送包含商品或網站介紹的電子郵件給潛在客戶；亦即任何擁有電子郵件帳號的人 [14]。值得注意的是，我們收集到的無名小站樣本中，有 46% 的使用者公開格式正確的電子郵件地址。

在垃圾郵件在網路上氾濫的同時，學術及產業界也研發出不少有效的反垃圾郵件機制。然而，以垃圾郵件謀利的人同時也花更多的時間去製作更有效的電子郵件地址名單。從社群網路服務中獲取更多的使用者資訊是有效的方法。垃圾郵件發送者可以透過郵件地址抓取程式 [14] 取得社群網路服務中使用者的資料以及交友關係來製作電子郵件名單。

一般來說，使用者會把好友的電子郵件地址設為白名單（white list，例如聯絡人或垃圾郵件過濾器的例外名單），所以從朋友郵件地址寄來的郵件不會被判定為垃圾郵件。如果垃圾郵件發送者利用使用者朋友的電子郵件地址來偽造郵件，所送出的垃圾郵件就可能獲得更多的信任，甚至可以騙過垃圾郵件過濾器。此外，如果垃圾郵件發送者在郵件的主題或內容中加入使用者好友的姓名或暱稱，即使該郵件已被判定為垃圾郵件，使用者仍可能會開啟它。在此情況下，垃圾郵件即使已被成功阻擋，也能對使用者造成困擾。目前類似案例數量已持續成長，成為反垃圾郵件研究者的新挑戰。

5.1.2 個人化的釣魚攻擊

近年來網路釣魚（phishing）也是一個非常迫切的安全議題。網路釣魚者（phisher）傳送包含釣魚網站連結的電子郵件給使用者，欺騙他們到釣魚網站輸入個人或財務資料如身份證號碼、銀行帳號及信用卡號碼等等 [5]，然後盜用這些資料謀取利益。在網路釣魚行為中，使用者對釣魚郵件及網站的信任為攻擊能否成功的關鍵。只要使用者信任釣魚網站，不論系統提供多可靠的加密方法保護資料的傳送，也無法完全保護使用者免受網路釣魚攻擊。保護使用者避免網路釣魚需要的是使用者與網站之間的安全通道，以供使用者確認網站的可信性；而非瀏覽器與網站之間的安全通道，例如資料加密 [18]。

不少成為網路釣魚目標的業者（例如 eBay.com 及 Paypal.com）及一些網路安全相關的網站提供如何辨認網路釣魚郵件的指引。常見的規則有「查看郵件內容是否提及到您的姓名」或「網路釣魚郵件無法取得個別使用者的姓名」等等 [4,19]。如果使用者相信網路釣魚郵件內容不會出現自己的姓名並以此作為判斷原則，那麼包含使用者姓名

的釣魚郵件就更可能取得使用者的信任。在網路釣魚多變的本質之下，現今的網路釣魚防治工具及方法仍不夠完美 [5, 8, 25]，此種個人化的網路釣魚攻擊可能會使更多的使用者受騙。

網路釣魚駭客也可能直接攻擊社群網路服務。駭客們可在熱門使用者的留言板上以其朋友的姓名貼上偽造的網路服務登入網頁，當其他使用者瀏覽該留言板並連結到釣魚登入網頁時，他們可能以為是偶然地被服務系統登出（例如 session 時間超出），立即輸入帳號密碼欲重新登入。得到某使用者的登入資料後，網路釣魚駭客就可以對其在社群網路服務的朋友使用更多的詐騙手法。

5.2 建議的解決方案

針對社群網路上的非自願姓名洩露，我們提出以下三個可能的解決方案：

1. **提供個人隱私性設定：**我們認為，社群網路服務提供者應該提供下列選項供使用者自訂：
   a. 隱藏個人資訊的選項
   b. 隱藏交友情況的選項（可以設定限制範圍，例如好友或好友的好友）。
   c. 是否同意被其他使用者描述的選項
   d. 是否允許連入朋友公開顯示朋友關係的選項。

   以上保護個人隱私的選項功能不應該只提供給收費的使用者。

2. **提供瀏覽限制範圍設定：**社群網路服務應該限制使用者的瀏覽範圍。例如，除非某些設為「任何人都能看到我」的使用者，在社群網路中不能無限制的從使用者的好友名單連結到其他使用者。另一個可行方法是把使用者以群組區分，欲瀏覽該群組的使用者必須接受認證或加入群組。這些機制可以防治惡意的人輕易抓取大量的使用者資料，進而取得使用者姓名。

3. **取得使用者確認：**所有好友關係設定的活動（包含朋友描述設定）都必須經過雙方的確認，而只有經過確認的朋友才會出現在好友名單之中。這樣可以避免資訊無意間被網友或陌生人洩露。

6. 結論

在本篇論文中，我們分析社群網路服務中非自願姓名洩露的程度、成因及解決方案。我們從國內最大的社群網路服務網站 - 無名小站 - 蒐集使用者樣本，發現 78% 使用者的真實姓名面臨非自願性洩露的風險。非自願姓名洩露的主因是使用者的連入朋友們在對該使用者的描述中使用其真實姓名。我們發現使用者的自我資訊透露程度並沒有

與使用者以真實姓名互稱的傾向有明顯的關係，代表使用者公開個人資訊的多寡無法影響非自願姓名洩露。

針對社群網路上的非自願姓名洩露，我們提出以下三個可能的解決方案，包括提供個人隱私性設定、提供瀏覽限制範圍設定以及取得使用者的確認。我們希望藉由此文拋磚引玉，引起國內產學界對於網路社群中資安問題的重視，共同建立一個可讓使用者安全無虞建立關係及讓網路創作者安心製作及交流數位內容的虛擬網路空間。

誌謝

由於 XDite（blog.xdite.net）在 2007 台灣駭客年會上演示無名小站可能有的個人隱私洩露問題，才有本篇研究的發想。在此特別感謝 XDite 以及與她一同進行研究的海洋大學資訊工程研究所研究生 CornGuo。

參考文獻

[1] Alessandro Acquisti, 「Privacy in electronic commerce and the economics of immediate gratification,」 Proceedings of the *5th ACM conference on Electronic commerce*, 2004, pages 21 – 29, 2004.

[2] Y.Y. Ahn, S. Han, H. Kwak, S. Moon, and H. Jeong, 「Analysis of topological characteristics of huge online social networking services,」 Proceedings of *the 16th international conference on World Wide Web*, pages 835 – 844, 2007.

[3] D. Boyd, 「Friendster and Publicly Articulated Social Networks,」 *Conference on Human Factors and Computing Systems*（*CHI 2004)*, Vienna, Austria, April, pages 24 – 29, 2004.

[4] Microsoft Corporation, 「Recognize phishing scams and fraudulent e-mails,」 September 2006.

[5] Lorrie Faith Cranor, Serge Egelman, Jason Hong, and Yue Zhang, 「Phinding phish: Evaluating anti-phishing tools,」 In Proceedings of the *14th Annual Network & Distributed System Security Symposium*（*NDSS 2007)*, March 2007.

[6] Duff, 「Social Engineering in the Information Age,」 The Information Society, 21(1):67–71, 2005.

[7] ENISA, 「Security issues and recommendations for online social networks,」 *ENISA Position Paper* No.1, October 2007.

[8] D Florencio and C Herley, 「Analysis and improvement of anti-phishing schemes,」 *Microsoft Research*, One Microsoft Way, Redmond, WA.

[9] Ralph Gross, Alessandro Acquisti, and III H. John Heinz,「Information revelation and privacy in online social networks,」 Proceedings of the *7th ACM SIGCOMM conference on Internet measurement*, pages 71 – 80, 2005.

[10] CipherTrust Inc, 「Spam Statistics」, http://www.ciphertrust.com/resources/statistics/index.php. 2006.

[11] Adam N. Joinson and Carina B. Paine, 「Self-disclosure, privacy and the internet. Institute of Educational Technology,」 The Open University, United Kingdom, 2006.

[12] S.M. Jourard, 「Self-disclosure: An experimental analysis of the transparent self,」 *New York: Krieger*, 1971.

[13] S.M. Jourard and P Lasakow, 「Some factors in self-disclosure,」 *Journal of Abnormal and Social Psychology*, 56(1):91–98, 1958.

[14] Paul Judge, Dmitri Alperovitch, and Weilai Yang, 「Understanding and reversing the profit model of spam,」 CipherTrust Inc, Alpharetta, GA 30022 USA, March 2005.

[15] Ravi Kumar, Jasmine Novak, and Andrew Tomkins, 「Structure and evolution of online social networks,」 Proceedings of the *12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 611–617, 2006.

[16] Alan Mislove, Massimiliano Marcon, Krishna P. Gummadi, Peter Druschel, and Bobby Bhattacharjee, 「Measurement and analysis of online social networks,」 Proceedings of the *7th ACM SIGCOMM conference on Internet measurement*, pages 29 – 42, 2007.

[17] Ina OMurchu, John G. Breslin, and Stefan Decker, 「Online social and business networking communities,」 Digital Enterprise Research Institute, National University of Ireland, August 2004

[18] Rolf Oppliger and Sebastian Gajek, 「Effective protection against phishing and web spoofing,」 Processing of *IFIP International Federation for Information 2005*, 2005.

[19] PayPal. 「Phishing guide part 2」, https://www.paypal.com/us/cgi-bin/webscr?cmd=xpt/cps/securitycenter/general/RecognizePhishing-outside

[20] ASHLEY PHILLIPS, 「Facebook code leaked on site,」 *ABCNews Internet Ventures*, August 2007.

[21] Russell Research, 「CA/NSCA Social Networking Study Report」, 2006.

[22] Chih-Hao Tsai, 「A review of chinese word lists accessible on the internet,」

http://technology.chtsai.org/wordlist/, January 2006.

[23] Chih-Hao Tsai, 「Common chinese names,」 http://technology.chtsai.org/namefreq/, August 2007.

[24] Chih-Hao Tsai, 「A list of chinese names,」 August 2007.

[25] Min Wu, Robert C. Miller, and Simson L. Garfinkel, 「Do security toolbars actually prevent phishing attacks?」 Proceedings of the *SIGCHI conference on Human Factors in computing systems*, pages 601–610, 2006.