Social Networking Site

By

Mayank Chandak | 071272
Rahul Tripathi | 071253
Ajay Kumar | 071343

Under the supervision of
Dr. Rajesh Siddavatam
&
Mr. Rahul Dwivedi

May – 2011

Submitted in partial fulfillment of the Degree of Bachelor of Technology

DEPARTMENT OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
WAKNAGHAT

TABLE OF CONTENTS

CERTIFICATE

This is to certify that the work titled **SOCIAL NETWORKING SITE** submitted by **MAYANK CHANDAK, RAHUL TRIPATHI & AJAY KUMAR** in partial fulfillment for the award of degree of Bachelor of Technology (C.S.E.) of **JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT** has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

Signature in full of Supervisor: _____

Name in Capital block letters: ***Dr. RAJESH SIDDAVATAM***

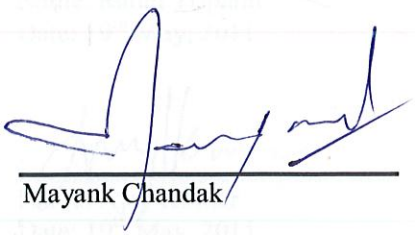
Designation: **Associate Professor, Department of Computer Science and IT, JUIT**

Date: **19th May, 2011**

ACKNOWLEDGEMENT

No venture can be completed without the blessing of Almighty. I consider it my bounded duty to bow to Almighty whose kind blessings always inspire us on the right path of life .This project is our combined effort and realizes the potential of team and gives us a chance to work in co-ordination.

Science has caused many frontiers so has human efforts towards human research. Our supervisor **Dr. Rajesh Siddavatam,** Associate Professor, Department of Computer Science and IT, JUIT, and our project guide **Mr. Rahul Dwivedi,** Lecturer Dept. of CSE & IT, have indeed acted as a light house showing us the need of sustained effort in the field of Web Development to learn more and more. So we take this opportunity to thank them, for lending us stimulating suggestions, innovative quality guidance and creative thinking. They provided us the kind of strategies required for the completion of a task. We are grateful to them for the support they provided us in doing things at our pace and for being patient for our mistakes.



Mayank Chandak

Rahul Tripathi

Ajay Kumar

Date: 19th May

SUMMARY

The project entitled as **Social Networking Site** is a system developed for the World Wide Web .This social network is one proposed system which is made to run on any Web server but we have executed it on Apache Server (PHP). This system is developed using CakePHP framework.

The basic idea behind this project is to provide a good, efficient method for easy communication amongst users and provide them with basic networking features like status updates, photo sharing and creation of Public Profile with user data.

Name: Mayank Chandak
Date: 19th May, 2011

Name: Rahul Tripathi
Date: 19th May, 2011

Name: Ajay Kumar
Date: 19th May, 2011

Signature of Supervisor
Name: *Dr. Rajesh Siddavatam*
Date: 19th May, 2011

Signature of Project Guide
Name: Mr. Rahul Dwivedi
Date: 19th May, 2011

Chapter 1

Introduction

Wikipedia states that a social network is "...a social structure made of nodes (which are generally individuals or organizations) that are tied by one or more specific types of relations...".With the rapid growth of people who use or have access to the Internet, social networking websites are a must for the Internet community to stay in touch with each other. Social networking web sites help people keep in touch with old friends make new friends; distribute new data or product, and many more aspects of our everyday lives.

The first official social networking web site was Classmates.com which was founded in 1995 .What followed was a slow but steady growth in numbers of social networking websites to the overwhelming number of sites we have today. The reason that social networking websites work so well is that, like their inception, they start of small and then grow exponentially. The site starts off with a few people who then tell their friends about the site, then those friends tell their friends about the site and soon the site is a huge database of users connected by friends, acquaintances, or just random people. The web sites are made to allow users to create a "profile" describing themselves and to exchange public or private messages and list other users or groups they are connected to in some way.

Social Networking is a nice form of entertainment, great for meeting people with similar interests, and can be a very effective business technique for engineers like us. Most of us have hobbies, or things that we are keenly interested in such as books, television, video games or movies. Social networks allow us to reach out to others that have the same interests.

For example, if you like movies, Flixster can help you decide if you might like a new movie just out at the theaters. If you like music, Last.FM can help you find new artists that are similar to your favorite bands.

1.1. Scope

In current world, social networking is tremendously successful and social media marketing is very essential to a business due to the perpendicular number of people who access social media sites frequently. There are numerous logics why this sort of marketing is so useful, or rather, necessary for a business.
First of all it is a low costly as compared to the other facilities available. Social media is normally free to use however advertising the similar thing through usual methods would cost us plenty of dollars. These sites attract a significant traffic and to your site. In current days climate, business and contentment are very narrowly linked and what I mean to say this is, people will only do transaction with you if they believe they KNOW, LIKE and have FAITH on you

Social media is a way, by which you are build a relationship and that will raise a natural interest in the people you are conversing with you would be developing an interest in yourself as well. People will afterward naturally like to know further about you and will inquire. They will notice you as a person with authority.

1.2. Problem Statement

To create a Social Networking site that can be used by multiple users concurrently. The project should provide users with basic networking functionalities like status updates, public profiles, photo sharing. The project should also look into integrating existing social network and new-media for user convenience.

Chapter 2

History

USERNETS

Usenet systems were first conceived of in 1979 by Tom Truscott and Jim Ellis. Usenets let users post articles or posts (referred to as "news") to newsgroups. Usenets have no centralized server or dedicated administrator, setting them apart from most BBSs and forums. Usenets are mostly responsible for the development of newsreader clients, which are the precursor to RSS feed readers so commonly used to follow blogs and news sites today. Group sites such as Google Groups and Yahoo! Groups use many of the conventions established by the original usenet systems.

BBSS (BULLETIN BOARD SYSTEMS)

The first BBSs came online in the late 70s. Originally these were primarily hosted on personal computers and users had to dial in through the host computer's modem. Only one person at a time could gain access to the BBS.

While there were legitimate BBSs, most were at least somewhat involved in illicit, illegal, or other shady practices. Adult material, virus code, information and instructions for hacking and phreaking (phone hacking), and materials like The Anarchist's Cookbook were commonly hosted on BBSs. But BBSs were the first type of sites that allowed users to log on and interact with one another, albeit in a much slower fashion than we currently do.

ONLINE SERVICES

After BBSs came "online services" like CompuServe and Prodigy. These were the first real "corporate" attempts at accessing the Internet. CompuServe was the first company to incorporate a chat program into their service. Prodigy was responsible for making online service more affordable (CompuServe had been prohibitively expensive for many, with charges of $6/hour plug long-distance fees that often made the service run $30/hour or more).

Genie was an early online service created by a General Electric subsidiary (GEIS) in 1985. It ran through 1999 and was one of the earliest services available. It was a text-based service, and considered the first viable commercial competition to CompuServe. The service was created to make use of idle time-sharing mainframes after normal U.S. business hours. GEnie offered games, shopping, mail, and forums (called RoundTables). There was even a print magazine associated with the service at one time.

AOL started as an online service too and made great strides at making the Internet more universally accessible in the U.S.

IRC, ICQ, and Instant Messaging

IRC (Internet Relay Chat) was developed in 1988 and used for file sharing, link sharing and otherwise keeping in touch. It was really the father of instant messaging as we know it today. IRC was mostly UNIX-based though, limiting access to most people. ICQ was developed in the mid-90s and was the first instant messaging program for PCs. It was at least partly responsible for the adoption of avatars, abbreviations (LOL, BRB) and emoticons. Other IM clients soon followed.

Forums

Online forums also played a large part in the evolution of the social web. These were really descendants of the BBSs popular in the 70s and 80s, but usually came with a more user-friendly interface, making them easier for non-technical visitors to use. Various forum platforms, including vBulletin and phpBB, were developed, many of which are still used for forums. Forums remain a popular part of online culture, and many have made strides to add more social networking-type features (like profiles).

While many people consider dating sites or sites like Classmates.com to be the first social networks, they don't really fit the definition. Dating sites rarely allowed you to keep a friends list neither did Classmates in its early years (and profiles were severely limited).

Six Degrees

Six Degrees was launched in 1997 and was the first modern social network. It allowed users to create a profile and to become friends with other users. While the site is no longer functional, at one time it was actually quite popular and had around a million members at its peak. In 2000 it was purchased for $125 million and in 2001 it was shut down.

AsianAvenue, MiGente, BlackPlanet

These sites cropped up in the years following SixDegrees' launch, between 1997 and 2001.They allowed users to create profiles and add friends (generally without needing approval to add people). Users could create professional, personal and dating profiles on these sites. While they were some of the earliest social networks, there were few innovations among them.

LiveJournal

LiveJournal started in 1999 and took a different approach to social networking. While Six Degrees allowed users to create a basically-static profile, LiveJournal was a social network built around constantly-updated blogs. LiveJournal encouraged its users to follow one another and to create groups and otherwise interact. It was really the precursor to the live updates we see in social networks currently.

World of Warcraft / MMORPGS

MMORPGS (Massively multiplayer online role-playing games) have become social networks in their own right. The most famous of these is World of Warcraft, where players interact both in the game world and on related forums and community sites. Social interaction within the games ranges from teams set up specifically for tactical reasons within the game to friendships to romances. MMORPGS became popular in the early 2000s, though there were other online role-playing and other games prior to that.

Major Advances in Social Networking

The early 2000s brought some huge developments in social networking and social media. Friendster. Friendster was really the first modern, general social network. Founded in 2002, Friendster is still a very active social network, with over 90 million registered users and 60+ million unique visitors each month. Most of Friendster's traffic comes from Asia (90% of it). Friendster operated by allowing people to discover their friends and then friends-of-friends, and so on to expand their networks.

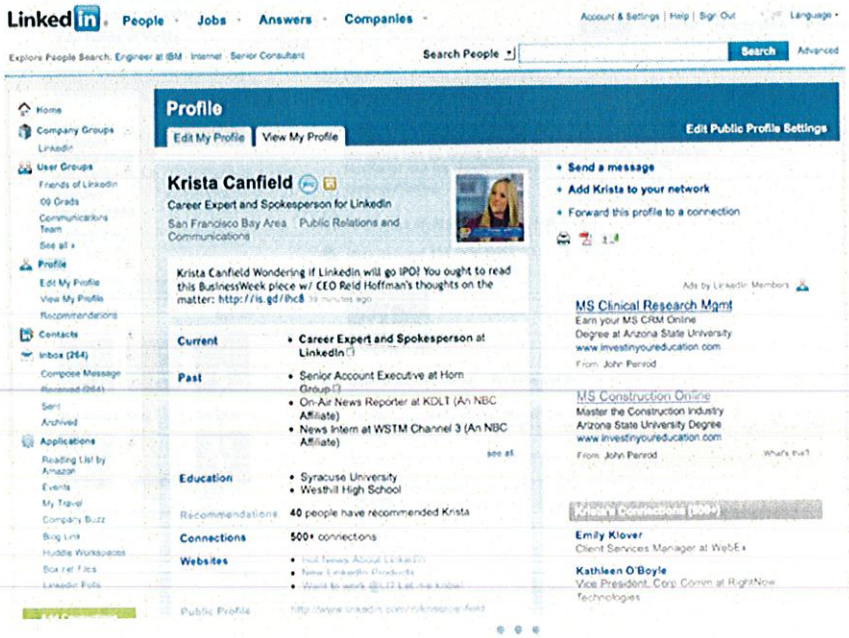
Its goal was to be a safer place to meet new people than in real-life, as well as being faster. Friendster was, in part, a new kind of dating site. Instead of matching complete strangers based on shared interests, it operated on the assumption that people with shared friends and acquaintances would have a better chance than those who had no shared connection. Friendster was most popular with three different groups: gay men, attendees of Burning Man and bloggers.

Hi5

Hi5 is another major social network, established in 2003 and currently boasting more than 60 million active members according to their own claims. Profile privacy works a bit differently on Hi5, where a user's network consists of not only their own contacts, but also second (friends of friends) and third (friends of friends of friends) degree contacts. Users can set their profiles to be seen only by their network members or by Hi5 users in general. While Hi5 is not particularly popular in the U.S., it has a large user base in parts of Asia, Latin America and Central Africa.

LinkedIn

LinkedIn was founded in 2003 and was one of the first mainstream social networks devoted to business. Originally, LinkedIn allowed users to post a profile (basically a resume) and to interact through private messaging. They also work on the assumption that you should personally know the people you connect with on the site.

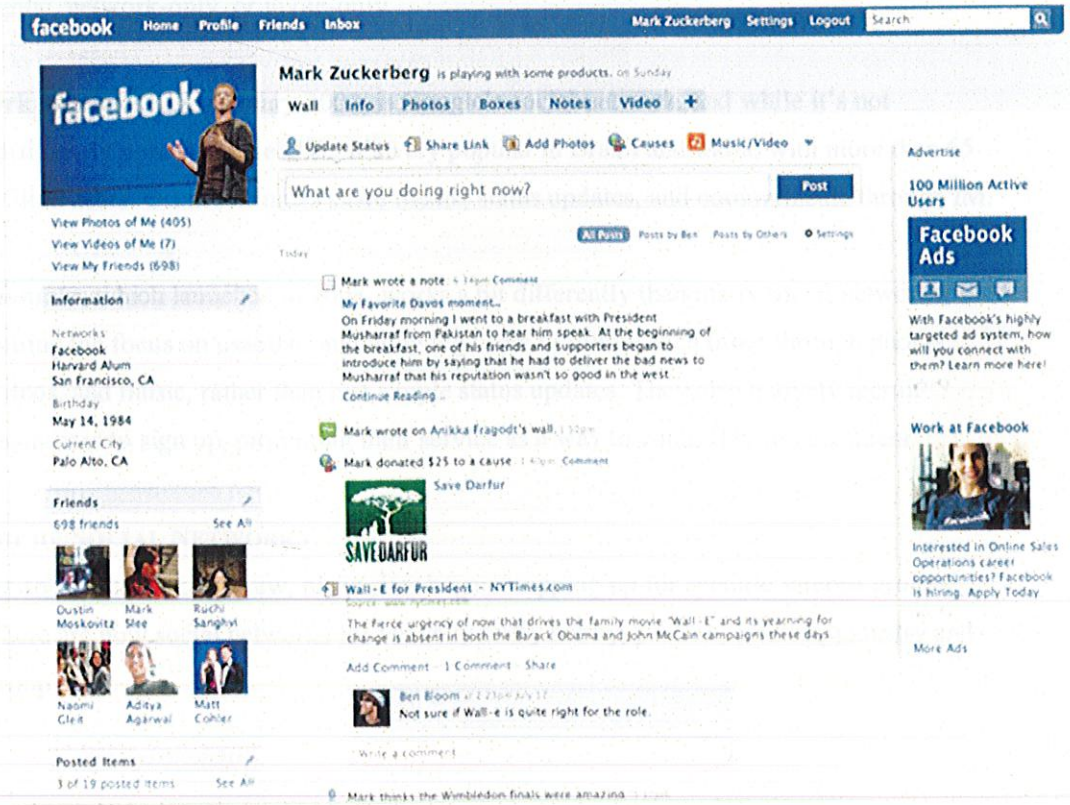


MySpace

MySpace was founded in 2003 and by 2006 had grown to be the most popular social network in the world. MySpace differentiated itself from competitors by allowing users to completely customize the look of their profiles. Users could also post music from artists on MySpace and embed videos from other sites on their profiles. Originally MySpace allowed communication through private messages, public comments posted to a user's profile, and bulletins sent out to all of your friends. Blogs are also a big part of MySpace profiles, with each member automatically getting a blog. In 2006 MySpace introduced MySpace IM, an instant messaging client that lets users chat with their friends.
Other recent additions to MySpace's functionality include the addition of real-time status updates and news feed showing friend activity.

Facebook

While Facebook started out as a Harvard-only social network back in 2004, it quickly expanded to other schools, then to high schools, businesses and eventually everyone (by 2006).In 2008 Facebook became the most popular social networking site, surpassing MySpace, and continues to grow.



Facebook doesn't allow the same kind of customization that MySpace does. Facebook does, however, allow users to post photos, videos and otherwise customize their profile content, if not the design. Facebook has added a number of features over the past few years, including instant messaging/chat and apps (and their developer platform). Users have a few different methods of communicating with one another. Private messaging is available as well as writing on another user's wall. Wall posts are visible to that user's friends, but usually not to the general public. Users can easily change their privacy settings to allow different users to see different parts of their profile, based on any existing relationships (the basic privacy settings are "only friends", "friends of friends", and "everyone").Users can post notes that are visible to all of their friends. Users can also comment on or, more recently, "like" the posts of their friends, and conversations often occur within the comment sections among multiple people.

Other Major Social Networks

Multiply, a "family-friendly" social network and media sharing site was established in 2004 and puts much more emphasis on security and privacy than many other networks. Multiply users have the option to set security levels on each item they post, making things public, network-only, or invite-only.

Orkut, launched in January 2004, is Google's social network, and while it's not particularly popular in the U.S., it's very popular in Brazil and India, with more than 65 million users. Orkut lets users share media, status updates, and communicate through IM.

Kontain, which launched in 2008, works a bit differently than many social networks, putting the focus on usability and allowing users to follow each other through photos, videos, and music, rather than just simple status updates. They also actively recruit businesses to sign up, promoting their service as a way to connect with customers.

Niche Social Networks

As social networking grew, niche sites began cropping up for specific interest groups. There are now social networks for virtually every hobby, passion, interest, industry and group

NING

Ning is a platform for creating niche social networks. Networks are hosted by Ning but can take on their own personality and can even pay to have their own branding instead of the Ning brand.

Ning was the first widely-used social networking platform. Its biggest advantage in the market was that it made it incredibly simple for even non-technical users to set up their own social network. While most other social networking platforms required coding and programming knowledge, Ning required neither of those.

MEDIA SHARING

Social media isn't just limited to social networking sites. Sharing photos, videos, and other multimedia content is also a popular social media activity.

PHOTOBUCKET

Photobucket was the first major photo sharing site, launched in 2003.Photobucket allows users to share photos publically or in password-protected albums. They allow users 500MB of storage (lowered from 1GB in August of 2009).Pro accounts get 10GB of storage (lowered from 100GB to 25GB in July of '08 and then to 10GB in August of '09). Photobucket also hosts video content.

In 2007, Photobucket was purchased by Fox Interactive Media (a News Corporation subsidiary). It was rumored to have sold for as much as $250 million, though terms of the sale were never disclosed.

FLICKR

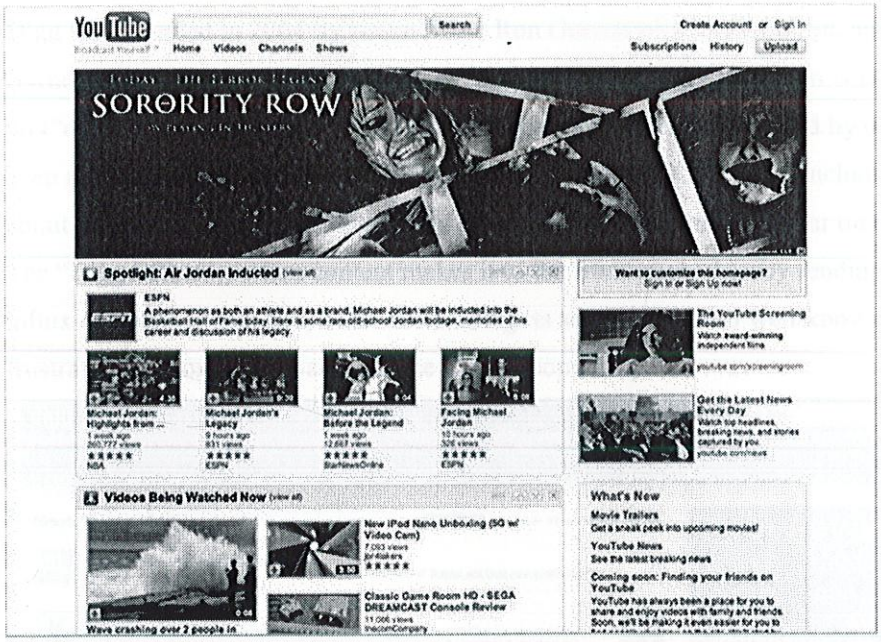
Flickr has become a social network in its own right in recent years. They claim to host more than 3.6 billion images as of June 2009.Flickr also has groups, photo pools, and allows users to create profiles, add friends, and organize images and video into photo sets/albums.

One of Flickr's major advantages is that they allow users to license their photos through Creative Commons, as well as retaining all copyrights. Flickr has also recently launched a collection called "The Commons", which features archived photos and images from a

variety of museums and other institutions under a "no known restrictions" license (basically meaning the photos are believed to be in the public domain).

YouTube

YouTube was the first major video hosting and sharing site, launched in 2005. Users can upload videos up to 10 minutes long and share them through YouTube or by embedding them on other websites (social networks, blogs, forums, etc.). YouTube now allows users to upload HD videos and recently launched a service to provide TV shows and movies under license from their copyright holders.



YouTube's major social features include ratings, comments, and the option to subscribe to the channels of a user's favorite video creators.

Revver

Revver took a slightly different approach to video hosting and sharing. While YouTube, Metacafe, and most other video sharing sites let you post videos for free and didn't pay content creators for any advertising revenues their videos generated, Revver has been sharing revenue from the start.

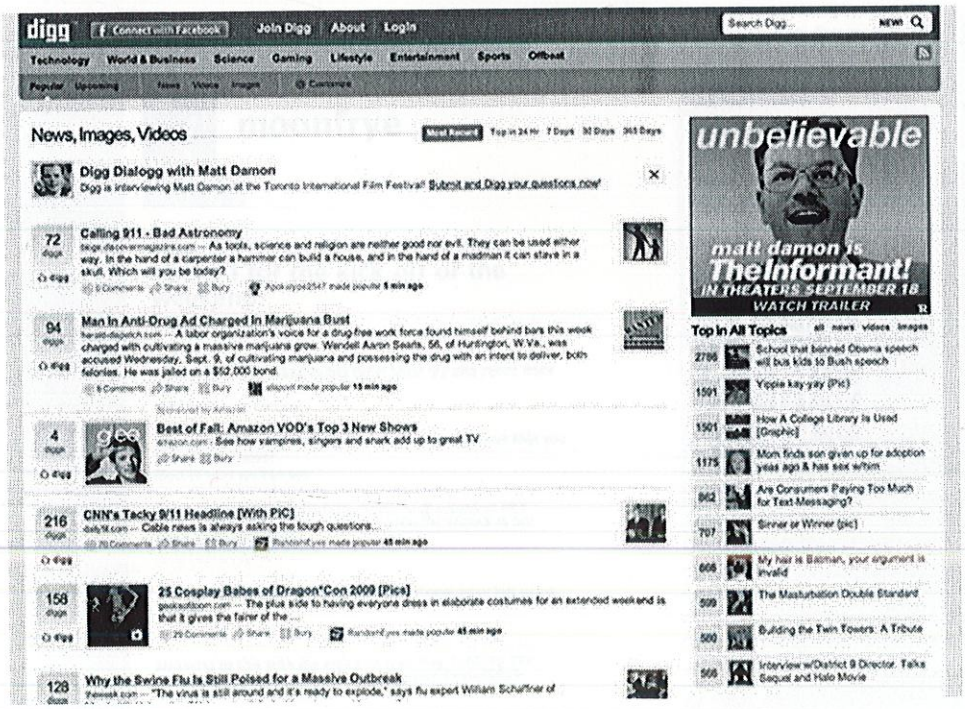
Social News and Bookmarking

Delicious

Delicious (aka, Del.icio.us) is a social bookmarking site founded in 2003. It allows its users to bookmark any content they find online, tag that content, and then share it with other users. Users can search for bookmarks or browse for them via tags. Delicious also allows users to view the most popular content among other users, as well as up-and-coming content, not unlike most social news sites.

Digg

Digg was founded in 2004 by Kevin Rose, Ron Gorodetzky, Jay Adelson, and Owen Byrne. Digg users can share links to anything online and other users can vote that content up ("dig") or down ("bury"). Users can also comment on content posted by others and keep a friends list. Digg has undergone a lot of controversy in its day; including criticism about the power the top 100 Digg users have over what becomes popular on the site. The "Digg Effect"—when content makes it to the front page, thereby sending a huge influx of traffic to that site, often overloading its servers—is also well-known and often frustrating to those unprepared for the sudden popularity.



Reddit

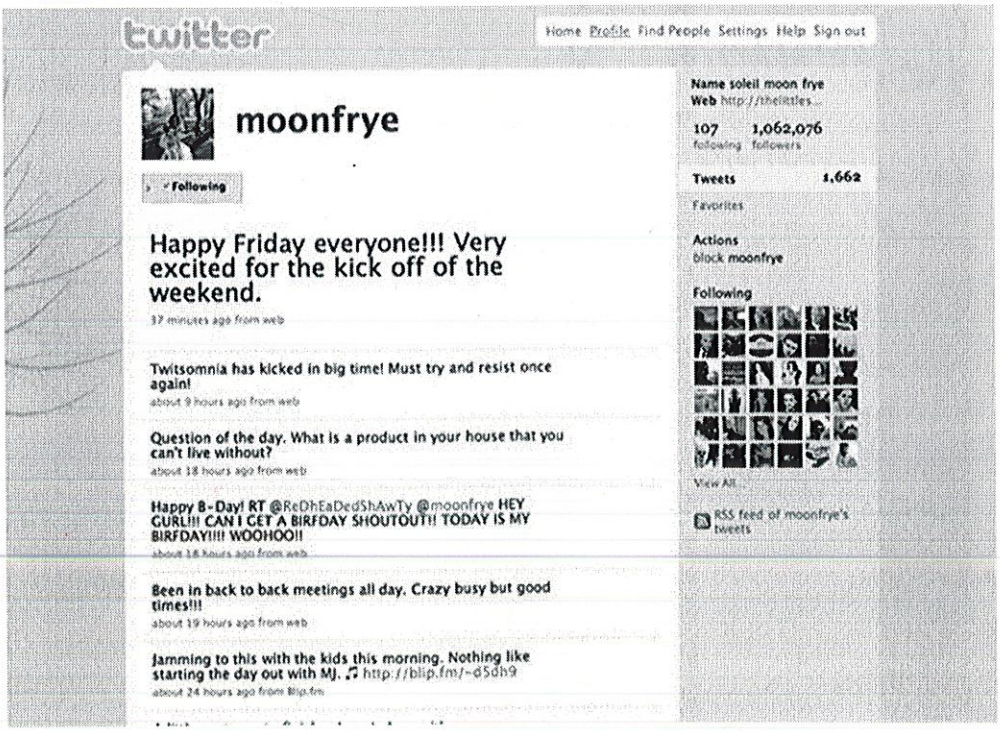
Reddit is another social news site founded in 2005. Reddit operates in a similar fashion to Digg, allowing users to vote content up or down. Users can view popular items, new items, and "controversial" items (presumably those items that have received a lot of both up and down votes). Reddit, like Digg, also allows users to comment on posted items.

Real Time Updates

Real-time updates have become the new norm in social media. With the advent of Twitter in 2006, status updates have become the new norm in social networking. Virtually all major social networks now allow real-time updates.

Twitter

Twitter was founded in 2006 and gained a lot of popularity during the 2007 SxSW (South by Southwest) conference. Tweets tripled during the conference, from 20k per day to 60k. Twitter has developed a cult-like following and has a number of famous users (Ashton Kutcher, Demi Moore, Soleil Moon Frye, MC Hammer, Oprah, Martha Stewart, and many, many more).



POSTEROUS

Posterous is the newest major microblogging application, started by Y Combinator in May 2009.Users post content via email. Emails can include attached photos, MP3s and other file types that are also posted. No initial signup is needed, setting it apart from most other social media services.

TUMBLR

Tumblr is sort of a cross between a lifestreaming application and a microblogging platform. Tumblr was founded in 2007 and had around 75,000 tumblebloggers switch to the service immediately. The site lets users post photos, video, text, audio, links, conversations, and other content on blog-like sites. There are mobile applications available for posting to Tumblr, making it ideal for lifestreaming. Tumblr is also very easy to use, making it well-suited to less technical users. It's similar to Twitter and other microblogging platforms in the way that it lets you follow other Tumblr users and see their updates in a specialized dashboard feed. Users can also "heart" (favorite) other Tumblr users' content and reblog posts from other users, keeping the original credit intact.

Chapter 3

3.1. PHP

PHP is a popular, open source programming language. Also, unlike some other languages, it isn't a framework in its own right, which means we can structure our application however we wish. Most modern web hosts support PHP and the database platform we will be using with it (MySQL) and although some other languages are gaining popularity (such as the Ruby on Rails framework), hosting for this isn't as common. Facebook, the World's largest social networking website, is written using PHP (albeit with countless customizations, improvements, and extras), as does Yahoo!, which operates a search engine, news portal websites, and various social websites too. Yahoo! also, until recently, employed Rasmus Lerdorf, the creator of the original PHP engine

3.2. Apache Web Server

The **Apache HTTP Server** commonly referred to as **Apache**, is web server software notable for playing a key role in the initial growth of the World Wide Web. In 2009 it became the first web server software to surpass the 100 million website milestone.[3] Apache was the first viable alternative to the Netscape Communications Corporation web server (currently known as Oracle iPlanet Web Server), and has since evolved to rival other web servers in terms of functionality and performance. Typically Apache is run on a Unix-like operating system.

Apache is developed and maintained by an open community of developers under the auspices of the Apache Software Foundation. The application is available for a wide variety of operating systems, including UNIX, GNU, FreeBSD, Linux, Solaris, Novell NetWare, AmigaOS, Mac OS X, Microsoft Windows, OS/2, TPF, and eComStation. Released under the Apache License, Apache is characterized as open-source software. Since April 1996 Apache has been the most popular HTTP server software in use. As of May 2011 Apache was estimated to serve 63% of all websites and 66% of the million busiest

FEATURES

Apache supports a variety of features, many implemented as compiled modules which extend the core functionality. These can range from server-side programming language support to authentication schemes. Some common language interfaces support Perl, Python, Tcl, and PHP. Popular authentication modules include mod_access, mod_auth, mod_digest, and mod_auth_digest, the successor to mod_digest. A sample of other features include SSL and TLS support (mod_ssl), a proxy module (mod_proxy), a URL rewriter (also known as a rewrite engine, implemented under mod_rewrite), custom log files (mod_log_config), and filtering support (mod_include and mod_ext_filter).
Popular compression methods on Apache include the external extension module, mod_gzip, implemented to help with reduction of the size (weight) of web pages served over HTTP. ModSecurity is an open source intrusion detection and prevention engine for web applications. Apache logs can be analyzed through a web browser using free scripts such as AWStats/W3Perl or Visitors.
Virtual hosting allows one Apache installation to serve many different actual websites. Apache features configurable error messages, DBMS-based authentication databases, and content negotiation. It is also supported by several graphical user interfaces (GUIs).
It supports password authentication and digital certificate authentication. Apache has a built in search engine and an HTML authorizing tool and supports FTP.

PERFORMANCE

Although the main design goal of Apache is not to be the "fastest" web server, Apache does have performance comparable to other "high-performance" web servers. Instead of implementing a single architecture, Apache provides a variety of MultiProcessing Modules (MPMs) which allow Apache to run in a process-based, hybrid (process and thread) or event-hybrid mode, to better match the demands of each particular infrastructure. This implies that the choice of correct MPM and the correct configuration is important. Where compromises in performance need to be made, the design of Apache is to reduce latency and increase throughput, relative to simply handling more requests, thus ensuring consistent and reliable processing of requests within reasonable time-frames.

The Apache version considered by the Apache Foundation as providing high-performance is the multi-threaded version which mixes the use of several processes and several threads per process

While this architecture works faster than the previous multi-process based topology (because threads have a lower overhead than processes), it does not match the performances of the event-based architecture provided by other servers, especially when they process events with several worker threads.

This difference can be easily explained by the overhead that one thread per connection brings (as opposed to a couple of worker threads per CPU, each processing many connection events). Each thread needs to maintain its own stack, environment, and switching from one thread to another is also an expensive task for CPUs.

3.3. MySQL

MySQL is the most popular SQL database in the open source community and is used almost universally by web sites running on open source systems. As powerful and flexible as it is lightweight and efficient, MySQL packs a large feature set into a very small and fast engine that now runs on more than 500,000 servers. This renowned online manual that has supported MySQL administrators and database developers for years is now available in paperback format. This book is an exact reproduction of the MySQL Reference Manual from the MySQL development team's Web site, minus some non-technical appendices. This version covers MySQL 4.0. Many sophisticated topics appear in this comprehensive manual, ranging from the hitches you may run into when first installing MySQL to internals that will help you tune your queries. *MySQL Reference Manual* contains all the comprehensive reference material one would expect for building the product, running administrative utilities, and using various API as well as MySQL's rich version of SQL. In addition, you can turn a page and find such unexpected riches as:

- A thorough comparison of MySQL with SQL standards and other databases
- A discussion of privileges and suggested uses of privileges to enhance security
- Directions for replicating a database and for running several MySQL servers on a single system
- Directions for initializing a database from a flat file

- Guidelines for estimating the performance of different queries
- A far-reaching discussion of optimization, with reference to the implementation of MySQL
- An extended inquiry into the effects of using delayed inserts
- A candid explanation of why various errors occur and how to recover from them
- Tips for weighted, full-text searches
- Detailed descriptions of the features, strengths, and weaknesses of available table formats
- A guide to adding new functions to MySQL

3.4. Framework

This will be a small, light-weight framework, as our focus is on social networking, and the purpose of the framework is purely to help us do this.

Important things to take care:

• About some common design and architectural patterns that solve common programming problems, including:

  o MVC: The Model-View-Controller architecture

  o The Registry pattern

  o The Factory pattern

  o The Front Controller pattern

• How to effectively structure files within a development framework

• How to build the framework, including:

  o How to handle user authentication

  o How to abstract database access functions

  o Template management

• Providing a single point of access to the site

It is also important for us to think about how we might allow users to change their own privacy settings, including which profile information they would like to make public, public only to their network, or completely private—particularly with regards to contact details and dates of birth. Some countries also have legislation in place governing the

management of user data, such as the Data Protection Act in the UK. This covers issues such as:

- Security—ensuring data is held securely, and isn't easy for others to access, unless the user's permission has been given
- Relevancy—ensuring data held is kept up to date and is relevant
- Removal—allowing users to request full removal of their data
- Access—allowing users to request copies of all data held about them

3.5. CakePHP

In 2005, Michal Tatarynowicz wrote a minimal version of a Rapid Application Framework in PHP. He found that it was the start of a very good framework. Michal published the framework under the MIT license, dubbing it Cake, and opened it up to a community of developers, who now maintain Cake under the name CakePHP.
CakePHP is a free, open-source, rapid development framework for PHP. It's a foundational structure for programmers to create web applications. Our primary goal is to enable you to work in a structured and rapid manner–without loss of flexibility.
CakePHP takes the monotony out of web development. We provide you with all the tools you need to get started coding what you really need to get done: the logic specific to your application. Instead of reinventing the wheel every time you sit down to a new project, check out a copy of CakePHP and get started with the real guts of your application.
CakePHP has an active developer team and community, bringing great value to the project. In addition to keeping you from wheel-reinventing, using CakePHP means your application's core is well tested and is being constantly improved.
Here's a quick list of features you'll enjoy when using CakePHP:

- Active, friendly community
- Flexible licensing
- Compatible with versions 4 and 5 of PHP
- Integrated CRUD for database interaction
- Application scaffolding
- Code generation
- MVC architecture
- Request dispatcher with clean, custom URLs and routes
- Built-in validation

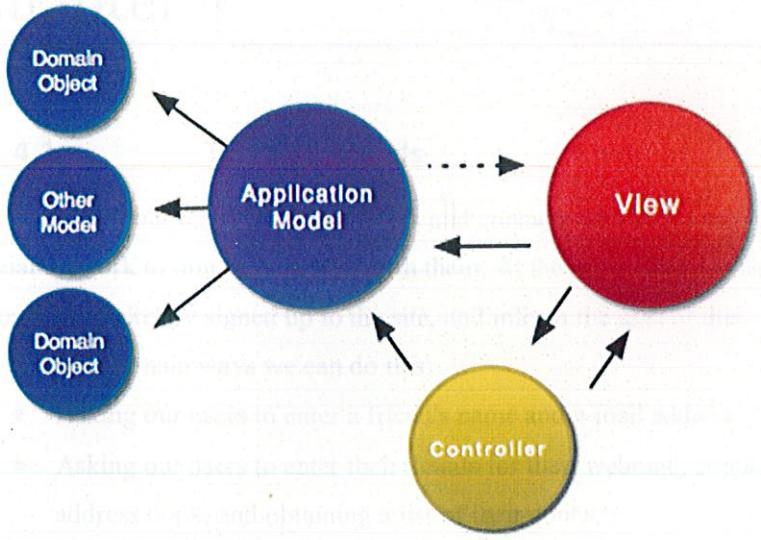
- Fast and flexible templating (PHP syntax, with helpers)
- View Helpers for AJAX, JavaScript, HTML Forms and more
- Email, Cookie, Security, Session, and Request Handling Components
- Flexible ACL
- Data Sanitization
- Flexible Caching
- Localization
- Works from any web site directory, with little to no Apache configuration involved

3.6. Salting Passwords

Our passwords are stored in the database as an MD5 one-way hash. This means we don't keep a copy of the user's password; instead, we hash the password when they try to log in, and compare the hash to the password in the database. If our database was compromised, our users' passwords should be safe. This hashing cannot be reversed, but there are dictionaries available for common words or phrases, which means it is possible to work out some passwords from the hashes. We can prevent this further by salting the password; this involves adding a "salt" to the password and then hashing it. This is typically done by creating a random string for each user and storing it in their row in the users table.

3.7. Model View Controller(MVC)

The Model-View-Controller pattern is an architectural design pattern designed to separate the user interface from the business logic of an application. The user interface (view) uses the controller to interact with the logic and data of the application (model). When they click the appropriate button to add the user as a friend, the controller processes this request from the user, and passes the request to the model, which updates the user's friends list, and where appropriate, sends any notifications. The view then updates, via instructions from the controller, to inform the user of the outcome of their request.


Domain Object
Other Model
Domain Object
Application Model
View
Controller

Chapter 4

4.1. Inviting friends

We can allow our users to invite friends and contacts who they know from outside of the social network to join and connect with them. At the same time, we can also see if these people have already signed up to the site, and inform the user of this.

There are two main ways we can do this:

- Asking our users to enter a friend's name and e-mail address
- Asking our users to enter their details for their webmail, connecting to their address book, and obtaining a list of their contacts

Once we have a name and an e-mail address we can either:

- Send the user an e-mail inviting them to join the site to connect with their friend
- Inform the user that someone with that e-mail address has already signed up and suggest they connect with them directly, or automatically connect them

4.2. Manually inviting friends

To allow users to invite other users we need to:

1. Request User to enter Friends' name and e-mail address.
2. Check to see if Friend's e-mail address exists in the website (that is, friend already a member).
3. If Friend is already a member, we suggest that User connect with him and show him Friend's profile.
4. If Friend hasn't already joined the site, we want to validate this data and display a template message to User showing the invitation message, which he can edit and personalize.
5. We allow User to edit the message.
6. Once User clicks on Send, we e-mail the invitation to Friend dynamically inserting User's details so Friend knows who it was that invited him.

4.3. Automatically inviting friends

Most social networking websites offer the user the chance to enter their details for their webmail login, to have the site automatically invite their contacts to use the site. In the past, this would be done by scripts that would connect to the various websites using libraries such as cURL, pretending to be a user, to obtain the contacts list. This technique isn't ideal, as the code obviously needs to be updated each time the site changes how it works.

Thankfully, most e-mail providers realize this is a useful feature, and so they have provided APIs that developers can interact with to obtain a list of contacts to e-mail. Of course, APIs change, but changes are generally announced in advance, and there is normally a wealth of resources for developers.

4.4. Listing users

User lists have a limitation with large social networks—they end up being large lists of users that are irrelevant to the user viewing the list. This can be overcome by listing a subset of users; for instance, those in a particular group, contact sphere, or network. For example, when Facebook started, users joined up to two networks, which was generally their university, school, workplace, or city. This could be used to segregate groups of users when listing them; obviously we wouldn't want to segregate users from each other, but this could make lists more meaningful. At this stage we don't have this concern; we can simply provide a paginated list of our users.

Chapter 5 Privacy

Privacy has been recognized as an important topic in the Internet for along time, and technological developments in the area of privacy tools are ongoing. However, their focus was mainly on the individual. With the proliferation of social network sites, it has become more evident that the problem of privacy is not bounded by the perimeters of individuals but also by the privacy needs of their social networks. The objective of this paper is to contribute to the discussion about privacy in social network sites, a topic which we consider to be severely under-researched. We propose a framework for analyzing privacy requirements and for analyzing privacy-related data. We outline a combination of requirements analysis, conflict-resolution techniques, and a P3P extension that can contribute to privacy within such sites.

Technically, the use of SNS data for novel purposes is even simpler than in traditional eCommerce. The very essence of social media is that user-profile information is public (as opposed to, for example, Amazon's usage data which are an important and secret business asset of the company). Moreover, the data often carries semantic markup and/or is presented in a uniform (hence easily minable) manner, for example as RSS feeds. Thus, while the legal issues at this level are the same in SNSs as in other sites, technical uses become simpler

Private data is disclosed to the SNS operator for its internal purposes only. This data must not be disclosed unless explicit consent is given. An example is the user's email address provided upon registration. Group data is disclosed to the SNS operator and can be accessed by other users of the same SNS that are also in the same group as the user: data disclosure is limited to the group. Here we imagine messages shared among a certain group, almost like a closed mailing-list. Community data has been disclosed to the SNS operator and is available to all registered and logged-in users of the SNS. The data is not accessible for anonymous SNS visitors. Examples are the user's online status, her contacts, her member page details, photos, etc.

Public data has been disclosed to the SNS operator and is made accessible for all SNS visitors, including anonymous visitors: this may include the fact that the user is registered in the SNS, her user name, or her guestbook.

5.1. Identifying privacy conflicts in the interaction of requirements for social network sites

The main idea of the **Multilateral Security Requirements Analysis (MSRA)** method is to consider the security and privacy interests or needs of all stakeholders related to the system. An important aspect of the method is to identify interest conflicts among these stakeholders and develop mechanisms for negotiating these conflicts. Here we introduce aspects of conflict identification and negotiation mechanisms in multilateral security requirements analysis.

5.2. Stakeholders and their privacy interests

In MSRA, stakeholders of a system are all persons who have some functional, knowledge, security or privacy interest in the system. This encompasses all persons involved in the conception, production, use and maintenance of the system. Stakeholders encompass more than users (those who will use the functionality of the system). Stakeholders, for example, include all persons who have a privacy interest in the system. This could be stakeholders representing legal requirements as well as non-users whose data is processed by the system – i.e. patients in a Hospital Information System or customers in a Customer Relationship Management System. The sender of an email to a Gmail account may count as a stakeholder of the Gmail platform, although she is not a user of that platform. This stakeholder is likely to have different privacy interests towards the Gmail platform than a user or provider of the platform.

The inclusion of an external sender of an email as a stakeholder of an email platform also points to the fact that further stakeholders may be acquired once the system is running. Subsuming the privacy interests of all prospective stakeholders is not possible during the development of the system. Nevertheless, the potential of discovering new stakeholders requires the conception of negotiation mechanisms during the development process that anticipate potential divergences in privacy interests during run-time. Moreover, the

introduction of new stakeholders and their requirements often demands a review of all security and other requirements and hence an iterative approach.

5.3. Enhancing privacy in social network sites using P3P

Appropriate measures need to be taken to satisfy privacy requirements in an operational SNS. This includes the conception and adaptation of technologies and processes, mainly privacy languages and tools to interpret and enforce these languages. The design goal is Twofold:

First, we need mechanisms to ensure that data/information of one privacy level must not be made accessible via data/information of a lower privacy level. For example one should not be able to perform (data) inferences towards personal information that is private on a "community level", from personal information that is private on a "public level". The AOL privacy breach gives evidence that trivial anonymization is insufficient for preventing data inferences that may even lead to the identification of individuals.

Second, we need mechanisms that prevent users from disclosing personal information about other users inside an SNS.

Both objectives should be addressed within the existing technological and legal infrastructure of Privacy Enhancing Technologies (PETs), Privacy Protocols (especially P3P and APPEL / XPref), and mandatory legislation.

P3P, the Platform for Privacy Preferences, is a protocol designed to inform Web users about the data-collection practices of Web sites. It provides a way for a Web site to encode its data-collection and data-use practices in a machine-readable XML format known as a P3P policy. Moreover, P3P enables Web users to understand what data will be collected by sites they visit, how that data will be used, and what data/uses they may "opt-out" of or "opt-in" to. An SNS operator will post a P3P policy on its Web site to communicate its data handling practices. Visitors and users can receive this policy in a textually presented format. Their decision whether to send data to the site or not can be supported by APPEL rules: APPEL, A P3P Preference Exchange Language, allows a user to express her preferences in a set of preference rules, interpreted by her user agent to

make automated or semi-automated decisions regarding the acceptability of P3P Privacy Policies

XPref is a newer privacy preference language, more expressive than APPEL yet easier to use. In a P3P policy, one or several statements describe data practices that are applied to particular types of data. A statement indicates recipients, usage purposes, and a retention-time for data elements. Every potential data usage must be indicated by an appropriate statement; hence statements span a superset over the actually implemented data usage. P3P hereby translates the privacy concepts of, e.g., European privacy legislation and the OECD Fair Information Practices into a machine-readable policy.

A professional SN collects the username, publicly accessible, and the details about the user's job, the latter being secured. Users may join special interest groups based on their industrial and departmental focus, e.g. "Helpdesk Professionals Group", "Data Protection Officers Group", or "CEO VIP Club".

Chapter 6

Since the success of an SNS depends on the number of users it attracts, there is pressure on SNS providers to encourage design and behavior which increase the number of users and their connections. As with every fast-growing technology, however, security and privacy have not been the first priority in the development of SNSs. As a result, along with the above benefits, significant privacy and security risks have also emerged

6.1. Digital Dossier Aggregation

Profiles on SNSs can be downloaded and stored over time and incrementally by third parties, creating a digital dossier of personal data. Information revealed on an SNS can be used for purposes and in contexts different from the ones the profile owner had considered. Due to the greatly diminished costs associated with disk storage and Internet downloads, it is feasible to take regular snapshots of an entire network and store the profiles of its members indefinitely. On the other hand, it is costly and technically challenging to manage the complete deletion of data when it is no longer necessary. The information contained in individual profiles can be accumulated easily in order to track and highlight changes (e.g. list of girlfriends over time). This could be done either overtly or covertly by internal employees or external applications which can access profile information in bulk (e.g. through search features).

A common vulnerability is that more private attributes which are directly accessible by profile browsing can be accessed via search (e.g. a person's name and profile image is accessible via search on MySpace, Facebook and others, unless default privacy settings are changed).

6.2. Secondary Data Collection

In addition to personal data knowingly disclosed in a profile, an SNS member discloses personal information to the network operator using the network itself: data such as time and length of connections, location (IP address) of connection, other users' profiles

visited, messages sent and received and so forth. While this in itself is not specific to SNSs, in other contexts information on a user's behavior is protected to some degree by being spread across multiple websites, e-mail accounts (e.g. work e-mail and one or multiple private e-mail accounts), one or more instant messaging systems and multiple Internet access points (at work, at home, through a cell phone etc.).
Since SNSs are not currently portable (there is no widely used standard for exchanging SNS data and there is a significant overhead in joining a new network: defining relationships, profiles, inviting contacts and so forth), there is a strong tendency towards amalgamating all SNS activity under a single provider. This is then a powerful data warehouse for the owners of the SNS. Despite this, there is currently a lack of transparency about certain data collection practices. For example, it is not made clear to Users how and to whom the data about visits to other profiles are made public. Privacy policies tend to be vague in specifying what is and what not personal information is. This does not specify which elements of 'profile information' are disclosed to third parties. While some elements of a person's profile, such as their hobbies, are arguably not classified as personally identifiable, the high value paid for SNSs (e.g. $35 per user for MySpace in 2006) strongly suggests that the data they contain is being used to considerable financial gain.

6.3. Face Recognition

User-provided digital images are an integral and exceedingly popular part of profiles on SNSs. As an example, Facebook hosts in excess of 1.7 billion user photos (as of 21 May 2007), a database growing at a rate of more than 60 million per week. Since images are tied to individual profiles and often either explicitly (through, for example, labeled boxes on images) or implicitly (through recurrence) identify the profile holder, they constitute a data source suitable for correlating profiles across services using face recognition. The efficiency of face recognition algorithms has improved dramatically over the last decade. While early systems performed well only on carefully controlled images, newer systems can handle a wide variety of image conditions. The combination of better algorithms with faster computing hardware and essentially unlimited storage enables comparisons of large numbers of images

.

Until recently, the use of face recognition software has been the domain of law enforcement and border-control agencies. Recently, however, some mainstream web service providers have announced work on integrating face recognition technologies into their applications

.

6.4. CBIR (Content-based Image Retrieval)

Related to face recognition, Content-based Image Retrieval (CBIR) was originally developed for digital forensics. CBIR is an emerging technology which is able to match features, such as identifying aspects of a room (e.g. a painting) in very large databases of images. Traditional search terms are replaced with a reference image or image template. Search is designed to be resilient to cropping, resizing, rotation and quality adjustment (e.g. for JPEG)

. Currently, privacy controls on images uploaded and the advice given on SNSs do not take into account the possibility of CBIR, and few people are aware of the consequences of posting images with location-specific content online.

6.5. Linkability from Image Metadata, Tagging and Cross-profile Images

Many SNSs now allow users to tag images with metadata such as the name of the person in the photo, a link to their SNS profile (even if they are not the owner/controller of that profile), or even their e-mail address. As an example, shown in the following extract from the Facebook help pages, Facebook allows tagging of images with profile data and even e-mail addresses.

Even if users exercise caution over which images they post of themselves and their location, their privacy may be under even greater threat from images posted by others.

While profile links can usually only be included for profiles in a person's friend list, given the low trust threshold for inclusion in this list, this does not offer much consolation. Very few SNSs offer privacy tools to control the tagging of images with links to their profile or the accessibility of tagged images in search results. Another aspect of image metadata is that many cameras embed metadata about the camera in the image

including, in many cases, the serial number of the camera. Given that many cameras are linked to address data through warranty registration cards, this constitutes a threat to the user's privacy. An interesting recent case was the posting of a full illegal copy of Harry Potter and the Deathly Hallows which included embedded versions of the serial number of the camera used to take it, as well as the exact date and time the images were taken

6.6. Difficulty of Complete Account Deletion

Users wishing to delete accounts from SNSs will find that, although it is usually very easy to remove their primary pages, secondary information such as public comments they have made on other accounts using their identity will remain online. Moreover, in general there is ambiguity as to whether information will be deleted upon account closure. As an example, the Facebook privacy policy makes the statement:

"Removed information may persist in backup copies for a reasonable period of time but will not be generally available to members of Facebook."

Upon 'deactivating' an account, users of some providers such as Facebook receive an e-mail telling them how to 'reactivate' their account –implying that a copy is kept of personal data. Furthermore, personal data cannot be completely deleted unless users manually remove all public notes or comments on other profiles. This is usually not feasible due to the large number of steps involved.

6.7. SN Spam

SN spam is unsolicited messages propagated using SNSs. Many spammers have sought to capitalize on the exponential growth of SNSs and the free traffic they provide. This is a very serious issue since statistics suggest that SNSs are replacing e-mail in some circles as a method of communication. This means that the same scale of spam problems which have affected e-mail communications systems could soon affect SNSs.

Common techniques used by spammers include:

• The use of specialized SNS spamming software such as FriendBot to automate friend invitations and note/comment posting. Such tools use the SNSs' search tools to target a

certain demographic segment of the users and communicate with them from an account disguised as that of a real person.

• The sending of notes typically including embedded links to pornographic or other product sites designed to sell something.
• Friend invitations, using an attractive profile which is likely to persuade someone to accept the invitation. The profile or the invitation then contains links to external sites advertising products or even phishing for passwords.
• The posting of spam comments on public notes or comments areas of 'friends'. Typically, spammers will create as many 'friends' as possible, focusing on those with public notes or comments areas or message boards and fitting a specific demographic profile (Friendbot has features to automate this) and then post spam messages on their public notes or comments areas.
• Stealing members' passwords to insert and promote their offers on another profile.

Until recently no filters were available for notes or friend requests. The best a user could do was block notes from the sender's address. MySpace, for example, now includes an option for users to 'report spam/abuse' addresses. Spammers, however, frequently change their address from one throw-away account to another.

6.8. SNS Aggregators

Social Aggregators such as Snag, ProfileLinker and many others are relatively new applications, which address the problem of having to set up social networks on multiple platforms by integrating data from various SNSs into a single web application. This unfortunately multiplies the vulnerability of accounts by giving read/write access to several SNS accounts based on a single weak username/password authentication. Social Aggregators also carry the risk of increasing the potential for attackers to mine data across sites since many offer cross-SNS search features.

6.9. Spear Phishing using SNSs and SN-specific Phishing

Spear phishing describes any highly targeted phishing attack. The existence of easily accessible self-created 'profiles' and self-declared 'circles of friends' on SNSs allows a phisher to harvest large amounts of reliable social network information which may be

used for a highly personalized phishing attack. An experiment conducted by researchers at the University of Indiana showed that, using data available on SNSs, e-mail phishing attacks can achieve a hit rate of 72%, compared with a control of 15%.

A related threat is the use of SNSs for the phishing attack itself (rather than just gathering data to be used elsewhere). The worm JS/Quickspace.A was designed to spread through MySpace profile pages. Pages were infected with links to a phishing site which then asked for the user's logon details and used these to embed a link to the phishing site in the stolen profile.

Although this is merely a new modality of an existing threat, the extra trust created by the 'circle of friends' can make this a particularly effective form of phishing attack.
The spread of such phishing attacks is greatly increased by the vulnerability of SNSs to social engineering techniques based on the infiltration of SNSs with low entry thresholds. Another factor is the prevalence of scripting attacks allowing the automated injection of phishing links.

6.10. Infiltration of Networks Leading to Information Leakage

Some information is only available to 'friends' or members of a restricted group and this is the first line of defense in protecting privacy on SNSs. Since it is often very easy to become someone's 'friend' under false pretenses, this mechanism is not very effective. Currently it is even possible to use scripts to invite friends on MySpace, and the growing amount of specialized commercial software such as Friendbot and FriendBlasterPro, created for exactly this purpose, suggests that this is increasingly common.
CAPTCHAs are not implemented by default and in fact a significant number of 'SNS spam' invitations do get sent.

Some SNSs have extremely broad criteria for membership of a network (and access to data within that network). For example, currently anyone with an appropriate e-mail address can join any geographical (i.e. city) network on the Facebook and gain access to the public profiles of this network.

Social and commercial pressure to get as many friends as possible means there is often a tendency to accept friend requests without checking their authenticity or suitability. In a recent experiment, antivirus company Sophos created a profile page for 'Freddi Staur' (an anagram of 'ID Fraudster'), a green plastic frog with only minimal personal information in his profile. They then sent out 200 friend requests to see how many people would respond, and how much personal information could be gleaned from the respondents. The following are some of the results:

- 87 of the 200 users contacted responded to Freddi, with 82 leaking personal information (41% of those approached)
- 72% of respondents divulged one or more e-mail address
- 84% of respondents listed their full date of birth

Chapter 7

SYSTEM REQUIREMENTS

HARDWARE REQUIREMENTS (Minimum)

CPU	: 1.6 GHz
RAM	: 384 MB
DISPLAY	: 1024*768 Monitor
Hard Disk	: 8 GB

SOFTWARE REQUIREMENTS

OS	: Windows XP/Vista/7/Linux/ OSX
Software	: Any Web Browser (Firefox, Chrome, Safari)
	Apache Server with PHP extension
	MySQL Database

Coding: PHP

Conclusion

During the implementation of project, we shall learn and implement various Security Measures like SSL Security and HTTPS. While doing a project which deals with personal data of millions of users, Security of user data is a prime concern and is one of the priorities of our project.

Future Scope

The social networking site will be enhanced with a version optimized for mobile devices. For doing this we need to embed a JavaScript for detecting browser version and usage of different CSS Style sheet for presentation on mobile.

We plan to develop on RSS feeds for easy data syndication. Web syndication and RSS feeds offer users easier access to relevant information without having to visit numerous web sites. Web syndication also allows web designers to flow content into their website from relevant locations across the Internet. This technological advance typifies the evolution from the first generation world wide web to the current social networking web. Another advance in this collaborative direction is the ability to re-use and re-purpose digital files for different uses.

References

- CakePHP Documentation (http://book.cakephp.org/view/3/The-Manual)
- R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu. XPref: a preference language for P3P. Computer
- Facebook Privacy Policy www.facebook.com/policy.php
- Safefamilies, Safety Tips for Social Networking www.safefamilies.org/socialnetworking.php
- Friendbot, automated friend adding software www.friendbot.com/
- Friendblasterpro, friend adding software www.addnewfriends.com/
- List of Social Network Aggregators http://mashable.com/2007/07/17/social-network-aggregators/

Appendix: Coding

INDEX.PHP

```
	if (!defined('WEBROOT_DIR')) {
		define('WEBROOT_DIR', basename(dirname(__FILE__)));
	}
	if (!defined('WWW_ROOT')) {
		define('WWW_ROOT', dirname(__FILE__) . DS);
	}
	if (!defined('CORE_PATH')) {
		if (function_exists('ini_set') && ini_set('include_path',
CAKE_CORE_INCLUDE_PATH . PATH_SEPARATOR . ROOT . DS . APP_DIR . DS .
PATH_SEPARATOR . ini_get('include_path'))) {
			define('APP_PATH', null);
			define('CORE_PATH', null);
		} else {
			define('APP_PATH', ROOT . DS . APP_DIR . DS);
			define('CORE_PATH', CAKE_CORE_INCLUDE_PATH . DS);
		}
	}
	if (!include(CORE_PATH . 'cake' . DS . 'bootstrap.php')) {
		trigger_error("CakePHP core could not be found.  Check the
value of CAKE_CORE_INCLUDE_PATH in APP/webroot/index.php.  It should
point to the directory containing your " . DS . "cake core directory and
your " . DS . "vendors root directory.", E_USER_ERROR);
	}
	if (isset($_GET['url']) && $_GET['url'] === 'favicon.ico') {
		return;
	} else {
		$Dispatcher = new Dispatcher();
		$Dispatcher->dispatch();
```

CSS.PHP

```
	function write_css_cache($path, $content) {
		if (!is_dir(dirname($path))) {
			mkdir(dirname($path));
		}
		$cache = new File($path);
		return $cache->write($content);
	}

	if (preg_match('|\.\.|', $url) || !preg_match('|^ccss/(.+)$|i',
$url, $regs)) {
		die('Wrong file name.');
	}

	$filename = 'css/' . $regs[1];
	$filepath = CSS . $regs[1];
	$cachepath = CACHE . 'css' . DS . str_replace(array('/','\\'), '-
', $regs[1]);

	if (!file_exists($filepath)) {
		die('Wrong file name.');
	}
```

STATUS.PHP

```
<?php
class Status extends AppModel
{
      var $name='Status';
      var $useTable='statuss';
}
```

USER.PHP

```
<?php
class User extends AppModel {
var $name='User';
var $useTable = 'users';
}
?>
```