Volume 8 • Issue 4 • August 2014

Editor-in-Chief
Dr. Chen-Chi Shing

INTERNATIONAL JOURNAL OF
COMPUTER SCIENCE AND SECURITY (IJCSS)

ISSN : 1985-1553

Publication Frequency: 6 Issues / Year

CSC PUBLISHERS
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND SECURITY (IJCSS)

VOLUME 8, ISSUE 4, 2014

EDITED BY
DR. NABEEL TAHIR

ISSN (Online): 1985-1553

International Journal of Computer Science and Security is published both in traditional paper form and in Internet. This journal is published at the website http://www.cscjournals.org, maintained by Computer Science Journals (CSC Journals), Malaysia.

IJCSS Journal is a part of CSC Publishers
Computer Science Journals
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND SECURITY (IJCSS)

Book: Volume 8, Issue 4, August 2014
Publishing Date: 10 - 08 - 2014
ISSN (Online): 1985 -1553

IJCSS Journal is a part of CSC Publishers
http://www.cscjournals.org

Published in Malaysia

Typesetting: Camera-ready by author, data conversation by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2014

EDITORIAL PREFACE

This is *Fourth* Issue of Volume *Eight* of the International Journal of Computer Science and Security (IJCSS). IJCSS is an International refereed journal for publication of current research in computer science and computer security technologies. IJCSS publishes research papers dealing primarily with the technological aspects of computer science in general and computer security in particular. Publications of IJCSS are beneficial for researchers, academics, scholars, advanced students, practitioners, and those seeking an update on current experience, state of the art research theories and future prospects in relation to computer science in general but specific to computer security studies. Some important topics cover by IJCSS are databases, electronic commerce, multimedia, bioinformatics, signal processing, image processing, access control, computer security, cryptography, communications and data security, etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 8, 2014, IJCSS appears with more focused issues. Besides normal publications, IJCSS intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

This journal publishes new dissertations and state of the art research to target its readership that not only includes researchers, industrialists and scientist but also advanced students and practitioners. The aim of IJCSS is to publish research which is not only technically proficient, but contains innovation or information for our international readers. In order to position IJCSS as one of the top International journal in computer science and security, a group of highly valuable and senior International scholars are serving its Editorial Board who ensures that each issue must publish qualitative research articles from International research communities relevant to Computer science and security fields.

IJCSS editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJCSS. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJCSS provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members
International Journal of Computer Science and Security (IJCSS)

EDITORIAL BOARD

Assistant Professor Vishal Bharti
Maharishi Dayanand University
India

Dr. Parvinder Singh
University of Sc. & Tech
India

Assistant Professor Vishal Bharti
Maharishi Dayanand University,
India

TABLE OF CONTENTS

Volume 8, Issue 4, August 2014

Pages

PI-Tool To Improve Performance of Application In Multi-core Architecture

P. Bala Subramanyam Raju *bsr3011@gmail.com*
Research Scholar, Dept. of Computer Science,
S.V University, Tirupathi
Chittoor (Dt) AP, India Pin 517501

P. Govindarajulu *PGovindarajulu@yahoo.com*
Professor, Dept. of Computer Science,
S.V University, Tirupathi,
Chittoor (Dt) AP, India Pin 517501

Abstract

Multi-core CPUs become increasingly popular on server-grade hardware as well as end-user desktop PCs or laptops. Necessary efforts are taken in the community towards developing applications optimized for multi-core architecture [1], still every user is expecting to run applications or programs with more speed than normal execution. In order to achieve maximum possible performance, there is a need for fine tuning in the areas of Operating System scheduling, Process Priority and CPU Affinity. So this paper presents PI-tool, to aid user by changing priority and CPU Affinity to his application or program at runtime to achieve better performance .The PI-tool is compared with Red Hat TUNA tool with NQueens program the results are presented for evaluation.

Keywords*:* CPU Affinity, Ping-Pong Effect, Chip Multi-core, Performance, Benchmark, Microprocessor, Software Challenges, OS (Operating System) High Performance Computing [HPC].

1. INTRODUCTION

Modern microprocessor is the most complex systems ever created by humans. A single chip, roughly fingernail size contains trillions of transistors to perform billions of operations per second. The advent of multi-core processors is able to still deliver more computing power. It is a complex, challenging task to take full advantage of today's high performance computer architectures. In order to achieve this, research is focused in the areas of Operating system(OS), OS Service Libraries ,Cache Management, Process Scheduling, Modifying Existing applications, and Other software's management[,2,3,4,5,6,7,8] to identify the performance bottlenecks to provide a way ,to full utilization of modern microprocessor effectively.

New process created will be assigned to processor Queue by the operating system in some order or priority Example FIFO, LIFO, Round Robin Scheduling. The process doesn't have any choice of choosing the processor on which it is going to execute. Linux Kernel by default provides soft affinity [9] (also called natural affinity), it is the tendency of the scheduler to keep processes on the same CPU as long as possible. In certain situations the scheduler bounces processes between multiple processors each time they are scheduled and rescheduled called Ping-Pong effect. Migrating processes from one core to another core can be costly due to cache invalidation, this problem is considered to be the bottleneck for performance. Low priority for the process is also another major bottleneck for the performance of the application. So this paper presents a tool which gives the user, capability to bind his application to less utilizing core to reduce Ping-Pong effect, along with this user can increase the priority of process. High priority processes are run before low-priority processes, they are also allowed to run longer period before being pre-empted**.**

The rest of the paper is organized as follows: Section 2 presents a brief survey of performance tuning tools, advantages of CPU Affinity is discussed in Section 3, and description of PI-Tool presented in Section 4. Section 5 specifies software and hardware details and tuning of a process is discussed in section 6. Section 7 presents the comparative study the effects of CPU Affnitity. Section 8 deals with positive and negative effects of using PI-Tools .Finally a conclusion with Future work is made in Section 9.

2. A SURVEY OF PERFORMANCE TUNING TOOLS

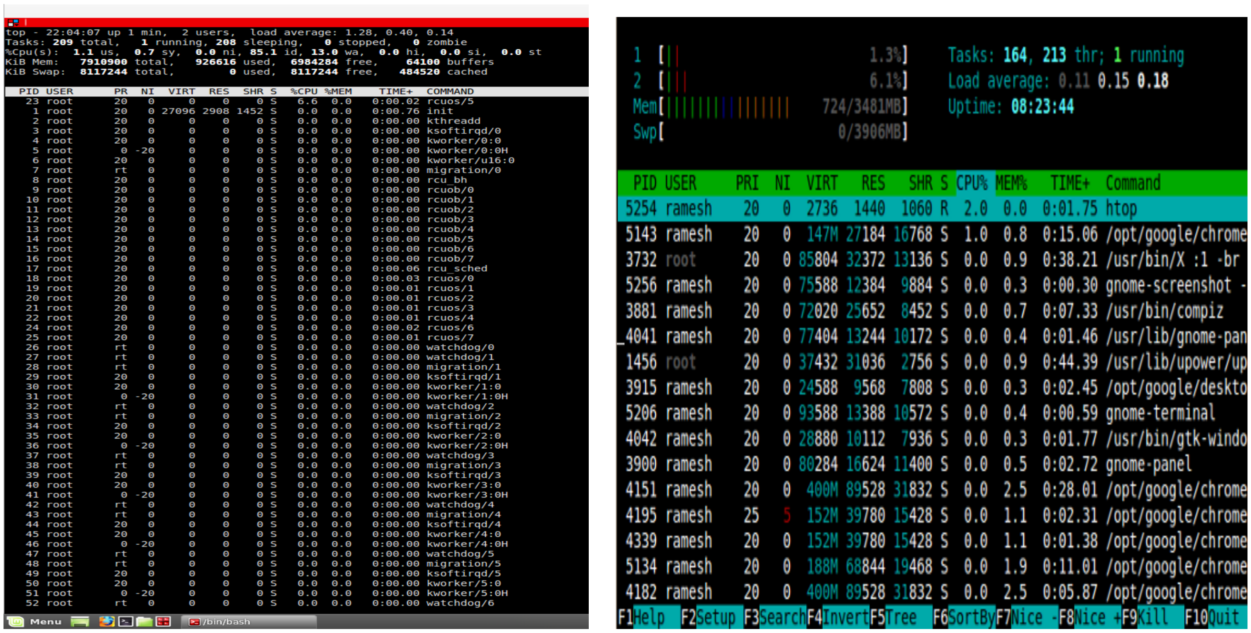
In this section, we present a historical survey on top 5 performance monitoring and tuning tools from early systems to today. We will discuss the goals, strengths and weakness of each tool.

2.1 TOP & HTOP

The top is a text-mode program provides a dynamic real-time view of a running system.it displays the most CPU-intensive tasks running on the server and updates the list every five seconds. It can sort the tasks by CPU usage, memory usage and runtime [10]

The htop program is very similar to top but has some extra features that make it an even better command line utility for process monitoring. It allows you to scroll vertically and horizontally. Tasks related to processes (killing, renicing) can be done without entering their PIDs [11].

The below figure shows the sample output top of htop program



```
top - 22:04:07 up 1 min,  2 users,  load average: 1.28, 0.40, 0.14
Tasks: 209 total,   1 running, 208 sleeping,   0 stopped,   0 zombie
%Cpu(s):  1.1 us,  0.7 sy,  0.0 ni, 85.1 id, 13.0 wa,  0.0 hi,  0.0 si,  0.0 st
KiB Mem:   7910900 total,   926616 used,  6984284 free,    64100 buffers
KiB Swap:  8117244 total,        0 used,  8117244 free,   484520 cached

  PID USER      PR  NI  VIRT  RES  SHR S  %CPU %MEM    TIME+  COMMAND
   23 root      20   0     0    0    0 S   6.6  0.0   0:00.02 rcuos/5
    1 root      20   0 27096 2908 1452 S   0.0  0.0   0:00.76 init
    2 root      20   0     0    0    0 S   0.0  0.0   0:00.00 kthreadd
    3 root      20   0     0    0    0 S   0.0  0.0   0:00.00 ksoftirqd/0
    4 root      20   0     0    0    0 S   0.0  0.0   0:00.00 kworker/0:0
    5 root       0 -20     0    0    0 S   0.0  0.0   0:00.00 kworker/0:0H
    6 root      20   0     0    0    0 S   0.0  0.0   0:00.00 kworker/u16:0
    7 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 migration/0
    8 root      20   0     0    0    0 S   0.0  0.0   0:00.00 rcu_bh
    9 root      20   0     0    0    0 S   0.0  0.0   0:00.00 rcuob/0
   10 root      20   0     0    0    0 S   0.0  0.0   0:00.00 rcuob/1
   11 root      20   0     0    0    0 S   0.0  0.0   0:00.00 rcuob/2
   12 root      20   0     0    0    0 S   0.0  0.0   0:00.00 rcuob/3
   13 root      20   0     0    0    0 S   0.0  0.0   0:00.00 rcuob/4
   14 root      20   0     0    0    0 S   0.0  0.0   0:00.00 rcuob/5
   15 root      20   0     0    0    0 S   0.0  0.0   0:00.00 rcuob/6
   16 root      20   0     0    0    0 S   0.0  0.0   0:00.00 rcuob/7
   17 root      20   0     0    0    0 S   0.0  0.0   0:00.06 rcu_sched
   18 root      20   0     0    0    0 S   0.0  0.0   0:00.03 rcuos/0
   19 root      20   0     0    0    0 S   0.0  0.0   0:00.01 rcuos/1
   20 root      20   0     0    0    0 S   0.0  0.0   0:00.01 rcuos/2
   21 root      20   0     0    0    0 S   0.0  0.0   0:00.01 rcuos/3
   22 root      20   0     0    0    0 S   0.0  0.0   0:00.01 rcuos/4
   24 root      20   0     0    0    0 S   0.0  0.0   0:00.02 rcuos/6
   25 root      20   0     0    0    0 S   0.0  0.0   0:00.01 rcuos/7
   26 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 watchdog/0
   27 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 watchdog/1
   28 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 migration/1
   29 root      20   0     0    0    0 S   0.0  0.0   0:00.00 ksoftirqd/1
   30 root      20   0     0    0    0 S   0.0  0.0   0:00.00 kworker/1:0
   31 root       0 -20     0    0    0 S   0.0  0.0   0:00.00 kworker/1:0H
   32 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 watchdog/2
   33 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 migration/2
   34 root      20   0     0    0    0 S   0.0  0.0   0:00.00 ksoftirqd/2
   35 root      20   0     0    0    0 S   0.0  0.0   0:00.00 kworker/2:0
   36 root       0 -20     0    0    0 S   0.0  0.0   0:00.00 kworker/2:0H
   37 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 watchdog/3
   38 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 migration/3
   39 root      20   0     0    0    0 S   0.0  0.0   0:00.00 ksoftirqd/3
   40 root      20   0     0    0    0 S   0.0  0.0   0:00.00 kworker/3:0
   41 root       0 -20     0    0    0 S   0.0  0.0   0:00.00 kworker/3:0H
   42 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 watchdog/4
   43 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 migration/4
   44 root      20   0     0    0    0 S   0.0  0.0   0:00.00 ksoftirqd/4
   45 root      20   0     0    0    0 S   0.0  0.0   0:00.00 kworker/4:0
   46 root       0 -20     0    0    0 S   0.0  0.0   0:00.00 kworker/4:0H
   47 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 watchdog/5
   48 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 migration/5
   49 root      20   0     0    0    0 S   0.0  0.0   0:00.00 ksoftirqd/5
   50 root      20   0     0    0    0 S   0.0  0.0   0:00.00 kworker/5:0
   51 root       0 -20     0    0    0 S   0.0  0.0   0:00.00 kworker/5:0H
   52 root      rt   0     0    0    0 S   0.0  0.0   0:00.00 watchdog/6
```

```
  1  [||                      1.3%]     Tasks: 164, 213 thr; 1 running
  2  [|||                     6.1%]     Load average: 0.11 0.15 0.18
  Mem[||||||||||||||||   724/3481MB]    Uptime: 08:23:44
  Swp[                     0/3906MB]

  PID USER     PRI  NI  VIRT   RES   SHR S CPU% MEM%   TIME+  Command
 5254 ramesh    20   0  2736  1440  1060 R  2.0  0.0  0:01.75 htop
 5143 ramesh    20   0  147M 27184 16768 S  1.0  0.8  0:15.06 /opt/google/chrome
 3732 root      20   0 85804 32372 13136 S  0.0  0.9  0:38.21 /usr/bin/X :1 -br
 5256 ramesh    20   0 75588 12384  9884 S  0.0  0.3  0:00.30 gnome-screenshot -
 3881 ramesh    20   0 72020 25652  8452 S  0.0  0.7  0:07.33 /usr/bin/compiz
 4041 ramesh    20   0 77404 13244 10172 S  0.0  0.4  0:01.46 /usr/lib/gnome-pan
 1456 root      20   0 37432 31036  2756 S  0.0  0.9  0:44.39 /usr/lib/upower/up
 3915 ramesh    20   0 24588  9568  7808 S  0.0  0.3  0:02.45 /opt/google/deskto
 5206 ramesh    20   0 93588 13388 10572 S  0.0  0.4  0:00.59 gnome-terminal
 4042 ramesh    20   0 28880 10112  7936 S  0.0  0.3  0:01.77 /usr/bin/gtk-windo
 3900 ramesh    20   0 80284 16624 11400 S  0.0  0.5  0:02.72 gnome-panel
 4151 ramesh    20   0  400M 89528 31832 S  0.0  2.5  0:28.01 /opt/google/chrome
 4195 ramesh    25   5  152M 39780 15428 S  0.0  1.1  0:02.31 /opt/google/chrome
 4339 ramesh    20   0  152M 39780 15428 S  0.0  1.1  0:01.38 /opt/google/chrome
 5134 ramesh    20   0  188M 68844 19468 S  0.0  1.9  0:11.01 /opt/google/chrome
 4182 ramesh    20   0  400M 89528 31832 S  0.0  2.5  0:05.87 /opt/google/chrome
F1Help  F2Setup F3SearchF4InvertF5Tree  F6SortByF7Nice -F8Nice +F9Kill  F10Quit
```

FIGURE 1: Sample Output Top of Htop Program.

2.2 SYSSTAT

Sysstat is a powerful logging and monitoring tool for Linux/Unix systems. It can be used to monitor system performance and troubleshoot problems. The sysstat package contains the sar, mpstat and iostat commands for Linux. The sar command collects, report, or save Linux system activity information like CPU, Memory utilization, Network, IO transfers, process activity, device activity and interrupts. However, analyzing information provided by sar can be difficult. The iostat command reports CPU utilization and I/O statistics for disks. The mpstat command reports global and per-processor statistics. The statistics reported by sar concern I/O transfer rates, paging activity, process-related activities, interrupts, network activity, memory and swap space utilization, CPU utilization, kernel activities and TTY statistics, among others[12,13].

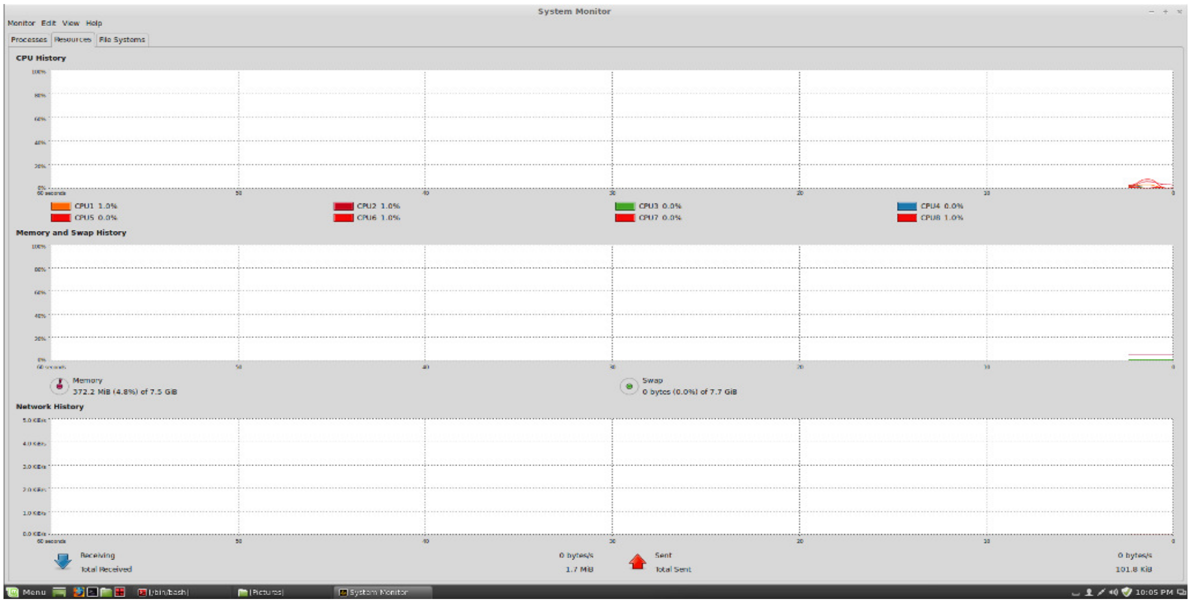
The sample output of sar command is

```
12:00:01 AM   CPU    %user   %nice   %system   %iowait   %steal   %idle
12:01:01 AM   all    73.28    0.00    1.25        0.00     0.00   25.47
12:02:01 AM   all     7.83    0.00    0.44        0.00     0.00   91.73
12:03:01 AM   all    61.65    0.00    0.70        0.00     0.00   37.66
12:04:01 AM   all    57.85    0.00    0.82        0.00     0.00   41.34
12:05:01 AM   all     4.25    0.00    0.41        0.00     0.00   95.34
12:06:01 AM   all     4.20    0.00    0.22        0.00     0.00   95.58
12:07:01 AM   all     5.05    0.00    0.33        0.00     0.00   94.63
12:08:01 AM   all     4.76    0.00    0.06        0.00     0.00   95.18
12:09:01 AM   all    37.57    0.00    0.37        0.00     0.00   62.05
12:10:01 AM   all    70.04    0.00    0.80        0.00     0.00   29.16
12:11:01 AM   all     5.03    0.00    0.12        0.00     0.00   94.84
```

2.3 KDE/GNOME SYSTEM MONITOR

The GNOME and KDE desktop environments both have graphical tools to assist you in monitoring and modifying the behavior of your system. System Monitor displays basic system information and allows you to monitor system processes, and resource or file system usage [14][15]

The below figure shows the sample out of GNOME/KDE System Monitor



🟧 CPU1 1.0% 🟥 CPU2 1.0% 🟩 CPU3 0.0% 🟦 CPU4 0.0%
🟥 CPU5 0.0% 🟥 CPU6 1.0% 🟥 CPU7 0.0% 🟥 CPU8 1.0%

Memory
372.2 MiB (4.8%) of 7.5 GiB

Swap
0 bytes (0.0%) of 7.7 GiB

Receiving 0 bytes/s
Total Received 1.7 MiB

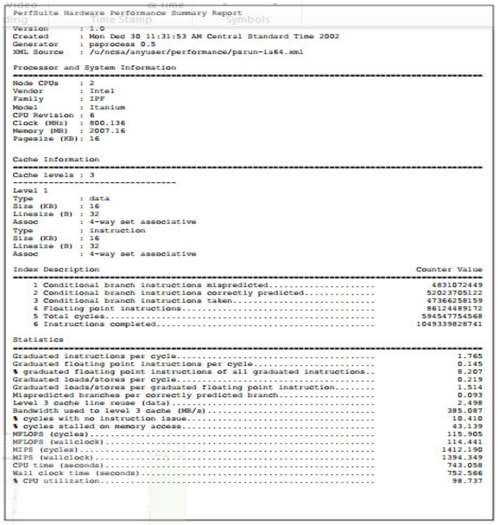
Sent 0 bytes/s
Total Sent 101.8 KiB

FIGURE 2: Sample of GNOME/KDE System Monitor.

2.4 PERFSUITE

PerfSuite is a collection of tools, utilities, and libraries for software optimization, benchmarking, and performance analysis where the primary design goals are ease of use, comprehensibility, interoperability, and simplicity. PerfSuite by itself gives you the information you need to focus your performance optimization efforts in the right direction without extensive source code changes to your application or complicated builds of the performance tools or libraries themselves. PerfSuite development is driven by people who do performance analysis on a daily basis with a wide variety of applications. Therefore the emphasis is on *practicality, robustness, simplicity, and utility* in real-world, everyday use on platforms ranging from single-processor workstations to terascale clusters and beyond [16][17].

Command-Line Utilities psinv,psprocess: psrun
Graphical Tools PSConfig: Web-Based Tools,AmdahlCalc



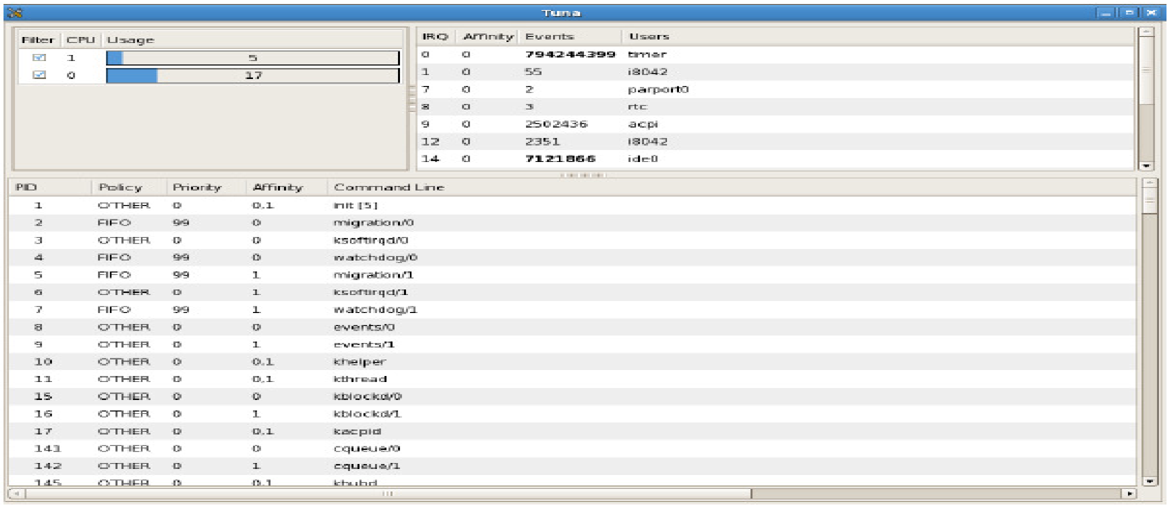
```
PerfSuite Hardware Performance Summary Report
Version       : 1.0
Created       : Mon Dec 30 11:31:53 AM Central Standard Time 2002
Generator     : psprocess 0.5
XML Source    : /u/ncsa/anyuser/performance/psrun-ia64.xml

Processor and System Information
=============================================================================================
Node CPUs     : 2
Vendor        : Intel
Family        : IPF
Model         : Itanium
CPU Revision  : 6
Clock (MHz)   : 800.136
Memory (MB)   : 2007.16
Pagesize (KB) : 16

Cache Information
=============================================================================================
Cache levels : 3
--------------------------------
Level 1
Type          : data
Size (KB)     : 16
Linesize (B)  : 32
Assoc         : 4-way set associative
Type          : instruction
Size (KB)     : 16
Linesize (B)  : 32
Assoc         : 4-way set associative

Index Description                                                              Counter Value
=============================================================================================
    1 Conditional branch instructions mispredicted.....................           4831072449
    2 Conditional branch instructions correctly predicted..............          52023705122
    3 Conditional branch instructions taken............................          47366258159
    4 Floating point instructions......................................          86124489172
    5 Total cycles.....................................................         594547754568
    6 Instructions completed...........................................        1049339828741

Statistics
=============================================================================================
Graduated instructions per cycle.......................................                1.765
Graduated floating point instructions per cycle........................                0.145
% graduated floating point instructions of all graduated instructions..                8.207
Graduated loads/stores per cycle.......................................                0.219
Graduated loads/stores per graduated floating point instruction........                1.514
Mispredicted branches per correctly predicted branch...................                0.093
Level 3 cache line reuse (data)........................................                2.498
Bandwidth used to level 3 cache (MB/s).................................              385.087
% cycles with no instruction issue.....................................               10.410
% cycles stalled on memory access......................................               43.139
MFLOPS (cycles)........................................................              115.905
MFLOPS (wallclock).....................................................              114.441
MIPS (cycles)..........................................................             1412.190
MIPS (wallclock).......................................................             1394.349
CPU time (seconds).....................................................              743.058
Wall clock time (seconds)..............................................              752.566
% CPU utilization......................................................               98.737
```

FIGURE 3: PerfSuite Tool.

2.5 RED HAT TUNA TOOL

Tuna was developed for tuning the MRG Realtime component of Red Hat Enterprise MRG [18], but can also be used to tune standard Red Hat Enterprise Linux systems. Tuna is a tool that can be used to adjust scheduler policy, Real-time priority and CPU affinity. It also allows the user to see the results of these changes. Threads and IRQ handlers are able to be tuned. It is also possible to isolate CPU cores and sockets, moving all threads away from them so that a new, more important set of threads can run exclusively. Tuna provides a graphical user interface (GUI), displays the CPU topology on one screen, which helps identify problems. It also allows changes to make for running threads, and see the results of those changes immediately .Most Tuna operations can be performed on either the command line, or in the GUI.

Below figure shows Design of Tuna tool



Filter	CPU	Usage
☑	1	5
☑	0	17

IRQ	Affinity	Events	Users
0	0	**794244399**	timer
1	0	55	i8042
7	0	2	parport0
8	0	3	rtc
9	0	2502436	acpi
12	0	2351	i8042
14	0	**7121866**	ide0

PID	Policy	Priority	Affinity	Command Line
1	OTHER	0	0.1	init [5]
2	FIFO	99	0	migration/0
3	OTHER	0	0	ksoftirqd/0
4	FIFO	99	0	watchdog/0
5	FIFO	99	1	migration/1
6	OTHER	0	1	ksoftirqd/1
7	FIFO	99	1	watchdog/1
8	OTHER	0	0	events/0
9	OTHER	0	1	events/1
10	OTHER	0	0.1	khelper
11	OTHER	0	0.1	kthread
15	OTHER	0	0	kblockd/0
16	OTHER	0	1	kblockd/1
17	OTHER	0	0.1	kacpid
141	OTHER	0	0	cqueue/0
142	OTHER	0	1	cqueue/1
145	OTHER	0	0.1	khubd

FIGURE 4: Design of Tuna Tool.

Top, Htop and Sysstat tools are Character User Interface (CUI) which is not user friendly and needs remember the commands, it doesn’t provide an opportunity to bind process to a core. The GNOME/KDE SYSTEM Monitor is GUI tool to help to monitor and change priority of a process but doesn’t have an option of utilizing the benefit of CPU affinity. The PerfSuite tool provides more information to user to understand the performance bottlenecks, but it requires lot of knowledge to use libraries in programming to get the relative information. The Red hat Linux TUNA tool is one of the best tools which provide the user a number of advanced options to obtain maximum performance improvement through graphical user interface. The tool has three drawbacks 1) the user has an option to the remove all the process from one processor and make it idle and assigning other process to other cores which results in overheating and overloading of one processor which results in improper load balancing among the CPUs 2) It is available through the MRG Real-time channels on the Red Hat Network (RHN),and requires registration to install tuna tool .3) It requires additional packages python-linux-procfs ,python-schedutils, python-ethtool for command line and pygtk2,pygtk2-libglade for GUI operations.

In order to avoid the existing problems and to provide a user friendly environment, to achieve maximum performance a new PI-Tool is designed by using the techniques of CPU affinity, and process priority.

3. Advantages of CPU Affinity

3.1. Optimizing Cache Performance

Multiprocessing computers go through a lot of trouble to keep the processor caches valid. Data can be kept in only one processor's cache at a time. Otherwise, the processor's cache may grow out of sync. Consequently, whenever a processor adds a line of data to its local cache, all the other processors in the system also must validate that data. This validation is costly and unpleasant [19][20]

.

But the real problem comes into play when processes bounce between processors: they constantly cause cache invalidations, and the data they want is never in the cache when they need it. Thus, cache miss rates grow very large. CPU affinity protects against this and improves cache performance.

3.2. Multithreading Performance

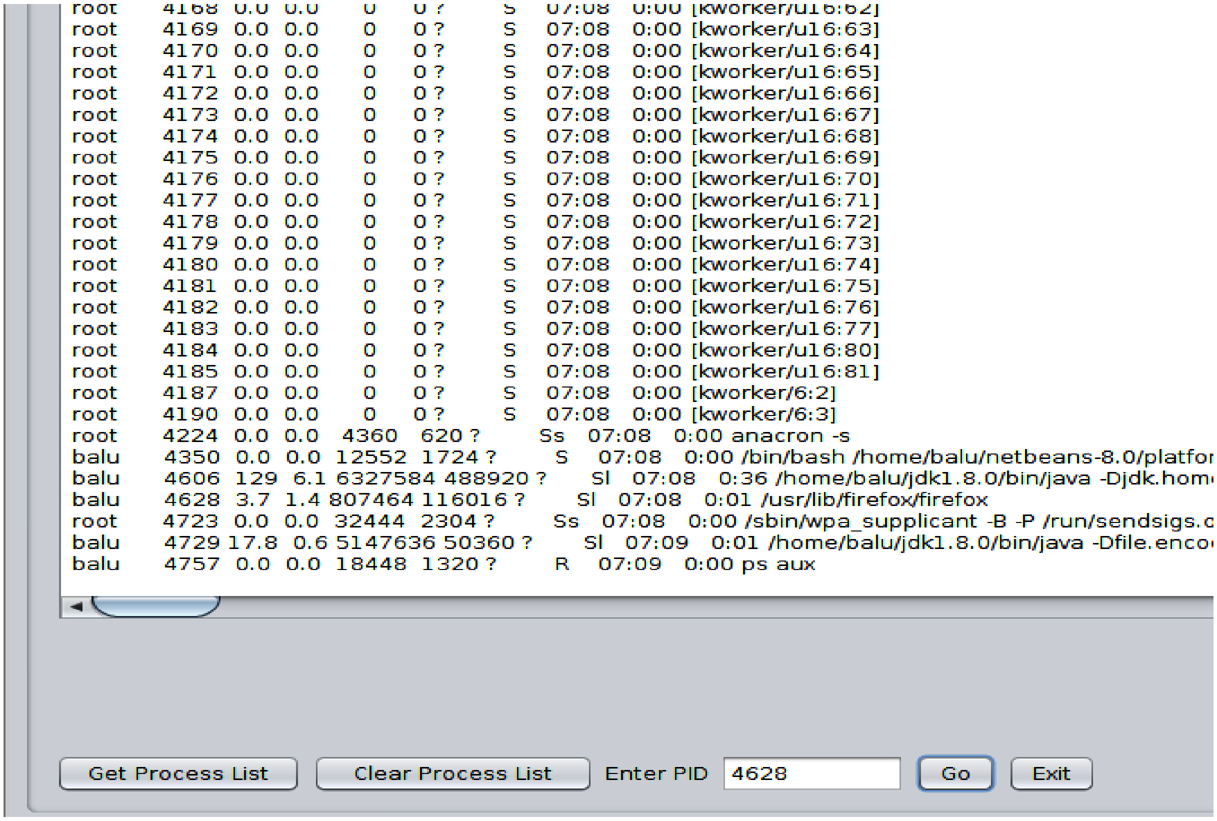
If multiple threads are accessing the same data, it might make sense to bind them all to the same processor. Doing so guarantees that the threads do not contend over data and cause cache misses. This does diminish the performance gained from multithreading on SMP. Thus binding a thread to a processor improves the multithreading performance.

3.3 Time Quantum

The third and final benefit is found in real-time or otherwise time-sensitive applications. In this approach, all the system processes are bound to a subset of the processors on the system. The specialized application then is bound to the remaining processors. Commonly, in a dual-processor system, the specialized application is bound to one processor, and all other processes are bound to the other. This ensures that the specialized application receives the full attention of the processor by receiving more time quantum.

4. DESCRIPTION OF PI-TOOL

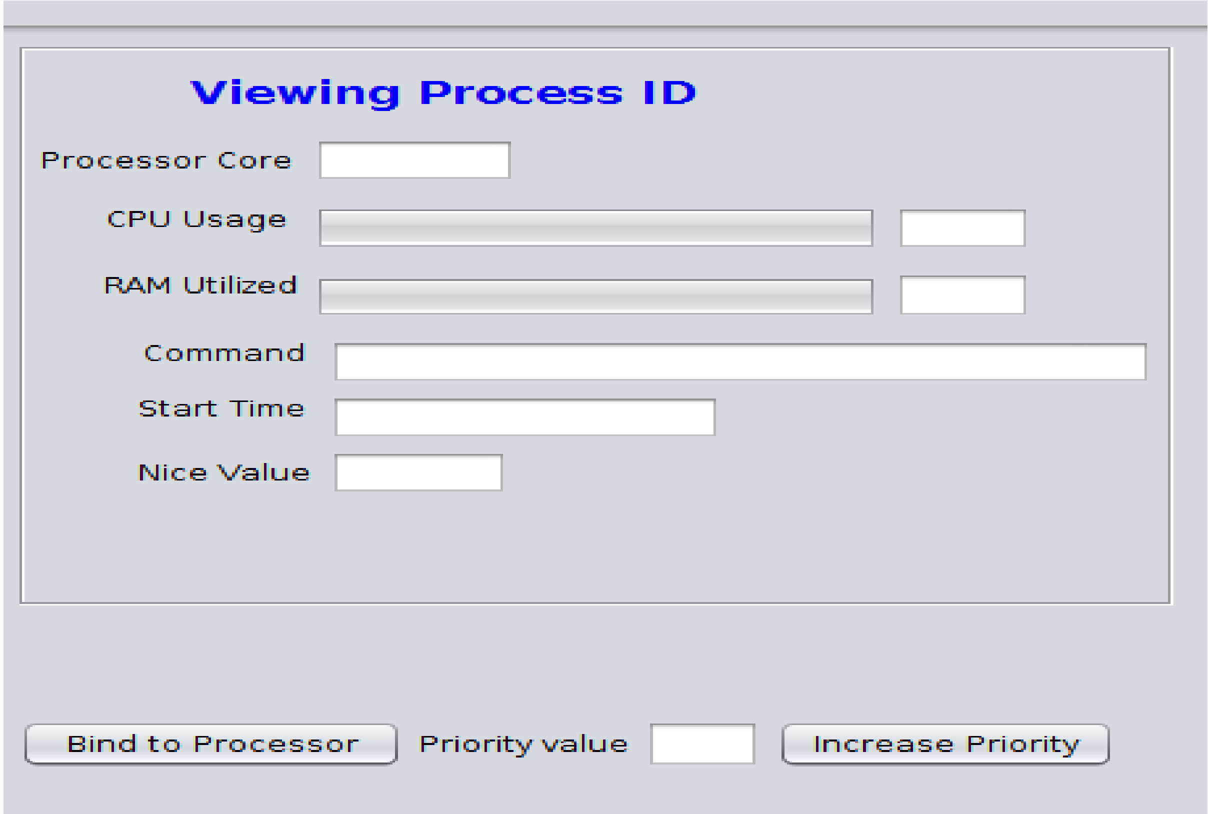
The PI-Tools divided into two sections; first section shows the information regarding the process executing in the system ,from which user can select one process by entering the PID of the process in the below text box.



```
root     4168  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:62]
root     4169  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:63]
root     4170  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:64]
root     4171  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:65]
root     4172  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:66]
root     4173  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:67]
root     4174  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:68]
root     4175  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:69]
root     4176  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:70]
root     4177  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:71]
root     4178  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:72]
root     4179  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:73]
root     4180  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:74]
root     4181  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:75]
root     4182  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:76]
root     4183  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:77]
root     4184  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:80]
root     4185  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/u16:81]
root     4187  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/6:2]
root     4190  0.0  0.0     0    0 ?       S    07:08   0:00 [kworker/6:3]
root     4224  0.0  0.0  4360  620 ?       Ss   07:08   0:00 anacron -s
balu     4350  0.0  0.0 12552  1724 ?      S    07:08   0:00 /bin/bash /home/balu/netbeans-8.0/platfor
balu     4606  129  6.1 6327584 488920 ?   Sl   07:08   0:36 /home/balu/jdk1.8.0/bin/java -Djdk.hom
balu     4628  3.7  1.4 807464 116016 ?    Sl   07:08   0:01 /usr/lib/firefox/firefox
root     4723  0.0  0.0 32444  2304 ?      Ss   07:08   0:00 /sbin/wpa_supplicant -B -P /run/sendsigs.c
balu     4729 17.8  0.6 5147636 50360 ?    Sl   07:09   0:01 /home/balu/jdk1.8.0/bin/java -Dfile.enco
balu     4757  0.0  0.0 18448  1320 ?      R    07:09   0:00 ps aux
```

Enter PID 4628

The second section shows the selected process information, CPU core on which process is running, processor, ram utilization, command name, process start time and priority.   Here user can find two buttons for obtaining CPU affinity and to increase process Priority.



The PI-Tool has divided into two procedures; 1.newprocess 2.updatevalues. The new process is responsible for reading and displaying all the process running in the system. When user enters a process ID and click button, the method is responsible for reading the processor, cpu,memory utilization values, starttime,command, nice values of selected process .

The updatevalues procedure is designed using a thread which continuously updates the information of selected process,

5. EXPERIMENTAL TEST BED

5.1 Hardware

Name	Details
Processor	Intel Core i7 4770K [23]
No of Cores	4
No of Threads	8
Clock Speed	3.5 GHz
Max Turbo Frequency	3.9 GHz
Intel Smart Cache	8 MB
RAM ,Speed	8 GB(1600 MHz)
Instruction Set	64 Bit

5.2 Software

- Linux Kernel 3.14.4-200 [22]
- Fedora OS -20 [21]
- GCC Compiler
- Java 1.7

- Net Beans IDE 7.4 [24]

6. PROCESS TUNING

- To select a process for tuning, click on get process list button, among the displayed list find the process ID and enter the process ID into text box then click go button, PI-Tool will display the core number on which program is executing, cpu, ram usage, command name, start time, and nice values of the of entered process ID

- If the process is exhibiting the Ping-Pong effect, then click bind to processor button, PI-Tool will automatically select the less utilizing core and binds to it.

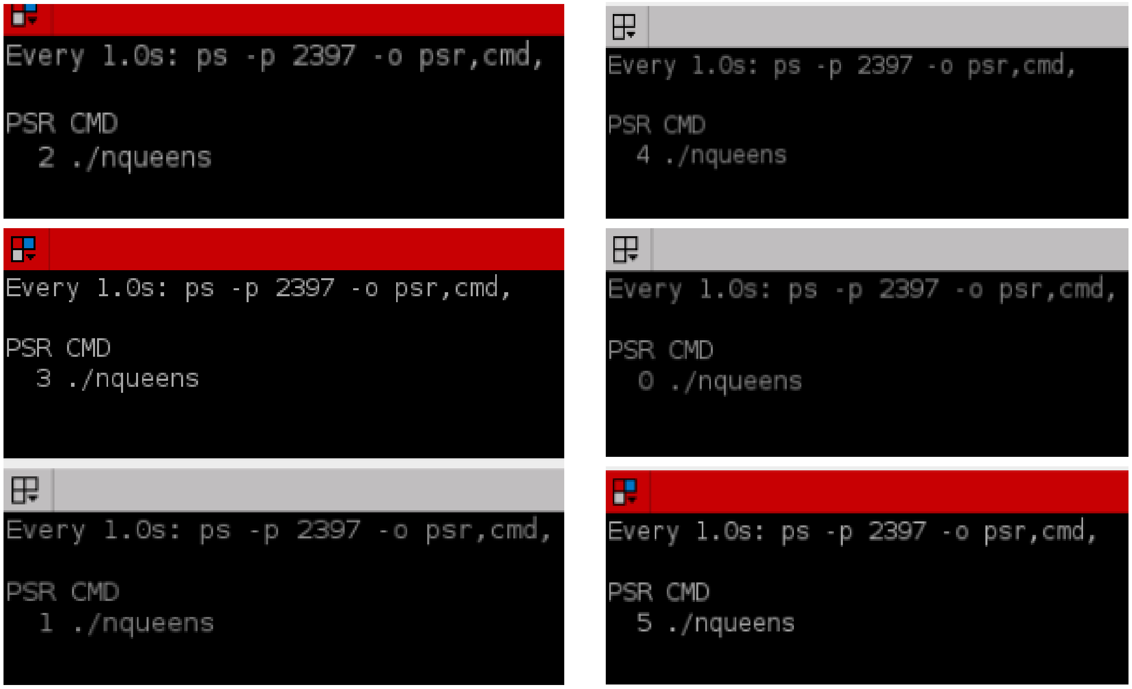
- To increase the priority of a process, enter new priority value into the textbox then click increase priority button, then PI-Tool will ask for password of root user, after entering the password, PI-Tool will increase the priority.

7. COMPARATIVE STUDY THE EFFECTS OF CPU AFFINITY

In order to study the effect of CPU affinity, with reference to CPU performance, the NQueens program has been executed through Linux shell. The NQueens program is executed in normal conditions and then the Tuna and PI tools are used with FIFO and Round Robin (RR) scheduling to study the CPU Performance.

7.1 Normal Execution

Linux provides natural CPU affinity, but when NQueens program is executed in normal conditions, the process bounces from one processor to another processor very frequently, during execution in a random way. This is shown in the figure 5.



```
Every 1.0s: ps -p 2397 -o psr,cmd,

PSR CMD
  2 ./nqueens
```

```
Every 1.0s: ps -p 2397 -o psr,cmd,

PSR CMD
  4 ./nqueens
```

```
Every 1.0s: ps -p 2397 -o psr,cmd,

PSR CMD
  3 ./nqueens
```

```
Every 1.0s: ps -p 2397 -o psr,cmd,

PSR CMD
  0 ./nqueens
```

```
Every 1.0s: ps -p 2397 -o psr,cmd,

PSR CMD
  1 ./nqueens
```

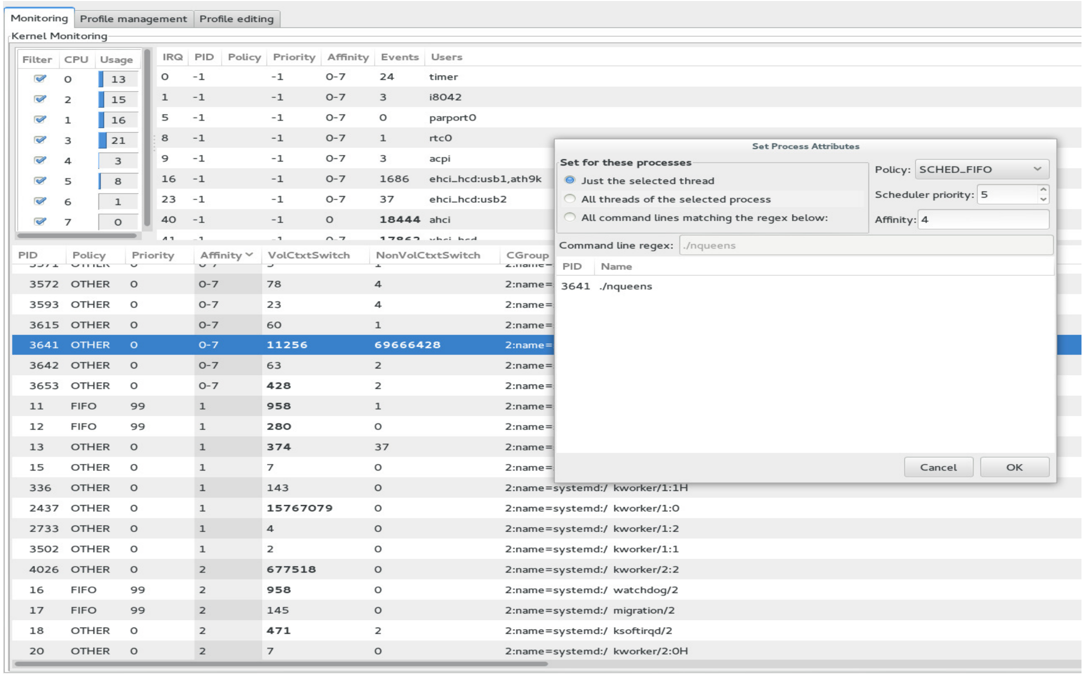
```
Every 1.0s: ps -p 2397 -o psr,cmd,

PSR CMD
  5 ./nqueens
```

FIGURE 5: Process Bouncing From One Processor To Another.

Migrating process from one processor to another will affect the performance by increasing the execution time. This is shown in the table 7.4.

7.2 Tuna-Tool Execution

When NQueens program is under execution through Linux shell, the Tuna tool can be activated under Super User Mode by executing the **sudo Tuna** command at shell prompt. Select the process (NQueens PID 3641), click the mouse right button, then a small dialog box will appear on the screen. The attributes of scheduling, priority, and CPU affinity can be modified in the dialogue box, as shown in the figure 6 (a). After modifying the attributes, the process has been bind to the selected processor and execution will be completed with given priority and scheduling policy as shown in the figure 6 (b). The results in the table 7.4 show the decrease in the execution time than the normal execution.



Filter	CPU	Usage
☑	0	13
☑	2	15
☑	1	16
☑	3	21
☑	4	3
☑	5	8
☑	6	1
☑	7	0

IRQ	PID	Policy	Priority	Affinity	Events	Users
0	-1		-1	0-7	24	timer
1	-1		-1	0-7	3	i8042
5	-1		-1	0-7	0	parport0
8	-1		-1	0-7	1	rtc0
9	-1		-1	0-7	3	acpi
16	-1		-1	0-7	1686	ehci_hcd:usb1,ath9k
23	-1		-1	0-7	37	ehci_hcd:usb2
40	-1		-1	0	18444	ahci

PID	Policy	Priority	Affinity	VolCtxtSwitch	NonVolCtxtSwitch	CGroup	
3572	OTHER	0	0-7	78	4	2:name=	
3593	OTHER	0	0-7	23	4	2:name=	
3615	OTHER	0	0-7	60	1	2:name=	
3641	OTHER	0	0-7	11256	69666428	2:name=	
3642	OTHER	0	0-7	63	2	2:name=	
3653	OTHER	0	0-7	428	2	2:name=	
11	FIFO	99	1	958	1	2:name=	
12	FIFO	99	1	280	0	2:name=	
13	OTHER	0	1	374	37	2:name=	
15	OTHER	0	1	7	0	2:name=	
336	OTHER	0	1	143	0	2:name=systemd:/	kworker/1:1H
2437	OTHER	0	1	15767079	0	2:name=systemd:/	kworker/1:0
2733	OTHER	0	1	4	0	2:name=systemd:/	kworker/1:2
3502	OTHER	0	1	2	0	2:name=systemd:/	kworker/1:1
4026	OTHER	0	2	677518	0	2:name=systemd:/	kworker/2:2
16	FIFO	99	2	958	0	2:name=systemd:/	watchdog/2
17	FIFO	99	2	145	0	2:name=systemd:/	migration/2
18	OTHER	0	2	471	2	2:name=systemd:/	ksoftirqd/2
20	OTHER	0	2	7	0	2:name=systemd:/	kworker/2:0H

Set Process Attributes

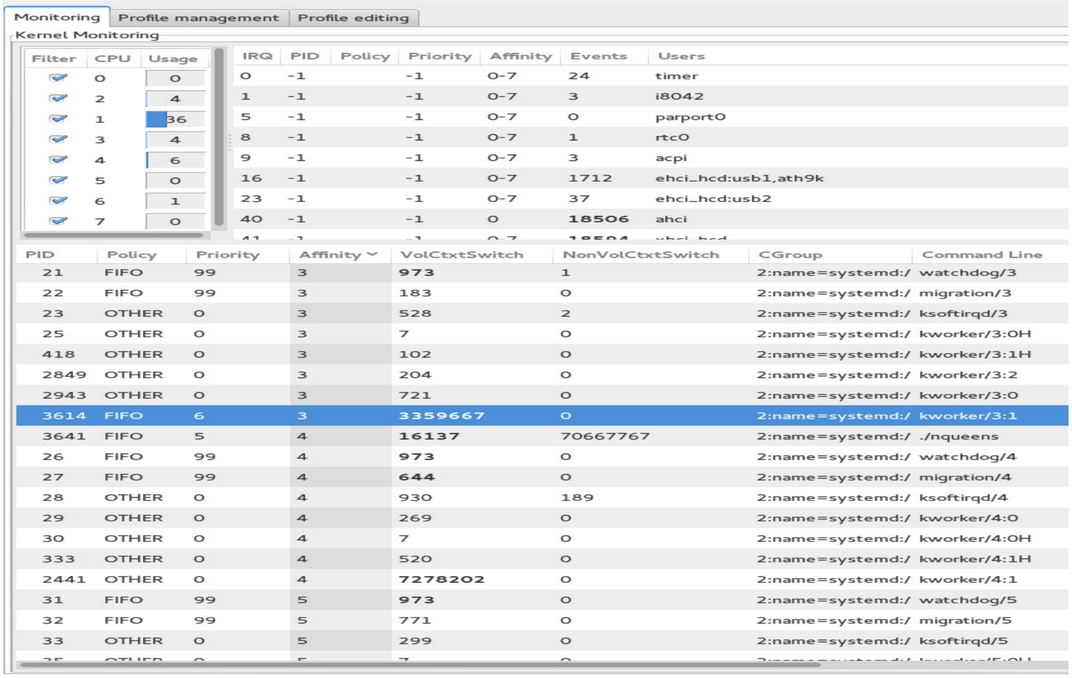
Set for these processes
● Just the selected thread
○ All threads of the selected process
○ All command lines matching the regex below:

Policy: SCHED_FIFO
Scheduler priority: 5
Affinity: 4

Command line regex: ./nqueens

PID	Name
3641	./nqueens

FIGURE 6 (a): Shows the selected process 3641 and dialogue box to change.



Filter	CPU	Usage
✓	0	0
✓	2	4
✓	1	36
✓	3	4
✓	4	6
✓	5	0
✓	6	1
✓	7	0

IRQ	PID	Policy	Priority	Affinity	Events	Users
0	-1		-1	0-7	24	timer
1	-1		-1	0-7	3	i8042
5	-1		-1	0-7	0	parport0
8	-1		-1	0-7	1	rtc0
9	-1		-1	0-7	3	acpi
16	-1		-1	0-7	1712	ehci_hcd:usb1,ath9k
23	-1		-1	0-7	37	ehci_hcd:usb2
40	-1		-1	0	18506	ahci

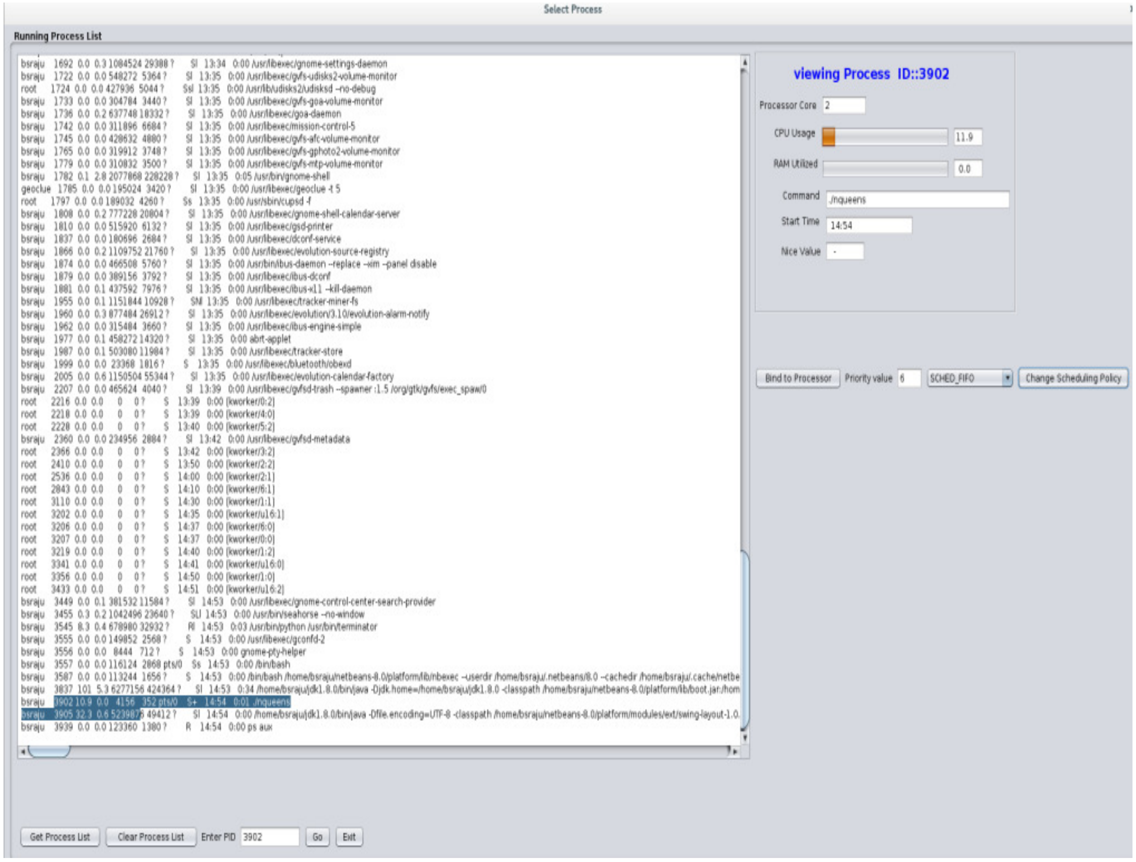
PID	Policy	Priority	Affinity	VolCtxtSwitch	NonVolCtxtSwitch	CGroup	Command Line
21	FIFO	99	3	973	1	2:name=systemd:/	watchdog/3
22	FIFO	99	3	183	0	2:name=systemd:/	migration/3
23	OTHER	0	3	528	2	2:name=systemd:/	ksoftirqd/3
25	OTHER	0	3	7	0	2:name=systemd:/	kworker/3:0H
418	OTHER	0	3	102	0	2:name=systemd:/	kworker/3:1H
2849	OTHER	0	3	204	0	2:name=systemd:/	kworker/3:2
2943	OTHER	0	3	721	0	2:name=systemd:/	kworker/3:0
3614	FIFO	6	3	3359667	0	2:name=systemd:/	kworker/3:1
3641	FIFO	5	4	16137	70667767	2:name=systemd:/	./nqueens
26	FIFO	99	4	973	0	2:name=systemd:/	watchdog/4
27	FIFO	99	4	644	0	2:name=systemd:/	migration/4
28	OTHER	0	4	930	189	2:name=systemd:/	ksoftirqd/4
29	OTHER	0	4	269	0	2:name=systemd:/	kworker/4:0
30	OTHER	0	4	7	0	2:name=systemd:/	kworker/4:0H
333	OTHER	0	4	520	0	2:name=systemd:/	kworker/4:1H
2441	OTHER	0	4	7278202	0	2:name=systemd:/	kworker/4:1
31	FIFO	99	5	973	0	2:name=systemd:/	watchdog/5
32	FIFO	99	5	771	0	2:name=systemd:/	migration/5
33	OTHER	0	5	299	0	2:name=systemd:/	ksoftirqd/5

FIGURE 6 (b): Show the changed attributes for the selected process 3641.

7.3 PI-Tool Execution

There is a delay in opening new dialog box, for setting attributes for the process, and there is also delay in applying the changes. Tuna tool is not available for all Linux variants.

When PI –Tool is opened, it display all the executing process, from which user can select the NQueens process. The attributes of selected process (NQueens) can be changed in the same window. PI-Tool is developed with light weight Multi-threading java,which takes less execution time. From the table 7.4 we conclude that PI-Tool reduces the execution time of selected process with reference to TUNA and normal execution. PI-Tool can be used in all Linux variants without any modifications, since it has been developed in platform independent language i.e. java.



```
bsraju   1692  0.0  0.3 1084524 29388 ?     Sl   13:34   0:00 /usr/libexec/gnome-settings-daemon
bsraju   1722  0.0  0.0 548272  5364 ?      Sl   13:35   0:00 /usr/libexec/gvfs-udisks2-volume-monitor
root     1724  0.0  0.0 427936  5044 ?      Ssl  13:35   0:00 /usr/lib/udisks2/udisksd --no-debug
bsraju   1733  0.0  0.0 304784  3440 ?      Sl   13:35   0:00 /usr/libexec/gvfs-goa-volume-monitor
bsraju   1736  0.0  0.2 637748 18332 ?      Sl   13:35   0:00 /usr/libexec/goa-daemon
bsraju   1742  0.0  0.0 311896  6684 ?      Sl   13:35   0:00 /usr/libexec/mission-control-5
bsraju   1745  0.0  0.0 428632  4880 ?      Sl   13:35   0:00 /usr/libexec/gvfs-afc-volume-monitor
bsraju   1765  0.0  0.0 319912  3748 ?      Sl   13:35   0:00 /usr/libexec/gvfs-gphoto2-volume-monitor
bsraju   1779  0.0  0.0 310832  3500 ?      Sl   13:35   0:00 /usr/libexec/gvfs-mtp-volume-monitor
bsraju   1782  0.1  2.8 2077868 228228 ?    Sl   13:35   0:05 /usr/bin/gnome-shell
geoclue  1785  0.0  0.0 195024  3420 ?      Sl   13:35   0:00 /usr/libexec/geoclue -t 5
root     1797  0.0  0.0 189032  4260 ?      Ss   13:35   0:00 /usr/sbin/cupsd -f
bsraju   1808  0.0  0.2 777228 20804 ?      Sl   13:35   0:00 /usr/libexec/gnome-shell-calendar-server
bsraju   1810  0.0  0.0 515920  6132 ?      Sl   13:35   0:00 /usr/libexec/gsd-printer
bsraju   1837  0.0  0.0 180696  2684 ?      Sl   13:35   0:00 /usr/libexec/dconf-service
bsraju   1866  0.0  0.2 1109752 21760 ?     Sl   13:35   0:00 /usr/libexec/evolution-source-registry
bsraju   1874  0.0  0.0 466508  5760 ?      Sl   13:35   0:00 /usr/bin/ibus-daemon --replace --xim --panel disable
bsraju   1879  0.0  0.0 389156  3792 ?      Sl   13:35   0:00 /usr/libexec/ibus-dconf
bsraju   1881  0.0  0.1 437592  7976 ?      Sl   13:35   0:00 /usr/libexec/ibus-x11 --kill-daemon
bsraju   1955  0.0  0.1 1151844 10928 ?     SNl  13:35   0:00 /usr/libexec/tracker-miner-fs
bsraju   1960  0.0  0.3 877484 26912 ?      Sl   13:35   0:00 /usr/libexec/evolution/3.10/evolution-alarm-notify
bsraju   1962  0.0  0.0 315484  3660 ?      Sl   13:35   0:00 /usr/libexec/ibus-engine-simple
bsraju   1977  0.0  0.1 458272 14320 ?      Sl   13:35   0:00 abrt-applet
bsraju   1987  0.0  0.1 503080 11984 ?      Sl   13:35   0:00 /usr/libexec/tracker-store
bsraju   1999  0.0  0.0  23368  1816 ?      S    13:35   0:00 /usr/libexec/bluetooth/obexd
bsraju   2005  0.0  0.6 1150504 55344 ?     Sl   13:35   0:00 /usr/libexec/evolution-calendar-factory
bsraju   2207  0.0  0.0 465624  4040 ?      Sl   13:39   0:00 /usr/libexec/gvfsd-trash --spawner :1.5 /org/gtk/gvfs/exec_spaw/0
root     2216  0.0  0.0      0     0 ?      S    13:39   0:00 [kworker/0:2]
root     2218  0.0  0.0      0     0 ?      S    13:39   0:00 [kworker/4:0]
root     2228  0.0  0.0      0     0 ?      S    13:40   0:00 [kworker/5:2]
bsraju   2360  0.0  0.0 234956  2884 ?      Sl   13:42   0:00 /usr/libexec/gvfsd-metadata
root     2366  0.0  0.0      0     0 ?      S    13:42   0:00 [kworker/3:2]
root     2410  0.0  0.0      0     0 ?      S    13:50   0:00 [kworker/2:2]
root     2536  0.0  0.0      0     0 ?      S    14:00   0:00 [kworker/2:1]
root     2843  0.0  0.0      0     0 ?      S    14:10   0:00 [kworker/6:1]
root     3110  0.0  0.0      0     0 ?      S    14:30   0:00 [kworker/1:1]
root     3202  0.0  0.0      0     0 ?      S    14:35   0:00 [kworker/u16:1]
root     3206  0.0  0.0      0     0 ?      S    14:37   0:00 [kworker/6:0]
root     3207  0.0  0.0      0     0 ?      S    14:37   0:00 [kworker/0:0]
root     3219  0.0  0.0      0     0 ?      S    14:40   0:00 [kworker/1:2]
root     3341  0.0  0.0      0     0 ?      S    14:41   0:00 [kworker/u16:0]
root     3356  0.0  0.0      0     0 ?      S    14:50   0:00 [kworker/1:0]
root     3433  0.0  0.0      0     0 ?      S    14:51   0:00 [kworker/u16:2]
bsraju   3449  0.0  0.1 381532 11584 ?      Sl   14:53   0:00 /usr/libexec/gnome-control-center-search-provider
bsraju   3455  0.3  0.2 1042496 23640 ?     SLl  14:53   0:00 /usr/bin/seahorse --no-window
bsraju   3545  8.3  0.4 678980 32932 ?      Rl   14:53   0:03 /usr/bin/python /usr/bin/terminator
bsraju   3555  0.0  0.0 149852  2568 ?      S    14:53   0:00 /usr/libexec/gconfd-2
bsraju   3556  0.0  0.0   8444   712 ?      S    14:53   0:00 gnome-pty-helper
bsraju   3557  0.0  0.0 116124  2868 pts/0  Ss   14:53   0:00 /bin/bash
bsraju   3587  0.0  0.0 113244  1656 ?      S    14:53   0:00 /bin/bash /home/bsraju/netbeans-8.0/platform/lib/nbexec --userdir /home/bsraju/.netbeans/8.0 --cachedir /home/bsraju/.cache/netbe
bsraju   3837  101  5.3 6277156 424364 ?    Sl   14:53   0:34 /home/bsraju/jdk1.8.0/bin/java -Djdk.home=/home/bsraju/jdk1.8.0 -classpath /home/bsraju/netbeans-8.0/platform/lib/boot.jar:/hom
bsraju   3902 10.9  0.0   4156   352 pts/0  S+   14:54   0:01 ./nqueens
bsraju   3905 32.3  0.6 5239876 49412 ?     Sl   14:54   0:00 /home/bsraju/jdk1.8.0/bin/java -Dfile.encoding=UTF-8 -classpath /home/bsraju/netbeans-8.0/platform/modules/ext/swing-layout-1.0.
bsraju   3939  0.0  0.0 123360  1380 ?      R    14:54   0:00 ps aux
```

viewing Process ID::3902

Processor Core: 2

CPU Usage: 11.9

RAM Utilized: 0.0

Command: ./nqueens

Start Time: 14:54

Nice Value: -

Priority value: 6

SCHED_FIFO

Enter PID: 3902

FIGURE 7: Show the PI-Tool Interface for changing the attributes of a process.

The table 7.4 presents the results of NQueen program, normal execution along with TUNA and PI-Tool with FIFO and Round Robin scheduling.

Queens	Solutions	Normal	Priority	TUNA		PI-tool	
				FIFO	RR	FIFO	RR
N=13	73712	00:00:57	6	00:00:57	00:00:57	00:00:55	00:00:55
N=14	3,65,596	00:05:45	6	00:05:36	00:05:31	00:05:33	00:05:29
N=15	22,79,184	00:40:35	6	00:40:30	00:39:46	00:39:39	00:39:15
N=16	1,47,72,512	05:02:57	6	04:50:56	04:38:15	04:46:00	04:37:30

TABLE 7.4: shows the results of NQueens program normal execution, TUNA tool and PI-Tool.

8. ADVANTGES AND DISADVANTAGES

Advantages

- PI-Tool is more user friendly , and can be operated by any user with little knowledge
- Helps to avoid Ping-Pong effects of process that results in performance improvements.
- Useful to increase priority of a process without knowing any shell commands
- No need of any additional subscription or downloading or installing any software's

- It can be used on any Linux system ,having JRE

Disadvantages

- With increase of priority to a selected process the other process may get delayed.

9. CONCLUSION and FUTURE WORK

The above results show the execution with the PI-tool gives marginally better results than TUNA Tool and normal execution. The marginally better results are achieved when small and medium size programs but may increase with the program size and execution time. This marginal advantage is mainly depending on the core selection by the PI-Tool. This PI tool is designed to provide the user a way to overcome the drawback of CPU Affinity, and to increase priority to applications to get more performance. So further research work may require by making the process to core mapping automatically by the Operating system.

10. REFERENCES

[1] "*Optimizing software for multi-core processors"* white paper Intel Corporation

[2] Donald Porter, Silas Boyd-Wickizer, Jon Howell, Reuben Olinsky, Galen Hunt,"*Rethinking the Library OS from the Top Down"* Microsoft Research,ASPLOS (March-2011)

[3] J Chen, LK John," Efficient *program scheduling for heterogeneous multi-core processors"* Design Automation Conference, 2009 DAC '09m 46 th ACM/IEEE

[4] Max Domeik. "*Optimizing Software for Multi-core Processors"*, 2007, Intel Corporation.

[5] Nandan Tripathi and Amrit Singh, "*Analyzing multithreaded applications—identifying performance bottlenecks on multicore systems"*, May 3, 2011 in San Jose, CA.

[6] Erik Hangersten, CTO Acumem, "Finding & Fixing Multicore Performance Bottlenecks in HPC Applications"

[7] Silas Boyd-Wickizer, Robert Morris, M.Frans Kaashoek, "Reinventing Scheduling for Multicore Systems". *Proceedings of the 12th conference on hot topics in operating systems, Berkeley, CA, USA, USENIX Association, (2009)*

[8] Jiang Lin; Qingda Lu; Xiaoning Ding; Zhao Zhang; Xiaodong Zhang; Sadayappan, P., "Enabling Software management for multicore caches with a lightweight hardware support," *High* Performance Computing Networking, Storage and Analysis, Proceedings of the Conference on, vol., no., pp.1, 12, 14-20 Nov.2009doi: 10.1145/1654059.1654074

[9] Robert Love (Jul 01, 2003) "CPU Affinity" Linux journal [On-line], Available: http://www.linuxjournal.com/article/6799 [May.13, 2014].

[10] Ravi Saive (March 4, 2013)" 12 TOP Command Examples in Linux" Tecmint.com. Available: http://www.tecmint.com/12-top-command-examples-in-linux/ [May, 13, 2014]

[11] RAMESH NATARAJAN (SEPTEMBER 14, 2011) "15 Practical Linux HTOP Examples" TheGeekStuff, Available: http://www.thegeekstuff.com/2011/09/linux-htop-examples/ [May 13, 2014]

[12] RAMESH NATARAJAN (*MARCH* 29, 2011) "10 Useful Sar (Sysstat) Examples for UNIX / Linux Performance Monitoring" TheGeekStuff, Available:http://www.thegeekstuff.com/2011/03/sar Examples/ [May 13, 2014]

[13] BLFS Development Team (Feb 20, 2014)" Beyond Linux from Scratch" chapter 12 Sysstat 10.2.1. Availability: http://www.linuxfromscratch.org/blfs/view/svn/general/sysstat.html [May 13, 2014]

[14] NIX CRAFT (JUNE 27, 2009 · UPDATED JANUARY 1,2014)" 20 Linux System Monitoring Tools Every SysAdmin Should Know" NIX CRAFT Availability:http://www.cyberciti.biz/tips/top-linux-monitoring-tools.html [may 13,2014].

[15] Don Domingo,Laura Bailey(2011) "Optimizing subsystem throughput in Red Hat Enterprise Linux 6" Performance Tunning Guide Availability: https://access.redhat.com/site/documentation/en-US/Red_Hat_Enterprise_Linux/6/html-single/Performance_Tuning_Guide/index.html [May 13,2014].

[16] Rick Kufrin Availability:http://perfsuite.ncsa.illinois.edu/[May 13, 2014].

[17] Rkufrin (Feb 20,2013) "PerfSuite" Sourceforge Availability: http://sourceforge.net/projects/perfsuite/[May 13, 2014].

[18] Lana Brindley, Alison Young(2011) "Red Hat Enterprise MRG 1.3" Tuna User Guide Avaialability:https://access.redhat.com/site/documentation/en-US/Red_Hat_Enterprise_MRG/1.3/html-single/Tuna_User_Guide/index.html [May 13, 2014].

[19] http://www.linuxjournal.com/article/6799.

[20] https://www.kernel.org/pub/linux/kernel/people/rml/cpu-affinity/.

[21] Availability: http://ark.intel.com/products/75123 [May 13, 2014].

[22] Availability: http:// www.kernel.org [May 13, 2014].

[23] Availability: http://fedoraproject.org [Jun 04, 2014].

[24] Availability: http://www.netbeans.org [May 13, 2014].

Comprehensive Social Media Security Analysis & XKeyscore Espionage Technology

Adam Ali.Zare Hudaib *adamhudaib@gmail.com*
Licensed Penetration Tester EC-Council
Certified Ethical Hacker EC-Council
Certified Security Analyst EC-Council
Wireshark Certified Network Analyst (Wireshark University)
CEH , ECSA , LPT , WCNA
Sweden

Abstract

Social networks can offer many services to the users for sharing activities events and their ideas. Many attacks can happened to the social networking websites due to trust that have been given by the users. Cyber threats are discussed in this paper. We study the types of cyber threats, classify them and give some suggestions to protect social networking websites of variety of attacks. Moreover, we gave some antithreats strategies with future trends.

Keywords: XKeyscore, Social Media Security, Privacy, TFC, Cyber Threats.

1. INTRODUCTION

In recent years, global online social network (OSN) usage has increased sharply as these networks become interwoven into people’s everyday lives as places to meet and communicate with others. Online social networks, such as Facebook, Google+, LinkedIn, Sina Weibo, Twitter, Tumblr, and VKontakte (VK), have hundreds of millions of daily active users. Facebook, for example, has more than 1.15 billion monthly active users, 819 million of which are active mobile Facebook users as of June 2013. Facebook users have been shown to accept friendship requests from people whom they do not know but who simply have several friends in common. By accepting these friend requests, users unknowingly disclose their private information to total strangers. This information could be used maliciously, harming users both in the virtual and in the real world. These risks escalate when the users are children or teenagers who are by nature more exposed and vulnerable than adults. As the use of online social networks becomes progressively more embedded into the everyday lives of users, personal information becomes easily exposed and abused. Information harvesting, by both the online social network operator itself and by third-party commercial companies, has recently been identified as a significant security concern for OSN users. Companies can use the harvested personal information for a variety of purposes, all of which can jeopardize a user’s privacy. For example, companies can use collected private information to tailor online ads according to a user’s profile, to gain profitable insights about their customers, or even to share the user’s private and personal data with the government. This information may include general data, such as age, gender, and income; however, in some cases more personal information can be exposed, such as the user’s sexual orientation and even if the user consumed addictive substances. These privacy issues become more alarming when considering the nature of online social networks: information can be obtained on a network user without even directly accessing that individual’s online profile; personal details can be inferred solely by collecting data on the user’s friends.

SOCIAL NETWORK SITES: METRICS, PRIVACY

2.1 Social Network Sites: Usage Metrics, Network Structure and Privacy

Social Network Sites (SNSs) exhibit wide popularity, high diffusion and an increasing number of features. Specifically, Facebook, which currently holds a prime position among SNSs, has a continuously evolving feature set and one billion monthly active users, approximately 81% of whom are from outside the U.S. and Canada, and 604 million of whom access the site via mobile devices [1]. Given this diversity, an effective way of understanding Facebook is by exploring motives for using the service via theoretical frameworks such as Uses and Gratifications (U&G) [2]. A good understanding A good understanding of these motives can shed light onto the intricate mechanisms behind important aspects of SNSs, such as site adoption, participation [4], information seeking [5], and the privacy of users [2]. Privacy, in particular, is a major concern since it dictates the usage decisions of many SNS users and as Facebook, specifically, has found itself under harsh criticism regarding the enactment of highly contentious privacy policies and privacy-sensitive features [3]. The emergence of social sites also represents a valuable research resource. Indeed, scholars have highlighted the enormous potential of taking advantage of data that are generated electronically when people use online services [3]. Furthermore, compared to the methods and data available to traditional social scientists, online information can be accessed and analyzed computationally in ways that are both efficient and accurate [3]. In particular, in the case of Facebook, a rich, robust Application Programming Interface (API) allows researchers to collect large volumes of data relating to issues such as site feature use and personal network structure with unprecedented accuracy, granularity and reliability.

Media is consumed for a wide range of purposes and individuals utilize different media channels to achieve very different ends [6]. U&G is a theoretical framework for studying these motives and outcomes – fundamentally, the “how” and “why” of media use [2]. A key strength of the approach is its established and broadly applicable frame of analysis (covering media as diverse as tabloids, reality TV as entertainment or social connection) with social and psychological *antecedents* (such as demographics) and cognitive, attitudinal, or behavioral *outcomes* (such as usage patterns) [3]. U&G has recently proven valuable in exploring and explaining a wide variety of social media phenomena including topics as diverse as the motivations for contributing content to an online community [4], explaining why political candidates are befriended [1], and cataloguing the psychosocial well-being of teenage girls. U&G studies have explored behavior on most common forms of social media including content sharing sites (e.g., YouTube), SNSs (e.g., Myspace [1]), media sharing communities, and blogs [3]. Taken together, this work highlights the importance of eliciting and understanding users’ motives in social media, as well as the value of employing data from a natural research instrument, like Facebook, for social studies. particular, has most commonly been captured by self-report methods using surveys. Typical questions include time spent on site and visit frequency [3]. Acknowledging the lack of rigor in such ad-hoc methods, the Facebook Intensity Scale was introduced to capture the extent to which a user is emotionally connected to Facebook and the extent to which Facebook is integrated into their daily activities. The scale has been subsequently adopted in a number of other studies.

The advent of SNSs has greatly facilitated the capture of personal social network data and a wide range of useful metrics can now be calculated automatically and in real time [2]. Commonly used metrics include:

- *Network Size*: The number of nodes in a participant’s egocentric network, i.e., the number of friends that an individual has. Correlations have been shown between network size and personality and social capital.
- *Network Density*: The extent that nodes in an egocentric network are interconnected – essentially, how many of an individuals’ friends know each other. This is calculated as the ratio of the number of ties to the number of possible ties.
- *Average Degree*: Mean number of mutual friends in an egocentric network. Higher values on this statistic have previously been associated with bonding social capital and higher socioeconomic status.

- *Average Path Length*: The average geodesic distance between all pairs of nodes in a network.
- *Diameter*: The longest geodesic distance within the network, i.e., maximum distance between two nodes.
- *Network Modularity*: A scalar value between −1 and 1 that measures the density of links inside communities as compared to links between communities.
- *Number of Connected Components*: The number of distinct clusters within a network. This has been interpreted as the number of an individual's social contexts and associated with bridging social capital and social contagion.
- *Average Clustering Coefficient*: The clustering coefficient is a measure of the embeddedness of a node in its neighborhood. The average gives an overall indication of the clustering in the network, and high values are associated with a "small-world" effect [3].

Users often make decisions about whether and how they use a SNS based on the perceived privacy implications of their actions. However, privacy is a complex concept that has presented challenges to the social media ecosystem. One key issue is the tradeoff between providing users with advanced new features that mine their data to provide relevant content but lead to negative effects in terms of how users perceive their privacy [6]. Attempting to understand this topic further, boyd [5] argues that in the context of the social web, privacy violations are common because mediated publics exhibit certain properties that are not present in unmediated publics, namely persistence, searchability, replicability, and invisible audiences. Researchers studying the social implications of privacy have concluded that the right to privacy can be considered a social stratifier that divides users into classes of haves and have-nots, thus creating a privacy divide [3].
The privacy of SNS information is a particularly pertinent topic of study because of research reporting that users find it challenging to understand the privacy implications of SNSs. For instance, recent research has shown that the current Facebook privacy controls allow users to effectively manage threats from outsiders, but are poor at mitigating concerns related to members of a user's existing friend network [3]. Similarly, a study on Facebook apps found abundant misunderstandings and confusion about how apps function and how they manage, use, and share profile data [4].

Investigating the uses and gratifications of a social network site can provide powerful descriptive and explanatory insights into the mechanisms that drive users' behaviors.

Nationality showed a significant effect on the regression model for the first privacy question, with participants from the USA being less concerned about their privacy on Facebook, possibly due to the fact that they are more tech savvy and comfortable with this online media. On the other hand, nationality did not have a significant effect on the second privacy question, but two of the motives for use did. Specifically, users that were motivated by communication opportunities with like-minded people were found to be more likely to report tweaking their privacy settings. From the factor's description we know that these people tend to be more enthusiastic about organizing or joining events and groups. This may be because they feel more comfortable in familiar settings and therefore have increased suspicion of strangers or companies on Facebook. Furthermore, since events predominantly take place offline and a popular use of groups is to organize offline meetings, it may be that these people have greater experience of the implications of Facebook privacy settings to offline social interaction. The fact that the *Content* motive was positively associated with frequently changing privacy settings may be due to the fact that people who frequently use applications and interactive content on Facebook have taken the time to understand the privacy implications of installing such dynamic features. Interestingly, the newsfeed feature, which caused a large backlash with regards to privacy when it was first introduced [6], does not show a significant effect on users' perceived privacy. Furthermore, a substantial discrepancy was observed in the motives of people that report to be concerned about their privacy on Facebook and those that engage in changing their privacy settings.

2.2 A New Social Media Security Model (SMSM)

Previous security methods have been proposed to reduce the vulnerabilities encountered by most social networks without altering the usability of the system. These models emphasised user education and policy implementation. However, these have very limited success as users are notoriously non-compliant with such methodology. Some work has been done to explore the effects of social networks in disseminating information about successful social engineering attacks. Coronges et al, 2012, developed an experiment where phishing emails were created and sent to a controlled user population. Their results showed that local leadership within an organisation appeared to influence security vulnerability but had no influence on security resilience or phishing prevention. Joshi and Kuo (2011) investigated mathematical formulation and computational models for security and privacy of social networks. In their empirical study, they presented several ways an attacker can take advantage of the security vulnerabilities in an online social network. Some of the vulnerabilities discussed were Social Phishing Attacks, and Neighbourhood Attacks which is a form of privacy attack. This gives the attacker sufficient information about the target. Furthermore, their work presented theories that recent work in programming language techniques demonstrates that it is possible to build online services that have strict privacy policies. Unfortunately, service providers need private data to generate revenue from their 'free' service, limiting the application of strict security and privacy measures. It is important to clearly identify the main vulnerabilities of current social media [7].

Key vulnerabilities of social media:

- Single Factor Authentication. Single factor authentication is the conventional security procedure which requires a username/email and password before a user is granted access to the system. The security methods involved in authenticating account owners before gaining access into a system started as a debate in the industry and has now become the greatest cause of concern (Altinkemer, 2011). The sophistication of social engineering tools used by attackers poses a serious threat to the integrity, reputation and lives of users which may result in a sudden decline of social media usage (Parwani et al, 2013) [8].
- Profile Cloning. The easiest technique used in stealing the identity of a social network's user is called profile cloning. With the current structural outlook of profile pages on Facebook, it is very easy to clone an original user and perform malicious activities with the cloned account before the victim or the platform providers discovers. The profile pages of Facebook have no particular unique separation from one another, and in cases were two users co-incidentally share the same first and surnames as well as the same profile image, the true owner of the account is left for the viewer to identify (Kumar et al, 2013). There are two types of profile cloning: existing profile cloning, cross site profile cloning.
- Phishing Attacks. Phishing could be described as a malicious craft used by social engineers with the aim of exploring the vulnerabilities of a system which are made easy through the ignorance of end-users (Khonji et al, 2013).
- Watering Hole Malware Attack. This method of malware attack involves the attacker guessing which websites members of an organization normally go, and then it infects these websites with malware with the hope that a user‘s computer within the target organisation will be infected when an infected site is visited. It was discovered in January 2013, that a watering hole attack was targeted at developers working at Facebook [10].
- Multiple Login Sessions. The security vulnerabilities created by allowing multiple login sessions can result in both data theft and economic consequences (Imafidon & Ikhalia, 2013). In websites that make it compulsory that users‘ pay before using their service e.g. an online movie store; by allowing the creation of multiple login sessions on one account to run concurrently, the user and other users can share the benefits on one subscription regardless of their remote locations.
- Weak Password Creation. uessed, cracked, deliberately shared or stolen. Most social networking sites require users to create passwords of no less than 6 characters during account registration. The passwords created are usually hashed or encrypted in the database using MD5 or Shal1 hash functions (Kioon et al, 2013). Unfortunately, users are meant to believe that when the passwords are encrypted in the database, it is impossible to decrypt them if the system is compromised [9].

The proposed social media security model (SMSM):

- Modification of Password Classification Algorithm: This model proposes that passwords should be classified according to their various levels of security using more declarative terms such as ‗very unsafe', ‗unsafe', ‗not secure', ‗a little secure', ‗secure' and ‗very secure'. This can be done by ensuring that passwords created by users must consist of combinations of uppercases, lowercases, numbers and at least two special characters. Furthermore, this new security feature will display an instant message to the user using any of the aforementioned declarative terms based on the character combinations selected (Vijaya et al, 2009) [11].
- Embedding Unique Usernames on Profile Pages: This security feature is proposed to solve the problem of profile cloning within the site. Many social networking sites only display the first and surnames of users on their profile pages which opens up certain security vulnerabilities when more than one user coincidentally share the same first and surnames. The vulnerabilities involved may allow the malicious user exhort money or commit an online crime under the guise of the victim which could take time to detect. In addition, embedding unique usernames on profile pages within the site will make it very difficult for a malicious user to steal the identity of a genuine user and act on their behalf (Chen et al, 2009).
- Re-authenticating Registered Users During Activation: This method will ensure that social media accounts must be activated before full access is granted. The application of this approach is generating a 25 character code for each user and the system sends the code to their email addresses to confirm the ownership of the email used for registration. When the users attempt to activate the account by clicking the activation link sent to their email address, they are redirected to another page, and are required to provide the 25 character code, the username and password before full activation is complete. This security mechanism will help in preventing web robots from undermining the validation method of the social network. Furthermore, it will increase security consciousness in every user subscribing for the service and protect a user whose email address has been compromised (Fu, 2006) [12].
- Email Based Two Factor Authentication: The technique used to implement two factor authentication in this model, is an email based approach. When a user initiates a login session and passes the first stage of authentication (traditional username/email and password), the system sends a randomly generated one time password token to the user‗s email address and redirects the user to another page within the site which requires the onetime password concurrently (Ikhalia & Imafidon, 2013). The user must navigate to the email address containing the randomly generated password token and supply it before access to the system is granted. This security model is now in huge demand by the industry and implementing an email based two factor authentication method is feasible and cost effective when compared with the SMS approach. Retrospectively, this new security enhancement will reduce security vulnerabilities faced by social media victims of spear phishing, session hijacking, identity and data theft (Yee, 2004).
- Make Private Message Notification 'Private': The importance of this security functionality is to protect the confidentiality of a user whose email address is being compromised. This model only allows a message notification sent to a user‗s email when a private message is sent from the social networking site to the user. In other words, the user must navigate to their private message inbox within the site to read the content of their messages. The necessity of this is to protect users whose emails have been compromised and users who have lost their email accounts (Joshi & Kuo, 2011).
- Restrict Unauthorised Access To Users' Profile Information: One of the keywords used to define social media by (Dabner, 2012) is ―A web-based service that allows individuals to construct a public or semi-public profile within a bounded system‖; therefore, only registered users must have access to the information they share within the site. This new security enhancement proposed will prevent external users from viewing profile information of registered users within the site (Ijeh, Preston & Imafidon, 2009). The major benefit of this security feature prevents the difficulties involved by digital forensic investigators in tracking

down malicious attackers hijacking users‘ information for cross site profile cloning (Ademu & Imafidon, 2012). When this security model is implemented an attacker must register in the system before a malicious activity can be performed, therefore making it easy for the attacker to be traced as most social networks store the IP address of registered users and updates them when they login. (Joshi & Kuo, 2011) [7].

- Prevent Multiple Login Sessions: The system prevents users from creating multiple sessions at the same time. The implementation of this security mechanism will ensure that users have control over their accounts. This new enhancement will also make users know when a cyber intruder is accessing their accounts because a message will be shown to the user if a session has been established on the account (Choti et al, 2012).

2.3 Privacy Issues of Social Network Based on Facebook Example

Helen Nissenbaum’s work offers an analytical framework for understanding the commonplace notion that privacy is ‘‘contextual.’’ Nissenbaum’s account of privacy and information technology is based on what she takes to be two non-controversial facts. First, there are no areas of life not governed by context-specific norms of information flow. For example, it is appropriate to tell one’s doctor all about one’s mother’s medical history, but it is most likely inappropriate to share that information with casual passersby. Second, people move into and out of a plurality of distinct contexts every day. Thus, as we travel from family to business to leisure, we will be traveling into and out of different norms for information sharing. As we move between spheres, we have to alter our behaviors to correspond with the norms of those spheres, and though there will always be risks that information appropriately shared in one context becomes inappropriately shared in a context with different norms, people are generally quite good at navigating this highly nuanced territory. On the basis of these facts, Nissenbaum suggests that the various norms are of two fundamental types. The first, which she calls norms of ‘‘appropriateness,’’ deal with ‘‘the type or nature of information about various individuals that, within a given context, is allowable, expected, or even demanded to be revealed’’ (2004, 120). In her examples, it is appropriate to share medical information with doctors, but not appropriate to share religious affiliation with employers, except in very limited circumstances. The second set are norms of ‘‘distribution,’’ and cover the ‘‘movement, or transfer of information from one party to another or others’’ [17]. She grounds these norms partly in work by Michael Walzer, according to which ‘‘complex equality, the mark of justice, is achieved when social goods are distributed according to different standards of distribution in different spheres and the spheres are relatively autonomous’’ [14]. Information is such a social good. Thus, in friendship, confidentiality is a default rule: it may be appropriate to share the details of one’s sex life with a friend, but it is not appropriate for that friend to broadcast the same information on her radio show. In medical situations, on the other hand, a certain sharing of information—from a radiologist to a primary care physician, for example—is both normal and expected [13].

The use of Nissenbaum’s account extends well beyond the public surveillance context to which she applies it. Evidence from social networking sites, and Facebook in particular, suggests that contextual integrity will be an apt analytical framework in that context as well. First, and unlike popular perceptions and perhaps unlike other SNS, almost all of the evidence suggests that Facebook users primarily use the site to solidify and develop their offline social relationships, rather than to make new relationships online. Information on Facebook would predictably tend to mirror the offline tendency to contextual situatedness.3 Second, evidence about social conflicts in Facebook suggests that one of the greatest sources of tension is when information that would remain internal to one context offline flows to other contexts online. This is in large part an interface issue—there is no easy way for a user to establish and maintain the many separate and sometimes overlapping social spheres that characterize offline life. If these fluid spheres are difficult to explicitly create on Facebook, then managing information flows among them is likely to be very difficult. Indeed, the underlying architecture of Facebook is largely insensitive to the granularity of offline social contexts and the ways norms differ between them. Instead, the program assumes that users want to project a single unified image to the world. These contextual gaps are endemic to SNS, and a substantial source of privacy problems.4 For example, the ‘status line,’ generally gets projected indiscriminately to all friends. One can think of the status line as an answer to the questions ‘‘How’s it going?’’ or ‘‘What’s up?’’ However, in offline life one tailors responses to the audience one is with. While Facebook now allows users to create multiple friend groups and then control which groups can see their status, this requires explicit, cumbersome effort and falls far short of reflecting the many complex and overlapping spheres of offline life [15].

Two recent Web phenomena, which pre-date social networking sites, highlight some of the privacy issues that are relevant to Facebook and other social networking sites. The first is blogging. The number of people who post their thoughts online in Web logs is staggering, and growing rapidly.

Information posted to the Internet is potentially visible to all. For most people, such universal broadcast of information has no parallel offline. In other words, offline personal information is seldom communicated to a context anywhere near as broad as the entire Internet. As social networking sites continue to grow and increasingly integrate with the rest of the Internet, they should be expected to inherit some of these issues. At the same time, the contextual situation for social networking is made both more complex and more difficult by the import of the offline concept of ‘‘friend.’’ Information flows on social networking sites are mediated not just by the global nature of Internet communication, but by the ways that those sites and their users interpret the meaning of online friendship and the social norms that go with it. Not surprisingly, empirical research indicates that SNS users (possibly excepting younger teens on Myspace) tend to construct their identities relationally, becoming and expressing who they are by way of highlighting whom they associate with (Papacharissi 2009; Livingstone 2008) [16].

The precursors to social networking (blogging and webcams) demonstrated two phenomena—the risk of collateral privacy damage, and the plasticity of norms—and these phenomena are also evident in recent changes to Facebook. We will examine the case of collateral damage first, because its analysis is more straightforward than the plasticity of norms. In both cases, we will suggest that the difficulties lie in the tensions between the tendency of offline social contexts to be highly granular and specific, and the tendency of social network sites to mirror the more global context of the Internet, even as they substantially nuance that context through the emerging norms of online friendship. This granularity gap needs to be addressed through interface and design decisions that highlight, rather than obscure, the new ways that information travels online [17].

Another serious violation to the norms of distribution is that the developers behind the applications are also largely invisible, obscuring the fact that the information is in fact leaving the confines of Facebook and not just going to the user’s friends. A college student in Germany can create a quiz application that matches people with the beer they are most like, running it on his computer in his dorm. Users accessing this application will likely be completely unaware that their profile information, and their friends’ profile information, can be accessed by a student in a dorm room in another country. While a third party application developer, such as the student in the dorm room in Germany, may be acting completely ethically and only accessing the profile information needed for the application, there is no guarantee of this, and little vetting process—anyone can become a Facebook application developer [15].

The invisibility of information flows presents a particular problem because when we do not know what is being done with our information, we have no ability to contest it. If the architecture and interface of Facebook essentially hides the amount of information that is shared to third parties, then there is little that a user can do about that sharing. Indeed, there is little that she can do to avoid the situation, other than decline to use the third party applications. The choice for users is binary: install the application and give full access to their own and their friends’ personal information, or don’t use the application at all.16 We can imagine designing a number of different mechanisms that could provide greater transparency (and in some cases, greater control) of these information flows.

The backlash was in part generated by a sense that the News Feed violated the privacy of users. In the terms of Nissenbaum's framework, the analogy would be to the case of public records being placed online.20 For public records, she suggests, the privacy problem is one of the scope of the distribution of the information. It is not that there is any more information publicly available; it is that it is a lot easier to get to the information. It is no longer necessary to travel to a (possibly remote) local courthouse. When an activity is easier, more people will do it. Thus, for example, idle curiosity about 'who owns that house, and what they paid for it' becomes a topic that one might actually pursue online, when the same topic would be too much like work to pursue offline. As a result, people have less de facto privacy. Facebook similarly increased the efficiency of accessing information that was already in principle accessible; in so doing, it reduced the de facto privacy of its users. A more significant difference occurs on the other end, as newsfeeds do not just automate the process of receiving updates; they automate the process of sending them. Here a significant difference from holiday letters emerges: the process of sending those letters is inefficient, in that it involves individually addressing and stuffing envelopes for every update. If the letter says something that a recipient might find offensive, he gets a card but not the letter. If the letter is going to someone that a person no longer cares to keep up with, she might not send it, and might even remove the person from her list. In other words, being offline in this case introduces a necessary pause between composing the letter and sending it, and this encourages the list itself being subject to at least a casual vetting procedure. No such procedural encouragement exists for Facebook updates. Nothing in the process of sending updates encourages one to even notice who the full audience is, much less to assess who should be in it. Retention of updates, combined with reminders to that effect, would further encourage users to view their identities online as constructed and performative, moving Facebook closer to the world of Jennicam.25 This encouragement would be magnified even more by a ''Mary has just looked at all of your updates'' option; users would increasingly view their Facebook identities as subject to constant surveillance, and modify them accordingly. If I knew that Mary always looked at all my updates, I might update with her in mind. More generally, users would be encouraged to have updates resemble the news wires from which the model was taken. Whether or not this is a ''good'' model for online friendship on a social networking site could then be discussed by users who were encouraged to form a mental model that accurately reflected the various flows of information on the site.

An intellectual property regime that grants strong proprietary rights and few exceptions favors mass media over independent producers of speech. Even the placement of advertising on portal sites matters. Recent work comparing Facebook with other SNS suggests that the architectural features of those sites will tend to facilitate certain kinds of social interactions and not others.26 The preceding analysis underscores the important relations between privacy norms and application and interface design. Two points bear emphasis. First, at the core of the privacy issues discussed here is a gap between the granularity of offline social contexts and those on Facebook. One of the greatest strengths of Nissenbaum's contextual integrity framework is that it highlights the importance of social contexts to privacy norms. The application of Nissenbaum's work to Facebook underscores the extent to which the granularity differences are driving many of the surface-level privacy issues. Facebook, as we have argued, does not adequately reflect this insight. As a result, offline contexts are more easily differentiated and kept separate than those on Facebook. To the extent that Facebook users use the program to develop offline social ties, this granularity gap will inevitably produce conflicts. Second, this gap shows that part of what is at stake is the definition of what it means to be a ''friend'' on Facebook. The switch to the News Feed shows that the norms for online friendship are in a state of considerable flux. While no-one would claim that a Facebook friend is the same as an offline friend, the fact that there are many overlapping cases means that the shared vocabulary sometimes leads us to a misperception of sameness. In fact, the concepts of a 'friend' on Facebook and an offline friend may be diverging even more, making this difference bigger. The News Feed makes updates a lot like blogging, and Applications put users in the position of treating their friends' information as commodities to exchange for access to whatever the application in question does. If Facebook norms for friendship move further away from those offline, we would expect that users' would model their online behavior accordingly. While few of them might endorse the sense of complete publicity that

Jennicam embraced, publicity at nearly that level is enabled by the combination of the News Feed and a user who posts constant updates.

2.4 Social Media, Privacy and Security

Dr Ackland [18] explained that social media use leads to vast digital repositories of searchable content such as web pages, blog posts, newsgroup posts, Wikipedia entries. Such repositories can be thought of as information public goods that have economic value to business because they enable people to find information efficiently – for example, by using Google to search for answers to technical questions. For governments, the value of social media lies in the potential for direct dialogue between citizens and government, promoting participatory democracy (that is, broader participation in the direction and operation of government). However, people are often not involved in social media for the instrumental reasons of creating information public goods or strengthening democracy but for social reasons and to express their individual or collective identity. Much of the time people's use of social media is expressive (like voting, or cheering at the football) rather than instrumental and people tend to be intrinsically motivated to participate – that is, for the inherent enjoyment of the activity – rather than being extrinsically motivated (for example, by the expectation of external rewards).

Two issues that may reduce the value of social media to business and government by adversely affecting people's motivation to participate.

1. Game mechanics, was not covered in detail but refers to the use of rewards such as badges on foursquare, which may encourage some users but deter others, so may have the effect of selecting particular types of users to participate, and may also change the way in which people engage with online environments.
2. Privacy. The second issue identified as potentially diminishing people's motivation to participate in social media, and the issue of particular relevance in the context of this workshop, was privacy. Existing research [20] is mostly survey research focusing on a particular demographic: youth. e.g.;

- Sonia Livingstone [21] found that younger teenagers use social media for creating 'identity through display' while older teenagers create 'identity through connection'. Different risks are associated with each of these behaviours.
- Zeynep Tufecki's quantitative study of US college students [18] looked at how privacy concerns affect students' willingness to disclose information online. This research suggested that 'treating privacy within a context merely of rights and violations, is inadequate for studying the Internet as a social realm'.

Further research needed. Most existing research has focused on youth in wealthy countries but there is a need for research on other demographic groups. For understanding how people's participation in Government 2.0 might be affected by privacy and security concerns, for example, it will be important to include demographic groups more relevant to these initiatives as youth tend not be involved in political activities online [18]. Dr Ackland drew attention to a study of older Australians and their social media use (funded by an ARC Linkage Grant) being undertaken by the Australian Demographic and Social Research Institute at The Australian National University in partnership with National Seniors Australia. He noted that research into privacy and Government 2.0 is generally not based on data and is mainly speculation at this stage. Little is known about people's attitudes towards the potential for governments to use their data for purposes other than policymaking, such as surveillance and law enforcement. Developments in network science and mining social graphs (that is, the social connections between people or organisations). Social media use leads to digital traces of activity that are permanent, searchable and crossindexable. New tools are being developed to support this type of research through data mining, network visualisation and analysis, including:

- The Virtual Observatory for the Study of Online Networks (VOSON) System (http://voson.anu.edu.au) developed by Dr Ackland.

- SAS Social Media Analytics.
- NodeXL, which enables social network data such as that gathered by VOSON to be imported into Excel.

It was noted [23] that there is an important difference between individual privacy and social graph security. Network scientists increasingly have the ability to mine social graphs and place people in groups by identifying clusters in social networks and predicting the existence of social connections. Protecting the social graph is more important –and more difficult – than protecting personal data.

Professor Bronitt [21] drew attention to the work of the Centre of Excellence in Policing and Security (CEPS), which is a research partnership between The Australian National University, Griffith University, The University of Queensland and Charles Sturt University. The Centre also brings together a range of industry partners (refer to the CEPS website for a full list) who are committed to innovation and evidence-based research in policing and security. The Centre's goal is 'to gain a better understanding of the origins, motivations and dynamics of crime and security threats. Its objectives are to create significant transformations in crime, crime prevention and security policies, and to bring about evidence-based reform in the practices of policing and security to enhance Australia's social, economic and cultural wellbeing.' Professor Bronitt discussed the issues of vulnerability and resilience. In relation to infrastructure, he noted that much vulnerable infrastructure is in the hands of private companies and it will be necessary to engage with policymakers and private companies before starting research. Social media can make you vulnerable but also resilient. For example, in the Brisbane floods social media was used for self-organising [20]. A key role for 'foresight forums', which have not previously been used in Australia much. Laws and regulations don't yet cover all aspects of what is now appearing in social media – eg video. He noted that privacy in the modern context clearly goes into the realm of the public – physically and virtually – related to human dignity. Social media has no borders but laws are tied to territory, so addressing interjurisdictional issues poses significant challenges.

2.5 Information Sharing and Privacy on the Facebook

Nobody is literally forced to join an online social network, and most networks we know about encourage, but do not force users to reveal - for instance - their dates of birth, their cell phone numbers, or where they currently live. And yet, one cannot help but marvel at the nature, amount, and detail of the personal information some users provide, and ponder how informed this information sharing is. Changing cultural trends, familiarity and confidence in digital technologies, lack of exposure or memory of egregious misuses of personal data by others may all play a role in this unprecedented phenomenon of information revelation. Yet, online social networks' security and access controls are weak by design - to leverage their value as network goods and enhance their growth by making registration, access, and sharing of information uncomplicated. At the same time, the costs of mining and storing data continue to decline. Combined, the two features imply that information provided even on ostensibly private social networks is, effectively, public data, that could exist for as long as anybody has an incentive to maintain it. Many entities - from marketers to employers to national and foreign security agencies - may have those incentives.

At the most basic level, an online social network is an Internet community where individuals interact, often through profiles that (re)present their public persona (and their networks of connections) to others. Although the concept of computer-based communities dates back to the early days of computer networks, only after the advent of the commercial Internet did such communities meet public success. Following the SixDegrees.com experience in 1997, hundreds of social networks spurred online (see [24] for an extended discussion), sometimes growing very rapidly, thereby attracting the attention of both media and academia. In particular, [30] have taken ethnographic and sociological approaches to the study of online self-representation; [25] have focused on the value of online social networks as recommender systems; [25] have discussed information sharing and privacy on online social networks, using FB as a case study; [26] have demonstrated how information revealed in social networks can be exploited for “social” phishing; [28] has studied identity-sharing behavior in online social networks.

FB requires a college’s email account for a participant to be admitted to the online social network of that college. As discussed in [28], this increases the expectations of validity of the personal information therein provided, as well as the perception of the online space as a closed, trusted, and trustworthy community (college-oriented social networking sites are, ostensibly, based “on a shared real space” [29]). However, there are reasons to believe that FB networks more closely resemble imagined [30] communities (see also [28]): in most online social networks, security, access controls, and privacy are weak by design; the easier it is for people to join and to find points of contact with other users (by providing vast amounts of personal information, and by perusing equally vast amounts of data provided by others), the higher the utility of the network to the users themselves, and the higher its commercial value for the network’s owners and managers. FB, unlike other online networks, offers its members very granular and powerful control on the privacy (in terms of searchability and visibility) of their personal information. Yet its privacy default settings are very permeable: at the time of writing, by default participants’ profiles are searchable by anybody else on the FB network, and actually readable by any member at the same college and geographical location. Based on research [27] we have some conclusions: a total of 506 respondents accessed the survey. One-hundred-eleven (21.9%) were not currently affiliated with the college Institution where we conducted our study, or did not have a email address within that Institution’s domain. They were not allowed to take the rest of the survey. A separate set of 32 (8.2%) participants had taken part in a previous pilot survey and were also not allowed to take the survey. Of the remaining respondents, 318 subjects actually completed the initial calibration questions. Out of this set, 278 (87.4%) had heard about FB, 40 had not. In this group, 225 (70.8%) had a profile on FB, 85 (26.7%) never had one, and 8 (2.5%) had an account but deactivated it. Within those three groups, respectively 209, 81, and 7 participants completed the whole survey. We focus our analysis on that set - from which we further removed 3 observations from the non-members group, since we had reasons to believe that the responses had been created by the same individual. This left us with a total of 294 respondents. Age and student status are correlated with FB membership - but what else is? Well, of course, having heard of the network is a precondition for membership. Thirty-four participants had never heard of the FB - nearly half of the staff that took our survey, a little less than 23% of the graduate students, and a negligible portion of the undergraduate students (1.59%). However, together with age and student status (with the two obviously being highly correlated), another relevant distinction between members and non-members may arise from privacy attitudes and privacy concerns. Before we asked questions about FB, our survey ascertained the privacy attitudes of participants with a battery of questions modelled after the Alan Westin’s studies [26], with a number of modifications. In particular, in order not to prime the subjects, questions about privacy attitudes were interspersed with questions about attitudes towards economic policy and the state of the economy, social issues such as samesex marriage, or security questions related to the fear of terrorism. In addition, while all instruments asked the respondent to rank agreement, concern, worries, or importance on a 7-point Likert scale, the questions ranged from general ones (e.g., “How important do you consider the following issues in the public debate?”), to more and more specific (e.g., “How do you personally value the importance of the following issues for your own life on a day-to-day basis?”), and personal ones (e.g., “Specifically, how worried would you be if” [a certain scenario took place]). “Privacy policy” was on average considered a highly important issue in the public debate by our respondents (mean on the 7-point Likert scale: 5.411, where 1 is “Not important at all” and 7 is “very important”; sd: 1.393795). In fact, it was regarded a more important issue in the public debate than the threat of terrorism (t = 2.4534, Pr>t = 0.0074; the statistical significance of the perceived superiority was confirmed by a Wilcoxon signed-rank test: z = 2.184 Pr>|z|= 0.0290) and same sex marriage (t = 10.5089, Pr>t = 0.0000; Wilcoxon signed-rank test: z = 9.103 Pr>|z|= 0.0000); but less important than education policy (mean: 5.93; sd: 1.16) or economic policy (mean: 5.79; sd: 1.21) [24]. The slightly larger mean valuation of the importance of privacy policy over environmental policy was not significant. (These results are comparable to those found in previous studies, such as [29].) The same ranking of values (and comparably statistically significant differences) was found when asking for “How do you personally value the importance of the following issues for your own life on a day-to-day basis?" The mean value for the importance of privacy policy was 5.09. For all categories, subjects assigned slightly (but statistically significantly) more importance to the issue in the public debate than in their own life on a day-to-day basis (in the privacy policy case, a Wilcoxon signed-rank test returns z = 3.62 Pr > |z| = 0.0003 when checking the higher valuation of the issue in the public debate)[24]. Similar results were also found when asking for the respondents' concern with a number of issues directly relevant to them: the state of the economy where they live, threats to their personal privacy, the threat of terrorism, the risks of climate change and global warming. Respondents were more concerned (with statistically significant differences) about threats to their personal privacy than about terrorism or global warming, but less concerned than about the state of the economy. Privacy concerns are not equally distributed across FB members and non-members populations: a two-sided t test that the mean Likert value for the "importance" of privacy policy is higher for non-members (5.67 in the non-members group, 5.30 in the members group) 6 is significant (t = -2.0431, Pr<t = 0.0210) [24]. Similar statistically significant differences arise when checking for the level of concern for privacy threats and for worries associated with the privacy scenarios described above. The test becomes slightly less significant when checking for member/non-member differences in the assigned importance of privacy policy on a day-to-day basis. Importantly, in general no comparable statistically significant differences between the groups can be found in other categories. For example, worries about the global warming scenario gain a mean Likert valuation of 5.36 in the members sample and 5.4 in the non-members sample. (A statistically significant difference can be found however for the general threat of terrorism and for the personal worry over marriage between two people of same sex: higher values in the non-members group may be explained by their higher mean age.) In order to understand what motivates even privacy concerned individual to share personal information on the Facebook, we need to study what the network itself is used for. Asking participants this question directly is likely to generate responses biased by self-selection and fear of stigma. Sure enough, by far, FB members deny FB is useful to them for dating or self-promotion. Instead, members claim that the FB is very useful to them for learning about and finding classmates (4.93 mean on a 7-point Likert scale) and for making it more convenient for people to get in touch with them (4.92) [24], but deny any usefulness for other activities. Other possible applications of FB - such as dating, finding people who share one's interests, getting more people to become one's friends, showing information about oneself/advertising oneself - are ranked very low. In fact, for those applications, the relative majority of participants chooses the minimal Likert point to describe their usefulness (coded as "not at all" useful). Still, while their mean Likert value remains low, male participants find FB slightly more useful for dating than female.

Often, survey participants are less privacy conscious than non participants. For obvious reasons, this self-selection bias is particularly problematic for survey studies that focus on privacy. Are our respondents a biased sample of the Institution's FB population - biased in the sense that they provide more information than the average FB members? We did not find strong evidence of that. Since we mined the network before the survey was administered, we were able to compare information revelation by survey participants and non survey participants. It is true that, on average, our survey takers provide slightly more information than the average FB member. However, the differences in general do not pass a Fisher's exact test for significance, except for personal address and classes (where non participants provide statistically significant less information) and political views (in which the difference is barely significant). Similarly, around 16% of respondents who expressed the highest concern for the scenario in which someone 5 years from now could know their current sexual orientation, partner's name, and political orientation, provide nevertheless all three types of information - although we can observe a descending share of members that provide that information as their reported concerns increase. Still, more than 48% of those with the highest concern for that scenario reveal at least their current sexual orientation; 21% provide at least their partner's name (although we did not control for the share of respondents who are currently in relationships); and almost 47% provide at least their political orientation. How knowledgeable is the average FB member about the network's features and their implications in terms of profile visibility? By default, everyone on the Facebook appears in searches of everyone else, and every profile at a certain Institution can be read by every member of FB at that Institution. However, the FB provides an extensive privacy policy and offers very granular control to users to choose what information to reveal to whom. As mentioned above, relative to a FB member, other users can either be friends, friends of friends, non-friend users at the same institution, non-friend users at a different institution, and non-friend users at the same geographical location as the user but at a different university (for example, Harvard vs. MIT). Users can select their profile visibility (who can read their profiles) as well as their profile searchability (who can find a snapshot of their profiles through the search features) by type of users. More granular control is given on contact information, such as phone numbers. And yet, among current members, 30% claim not to know whether FB grants any way to manage who can search for and find their profile, or think that they are given no such control. Eighteen percent do not know whether FB grants any way to manage who can actually read their profile, or think that they are given no such control. These numbers are not significantly altered by removing the 13 members who claim never to login to their account. In fact, even frequency of login does not explain the lack of information for some members. On the other hand, members who claim to login more than once a day are also more likely to believe that they have “complete” control on whom can search their profile. Awareness of one’s ability to control who can see one’s profile is not affected by the frequency of login, but is affected by the frequency of update (a Pearson 2(12) = 28.9182 Pr = 0.004 shows that the distribution is significant). Note the difference between the two graphs and, specifically, the distribution by frequency of update for respondents who answered “Do not know” or “No control” (graph on the right). Twenty-two percent of our sample do not know what the FB privacy settings are or do not remember if they have ever changed them. Around 25% do not know what the location settings are. To summarize, the majority of FB members claim to know about ways to control visibility and searchability of their profiles, but a significant minority of members are unaware of those tools and options. More specifically, we asked FB members to discuss how visible and searchable their own profiles were. We focused on those participants who had claimed never to have changed their privacy settings (that by default make their profile searchable by everybody on FB and visible to anybody at the same Institution), or who did not know what those settings were. Almost every such respondent realizes that anybody at their Institution can search their profile. However, 24% incorrectly do not believe that anybody on FB can in fact search their profile. Misunderstandings about visibility can also go in the opposite direction: for instance, 16% of current members believe, incorrectly, that anybody on FB can read their profile. 12 In fact, when asked to guess how many people could search for their profile on FB (respondents could answer by selecting the following possible answers from a drop-box: a few hundred, a few thousands, tens of thousands, hundreds of thousands, millions), the relative majority of members who did not alter their default settings answered, correctly, “Millions.” However, more than half actually underestimated the number to tens of thousands or less. In short, the majority of FB members seem to be aware of the true visibility of their profile - but a significant minority is vastly underestimating the reach and openness of their own profile. Does this matter at all? In other words, would these respondents be bothered if they realized that their profile is more visible than what they believe? The answer is complex. First, when asked whether the current visibility and searchability of the profile is adequate for the user, or whether he or she would like to restrict it or expand it, the vast majority of members (77% in the case of searchability; 68% in the case of visibility) claim to be satisfied with what they have - most of them do not want more or less visibility or searchability for their profiles (although 13% want less searchability and 20% want less visibility) than what they (correctly or incorrectly) believe to have. Secondly, FB members remain wary of whom can access their profiles, but claim to manage their privacy fears by controlling the information they reveal.

While respondent are mildly concerned about who can access their personal information and how it can be used, they are not, in general, concerned about the information itself, mostly because they control that information and, with less emphasis, because believe to have some control on its access. Respondents are fully aware that a social network is based on information sharing: the strongest motivator they have in providing more information are reported, in fact, as “having fun” and “revealing enough information so that necessary/useful to me and other people to benefit from FaceBook.” However, psychological motivations can also explain why information revelation seems disconnected from the privacy concerns. When asked to express whether they considered the current public concern for privacy on social network sites such as the FaceBook or MySpace to be appropriate (using a 7-point Likert scale, fact, the majority of respondents agree (from mildly to very much) with the idea that the information other FB members reveal may create privacy risks to those members (that is, the other members; average response on a 7-point Likert scale: 4.92) - even though they tend to be less concerned about their own privacy on FB (average response on a 7-point Likert scale: 3.60; Student's t test shows that this is significantly less than the concern for other members: t = -10.1863, P <t = 0.0000; also a Wilcoxon matched pair test provides a similar result: z = -8.738, Pr <|z| = 0.0000). In fact, 33% of our respondents believe that it is either impossible or quite difficult for individuals not affiliated with an university to access FB network of that university. "Facebook is for the students" says a student interviewed in [29]. But considering the number of attacks described in [30] or any recent media report on the usage of FB by police, employers, and parents, it seems in fact that for a significant fraction of users the FB is only an imagined community.

In order to gauge the accuracy of the survey responses, we compared the answers given to a question about revealing certain types of information (specifically, birthday, cell phone, home phone, current address, schedule of classes, AIM screenname, political views, sexual orientation and the name of their partner) with the data from the actual (visible) profiles. We found that 77.84% of the answers were exactly accurate: if participants said that they revealed a certain type of information, that information was in fact present; if they wrote it was not present, in fact it was not. A little more than 8% revealed more than they said they do (i.e. they claim the information is not present when in fact it is). A little more than 11% revealed less then they claimed they do. In fact, 1.86% claimed that they provide false information on their profile (information is there that they claim is intentionally false or incomplete), and 0.71% have missing false information (they claimed the information they provide is false or incomplete, when in fact there was no information). We could not locate the FB profiles for 13 self-reported members that participated in the survey. For the participants with CMU email address, 2 of them did mention in the survey that they had restricted visibility, searchability, or access to certain contact information, and 3 wrote that not all CMU users could see their profile.

2.6 Social Media Compliance Policy

Social media are powerful communication tools but they carry significant risks to the reputation of the University and its members. A prominent risk arises from the blurring of the lines between personal voice and institutional voice. For the purposes of this policy, social media is defined as media designed to be disseminated through social interaction, created using highly accessible and scalable publishing techniques using the internet.

Examples of popular social media sites include but are not limited to:

- LinkedIn
- Twitter
- Facebook
- YouTube
- MySpace
- Flickr
- Yammer
- Yahoo/MSN messenger
- Wiki's/Blogs

All members of the University using social media tools, including via personal accounts, must be aware that the same laws, professional expectations, and guidelines apply at all times.

Any social media announcements not issued by the Corporate Communications team, which may include (but are not limited to) content, comments and opinions must not give the impression they are in anyway the explicit positioning of the University of Liverpool. Any official University position and comment must be approved by the University Corporate Communications Director or his or her representative.

The Social Media Compliance Policy helps members of the University to use social media sites without compromising their personal security or the security of University information assets.

The University has adopted the following principles, which underpin this policy:

- All information assets must be appropriately handled and managed in accordance with their classification.
- University information assets should be made available to all who have a legitimate need for them.
- The integrity of information must be maintained; information must also be accurate, complete, timely and consistent with other related information and events.
- All members of the University, who have access to information assets, have a responsibility to handle them appropriately and in accordance with their classification.
- Information asset owners are responsible for ensuring that the University classification scheme (which is described in the Information Security Policy) is used appropriately [31].

Below is a table of examples of threats and risks associated with the use of social media:

Threats: Introduction of viruses and malware to the University network; Exposure of the University and its members through a fraudulent or hijacked presence e.g. unofficial social media accounts; Unclear or undefined ownership of content rights of information posted to social media sites (copyright infringement); Use of personal accounts to communicate University owned information assets; Excessive use of social media within the University.

Risks:

- Data leakage/theft.
- System downtime.
- Resources required to clean systems.
- Customer backlash/adverse legal actions.
- Exposure of University information assets.
- Reputational damage.
- Targeted phishing attacks on customers or employees .
- Exposure of customer information.
- Privacy violations.
- Network utilisation issues.
- Productivity loss.
- Increased risk of exposure to viruses and malware due to longer duration of sessions.

The University’s Information Security Policy defines the categories which are assigned to University information assets. Those assets which are classified as Confidential, Strictly Confidential or Secret must not be posted on social media sites. Postings must not contain personal information concerning students, employees, or alumni. In addition, members of the University must be aware of copyright and intellectual property rights of others and of the University.

Members of the University should be aware of the social media terms and conditions and ownership of content before submitting. Members should familiarise themselves with key University policies including:

- University IT Regulations
- Information Asset Classification Policy

- Copyright Policy
- Data Protection Policy
- Acceptable Use of Electronic Resources [31]

Communication via social media sites and tools must protect the University's institutional voice by remaining professional in tone and in good taste. Members of the University who use personal social media accounts must not give the impression that their social media site represents the explicit positioning of the University of Liverpool. This should be considered when:

- Naming pages or accounts
- Selecting a profile picture or icon
- Selecting content to post [31]

Names, profile images, and posts should all be clearly linked to the particular Faculty, School, Institute and Professional Service Department. In addition, All University pages must have an associated member who is identified as being the information asset owner and who is responsible for its official affiliation of the University. If you are responsible for representing official social media accounts on behalf of the University of Liverpool when posting on a social media platform, clearly acknowledge this.

Members of the University must be aware that the University has the right to request the removal of content from an official social media account and from a personal account if it is deemed that the account or its submissions pose a risk to the reputation of the University or to that of one of its members.

All members of the University are directly responsible and liable for the information they handle. Members of staff are bound to abide by the University IT regulations by the terms of their employment. Students are bound to abide by the University IT Regulations when registering as a student member of the University.

Authorised members of the University may monitor the use and management of information assets to ensure effective use and to detect unauthorised use of information assets.

2.7 Threats and Anti-threats Strategies for Social Network Sites

Online Social Networks (OSN) such as Facebook, Tweeter, MySpace etc. play an important role in our society, which is heavily influenced by the Information Technology (IT) [32]. In spite of their user friendly and free of cost services, many people still remain reluctant to use such networking sites because of privacy concerns. Many people claim, these sites have made the world more public and less private – consequently a world with less morale is evolving. Some consider this social change as positive because people are more connected to each other. Nowadays almost all people use the social networking websites to share their ideas photos, videos, with their friends and relatives and even discuss many things about their daily life not only social issues but also others like politics which help lately to change the regime status in many countries such as Egypt Libya and Tunisia. in 1971 started the first communications in the form of social networks this happened in the past when the first email was send from one computer to another connected one. The data exchanged over the phone lines was happened in 1987 where bulletin board system set and web browsers were used for the first time to make to establish the principle of communication. Even, social networking websites have many advantages for users to communicate and exchanges information as we mention above, unfortunately, they have their negative impact! Most people spend all their time using such websites and forget about their duties and sometimes many use loss their life using such websites due to the illegal contents like pornographic, terrorism, religiolism and many other. Social networking websites must satisfy many security issues to work successfully and to be used by people who unfortunately trust most of websites. Social networking websites should have save storage for personal data and tools for managing and securing data and control access to the saved data according some limitations. They must add some features to restrict individual data access from other users. Unfortunately most of social networks have not had applied this issues actually There are many attacks on social networks such as worms, viruses, Trojan horses and fishing websites. Malicious programs vulnerabilities and many others such as sql injection top attacks users should be worried about all types of attacks and have the right firewalls and software which protect them of being hack by cyber criminal. Social networking features:

- Global social network where geographical and spatial barriers are cancelled.
- Interaction where sites gives space for the active participation for the viewer and reader.
- Easy to use most social networks can be use easily and they contains images and symbols that make it easier for user interaction [34].

Types of social networks divisions depending on the service provided or targets from its inception to the following types: personal networks; cultural networks.

Social networks also can be divided according to the way we communicate and services into three types:

- Networks Allow For Written Communication.
- Network Allow Voice Communication.
- Network allow for visual Communication [32].

Social networks attract thousands of users who represent potential victims to attackers from the following types (Ref: Figure 1) [34].

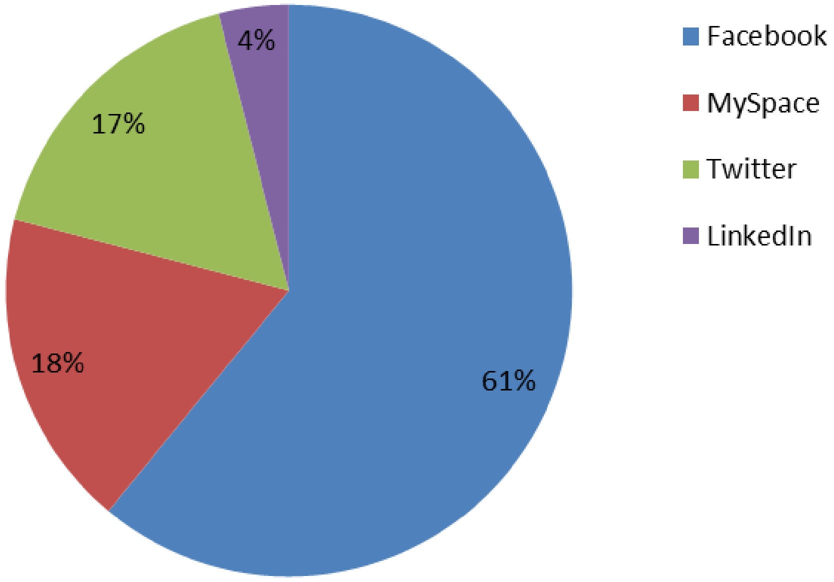


FIGURE 1. Threats percentage on social networks (Sophos 2010 Security Threat Report).

There are many procedures, which help us to protect as much as possible our privacy when using social networks.

- Be careful and don’t write any sensitive information in your profile page bulletin board, instant messaging or any other type of participation and electronic publishing in internet so that the identity could be protected against the thefts or security threats.
- Be skeptical because social networking websites are full of illegal users, hackers and cyber criminals.
- Be wise man and thinks twice before you write any think when using social networking websites.

- Be polite and do not publish any illegal picture and video and even don’t write any abnormal messages and also reflects your personal impacts and be ambassador to all others on the internet.
- Read the privacy policy of all social networks before using them [35].

Cyber threats that might the users face can be categorized into two categories.

Privacy concerns demand that user profiles never publish and distribute information over the web. Variety of information on personal home pages may contain very sensitive data such as birth dates, home addresses, and personal mobile numbers and so on. This information can be used by hackers who use social engineering techniques to get benefits of such sensitive information and steal money. Generally, there are two types of security issues: One is the security of people. Another is the security of the computers people use and data they store in their systems. Since social networks have enormous numbers of users and store enormous amount of data, they are natural targets spammers, phishing and malicious attacks. Moreover, online social attacks include identity theft, defamation, stalking, injures to personal dignity and cyber bulling. Hackers create false profiles and mimic personalities or brands, or to slander a known individual within a network of friends[36].

Anti threats strategies. Recent work in programming language techniques demonstrates that it is possible to build online services that guarantee conformance with strict privacy policies. On the other hand, since service providers need private data to generate revenue, they have a motivation to do the opposite. Therefore, the research question to be addressed is to what extent a user can ensure his/her privacy while benefiting from existing online services.

There are recommended the following strategies for circumventing threats associated with social website [33]:

a) Building awareness the information disclosure: users most take care and very conscious regarding the revealing of their personal information in profiles in social websites.
b) Encouraging awareness -raising and educational campaigns: governments have to provide and offer educational classes about awareness -raising and security issues.
c) Modifying the existing legislation: existing legislation needs to be modified related to the new technology and new frauds and attacks.
d) Empowering the authentication: access control and authentication must be very strong so that cybercrimes done by hackers, spammers and other cybercriminals could be reduced as much as possible.
e) Using the most powerful antivirus tools: users must use the most powerful antivirus tools with regular updates and must keep the appropriate default setting, so that the antivirus tools could work more effectively.
f) Providing suitable security tools: here, we give recommendation to the security software providers and is that: they have to offers some special tools for users that enable them to remove their accounts and to manage and control the different privacy and security issues.

2.8 Facebook’s Privacy Policy in the Law

Facebook has apparently decided to delay a proposed new privacy policy after a coalition of privacy groups asked the Federal Trade Commission on Wednesday to block the changes on the grounds that they violated a 2011 settlement with the regulatory agency.

A spokeswoman for the F.T.C. confirmed Thursday that the agency had received the letter but had no further comment.

In a statement published by The Los Angeles Times and Politico on Thursday afternoon, Facebook said, “We are taking the time to ensure that user comments are reviewed and taken into consideration to determine whether further updates are necessary and we expect to finalize the process in the coming week.”

Asked about the delay, a Facebook spokesman said he was unaware of the latest developments. When it first announced the changes on Aug. 28, Facebook told its 1.2 billion users that the updates were “to take effect on September 5.” [37]

The changes, while clarifying how Facebook uses some information about its users, also contained a shift in legal language that appeared to put the burden on users to ask Facebook not to use their personal data in advertisements. Previously, the company’s terms of use, and its settlement with the F.T.C., had indicated that it wouldn’t use such personal data without explicit consent. Facebook’s new terms would also allow it to use the names and photos of teenagers in advertising, an area of particular concern to privacy advocates because children may be unaware of the consequences of actions such as liking a brand page.

The original proposal has drawn tens of thousands of comments from Facebook users, most of them opposed to the changes [37].

Individual differences has been found to greatly impact information disclosure and perceptions of privacy. The traditional division of users according to their privacy attitudes was proposed by Hofstede (Hofstede 1996): users were classified into two extreme groups, which were reminiscent of the classic sociological concepts of individualist and collectivist societies. More recently, ethnographic studies have explored privacy attitudes of social-networking users, especially of Facebook ones. After surveying the same group of students twice, once in 2009 and later in 2010, Boyd et al. revealed that users’ confidence in changing their privacy settings strictly depended on frequency of site use and Internet literacy (Boyd and Hargittai 2010). Also, gender and cultural differences seem to affect privacy attitudes. Lewis et al. found that women’s profiles are more likely than men’s to be private and that there is a relationship between music taste and degree of disclosure of one’s profile (Lewis, Kaufman, and Christakis 2008). Finally, ethnicity plays a role. Chang et al. studied the privacy attitudes of U.S. Facebook users of different ethnicities and found that Hispanics tend to share more pictures, Asians more videos, and Afro-Americans more status updates (Chang et al. 2010). discloses personal information partly depends on one’s personality traits. In particular, it has been found to depend on the big five personality traits and the self-monitoring trait, all of which are discussed next. The five-factor model of personality, or the big five, consists of a comprehensive and reliable set of personality concepts (Costa and Mccrae 2005; Goldberg et al. 2006). The idea is that an individual can be associated with five scores that correspond to five main personality traits. Personality traits predict a number of real-world behaviors. They, for example, are strong predictors of how marriages turn out: if one of the partner is high in Neuroticism, then divorce is more likely (Nettle 2007). Research has consistently shown that people’s scores are stable over time (they do not depend on quirks and accidents of mood) and correlate well with how others close to them (e.g., friends) see them (Nettle 2007). The relationship between use of Facebook and the big five personality traits has been widely studied, yet contrasting results have emerged. To recap, the trait of Openness has been found to be associated with higher information disclosure, while Conscientiousness has been associated with cautious disclosure. Contrasting findings have emerged from the traits of Extraversion, Agreeableness, and Neuroticism. Schrammel et al. wrote: “personality traits do not seem to have any predictive power on the disclosure of information in online communities” (Schrammel, K¨offel, and Tscheligi 2009). In the same vein, Ross et al. concluded that “personality traits were not as influential as expected, and most of the predictions made from previous findings did not materialise” (Ross et al. 2009). These researchers, however, conceded that they only considered the big five personality traits (while other traits might better explain the use of Facebook) and that their samples consisted of only undergrad students. Current studies might tackle these two limitations by, for example: 1) considering personality traits other than the big five; and 2) having participants who are not only undergrad students, who might well be more Internet savvy than older people. Next, we will describe how we partly tackle the first problem by considering the personality trait of “self-monitoring”. In Section 5, we will describe how we tackle the second problem by collecting a representative sample of Facebook users in the United States. Having identified the personality traits that have been associated with information disclosure in the literature, we now need to model the process of disclosure itself. By doing so, we will be able to quantify a user's disposition to disclose her/his personal information. One measure of information disclosure is the count of fields (e.g., birthday, hometown, religion) a user shares on her/his profile. The problem of this approach is that some information fields are more revealing than others, and it is questionable to assign arbitrary (relevance) weights to those fields. An alternative way is to resort to a psychometric technique called Item Response Theory (IRT), which has already been used for modeling information disclosure in social media (Liu and Terzi 2009). Next, we detail what IRT is, why choosing it, and when and how it works. What IRT is. IRT is a psychometric technique used to design tests and build evaluation scales for those tests. It extracts patterns from participants' responses and then creates a mathematical model upon the extracted patterns, and this model can then be used to access an estimate of user attitude (de Bruin and Buchner 2010) (MacIntosh 1998) (Vishwanath 2006).

Limitations of our data. Critics might rightly put forward one important issue with our data: some users might have responded untruthfully to personality tests. However, validity tests on the personality data suggest that users have responded to the personality questions accurately and honestly, and that is likely because they installed the application motivated primarily by the prospect of taking and receiving feedback from a high quality personality questionnaire. Critics might then add that, since our sample consists of self-selected users who are interested in their personality, these users might be more active than the general Facebook population. We have looked into this matter and found that our users do not seem to be more active than average - we have reported that the number of contacts for the average user in our sample is 124 whereas Facebook reports an average of 1303. Limitation of our study. Our study has four main limitations. First, we have computed exposure scores only upon user-specified fields, while we should have ideally considered additional elements (e.g., photos, wall comments, fan page memberships, segmentation of one's social contacts into friend groups). We were not able to do so as we did not have full access to user profiles - we could access only the elements that we have been studying here. So we looked at sharing of profile fields, which is the simplest instance of sharing, making it an ideal starting point. Yet, even by considering only profile fields, we have learned that our users are far from behaving in the same way - very different privacy attitudes emerge (Figure 1). Second, we have considered Facebook users who live in the United States. Since cultural guidelines clearly exist about what is more and less private and public, one should best consider that our results are likely to hold for Facebook users in the United States. We would refrain from generalizing our results to other social-networking platforms (e.g., Twitter) or to any other society - or even to certain subcultures within the United States or within Facebook. To partly tackle this limitation, we are currently studying users of countries other than United States and platforms other than Facebook (i.e., Twitter). Third, information disclosure and concealment might be likely confounded by Facebook activity and one thus needs to control for it. We did not do so because activity information is not readily available from Facebook's API, and future studies should propose reasonable proxies for Facebook activity4. Fourth, information disclosure might be confounded by the general desire to complete one's profile: fields like "interested in" are quick to fill out, whereas "work information" just takes more effort and is not generally relevant to Facebook interactions. Yet, this reflects what the "relationship currency" of Facebook is (Nippert-Eng 2010) - that is, it reflects which fields are used to maintain relationships and which are instead concealed without causing any harm.

Novelty of Information Disclosure Model. IRT has already been used to model privacy attitudes of social-networking users (Liu and Terzi 2009). Therefore, our methodology closely followed what has been proposed before. However, since IRT has never been applied to large-scale socialnetworking data, we have run into a number of problems. First, algorithms used for estimating IRT's parameters turned out to have a greater impact on result accuracy than what one might expect.We found that best results are achieved if scores are computed in ways different than those proposed in (Liu and Terzi 2009) and similar to those proposed in psychometrics research (Baker 2001). Second, for a large number of users, errors associated with their exposure scores are unacceptably high. We had to filter those users out, yet we worked with a sample of 1,323 users. However, for studies starting off with a smaller initial sample, this problem should be addressed. Theoretical Implications. Studies that have attempted to understand how individuals conceptualize privacy offline and online have often resorted to in-depth interviews. This methodology has enabled scholars to better ground their work on exactly what people say about privacy in their daily experiences. We have proposed the use of a complementary methodology: we have studied what people do on Facebook by modeling the work of disclosure and concealment which their profiles reflect. As a result, we have quantified the extent to which personality traits affect information disclosure and concealment, and we have done so upon large-scale Facebook data that has been collected unobtrusively as the users were on the site. We have found that Openness is correlated, albeit weakly, with the amount of personal information one discloses. By contrast, after controlling for Extraversion, the self-monitoring trait's contribution disappears, suggesting that high self-monitors might present themselves in likable ways, as the literature would suggest, but they do not tend to share more private information or to make that information more visible. In addition to differentiating users based on their personality traits, we have also found important differences among profile fields and quantified the extent to which not all private fields are equally private: for example, work-related (job status) information is perceived to be more private than information on whether one is looking for a partner. Practical Implications. There are two areas in which our findings could be practically applied in the short term. The first is social media marketing. Marketing research has previously found that individuals high in Openness tend to be innovators and are more likely to influence others (Amichai- Hamburger and Vinitzky 2010). Since we have found that these individuals also tend to have less restrictive privacy settings, social media marketing campaigns would be able to identify influentials by determining which users have less restrictive privacy settings. The second area is privacy protection. Our results suggest that, by simply using age and gender, one could offer a preliminary way of personalizing default privacy settings, which users could then change. One could also imagine building privacy-protecting tools that inform users about the extent to which they are exposing information that is generally considered to be sensitive by the community. This tool could also exploit the independence assumption behind IRT to parallelize the estimation of the model's parameters and thus be able to monitor who is exposing privacy sensitive information in real-time.

2.9 Social Media Report and the Law

Social media is an increasingly important means of communicating and connecting with customersEven if you do not directly engage with social media, you should consider monitoring activity on social media that relates to your organisation. Not all interaction on social media will be positive. You should be prepared for active and sometimes critical debate with social media users. Attempting to suppress it may backfire. Garner support from the social media community and argue your case on the merits. Care must be taken over social media postings. They could lead to serious embarrassment or even be actionable. Social media is subject to the same rules as information published by other means. If you have paid or otherwise arranged for others to post on your behalf, you should make this clear. Do not try to pass these posts off as genuine user-generated content. In some cases, you may become liable for user-generated content. You should consider monitoring usergenerated content or, at the very least, putting an effective notice and take down procedure in place. If your organisation is listed, you should ensure you comply with relevant disclosure rules for inside information. In most cases, information should be disclosed via a regulatory information service and not just via social media. Those using social media to communicate on behalf of your organisation should be given very clear guidelines and training [39].

For most companies, some engagement with social media is unavoidable. Regardless of whether or not a company has an "official" social media page or account, it is very likely that it is the subject of a number of unofficial pages, its employees are interacting with social media and it is the subject of lively discussion online. For example, our review of the FTSE 100 Companies' use of social media also looked at unofficial Facebook pages, i.e. pages that do not appear to have been authorised by that company7. These have typically been set up by employees, ex-employees, customers or pressure groups and are not always complimentary. A key reason to become involved in social media is engagement. Social media provides a means for that organisation to directly interact with its customers and to obtain feedback from them. Handled properly, this can help to build a brand and create a buzz and awareness to help generate new marketing leads and attract new customers. Perhaps the most beneficial outcomes from this interaction is social validation – users who start to “organically” advocate your organisation or its aims and ideals. This “electronic word of mouth” can be a very powerful marketing and influencing tool. Engagement with users might include commenting on topical issues and the organisation using social media to put their side of the story across. One example comes from energy companies who have used their social media presence to comment on rising fuel costs and discuss the link between costs and wholesale energy prices, as well as providing information on related topics such as energy efficiency. One exercise most organisations will want to do, even if they do not directly engage with social media, is to listen to social media conversations to find out what users are saying about them. There are a number of off-the-shelf and cloud-based tools that can be used for this purpose. They not only allow monitoring of individual conversations but also monitoring on a macro-level, for example through sentiment analysis. These tools automatically analyse the 500 million tweets sent every day and report on how often a topic is mentioned and whether it is mentioned in a positive or negative context[39]. Finally, social media is an increasingly important tool for advertising. Some social media platforms are able to build detailed profiles about their users including their sex, age, location, marital status and connections. This allows advertising to be targeted precisely. One risk of social media is too much engagement. Some users are more than happy to share their views in very uncompromising terms. A thick skin is vital. For example, energy companies using social media accounts to comment on rising fuel costs have received direct and frank feedback from their customers. The same is true in financial services and many other industry sectors. Attempting to suppress critical comment may backfire. Instead you need to engage with customers and argue your case on the merits. This means your social media team need to be properly resourced. They will also need to update your social media presence to ensure it remains fresh and manages to respond to user comments. This engagement with users is a vital part of any social media strategy, either to build a distinctive brand (as is the case with Tesco Mobile) or to mollify users’ concerns (as is the case with energy companies). Another problem is embarrassing or ill-judged posts. There are numerous topical examples. For example, Ian Katz, the editor of Newsnight, sent a Tweet in September referring to *“boring snoring rachel reeves*” (Shadow Chief Secretary to the Treasury) following her appearance on the programme. He had intended to send a private message to a friend. The circumstances in which the Tweet was sent are not clear but there have certainly been other cases in which late night rants on Twitter have caused significant embarrassment[39]. Some postings may not just be stupid but also actionable. One example is the notorious Tweet by Sally Bercow: “*Why is Lord McAlpine trending? *innocent face**”. That Tweet was made when there was significant speculation about the identity of a senior unnamed politician who had been engaged in child abuse. The Tweet was false and found to be seriously defamatory11. It highlights a number of risks with social media. Firstly, the repetition rule applies such that those repeating a defamatory allegation made by someone else are treated as if they had made it themselves. This is relevant when retweeting or reposting content. Secondly, while the courts provide some leeway for the casual nature of social media and the fact that those who participate “*expect a certain amount of repartee or give and take*”12, that protection only extends so far. Civil and criminal liability for social media postings can also arise in a number of other ways. Paul Chambers was initially convicted for sending a “menacing” message13 after Tweeting: “*Crap! Robin Hood airport is closed. You’ve got a week and a bit to get your shit together, otherwise I’m blowing the airport sky high!!*”. He was fined J385 and ordered to pay J600 costs. However, this prosecution was overturned on appeal. As with more traditional formats, sales and marketing use of social media must be decent, honest and truthful. Most of the specific social media issues arise out of social validation, i.e. users who “organically” advocate that organisation or its aims and ideals. As this is such a powerful tool there is a real risk of it being misused. The Office of Fair Trading has already taken enforcement action against an operator of a commercial blogging network, Handpicked Media, requiring them to clearly identify when promotional comments have been paid for. This included publication on website blogs and microblogs, such as Twitter18. It is also important to comply with any sales and promotions rules for the particular platform being used for that promotion. In the case of Facebook, this includes prohibitions of promotions appearing on personal timelines, a complete release for Facebook by each entrant and an acknowledgement that it is not associated with Facebook [39]. Failure to comply can result in ejection from the platform. Liability can arise not only from content posted by an organisation itself, but that posted by its users. The sales and promotion rules will automatically include any content "adopted" by that organisation. What adoption means will vary depending on the circumstances, but a drinks company adopted the content of a third party website by linking to it20 and Amazon adopted a book description by including it as part of a product description21. However, organisations will not generally adopt organic-user generated reviews by simply allowing them on their sites22. An organisation might also become a data controller in respect of its personal data posted on its social media page23 or may become a publisher of that material for defamation purposes, particularly once the organisation has been notified that the material is defamatory[32]. Organisations should therefore consider if they wish to actively moderate content on their social media pages and, at the very least, should ensure that they have an effective notice and take down process to benefit from the various defences this affords.

Employees and social media

- There is no specific regulation of social media, so existing employment and data protection laws apply.
- Tell applicants if you intend to use social media for any preemployment vetting. That use should be proportionate, avoid decisionmaking on discriminatory grounds and steps should be taken to confirm the accuracy of any findings.
- There is considerable freedom for employers to dictate what constitutes acceptable use by employees through the use of an internal social media policy. It may be difficult to enforce appropriate use without such a policy.
- Social media policies should clearly state that they continue to apply to the use of social media in the employee's personal capacity, using their own computer equipment and outside of normal working hours.
- Tell employees if you intend to actively monitor their social media postings or usage. This should be included in your social media policy.
- The social media policy should be consistent with other policies and disciplinary rules.
- If any disciplinary action is taken in response to social media usage, it should follow approved procedures and be proportionate, recognising the individual's freedom of expression.

It is increasingly common for employers to review candidates' social media footprints as part of the recruitment process. An ACAS Research Paper in 201329 found that 61% of employers did so and 15% planned to start doing so in the future. There are a number of reasons why an employer would want to do this, especially for public-facing roles. A good example of this is the appointment of the 17-year old Paris Brown as Britain's first youth police and crime commissioner. After her appointment she was found to have sent a number of offensive, and potentially racist, Tweets. The subsequent media firestorm resulted in both her resignation and criticism of the Kent Police and Crime Commissioner for failing to adequately vet her appointment. However, applicants are not necessarily aware these checks are carried out and do not appear to agree with them. A 2011 ACAS Research Paper30 found that 58% of applicants surveyed would be angry, very angry or outraged if an employer refused them a job on the basis of social media research. Social media vetting also raises a range of risks for employers. Perhaps the key risk is that the employer obtains information about protected characteristics31 and an applicant subsequently claims the decision not to hire them was based on those characteristics and thus discriminatory. Privacy is also important. The Information Commissioner Employment Practices Code contains a range of general requirements for vetting of employees that are equally relevant to social media vetting. These include:

- informing applicants that social media vetting will take place. As much as anything this may encourage the applicant to clean up their social media accounts or alter their privacy settings to ensure their information is not publicly available;
- giving candidates the opportunity to comment on the accuracy of any findings. This is to mitigate the risk that some information about that individual may be inaccurate or may be about someone else with the same name;
- the search should be proportionate. Clearly, those in a prominent, publicfacing role will demand more scrutiny than those in less important roles; and
- the search should be undertaken as late in the process as possible.

For example, only at the point that the applicant is short-listed, or even conditionally appointed. Whilst not included in the Code, the Information Commissioner has also warned against gaining access to an applicant's social media profile by deception (for example, trying to be become a "friend" using a fake identity) or asking applicants for their username and password to conduct a full review of their social media account. There is anecdotal evidence of this sort of forced access in the US, and several States have legislated against it. Whilst there is no specific legislation in the UK, it is bad practice and is likely to be a breach of the Data Protection Act 1998.

Most issues have arisen where there has been damage to reputation. A typical example is Weeks v Everything Everywhere where Mr Weeks made several postings to his wall describing his place of work as "Dante's inferno" [39]. Everything Everywhere had a social media policy that expressly applied to postings in the employee's own time and included a requirement not to criticise Everything Everywhere. Mr Weeks was dismissed for gross misconduct, a decision subsequently upheld by the Employment Tribunal. However, disciplinary action against employees must be in accordance with established disciplinary policies and procedures. The employer must act fairly and the response must be one which a reasonable employer could have made. One implication of this is that the employer must consider actual impact on business rather than assumed or feared impact. Social media also provides a medium for online bullying and harassment. This could take a number of forms, including the posting of offensive photos or comments as well as the risk of social exclusion. This could lead to claims for claims for discrimination or constructive or unfair dismissal as employers are vicariously liable for the acts of one employee to another in the course of their employment. Again, it is important that social media and bullying policies are updated to clearly set out what sort of behaviour is acceptable and extend their scope to cover cyber-bullying outside of the workplace. However, any subsequent action against the employee must reflect the seriousness of the alleged behaviour. Excessive usage of social media by employees can lead to a loss of productivity and overburden the employer's computer systems. So this is one area in which the fact postings are made during normal working hours, or using the employer's computer systems, is relevant. Some employers have responded to this issue by blocking access to social media sites at work though this may be unpopular and does not prevent employees from using social media on their smartphones. Alternatively, employers might want to monitor their employee's use of social media, though that will need to comply with data protection laws as with any other employee monitoring (see below) and employees should be given clear guidance about what constitutes excessive use. These issues are potentially complicated by the overlap with human rights law including the right to privacy and freedom of speech, protected in the European Union by both the European Convention of Human Rights and the Charter of Fundamental Rights of the European Union. These rights might be relied upon by employees to claim that their social media postings are private and therefore should not be subject to their employer's disciplinary policy. Many social media accounts can be set up so that posting and other information are only available to that person's "friends" and, indeed, in many cases, the employer only becomes aware of the offending posting when the "friend" reports it to them. However, these arguments have fairly limited success. Equally, whilst most cases seem to arise from "friends" notifying the employer of offending content, if an employer wants to actively monitor its employees' social media postings it should respect their right to privacy and comply with data protection laws. Whilst there is no direct guidance from the Information Commissioner, this is likely to be subject to Part 3 of his Employment Practices Guide: Monitoring at work. This suggests, amongst other things, notifying employees through an appropriate policy and carrying out a privacy impact assessment. Finally, an employee might also argue that restrictions on his use of social media infringe his right to freedom of expression. This right was considered in Smith v Trafford House Trust [2012] EWHC 3221. Mr Smith was disciplined for setting out his negative views on the proposal to introduce gay marriage in the UK via his Facebook account. His comments, which were clear, reasoned and unaggressive in nature, caused upset to a fellow employee who was also a "friend" on Facebook. When the "friend" drew the employer's attention to Mr Smith's comments on gay marriage the employer took disciplinary action against him on the basis that the employer was a housing trust which included gay people among its clients. It therefore considered that his statements were inappropriate and that disciplinary action was justified as they breached its Code of Conduct that stated that employees should not "*promote their political or religious views*".

Corporate social media accounts Ownership of corporate social media accounts is relatively straightforward. The starting point is to identify what accounts your organisation currently owns and what accounts you want to own. Acquiring new accounts for more popular social media sites may give rise to "name-squatting" problems, similar to those that arise in relation to domain names. Whilst most social media sites have express squatting policies40, that does not help where there is genuine conflicting use. Certainly, if you are one of the many people to congratulate John Lewis on its recent #bareandhare Christmas adverts and you sent your Tweet to the handle @johnlewis, you might be surprised to get a response from the computer science professor, John Lewis of Blacksburg, Virginia, USA. It's also sensible to control passwords to those accounts and ensure they are not held by one person alone. This will help to manage the risk of an employee leaving with control of that corporate account or deliberately sending unauthorised messages. For example, this could have avoided mild embarrassment at the insolvent retailer HMV after its social media planner used its Twitter account to provide real-time updates on the dismissal of its staff. The Courts are also likely to be sympathetic to claims that a corporate account has been misused or misappropriated. For example, a company successfully obtained an injunction preventing its ex-employees from using a corporate LinkedIn group in a competing business[35].

Personal social media accounts Employee social media accounts are more difficult to deal with. The idea that a company "owns" its employees' social media accounts is conceptually difficult, as its contents, connections and interactions are normally personal to that employee. Quite apart from this conceptual difficulty of corporate ownership of personal social media accounts, there are also difficulties in identifying what legal rights exist in those accounts and therefore in protecting them in a meaningful way. An analysis of one professional social network (overleaf), LinkedIn, illustrates these difficulties.

Is social media compatible with privacy? New technology has challenged traditional concepts of privacy for well over a century. Samuel Warren and Louis Brandeis' seminal 1890 paper on "*The Right to Privacy*" grappled with the prospect of "*numerous mechanical devices*" and "*instantaneous photographs*" creating a world in which "*what is whispered in the closet shall be proclaimed from the house-tops*". Social media, smartphones and other wearable technology, such as Google Glass, has brought this threat to life. In the intervening hundred years, the law has evolved to provide generalised rights to privacy or specific data protection laws or both. New technologies are also redefining social attitudes to privacy. Many users disclose significant amounts of personal information about themselves on social media. Indeed, for many, the very purpose of social media is to provide an endless stream of information about themselves from the trivial, to the intimate to the tragic. However, and perhaps counterintuitively, users remain very concerned about their privacy and want to keep tight control of their information.

Big Data Social media generates huge volumes of information. Facebook alone generates 500 terabytes of data a day, including 2.7 billion new likes and 300 million new photos. This is fertile ground for Big Data analysis. To the extent that this involves personal information, it will be subject to privacy and data protection legislation, a question European privacy regulators grappled with earlier this year in its Opinion on purpose limitation45. For the regulators the key distinction is whether the analysis is just intended to detect general trends and correlations (for example, sentiment analysis) or is intended to support measures in respect of an individual. Unsurprisingly, the former is unlikely to be objectionable so long as there are proper safeguards in place. The regulators stress the need for “functional separation” such that the output of this analysis cannot be linked back to an individual. In contrast, if the analysis could be used to support measures in respect of an individual, then greater care will be needed. The regulators have an antipathy for profiling, e.g. direct marketing, behavioural advertisement, data-brokering, location-based advertising or trackingbased digital market research, and suggest it would “almost always” require specific, informed and unambiguous consent. The legitimacy of other uses will depend on the circumstances but, to a large degree, will depend on whether the new Big Data analysis is compatible with the purpose of the original social media posting.

Future of data protection regulation The European Union intends to deal with many of the challenges raised by social media through its proposed General Data Protection Regulation. A draft of the regulation was issued by the European Commission in January 2012 and is now being debated by the European Parliament and European Council, with the European Parliament voting through its draft of the regulation in October 2013. The regulation contains a number of provisions that are relevant to social media. For example, it contains restrictions on “profiling” that are likely to require consent for many types of profiling, mirroring the position already advocated by many European regulators. It also contains a “right to be forgotten”. This provides enhanced rights to ask that personal data be deleted. It is intended to deal with the problem that the internet may reveal information about individuals that is unfair, out of date or just plain wrong. However, this right is nuanced and is subject to a number of carve outs, such as where it would conflict with another person’s freedom of expression. This will make it difficult to apply in practice. For example, while it should be easier for an individual to remove material they have posted about themselves, forcing someone else to remove information they have posted about the individual will involve a harder tussle between competing fundamental rights.

2.10 Social Media Monitoring

Political communication has become a major focus in the growing field of social media studies. Re-searchers across disciplines and across the globe analyze political online communication with a spe-cific focus on elections – particularly since election campaigns increasingly take place in social media, a process that was prominently recognized during the US election campaign in 2008.

Our goal is to examine various aspects of the communication structures in online media and how such data can add new insights in comparison to existing data from surveys and (traditional) media anal-yses. By collecting data from both Twitter and Facebook we also add new dimensions to current stud-ies on social media use during elections. Our approach is situated in the broader framework of the German Longitudinal Election Study (GLES), a long term research project that examines the German federal elections in 2009, 2013, and 2017 with the aim to track the German electoral process over an extended period of time (Schmitt-Beck, Rattinger, Roßteutscher & Weßels, 2010). By collecting Twit-ter and Facebook data about the German Bundestag elections we can supplement traditional research data as used in the GLES. Specifically, the candidate study of the GLES (based on interviews with candidates) can be compared to the actual behavior of candidates on Twitter and Facebook. Similarly, the media corpus analysis of the GLES (analyzing different mass media channels) may be compared with the social media corpus.

The role of online communication and internet technology in German politics has also been studied from different perspectives (Albrecht & Schweizer, 2011; Jungherr & Schoen, 2013; Meckel et al., 2012). The German federal government structure allows to compare elections across the different states, like Elter (2013) has done for seven different German federal state elections. The project “Po-litical Deliberation on the Internet: Forms and Functions of Digital Discourse Based on the Mi-croblogging System Twitter” also monitors several regional as well as the various state elections and analyzes the broader impact of Twitter on political debates in Germany (Thimm, Einspänner & Dang-Anh, 2012). Siri and Seßler (2013) as well as Thamm and

Bleier (2013) focus on a set of politicians rather than on selected events like elections. Dang-Xuan et al. (2013) combine the two dimensions and take a closer look at influential individuals during an electoral event in order to investigate emotionali-ty and discussed topics. There are a number of ongoing projects collecting tweets around the 2013 German Bundestag election and some new publications can be expected in the near future.

Previous research has also been inspired by the challenge to use social media to predict election re-sults (e.g., Birmingham & Smeaton, 2011; Tumasjan et al., 2011) which has resulted in a considerable amount of skepticism and some counter examples (e.g., Jungherr, Jürgens & Schoen, 2012; Metaxas, Mustafaraj & Gayo-Avello, 2011). Predictions are a particular case that shows how selected methods and applied modes for data collection (e.g., based upon keywords vs. users, selection of time span) influence the potential outcome of a study. Much more research is conducted, however, not to predict election outcomes but to investigate the roles of politicians, media and publics from various perspec-tives, for example, by focusing on deliberation and participation. In all cases, however, the chosen methods highly influence what types of conclusions can be drawn. Current approaches comprise quan-titative analyses (e.g., number of interactions, network analyses), qualitative analyses (e.g., content analysis of posts) and combined methods – some of them automated, others carried out manually. In all approaches the modes of data collection also have an effect on the scope and limits of the study; if data collection from Twitter is, for example, based on one single hashtag, one needs to be aware that parts of the conversation are lost, as the same topic is likely to be discussed under the heading of different hashtags and not every user includes the same hashtag despite the fact that he or she is referring to the same discussion or topic [40].

Before actual social media data can be collected, researchers need to decide about the scope of the data corpus. Therefore, we had to construct a list of names of the relevant candidates. This list was the starting point for our search of the social media accounts for both corpus 1 and 2. Relevance was de-fined as the reasonable likelihood of becoming a member of the Bundestag (see appendix for more details). We refer to this list as the list of candidates although the complete number of overall candi-dates was higher. The data was collected in a two-stage process.

If the candidate had an account which he or she used for private purposes in addition to his profes-sional account3, only the professional account was included in our list. During our search for the ac-counts, this problem occurred primarily with Facebook accounts. Since a list of candidates of the 2009 Bundestag election was already available from the 2009 GLES candidate study, we also searched Facebook accounts for these candidates.

Since Twitter is a fast moving medium which takes up and redistributes new information quickly, it is likely that conventional media also use Twitter as a data source. We assume that conventional media select information from Twitter and refine and redistribute the topics over the more conventional media. Corpus 3 was designed to reflect this. We refer to the individuals who would follow such an information gathering approach as “gatekeepers” and searched for them among journalists and editors. In a first step, we identified journalists and editors working in internal political divisions of national daily newspapers and magazines (see appendix) and searched their Twitter accounts. The leading principle in selecting the media sources was whether they were included in the print media content analysis of GLES. The result of this first step is a list of all Twitter gatekeepers of conventional media.

In a second step, we retrieved all accounts that the gatekeepers followed. The assumption behind this approach is that the gatekeepers themselves track what we call “information authorities”. The infor-mation authorities push topics into Twitter and it is likely that they play a central role in shaping the agenda on Twitter. In order to be counted in the list of information authorities we introduced the crite-rion that at least 25 percent of the gatekeepers have to follow the account. The list is extended by accounts which are followed by at least 25 percent of the journalists or 25 percent of the editors.

Applying the list of candidate names which have an active professional Twitter account in the 2013 elections we used the Twitter streaming API4 to receive messages directly from these candidates as well as the retweets of and replies to their messages. We also collected the @messages/mentions and messages which included a hashtag from our lists. For that purpose we developed a software compo-nent called TweetObserver that is instantly reading the stream from Twitter resulting from our query in a stable manner. The software needs to register as a Twitter application in order to continuously receive update events for the requested items from the Twitter service. For each account the search query includes the account ID and the name, so that the application is geared towards receiving tweets from a certain account as well as any mentioning of its name. The software was implemented in Java and relied on the Twitter library twitter4j5. The software is connected to a MongoDB6 in which we store the data in JSON format. In the following we describe the data structure of the tweets in the Twitter data set.

Although various research projects are currently using social media data, and particularly data collect-ed from Twitter, almost no datasets are available for secondary analysis and replication. This is partly due to the complicated legal environment in which social media is situated. The terms of use of com-panies generating social media data especially add to this effect. In terms of sustainability and verifia-bility of research it seems desirable, however, to find feasible solutions for long-term archiving and sharing of social media datasets.

Consequently, we outline requirements and conditions for archiving and distribution of social media datasets. For GESIS this means assessing the scope and requirements for archiving a social media dataset and identifying the differences to survey data which the GESIS data archive is used to dealing with and for which we can rely on well-developed tools for archiving, documentation and modes of distribution.

In the following we outline three important areas in which the collected data pose challenges for ar-chiving and distribution.

Documentation, data structure, tools: Social media research so far is still in its beginnings and conse-quently lacks certain standards and methodologies. This applies to both data collection and data analy-sis. In order to enable re-use of archived datasets, one needs a good description of the available data and the data collection process has to be carefully documented, including the selected tools for gather-ing the data, data cleaning strategies and data formats. We expect that this document helps to under-stand what we collected and how we collected our data. We hope that it enables other researchers to compare our approach with others to understand the differences with other datasets that have recently been collected by other projects researching the German election through Twitter or Facebook.

2. XKEYSCORE TECHNOLOGY

3.1 Xkeyscore

XKeyscore or XKEYSCORE (abbreviated as XKS) is a formerly secret computer system first used by the United States National Security Agency for searching and analyzing Internet data it collects worldwide every day. The program has been shared with other spy agencies including Australia's Defence Signals Directorate, New Zealand's Government Communications Security Bureau and the German Bundesnachrichtendienst.

The program's existence was publicly revealed in July 2013 by Edward Snowden in The Sydney Morning Herald and O Globo newspapers, though the codename is mentioned in earlier articles, and like many other codenames can also be seen in job postings, and in the online resumes of employees.

XKeyscore is a complicated system and various authors have different interpretations about its actual capabilities. Edward Snowden and Glenn Greenwald explained XKeyscore as being a system which enables almost unlimited surveillance of anyone anywhere in the world, while NSA said that usage of the system is limited and restricted.

According to The Washington Post and national security reporter Marc Ambinder, XKeyscore is an NSA data-retrieval system which consists of a series of user interfaces, backend databases, servers and software that selects certain types of data and metadata that the NSA has already collected using other methods.

An NSA presentation about XKeyscore from 2008 says that it's a "DNI Exploitation System/Analytic Framework". DNI stands for Digital Network Intelligence, which means intelligence derived from internet traffic. In an interview with the German Norddeutscher Rundfunk, Edward Snowden said about XKeyscore: "It’s a front end search engine" [41].

Data sources. XKeyscore consists of over 700 servers at approximately 150 sites where the NSA collects data, like "US and allied military and other facilities as well as US embassies and consulates" in many countries around the world. Among the facilities involved in the program are four bases in Australia and one in New Zealand.

According to an NSA presentation from 2008, these XKeyscore servers are fed with data from the following collection systems:

- F6 (Special Collection Service) – joint operation of the CIA and NSA that carries out clandestine operations including espionage on foreign diplomats and leaders.
- FORNSAT – which stands for "foreign satellite collection", and refers to intercepts from satellites.
- SSO (Special Source Operations) – a division of the NSA that cooperates with telecommunication providers.

In a single, undated slide published by Swedish media in December 2013, the following additional data sources for XKeyscore are mentioned:

- Overhead – intelligence derived from American spy planes, drones and satellites.
- Tailored Access Operations – a division of the NSA that deals with hacking and cyberwarfare.
- FISA – all types of surveillance approved by the Foreign Intelligence Surveillance Court.
- Third party – foreign partners of the NSA such as the (signals) intelligence agencies of Belgium, Denmark, France, Germany, Italy, Japan, the Netherlands, Norway, Sweden, etc.

From these sources, XKeyscore stores "full-take data", which are indexed by plug-ins that extract certain types of metadata (like phone numbers, e-mail addresses, log-ins, and user activity) and index them in metadata tables, which can be queried by analysts. XKeyscore has been integrated with MARINA, which is NSA's database for internet metadata.

However, the system continuously gets so much Internet data that it can be stored only for short periods of time. Content data remain on the system for only three to five days, while metadata is stored for up to 30 days.[41] A detailed commentary on an NSA presentation published in The Guardian in July 2013 cites a document published in 2008 declaring that "At some sites, the amount of data we receive per day (20+ terabytes) can only be stored for as little as 24 hours."

Capabilities

For analysts, XKeyscore provides a "series of viewers for common data types", which allows them to query terabytes of raw data gathered at the aforementioned collection sites. This enables them to find targets that cannot be found by searching only the metadata, and also to do this against data sets that otherwise would have been dropped by the front-end data processing systems. According to a slide from an XKeyscore presentation, NSA collection sites select and forward less than 5% of the internet traffic to the PINWALE database for internet content.

Because XKeyscore holds raw and unselected communications traffic, analysts can not only perform queries using "strong selectors" like e-mail addresses, but also using "soft selectors", like keywords, against the body texts of e-mail and chat messages and digital documents and spreadsheets in English, Arabic and Chinese.

This is useful because "a large amount of time spent on the web is performing actions that are anonymous" and therefore those activities can't be found by just looking for e-mail addresses of a target. When content has been found, the analyst might be able to find new intelligence or a strong selector, which can then be used for starting a traditional search[42].

Besides using soft selectors, analysts can also use the following other XKeyscore capabilities:

- Look for the usage of Google Maps and terms entered into a search engine by known targets looking for suspicious things or places.
- Look for "anomalies" without any specific person attached, like detecting the nationality of foreigners by analyzing the language used within intercepted emails. An example would be a German speaker in Pakistan. The Brazilian paper O Globo claims that this has been applied to Latin America and specifically to Colombia, Ecuador, Mexico and Venezuela.[43]
- Detect people who use encryption by do searches like "all PGP usage in Iran". The caveat given is that very broad queries can result in too much data to transmit back to the analyst.
- Showing the usage of Virtual private networks (VPNs) and machines that can potentially be hacked via TAO.
- Track the source and authorship of a document that has passed through many hands.
- Most of these things cannot be detected by other NSA tools because they operate with strong selectors (like e-mail and IP addresses and phone numbers) and the raw data volumes are too high to forward them to other NSA databases.[42]

In 2008, it was planned to add a number of new capabilities in the future, like:

- VoIP
- More networking protocols[clarify]
- Exif tags, which often include geolocation (GPS) data.

The NSA slides published in The Guardian during 2013 claimed that XKeyscore had played a role in capturing 300 terrorists by 2008 [41]. This claim could not be substantiated as the redacted documents do not cite instances of terrorist interventions.

A 2011 report from the NSA unit in Griesheim (Germany) says that XKeyscore made it easier and more efficient to target surveillance. Previously, analysis often accessed data they were not interested in. XKeyscore allowed them to focus on the intended topics, while ignoring unrelated data. XKeyscore also proved to be an outstanding tool for tracking active groups associated with the Anonymous movement in Germany, because it allows for searching on patterns, rather than particular individuals. An analyst is able to determine when targets research new topics, or develop new behaviors[42].

To create additional motivation, the NSA incorporated various features from computer games into the program. For instance, analysts who were especially good at using XKeyscore could acquire "skilz" points and "unlock achievements." The training units in Griesheim were apparently successful and analysts there had achieved the "highest average of skilz points" compared with all other NSA departments participating in the training program.

A top secret National Security Agency program allows analysts to search with no prior authorization through vast databases containing emails, online chats and the browsing histories of millions of individuals, according to documents provided by whistleblower Edward Snowden. The NSA boasts in training materials that the program, called XKeyscore, is its "widestreaching" system for developing intelligence from the internet. The latest revelations will add to the intense public and congressional debate around the extent of NSA surveillance programs. They come as senior intelligence officials testify to the Senate judiciary committee on Wednesday, releasing classified documents in response to the Guardian's earlier stories on bulk collection of phone records and Fisa surveillance court oversight.

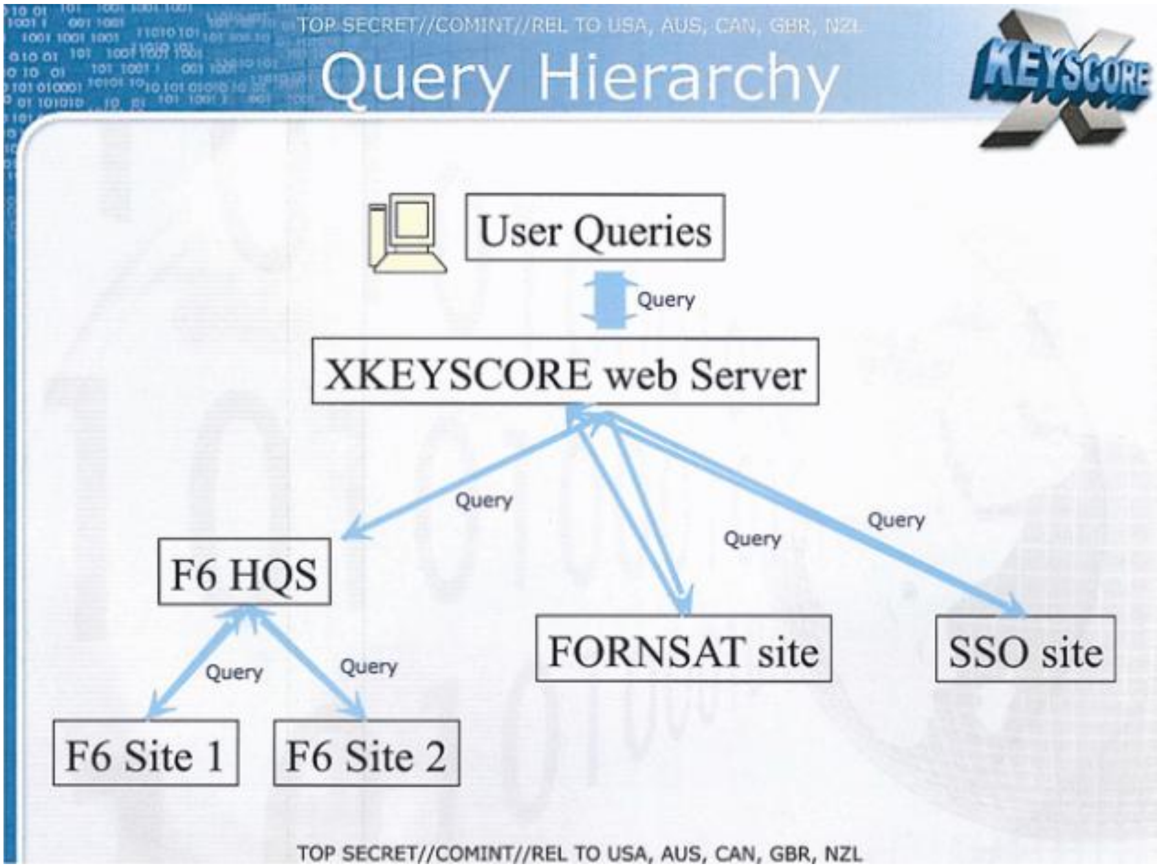
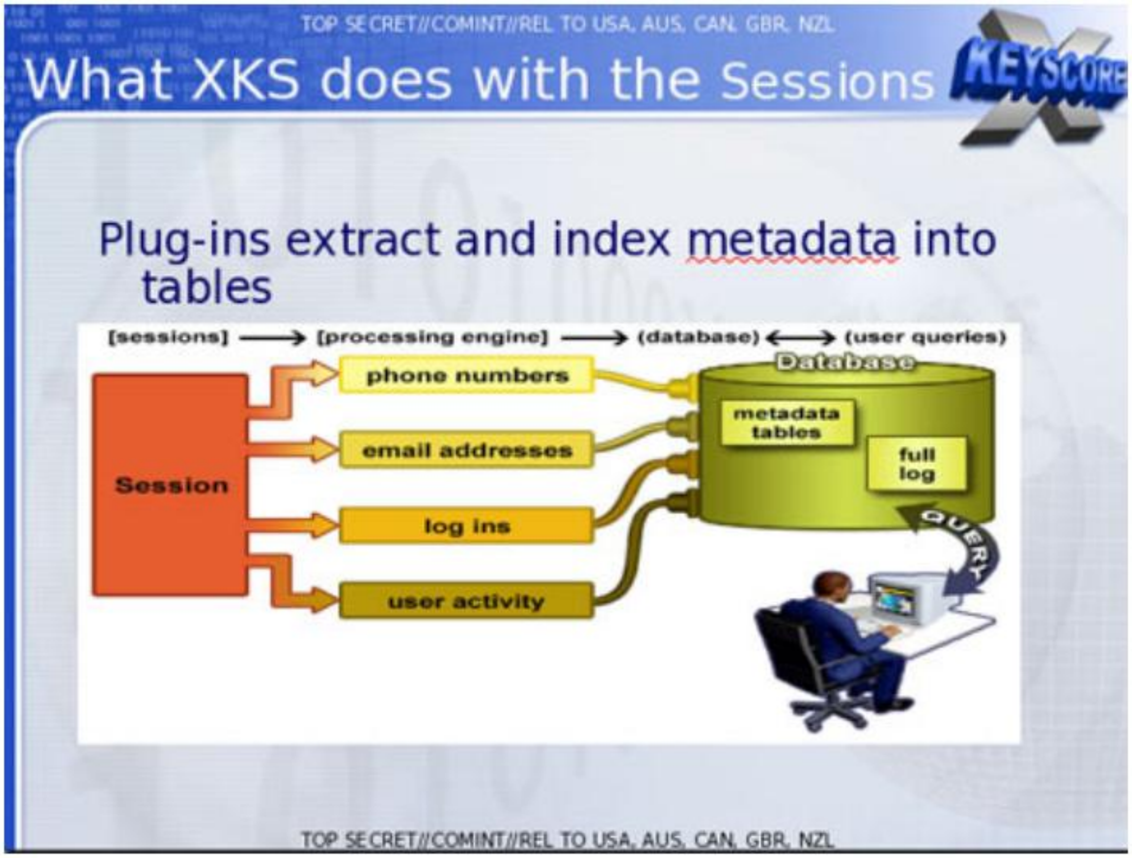


FIGURE 2: The Query Hierarchy [41].

But training materials for XKeyscore detail how analysts can use it and other systems to mine enormous agency databases by filling in a simple on-screen form giving only a broad justification for the search. The request is not reviewed by a court or any NSA personnel before it is processed. Analysts can also use XKeyscore and other NSA systems to obtain ongoing "real-time" interception of an individual's internet activity. Under US law, the NSA is required to obtain an individualized Fisa warrant only if the target of their surveillance is a 'US person', though no such warrant is required for intercepting the communications of Americans with foreign targets. But XKeyscore provides the technological capability, if not the legal authority, to target even US persons for extensive electronic surveillance without a warrant provided that some identifying information, such as their email or In a second Guardian interview in June, Snowden elaborated on his statement about being able to read any individual's email if he had their email address. He said the claim was based in part on the email search capabilities of XKeyscore, which Snowden says he was authorized to use while working as a Booz Allen contractor for the NSA. One top-secret document describes how the program "searches within bodies of emails, webpages and documents", including the "To, From, CC, BCC lines" and the 'Contact Us' pages on websites". To search for emails, an analyst using XKS enters the individual's email address into a simple online search form, along with the "justification" for the search and the time period for which the emails are sought. The analyst then selects which of those returned emails they want to read by opening them in NSA reading software. The system is similar to the way in which NSA analysts generally can intercept the communications of anyone they select, including, as one NSA document put it, "communications that transit the United States and communications that terminate in the United States". The quantity of communications accessible through programs such as XKeyscore is staggeringly large. One NSA report from 2007 estimated that there were 850bn "call events" collected and stored in the NSA databases, and close to 150bn internet records. Each day, the document says, 1-2bn records were added. William Binney, a former NSA mathematician, said last year that the agency had "assembled on the order of 20tn transactions about US citizens with other US citizens", an estimate, he said, that "only was involving phone calls and emails". A 2010 Washington Post article reported that "every day, collection systems at the [NSA] intercept and store 1.7bn emails, phone calls and other type of communications." The XKeyscore system is continuously collecting so much internet data that it can be stored only for short periods of time. Content remains on the system for only three to five days, while metadata is stored for 30 days. One document explains: "At some sites, the amount of data we receive per day (20+ terabytes). To solve this problem, the NSA has created a multi-tiered system that allows analysts to store "interesting" content in other databases, such as one named Pinwale which can store material for up to five years[41]. It is the databases of XKeyscore, one document shows, that now contain the greatest amount of communications data collected by the NSA can only be stored for as little as 24 hours." IP address, is known to the analyst (Fig.3).



TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

What XKS does with the Sessions

Plug-ins extract and index metadata into tables

TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

FIGURE 3: XKS and Sessions [41].

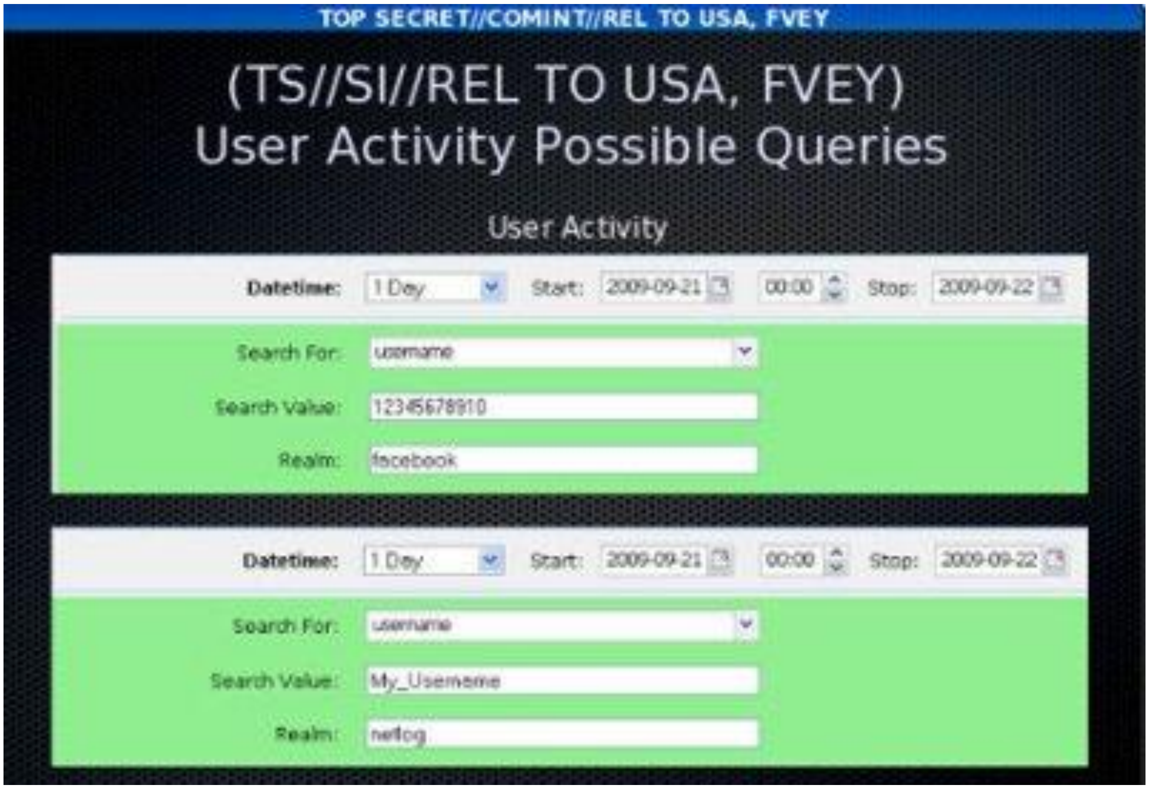
Legal v technical restrictions. While the Fisa Amendments Act of 2008 requires an individualized warrant for the targeting of US persons, NSA analysts are permitted to intercept the communications of such individuals without a warrant if they are in contact with one of the NSA's foreign targets. The ACLU's deputy legal director, Jameel Jaffer, told the Guardian last month that national security officials expressly said that a primary purpose of the new law was to enable them to collect large amounts of Americans' communications without individualized warrants.

Additional food for thought — that many have been seriously concerned about lately as well — is this all-powerful digital dragnet may explain why generals to senators to candidates have been compromised. It may well explain why votes and positions have become so erratic and mis-aligned with expectations. From McCain to Roberts to Petraeus — if you are on the wrong side of the power curve, you lose.

Regardless your political persuasion, your ideological opposite could be the current or next POTUS, or in any other position of power and use this ultimate spy power to destroy you and your position. Political donors can be easily destroyed. Operatives destroyed. No person, no entity, no one is safe as long as this power exists and is unchecked.

Therefore, our Republic hangs in the balance controlled by the few and the powerful, which means — we really do not have a Republic any longer — regardless the perception that our Republic remains. We, in fact, can only have a de-facto totalitarian oligarchy as long as this power remains.

Finally, by definition of my beliefs and convictions, the fact that the aforementioned is true means I am a de-facto enemy of this oligarchy as I want to remove its power — its ability to completely control. What does this say for my future as long as this power exists? What does it say for your future if you are a true American? It says the days of any true Americans — who claim their birthright of Constitutionally protected individual liberty, privacy and rights to due process — are numbered. And, it is only a matter of time.



TOP SECRET//COMINT//REL TO USA, FVEY

(TS//SI//REL TO USA, FVEY)
User Activity Possible Queries

User Activity

Datetime: 1 Day Start: 2009-09-21 00:00 Stop: 2009-09-22

Search For: username

Search Value: 12345678910

Realm: facebook

Datetime: 1 Day Start: 2009-09-21 00:00 Stop: 2009-09-22

Search For: username

Search Value: My_Username

Realm: netlog

FIGURE 4. Facebook Spying [41].

And, of course, the NSA is pushing back. Yet, their denial and later admission, denial and later admission and apology pattern, is not stellar. We must push for clear and complete oversight down to the gnat’s eyelash on any of NSA mining capabilities. Else, all trust in government is gone.

Knowing what I know, I highly doubt the complete veracity of these agencies in their responses. They are CYA oriented and have inordinate power. Without a clear, citizen-verified adherence to our Constitutional protections, I will have no confidence:

‘The implication that NSA’s collection is arbitrary and unconstrained is false,’ the agency said. ‘NSA’s activities are focused and specifically deployed against — and only against — legitimate foreign intelligence targets in response to requirements that our leaders need for information necessary to protect our nation and its interests.’ The agency said those with access to the system are trained on their ‘ethical and legal obligations.’ The agency complained that the ongoing leaks continue to jeopardize security. The statement said the programs as a whole have helped defend the nation, and that as of 2008, ‘there were over 300 terrorists captured using intelligence generated from XKEYSCORE. “[44].

“We will keep you safe”, said Hitler, Lenin, Mao, and now, our U.S. government. On the contrary, here in America, it is “We the People” who keep ourselves safe. First, with our 2nd amendment protected natural rights to self-defense and formation of state militias to repel government tyranny. Second, we keep ourselves safe with the volunteering of our own brave and courageous men and women, sons and daughters to serve in our armed forces only to fight real aggressors against our country. And, third, with those we elect to serve in public office and the agencies of public office, to uphold the line and letter of the U.S. Constitution that, first and foremost, protects each individual’s liberties and natural rights, and then, within its constrained war powers, authorizes the legitimate repelling of real aggressors against us.

No, we as a free society do not need the ‘government’s’ protection. As long as we are a free people, we never will. Unlike any other nation in history, our government serves at our will. We do, however, need those who would betray our only purpose as a nation — stepping beyond the boundaries of our individual liberties and natural rights — to be prosecuted as traitors to ‘We the People’ of America. “Thanks, but no thanks, for your offer,” say We the People, “we will keep ourselves safe.”

3.2 Case to Think: Mr. Snowden Interview Speaking about Xkeyscore

Main concepts of Snowden’s Interview are [45]:

- The Five Eyes alliance is sort of an artifact of the post World War II era where the Anglophone countries are the major powers banded together to sort of co-operate and share the costs of intelligence gathering infrastructure. So we have the UK’s GCHQ, we have the US NSA, we have Canada’s C-Sec, we have the Australian Signals Intelligence Directorate and we have New Zealand’s DSD. What the result of this was over decades and decades what sort of a supra-national intelligence organisation that doesn’t answer to the laws of its own countries.
- If you ask the governments about this directly they would deny it and point to policy agreements between the members of the Five Eyes saying that they won’t spy on each other’s citizens but there are a couple of key points there. One is that the way they define spying is not the collection of data. The GCHQ is collecting an incredible amount of data on British Citizens just as the National Security Agency is gathering enormous amounts of data on US citizens. What they are saying is that they will not then target people within that data. They won’t look for UK citizens or British citizens. In addition the policy agreements between them that say British won’t target US citizens, US won’t target British citizens are not legally binding. The actual memorandums of agreement state specifically on that that they are not intended to put legal restriction on any government. They are policy agreements that can be deviated from or broken at any time. So if they want to on a British citizen they can spy on a British citizen and then they can even share that data with the British government that is itself forbidden from spying on UK citizens. So there is a sort of a trading dynamic there but it’s not, it’s not open, it’s more of a nudge and wink and beyond that the key is to remember the surveillance and the abuse doesn’t occur when people look at the data it occurs when people gather the data in the first place.

- You could read anyone's email in the world. Anybody you've got email address for, any website you can watch traffic to and from it, any computer that an individual sits at you can watch it, any laptop that you're tracking you can follow it as it moves from place to place throughout the world. It's a one stop shop for access to the NSA's information. And what's more you can tag individuals using "XKeyscore". Let's say you saw you once and I thought what you were doing was interesting or you just have access that's interesting to me, let's say you work at a major German corporation and you want access to that network, you can track your username on a website on a form somewhere, you can track your real name, you can track associations with your friends and you can build what's called a fingerprint which is network activity unique to you which means anywhere you go in the world anywhere you try to sort of hide your online presence hide your identity, the NSA can find you and anyone who's allowed to use this or who the NSA shares their software with can do the same thing. Germany is one of the countries that have access to "XKeyscore".
- The contracting culture of the national security community in the United States is a complex topic. It's driven by a number of interests between primarily limiting the number of direct government employees at the same time as keeping lobbying groups in Congress typically from very well funded businesses such as Booze Allen Hamilton. The problem there is you end up in a situation where government policies are being influenced by private corporations who have interests that are completely divorced from the public good in mind. The result of that is what we saw at Booze Allen Hamilton where you have private individuals who have access to what the government alleges were millions and millions of records that they could walk out the door with at any time with no accountability, no oversight, no auditing, the government didn't even know they were gone.
- The NSA goes where the data is. If the NSA can pull text messages out of telecommunication networks in China, they can probably manage to get facebook messages out of Germany. Ultimately the solution to that is not to try to stick everything in a walled garden. Although that does raise the level of sophistication and complexity of taking the information. It's also much better simply to secure the information internationally against everyone rather than playing "let's move the data". Moving the data isn't fixing the problem. Securing the data is the problem.
- It's becoming increasingly clear that these leaks didn't cause harm in fact they served the public good. Because of that it will be very difficult to maintain sort of an ongoing campaign of persecution against someone who the public agrees serve the public interest.

3.3 US surveillance programs and their impact on EU citizens' fundamental rights

A careful analysis of US privacy laws compared to the EU Data Protection framework shows that the former allows few practical options for the individual to live their lives with selfdetermination over their personal data. However a core effect of Data Protection law is that if data is copied from one computer to another, then providing the right legal conditions for transfer exist, the individual cannot object on the grounds that their privacy risk increases through every such proliferation of "their" data5. This holds true if the data is copied onto a thousand machines in one organization, or spread onward to a thousand organisations, or to a different legal regime in a Third Country. The individual cannot stop this once they lose possession of their data, whereas for example if the data was "intellectual property", then a license to reproduce the data would be necessary by permission. We are all the authors of our lives, and it seems increasingly anomalous that Internet companies lay claim to property rights in the patterns of data minutely recording our thoughts and behaviour, yet ask the people who produce this data to sacrifice their autonomy and take privacy on trust. The EU Data Protection framework in theory is categorically better than the US for privacy, but in practice it is hard to find any real-world Internet services that implement DP principles by design, conveniently and securely. Privacy governance around the world has evolved around two competing models. Europe made some rights of individuals inalienable and assigned responsibilities to Data Controller organizations, whereas in the United States companies inserted waivers of rights into Terms and Conditions contracts allowing exploitation of data in exhaustive ways (known as the 'Notice-and-Choice" principle). The PRISM crisis arose directly from the emerging dominance over the last decade of "free" services operated from remote warehouses full of computer servers, by companies predominantly based in US jurisdiction, that has become known as Cloud computing. To explain this relationship we must explore details of the US framework of national security law[51]. After the terrorist attacks of September 11th 2001, privacy and data protection has been deeply challenged by exceptional measures taken in the name of security and the fight against terrorism. The USA PATRIOT Act of 2001 was enacted by the US Congress on October 26, 2001, and its primary effect was to greatly extend law enforcement agencies' powers for gathering domestic intelligence inside the US. The revised Foreign Intelligence Surveillance Amendment Act of 2008 (FAA) created a power of mass-surveillance specifically targeted at the data of non-US persons located outside the US. Numerous new surveillance programmes and modalities were further suggested to President Bush by NSA Director Gen. Hayden, without explicit authorization under statute, and approval was nevertheless given. Those programmes were retroactively deemed lawful in secret memoranda prepared by a relatively junior legal14 official, under the Authorisation to Use Military Force (AUMF) for the war in Afghanistan and associated War on Terror operations. Amongst these programmes was one codenamed Stellar Wind which involved placing fibreoptic cable "splitters" in major Internet switching centres, and triaging the enormous volumes of traffic in real-time with a small high-performance scanning computer (known as a deep-packet inspection box), which would send data filtered by this means back to the NSA. An AT&T technical supervisor in the San Francisco office was asked to assist in constructing such a facility ("Room 641A") and was concerned that this activity manifestly broke US Constitutional protections, because the cable carried domestic as well as international traffic. He took his story with documentation to the New York Times, which did not publish15 the story for a year, until 2005 after the re-election of President Bush [48]. The complexity of inter-related US legislation pertaining to 'foreign intelligence information', and its interpretations by secret courts and executive legal memoranda, has led to unlawful practices affecting both US citizens and non-US citizens. The consequences of this legal uncertainty, and lack of Fourth Amendment protection for non-US citizens, means that no privacy rights for non-Americans are recognized by the US authorities under FISA The accelerating and already widespread use of Cloud Computing further undermines data protection for EU citizens. A review of the mechanisms that have been put in place in the EU for data export to protect EU citizens' rights shows that they actually function as loopholes. entrusted to, or necessary to use a service provided by, a "third party" such as a bank or telephone company, there was no reasonable expectation of privacy, and therefore no warrant was required by the Fourth Amendment, which protects privacy against unreasonable searches without a particular warrant, issued on "probable cause" (meaning evidence of a 50% likelihood of criminality). Consequently such business records as credit-card transactions, bank statements, and itemized phone bills can be obtained by law enforcement authorities through administrative procedures authorized by the law enforcement agency rather than an independent judge, and no "probable cause" has to be evidenced. This doctrine has been subject to continuous criticism throughout the development of mobile communications which track individuals' location, Internet services which record website browsing and search-engine activity, and social networks in which merely the structure and dynamics of social interaction reveal intimate28 details of private life29. Obviously these conditions could not have been foreseen by courts in the 1970s, yet every challenge so far to overturn the doctrine has been unsuccessful. Such privacy concerns were increased by s.215 of the PATRIOT Act 2001, that attracted considerable controversy. It allows security authorities to obtain "tangible" business records from companies under a secret judicial order. Although secret non-judicial orders to obtain "non-content" data (i.e. "metadata") were already available under a procedure called a 'National Security Letter', s.215 is applicable to any kind of "tangible" data held by a great variety of private-sector businesses. After the first revelations about the PRISM programme, Gen. Alexander (Director of the NSA) confirmed over two public hearings of Congressional intelligence review committees that the NSA collects (both domestic and international) telephone call metadata from all major carriers and maintains a database of all such calls for five years30. By the NSA's own account it uses this data for the sole purpose of deciding whether there is a "reasonable articulable suspicion" of a connection to a terrorist investigation. The database is searched for whether a candidate target telephone number is within "three hops" (i.e. when there exists a "chain" of calls sometime over a 5 year period) to a nexus of numbers previously associated with terrorism. It is striking that so far in the evolution of the 'Snowden affair', domestic US political commentary has almost exclusively referred to the rights of

Americans. This is not a rhetorical trope and is meant literally - no reciprocity ought to be assumed (in law or popular discourse) which extends rights further40. The rights of non-Americans have scarcely been mentioned in the US media41 or legislature. It is even more surprising that careful analysis of the FISA 702 provisions clearly indicates that there exist two different regimes of data processing and protection: one for US citizens and residents ("USPERs"), another one without any protection whatsoever for non-US citizens and residents ("non- USPERs"). Cloud providers are transnational companies subject to conflicts of international public law. Which law they choose to obey will be governed by the penalties applicable and exigencies of the situation, and in practice the predominant allegiances of the company management. So far, almost all the attention on such conflicts has been focussed on the US PATRIOT Act, but there has been virtually no discussion of the implications of the US Foreign Intelligence Surveillance Amendment Act of 2008. §1881a of FAA for the first time created a power of mass-surveillance specifically targeted at the data of non-US persons located outside the US, which applies to Cloud computing. Although all of the constituent definitions had been defined in earlier statutes, the conjunction of all of these elements was new.....the most significant change escaped any comment or public debate altogether. The scope of surveillance was extended beyond interception of communications, to include any data in public cloud computing as well. This change occurred merely by incorporating "remote computing services" into the definition of an "electronic communication service provider". The EU/US Safe Harbour Agreement of 2000 implemented a process for US companies to comply with the EU Directive 95/46/EC on the protection of personal data. If a US company makes a declaration of adherence to the Safe Harbour Principles then an EU Controller may export data to that company (although a written contract is still required). Sometimes described as a 'simultaneous unilateral declaration', the Agreement left ambiguous whether it covered the situation of remote processing of data inside the US, on instruction from Controllers inside the EU. Especially in the case of Cloud computing, such remote processors were most unlikely to be capable of giving effect to the Safe Harbour Principles, which, the US argued, thus became void. Did the deal still apply, for unrestricted export of EU data for remote processing under essentially a self-regulatory framework? In 2000, the EU Commission over-ruled objections from civil society and some DPAs, to conclude a deal. The US negotiators in the Department of Commerce worked closely with US trade lobbies, on a series of "FAQs" for US companies to interpret the Agreement to marginalize EU privacy rights, building in loopholes on such questions as what counted as identifiable data, refusing rights of access, and avoiding any duty of finality or right-of-deletion. In this light, BCR-for-processors can be seen as an expedient strategy both for the Commission and for Data Protection Authorities (DPAs) who wish to maintain the semblance of legal control over EU data, and for the Cloud providers who find the existing EU Data Protection regime generally inconvenient, especially for tax purposes69. The Commission promoted the legal status of the BCR-for-processors concept in the text of the new draft Regulation. Subsequently, national DPAs would have no alternative but to accept their validity once issued. So far, only a few dozen of the existing Controller BCRs have been approved, and the standard of compliance already is not reassuring [49]. the mechanism of BCRs-for-processors, apparently tailor-made to ease the flow of EU data into Third Country cloud computing, is not sufficient to safeguard rights. It contains a loophole that condones unlawful surveillance. It is thus quite surprising that at various stages of development, the concept has been endorsed by the Article 29 Data Protection Working Party (WP29), the European Data Protection Supervisor79 (EDPS), and the French Commission Nationale de l'Informatique et des Libertés (CNIL) which led their formulation. No evidence has emerged that these DPAs understood the structural shift of data sovereignty implied by Cloud computing. Rather, an unrealistic and legalistic view has allowed the protection of EU citizens to be neglected [51]. Prominent notices should be displayed by every US web site offering services in the EU to inform consent to collect data from EU citizens. The users should be made aware that the data may be subject to surveillance (under FISA 702) by the US government for any purpose which furthers US foreign policy. A consent requirement will raise EU citizen awareness and favour growth of services solely within EU jurisdiction. This will thus have economic impact on US business and increase pressure on the US government to reach a settlement. Since the other main mechanisms for data export (model contracts, Safe Harbour) are not protective against FISA or PATRIOT, they should be revoked and re-negotiated. In any case, the requirement above for informed consent after a prominent warning notice should apply to any data collected, in the past or in the future, by a public or private sector EU controller, before it can be exported to the US for Cloud processing. A full industrial policy for development of an autonomous European Cloud computing capacity based on free/open-source software should be supported. Such a policy would reduce US control over the high end of the Cloud e-commerce value chain and EU online advertising markets. Currently European data is exposed to commercial manipulation, foreign intelligence surveillance and industrial espionage. Investments in a European Cloud will bring economic benefits as well as providing the foundation for durable data sovereignty. The PRISM scandal and Snowden's revelations have not been the first warnings to EU institutions in relation to EU citizens' rights. Privacy activists for instance warned the Commission in 2000 that the Safe Harbour Agreement contained dangerous loopholes88. More recently, the above-mentioned note produced on Cloud Computing for the European Parliament's LIBE Committee clearly highlighted the loopholes of FISA and their consequences on EU citizens' rights and protection. The Committee even held a hearing for the presentation of the Note, following a session on the EU Cybersecurity strategy on Feb 20th 2013. Afterwards MEPs asked for immediate proposals to meet the LIBE amendment deadline91 on the Data Protection Regulation. However, from March onwards, the level of interest in the Note declined, and there seemed only a remote possibility that Parliament would support fundamental revisions of the DP regulation. Thanks to the PRISM scandal and Snowden's revelations, such warnings and elections, or by the Parliament; - inclusion of a special technical Commissioner, nominated from the functional constituency of academic computer scientists specializing in privacy, and potentially another Commissioner from the field of Surveillance Studies, also with small independent staffs; - a requirement that DP Commissioners must be appointed by national Parliaments and not the executive; - a minimum quota for DPAs of 25% technical staff with suitable qualifications (or equivalent experience) with a career path100 to the most senior positions; - a subvention of funds to support the civil society sector, although great care must be taken to ring-fence this allocation. Funds should be distributed fairly and on merit, but avoiding the stifling effect of bureaucracy and the danger of institutional capture101. In the United States, the culture of philanthropy and mass-membership civil society supports four highly professional national NGOs102, with diverse approaches, which litigate test cases in privacy and freedom of information, and conduct world-class technical critique of government policies. In contrast, the EU still has a patchwork of dozens of NGOs, who with few resources and lacking the consistent capacity of a permanent research staff, did not campaign on FISA before Snowden.

3.4 An Idea to Defend Private Sphere against Xkeyscore

One of Edward Snowdens most important leaks was about XKeyscore. A software which allows intelligence services to grab and analyse the whole internet traffic. NSA analysts can see mostly everything of the whole Internet activity by grabbing data directly from important Internet nodes. E-Mails, search queries and private messages. Nothing is safe from NSA and its partners.

In the last weeks many people were frightened about that issue as the data analysis works out of legal borders of consitutional law. No judicial orders were needed for grabbing and analysing that data. This way a basic principle of modern democracies was bypassed: Separation of powers between judiciary and executive.

The good thing: People started fighting against this attack touching civil rights by encrypting their Mails, Messages, using Tor and informing about cryptographie techniques like PGP for example.

But this is not enough: Many people are still not aware of the issues a surveillance society has or are not willing to defend themself against this kind of total control as it is sometimes still complicated and not easy enough to use cryptography and surf anonymously through the internet.

Another idea to defend civil rights and to show civil disobedience against these anti democratic programs like XKeyscore might be interfering Signals that XKeyscore just receives white noise.
The idea behind that is as follows:

NSA is using data mining techniques to filter out interesting stuff from Internet communication. They are interested in if people are communicating about evil stuff by filtering keywords, or if one is communicating with wrong and evil people. The side effect: They know almost everything one might send over the internet. If we want to stop them spying out our private lifes, we have to disturb their eavesdropping techniques. This would be possible by sending lots of interfering signals all around the internet.

If lots of people start creating fake identities (for example mail and social network accounts) all over the internet and also every of this fake indentities is starting to communicate normal and also evil stuff to other fake identities, as well as to real people, maybe we can manage it to make XKeyscore absolutely useless. Lets set up mails, messages and search queries, the NSA might find interesting. If enough people are using and faking internet identities this way, the NSA will receive nothing but white noise with XKeyscore, as they cannot distuinguish between real and fake communication.

At this point you may note, that this would be a high manual effort for everyone, which want to be part of that protest. There is no need for that, if we can use intelligent botting techniques. If we can implement some softwares which are creating fake identities for themself and are communicating over the internet like real persons, one has just to install such a software on a home PC. And the best thing is: No one has to do something illegal. It is not illegal to create fake identities and it is also not illegal to let these fake identities communicate [52].

3.5 TFC

TFC is a proof-of-concept set of tools for Pidgin; a low cost solution that enables truly private conversations and a possibility to opt out of the dragnet surveillance. TFC aims to address major vulnerabilities present in current IM applications.

1. Weak encryption algorithms: messages are one-time-pad encrypted. OTP is the only algorithm proven to provide perfect secrecy; messages are unbreakable without the key.
2. Weak entropy of encryption keys: TFC utilizes hardware random number generator. This ensures all bits used to generate encryption keys are truly random.
3. Compromise of end point security: TFC uses data diodes that prevent exfiltration (theft) of plain-texts and encryption keys on hardware level[53].

To put the concept of TFC to practice, a prototype setup was built and simple programs were written in Python. The security of the platform was audited by analyzing the attack vectors remaining. TFC appears to be the only system that protects message confidentiality against the advanced attacks the Snowden leaks revealed NSA to utilize on massive scale; TFC if not unbreakable but it succeeds in making ubiquitous, unethical monitoring with implied presumption of guilt very expensive and by enforcing targeted surveillance, helps preserve the balance between individual privacy and collective security.

Unconvenient operation makes mainstream use difficult. The goal of TFC is to act as a proof-of-concept: a road sign for future high security IM applications. Google Talk w/o HTTPS offer no encryption to protect the messages. In the example, Bob wants to send Alice a password she requests. With Yahoo! Messenger (Yahoo warns the user about missing encryption), the conversation held would be viewable by any third party who has access to the raw TCP packets: Such parties include network administrators, ISP of either party, intelligence agencies and the staff in charge of the IM service's server. Packet inspection was performed with Wireshark: by searching packet contents for the string 'pwd' the password was found without effort. Thus, complicated attacks are not required for dragnet surveillance of unencrypted of communications. Every packet traveling through the borders of US that involves a foreigner is intercepted. NSA may have compromised Internet exchange points that manage the backbone of the Internet**.** This means that even if the packet does not cross US borders, it still might be intercepted. Communications intercepted in foreign outposts are not under the US jurisdiction and messages involving US citizens can be viewed. Most IM clients and servers today utilize SSL or similar encrypted connections that protect the messages from prying eyes. Such clients include Skype, Trillian, Apple's FaceTime, Kik Messenger, WhatsApp, Yullop and Pidgin. Contributing factors to security of connection and client side software include the type and implementation of encryption algorithm, source of randomness and availability of source code to verify non-existence of back-doors; After market products may be required by law to contain capability for lawful interception. law-abiding company must respect the ethically questionable gag-orders, data retention laws and lawful requests for user data. Additionally, the Snowden-leaks revealed some of the largest US-based web services are part of the PRISM program, that mirrors traffic traveling to and from servers for analysis. Encryption isn't an issue since it's dealt with pre-obtained private keys. PRISM compliant services include those offered by Facebook (chats, service), Apple (FaceTime, iCloud), Microsoft (Skype, Hotmail, SkyDrive) and Yahoo (email, messenger). In the case a service operates under foreign jurisdiction it's not required to comply with lawful requests unless the law-enforcement agencies are co-operating. In case lawful requests are not available, tactics such as coercion or bribery may be used to obtain data of interest. Alternatively NSA may resort to compromising the security of the connection between user and the server: By obtaining private keys of certificate authorities with any of the previously mentioned methods, NSA is able to perform man-in-the-middle attacks without detection. The client is unable to detect false certificate with valid CA signature. The Snowden-leaks revealed NSA has a cyber-espionage unit: Tailored Access Operations (TAO), the main purpose of which is to exploit vulnerabilities and gain access to systems of interest. War against terrorism has been used as the main excuse for proactive security, but the leaks indicate that NSA performs industrial espionage against foreign businesses as well as monitors the communications of it's allied nations' politicians**[2]**. In this example if the TAO would break into the Openfire server, access to log file at external database or in the embedded one at /var/lib/openfire/embedded-db/openfire.script would lead to the compromise of all stored conversations. Since the NSA is, quoting Bruce Schneier, “prioritizing techniques providing bulk payoff”, the server provides the easiest access to large number of log files and is more practical to subvert than the devices of end users[53].

Unknown vulnerability However, an article by New York Times claims NSA is able to “Unlock private communication such as end-to-end encryption (that cannot be decrypted even by the messaging service) such as the one used in Adium” The technique how NSA is able to subvert the flagship implementation of security remains unknown. Different attack methods may however be speculated:

OTR MITM against ignorant targets Not all users verify the hash of public key with their partner: in theory any unverified conversation could be MITM'd.

OTR hash swapping Unless hashes are verified face to face, they can be swapped to those calculcated from the key offered by adversary. Interception of common external routes used for hash verification compromise overall security. **Quantum computer** OTR uses Diffie-Hellman key exchange, the security of which relies on the discrete logarithm problem. There is no proof that easy solution to the problem doesn't exist. Furthermore, the problem can be solved with quantum algorithm designed by Peter Shor. According to the leaks NSA does not yet have a computer capable of running the algorithm. However, it should noted Snowden did not have access to files classified as ECI (extremely compartmentalized information). Information about existence of a computer that is able to break the cryptography present everywhere on the Internet, would most likely be classified as such[53].

Undermined RNG NSA may have undermined existing cryptographic algorithms such as the Dual EC random bit generator In other news, the academic world has learned it's possible to undermine functionality of RNG inside the Intel's Ivy Bridge Processors[1]. Compromised device outputs proper length bit strings but the truly random entropy pool used for seeding by the processors PRBG can be shrunk in a malicious way. The self test mechanism of the processor is unable to detect it's undermined performance. Cryptographic key generators rely on random numbers. If NSA has undermined the capability of computers to generate random numbers, implementations of cryptography that rely on kernel's cryptographic services (Pidgin mainly utilizes /dev/random) may be easy to break[53].

Encryption: One-time pad Asymmetric algorithms have not been proven to be secure Public key encryption relies on mathematical problems the true difficulty of which remain unknown. So far the intelligence community has been a decade ahead of the academic world and Shor's algorithm could in theory have been in use since quantum computers such as the D-Wave are slowly introduced to the public market. The use of post-quantum public-key cryptography such as the one based on lattices could in theory be secure but such algorithms should not be used before their security is properly audited.

Symmetric algorithms' keys are harder to contain Symmetric key cryptography with pre-shared-key is resilient against Shor's algorithm with large key lengths (256 bits or more). While it is reasonable to expect proper security from symmetric algorithms, the short key length is in theory much easier to exfiltrate through covert channel.

Auditability of crypto math applied OTP has been proven by Claude Shannon to be unbreakable. It is very easy algorithm to implement and requires less expertise from both the author and the user who wishes to audit the software. Whereas AES algorithm library is 1300 lines long, OTP requires two mathematical operations and the method used for encryption totals 12 lines of code.

Trust of every self-generated key relies on manual verification Public key crypto relies on users manually verifying the hash of public key. If key is transmitter over insecure platform such as the Internet, it must be digitally signed and thus security relies on passing manually the signing key's public portion so user can be sure of it's authenticity. Convenient methods such as public key hash verification over the phone places trust on the assumption voice content has not been tampered with. Another common way for adding trust in users' keys is a key signing party, the main purpose of which is to verify public keys of users. Secure symmetric key cryptography relies on users manually pre-sharing the key to their contact prior to conversation. OTP relies on users manually passing the entropy information used for message encryption prior to conversation as well.

Key generation The Vernam cipher requires large quantities of truly random bits for encryption (7 bits / char using the standard ASCII). For security reasons, messages of TFC are padded: the default length is 100 characters. The deskewing process and padding when short messages are used waste excessive amounts of bits; 700 deskewed (or ~2800 biased bits). The possibility that keys generated by processor come from low entropy pool cannot be skated over. Furthermore, without hardware acceleration, the /dev/random is very slow. Readily available peripheral devices for random bit generation also share the risk of undermined functionality such as was in the case of Crypto AG. In many cases, the high price makes after market solutions unavailable to the general audience. This calls for a cheap external device, where the simplicity of the design ensures that hidden features are unlikely to be present. The TRNG used in this paper generates randomness through quantum phenomenon called avalanche noise that is generated by a reverse biased transistor. The noise signal is linearly amplified by an op-amp. It is then converted into digital signal in a comparator and finally regulated to level of [-1.4, 1.4] volts in respect to the virtual ground of the TRNG device. This enables entropy collection through GPIO pin of a Raspberry Pi by measuring the voltage level at regular intervals. In author's initial tests, the Raspberry Pi collected two 256MB files from two separate TRNGs in less than 6 hours. The entropy files measured an average of 0.8% error at 'monte carlo pi' test when using 100 000 points. This paper does not dive into the aspects of testing since tools and instructions for that are available at the designer's web page.

Risk analysis

Software vulnerabilities The security approach of TFC relies on policies how hardware interacts and comes in contact. Security implementations in physical layer prevents software security vulnerabilities. Most important aspect of security in the software is the auditability of the source code. Probably the biggest challenge of TFC is to keep maintain minimal and easy to interpret source code.

Physical attack vectors are not assessed TFC addresses techniques for mass-scale compromise of end point devices. All other setups are just as vulnerable to following attacks; most of them require physical access to devices.

EMSEC Emissions security is mainly a concern of high value stationary targets known to the adversaries, such as governmental and military facilities, large businesses, or smaller ones in fields of high interest. In TFC the main emission sources are the cables of display units and keyboards connected to TxM and RxM: these emissions leak plain text information. Low power usage and heat radiation allows raspberry to be enclosed within a shielded case that block signals from escaping. Thermal conductance can be improved with submerging the device in oil and adding convection cooling to the outside surface of the case. Encapsulation of display and cables requires metallic mesh the density of which is suitable to blocks signals of frequency range specific to the used display device. The keyboard is much more difficult to protect. The use of retro reflectors as revealed by the leaks in EMSEC compromise might indicate TEMPEST type of monitoring of cable emissions is not effective enough. **Covert channels** Covert channels can in theory be injected to future versions of Raspberry Pi. Current devices may be subjected to such compromise via firmware/OS upgrade. This however, is unlikely, since data diode only transmits when messages are sent and existence of a hidden data transfer is easy to detect with multiple instruments ranging from oscilloscopes to spectrum analyzers.

Acoustic Cryptanalysis Most common approach is to listen to the sounds emitted by keyboard of TxM via a dedicated bugging device, the integrated microphone of the NH or a smart phone, or by using the accelerometer of the latter[53]. Thus, a bug sweep adds to the overall security.

Interdiction NSA was revealed to perform interdiction of hardware ordered online. Thus, ordering a Raspberry Pi online could result in receiving a compromised device with airgap capability defeated. What remains unclear is how the Pi would be able to connect to Internet: signals transmitted through data diode are easy to detect.

3.6 NSA & defend against NSA spying

The National Security Agency is running a surveillance program, named XKeyscore, that allows intelligence analysts to search databases of people's email, online chats and browsing histories without prior authorization, The Guardian reported on Wednesday.

Documents obtained by the newspaper from former government contractor Edward Snowden describe how analysts can obtain Internet data — including the content of email messages, Facebook chats, private messages and search histories — by filling out a simple on-screen form. The form asks for a justification for conducting the search, but the online request is not reviewed by a court or other NSA personnel before it's processed, allowing analysts to target U.S. citizens for electronic surveillance without a warrant, the newspaper claims.

In addition, the surveillance program allows analysts to access "ongoing" interception of a person's Internet activity in real-time[54].

The NSA has come under fire for broadly interpreting surveillance law to collect massive troves of phone and Internet data of Americans. The new revelations about XKeyscore are likely to spur calls from civil libertarians for greater oversight of the intelligence community.

Top intelligence officials and congressional leaders on the intelligence committees though argue that the NSA surveillance programs have thwarted potentially devastating terrorist attacks. They maintain that the spy agency has stayed within its legal bounds.

The leak about XKeyscore comes ahead of a speech that NSA Director Gen. Keith Alexander is scheduled to give on Wednesday to the Black Hat hacker conference in Las Vegas.

NSA documents credit the XKeyscore program with capturing 300 terrorists by 2008, according to The Guardian. The program allows analysts to monitor someone's email communications if they have the person's email or IP address alone.

The NSA told The Guardian in a statement that its XKeyscore program is only targeted at "legitimate foreign intelligence targets."

"NSA's activities are focused and specifically deployed against — and only against – legitimate foreign intelligence targets in response to requirements that our leaders need for information necessary to protect our nation and its interests," the spy agency said. "Allegations of widespread, unchecked analyst access to NSA collection data are simply not true."

White House press secretary Jay Carney said Wednesday that he did not know whether members of Congress had been informed about the XKeyscore program, but also declared "some of the claims made in that article are false."[55].

"As we've explained and the intelligence community has explained, allegations of widespread, unchecked analyst access to NSA collection data are false," Carney said during the daily White House press briefing. "Access to all of NSA's analytic tools is limited to only personnel who require access for their assigned tasks, and there are multiple technical, manual and supervisory checks within the system to prevent from those who don't have access from achieving that access."

In an Op-Ed piece published Tuesday in The New York Times, Eli Dourado, a research fellow at George Mason University, argued that companies using open hardware would be in a better position to detect backdoors or vulnerabilities planted by the NSA or any other government agency.

"To make the Internet less susceptible to mass surveillance, we need to recreate the physical layer of its infrastructure on the basis of open-source principles," wrote Dourado, who is with the technology policy program at George Mason's Mercatus Center.

Some experts were skeptical of the idea, saying the NSA would find other means to compromise systems, whether it was through the cooperation of software vendors or finding unknown vulnerabilities in the hardware.

"I don't see how this attempt at disintermediation would succeed," Al Pascual, analyst for Javelin Strategy & Research, said.

According to Dourado, success would come from the fact that anyone could fully audit the hardware, make changes and then distribute the modifications to others. This model has driven the success of open source software used across the Internet today. Such technology includes the Linux operating system and the Apache Web server.

Mistrust over the security of proprietary technology has been fed by revelations that the NSA collaborated with companies like Microsoft, Apple and Google to program encryption weaknesses into popular consumer products and services, which gave the agency the ability to siphon user data. The revelations are based on documents leaked to the media by former NSA contractor Edward Snowden.

The documents have also described how the NSA has been able to tap into the infrastructure of the Internet, intercepting traffic flowing through cables, routers and switches.

Such hardware would be much more difficult to tap undetected, if the companies using it could see all of the underlying technology, including the firmware, Dourado says.

"There is reason to be skeptical about the security of these networking products. The hardware firms that make them often compete for contracts with the United States military and presumably face considerable pressure to maintain good relations with the government. It stands to reason that such pressure might lead companies to collaborate with the government on surveillance-related requests," he wrote.

Examples of U.S. companies that make such hardware include Cisco, Hewlett-Packard and Juniper Networks. However, the same reasoning could apply to competitors based in foreign countries.

While the ability to fully audit hardware sounds good, the reality is many organizations do not have the people with the expertise to continuously examine updates of low-level code in hardware, Murray Jennex, a professor of information system security at San Diego State University, said.

"In principle a good idea, but in practice not so much," he said.

"Auditing code is always difficult, this will be low-level code that is difficult to follow. I think it will create an illusion of openness that will still be relatively easy to conceal backdoors and such in."
Dourado has his supporters. James W. Gabberty, a professor of information systems at Pace University, said "no other information security control trumps the importance of regular and comprehensive auditing."

"Moving towards an Internet infrastructure that is 100% auditable by both governments and companies alike makes the most sense since, after all, we live in an era of increasing paranoia exacerbated by highly publicized regular hacking incidents of our most important societal systems," he said.

Trust of U.S. technology in light of the NSA revelations has become a concern for vendors selling overseas. Malcolm Harkins, vice president and chief information security and privacy officer for Intel, recently told Network World that customers have expressed a lack of confidence in U.S.-based tech vendors[55].

3. SOCIAL MEDIA & CYBER SECURITY RISKS

4.1 Threats to Your Security when Using Social

As the use of online social networks becomes progressively more embedded into the everyday lives of users, personal information becomes easily exposed and abused. Information harvesting, by both the online social network operator itself and by third-party commercial companies, has recently been identified as a significant security concern for OSN users. Companies can use the harvested personal information for a variety of purposes, all of which can jeopardize a user's privacy. Today, many online social networks have tens of millions registered users. Facebook, with more than billion active users, is currently the biggest and most popular OSN in the world. Other well-known OSNs are Google+, with over 235 million active users [52]; Twitter, with over 200 million active users [41]; and LinkedIn, with more than 160 million active users [65]. While some experts insist that online social networks are a passing fashion and will eventually be replaced by another Internet fad, current user statistics concur that OSNs are here to stay. A recent survey by the Pew Research Center's Internet and American Life Project [17] revealed that 72% of online American adults use social networking sites, a dramatic increase from the 2005 Pew survey which discovered that just 8% of online American adults used social networking sites. Moreover, the survey revealed that 89% of online American adults between the ages of 18 to 29 use social network sites, while in 2005 only 9% of the survey participants in this age group used this type of site. These survey results are compatible with a previous report published by Nielsen in 2011 [56], disclosing that Americans spent 22.5% of their online time on OSNs and blogs, more than twice the time spent on online games (9.8%). Other common online activities that consume Americans' time include email (7.6%); portals (4.5%); videos and movies (4.4%); and searches and instant messaging at 4% and 3.3%, respectively. The amount of collective time spent on OSNs, especially on Facebook, is enormous and ever-growing. U.S. users spent a total of 53.5 billion minutes on Facebook during May 2011, 17.2 billion minutes on Yahoo [51], and 12.5 billion minutes on Google [50]. Mobile devices, or cellular phones, also serve as platforms for Internet usage statistics. The Nielsen report suggested that almost two out of five OSN users access online social networks from mobile devices. Social applications are the third most-used type of application by smartphone users who download applications. It should be noted that the use of OSNs on mobile devices not only allows for an even "closer relation" to social networks but also can pose additional privacy concerns, especially around the issues of collecting location data and providing opportunities for advertisers to identify specific types of users. Besides being popular among adults, online social networks have hugely attracted children and teenagers. A comprehensive 2010 study [68], carried out in 25 European countries with 25,000 participants, produced the following statistics: 60% of children 9-16 years old who access the Internet use it daily (88 minutes of use on average) and 59% of 9-16 year olds who use the Internet have a personal OSN site profile (26% of 9-10 year olds; 49% of 11-12 year olds; 73% of 13-14 year olds; 82% of 15-16 year olds). Note that the terms of use governing OSNs do not officially allow users under the age of 13. Furthermore, 26% of the children in the European study had their social network profile set to "public" (e.g., accessible to strangers), 14% reported having their address or phone number listed on their profile, and 16% admitted that their profile displayed an inaccurate age. Details of the study revealed that 30% of the children surveyed reported having an online connection with a person they had never met face to face, 9% reported having actually met face to face with someone with whom they had only an online connection, 9% reported experiencing a misuse of personal data, 21% reported encountering one or more types of potentially harmful user-generated content, and 6% reported receiving malicious or hurtful messages on the Internet. These findings reiterate our previous claim: the use of online social networks is embedded in the everyday lives of children and teenagers, and can result in personal information being exposed, misused, and potentially abused. Interestingly, about a third of the parents in this European study claimed that they filter their children's use of the Internet, while a quarter specifically stated that they use monitoring tools.

Classic Threats Classic threats have been a problem ever since the Internet gained widespread usage. Often referred to as malware, cross-site scripting (XSS) attacks, or even phishing (among others), they continue to be an ongoing issue. Though these issues have been addressed in the past, they have recently become viral due to the structure and nature of online social networks and have begun to spread more quickly among network users. Classic threats can take advantage of a user's personal information published in a social network to attack not only the user but also their friends simply by adjusting the threat to accommodate the user's personal information. For example, an attacker a user's details from his or her Facebook profile. Due to the personal nature of this crafted message, the chances that the innocent user will open the message and get infected are likely. In many cases, these threats target essential and everyday user resources, such as credit card numbers, account passwords, computing power, and even computer bandwidth (in order to send spam emails). The different classic threats are illustrated below, along with real-life scenarios, where these types of menaces have jeopardized a real user's privacy and security.

Malware. Malware is malicious software developed to disrupt a computer operation in order to collect a user's credentials and gain access to his or her private information. Malware in online social networks uses the OSN structure to spread itself between users and their friends in the network. For example, Koobface was the first malware to successfully propagate through OSNs. This malware succeeded in attacking users in many different online social networks, such as Facebook, MySpace [56], and Twitter. Upon infection, Koobface attempts to collect login information and join the infected computer in order to be part of a botnet [56], a so-called "zombie army" of computers, which in many cases is then used for criminal activities, such as sending spam messages and attacking other computers and servers in the Internet.

Phishing Attacks. Phishing attacks are a form of social engineering to acquire user-sensitive and private information by impersonating a trustworthy third party. A recent study [6] showed that users interacting on a social networking website are more likely to fall for phishing scams due to their innate social and trusting natures. Moreover, in recent years, phishing attempts within OSNs have increased sharply. According to the Microsoft security intelligence report [19], 84.5% of all phishing attacks target social networks sites. One such phishing attack occurred at Facebook. In 2009, a phishing scam lured users onto fake Facebook login pages. Then, the phishing attack spread among Facebook users by inviting the users' friends to click on a link posted on the original user's wall [79]. Fortunately, Facebook acted to stop this attack.

Spammers. Spammers are users who use electronic messaging systems in order to send unwanted messages, like advertisements, to other users. Online social network spammers use the OSN platform in order to send advertisement messages to other users by creating fake profiles [42]. The spammers can also use the OSN platform to add comment messages to pages which are viewed by many users in the network. An example of the prevalence of network spamming can be found in Twitter, which has suffered from a massive amount of spam. In August 2009, 11% of Twitter messages were spam messages. However, by the beginning of 2010, Twitter had successfully cut down the percentage of spam message to 1% [53].

Cross-Site Scripting (XSS). A cross-site scripting (XSS) attack is an assault against web applications. The attacker who uses the cross site-scripting exploits the trust of the web client in the web application and causes the web client to run malicious code capable of collecting sensitive information. Online social networks, which are types of web applications, can suffer from XSS attacks. Furthermore, attackers can use an XSS vulnerability combined with the OSN infrastructure to create an XSS worm that can spread virally among social network users [54]. In April 2009, such an XSS worm, called Mikeyy, rapidly transmitted automated tweets across Twitter and infected many users, among them celebrities like Oprah Winfrey and Ashton Kutcher. The Mikeyy worm used an XSS weakness and the Twitter network structure to spread through Twitter user profiles [56]. **Internet Fraud**. Internet fraud, also known as cyber fraud, refers to using Internet access to scam or take advantage of people. In the past, con artists used traditional in-person social networks, such as weekly group meetings, to gradually establish strong bonds with their potential victims. More recently, according to the North American Securities Administrators Association (NASAA) [84], with the rising popularity of online social networks, con artists have turned to OSNs in order to establish trust connections with their victims, and they now operate using personal data published in the victims' online profiles. In recent years, for example, fraudsters have been hacking into the accounts of Facebook users who travel abroad. Once they manage to log into a user's account, the scammers cunningly ask the user's friends for assistance in transferring money to the scammer's bank account. One victim of this type of fraud was Abigail Pickett. While travelling in Colombia, Abigail discovered that her Facebook account had been hijacked from someone in Nigeria and was being used to send requests for money to friends on the pretext that she was "stranded" [53].

Modern Threats Modern threats are additional current online threats that are typically unique to online social network environments. These threats use the OSN infrastructure to collect and expose personal information about a user and their friends and have even successfully lured users into clicking on specific malicious links. Today's attackers can also combine these two types of attacks - classic and modern - in order to create a more sophisticated attack. For example, an attacker can use a Trojan in order to collect a user's Facebook password and then post a message, which contains clickjacking attack, on the user's timeline in order to lure the user's trusting Facebook friends to click on the posted message and install a hidden virus onto their own computers. Another example is using cloned profiles to collect personal information about the cloned user's friends. Using the friends' personal information, the attacker can send uniquely crafted spam email containing a virus. Due to the usage of personal information, the virus is more likely to be activated. The recovery processes for these two types of threats are very distinct from each other. In order to recover from a classic attack, like a virus, in most cases it is possible to simply reinstall the operating system, change the current passwords, and cancel the affected credit cards. However, in order to recover from a modern OSN attack that “steals your reality” [5], more effort must be invested due to the fact that resetting personal information is excessively time consuming and not always 100% possible. For example, you can change your email address, but it takes a significantly greater amount of effort to change your home address. In the rest of this section, we illustrate the various modern threats and real-life scenarios where these types of threats have jeopardized an OSN user’s privacy and security.

Clickjacking. Clickjacking is a malicious technique which tricks users into clicking on something different from what they intended to click. By using clickjacking, the attacker can manipulate the user into posting spam messages on his or her Facebook timeline, performing “likes” to links unknowingly (also referred as likejacking), and even opening a microphone and web camera to record the user [56]. An example of a clickjacking attack occurred at Twitter in 2009 when Twitter was plagued by a “Don’t click” attack. The attacker tweeted a link with the message “Don’t click” along with a masked URL, in which the actual URL domain is hidden. When Twitter users clicked on the “Don’t click” message, the message automatically spread virally and was posted onto their Twitter accounts [56].

Threats Targeting Children Related to online social network threats is the issue of trust. Recent studies illustrate that OSN users tend to trust their friends in the network. In many cases, this trust can lead to further means of communication, possibly escalating to face-to-face encounters [56]. An additional major threat to children and teenagers is cyberbullying in which OSN users suffer from abuse initiated from people they know, such as classmates, or from strangers. The different threats that typically target children are illustrated below, along with studies’ findings on each threat. Naïve Trust. Children, in particular, are susceptible to trusting social network friends without question, yet the said friend’s profile may really be a mask for sinister intentions. Online Predators. The greatest concern for the personal information safety of children is reflected in the growing apprehension about Internet pedophiles, also referred to as online predators. Livingstone and Haddon [67] of EU Kids Online defined a typology in order to understand the risk and harm related to the following online activities: online harm from content (child’s exposure to pornography or harmful sexual content); harm from contact (a child who is contacted by an adult or another child for the purposes of sexual abuse); and harm from conduct (the child as an active initiator of abusive or risky behaviors). Behaviors considered to be Internet sexual exploitation of children include adults using children for the production of child pornography and its distribution, consumption of child porn, and the use of the Internet as a means to initiate online or offline sexual exploitation.

Solutions In recent years, social network operators, security companies, and academic researchers have tried to deal with the above-mentioned threats through proposing a variety of solutions.In this section we describe in detail possible solutions which can assist in protecting the security and privacy of online social network users. or logging into the online social network is a real person and not a socialbot or a compromised user account, OSN operators are using authentication mechanisms, such as CAPTCHA [56], photos-of-friends identification, multi-factor authentication [56], and in some cases even requesting that the user send a copy of his or her government issued ID [50]. As an example, Twitter recently introduced its two-factor authentication mechanism, requiring the user to not only insert a password when logging into Twitter but also provide a verification code that was sent to the user’s mobile device. This mechanism prevents a malicious user from logging in through hijacked accounts and publishing false information through those hijacked accounts. Such a mechanism would thwart incidents such as when hackers hijacked the Associated Press (AP) Twitter account, resulting in the rapid propagation of false information about explosions in the White House, which caused panic on Wall Street . Security and Privacy Settings. Many online social networks support a configurable user privacy setting that enables users to protect their personal data from others in the network [56]. For example, Facebook users can choose which other Facebook users are able to view their details, pictures, posts, and other personal information that appear in their Facebook profile [41]. However, many Facebook users simply maintain the default privacy settings, letting their data be exposed to strangers. Some online social networks also support extra security configurations which enable the user to activate secure browsing, receive login notifications, and establish other safety features [56]. Internal Protection Mechanisms. Several online social networks protect their users by implementing additional internal protection mechanisms for defense against spammers, fake profiles, scams, and other threats [56]. Facebook, for example, protects its users from malicious attacks and information collecting by activating the Facebook Immune System (FIS). The FIS is described as an adversarial learning system that performs realtime checks and classifications on read and write actions on Facebook's database [56].

4.2 Risk of Social Media for Corporate Reputation

It is no understatement to say that there is a worldwide obsession with social media – Facebook, Twitter, and Foursquare are all the rage and while the social network of the moment may change, social media as a communications channel is here to stay.

While consumers may have been first to jump on the social media bandwagon, the number of businesses utilizing social media is growing, with 83% of Inc. 500 companies using at least one social network in 2010 according to a study by the Center for New Communications Research and they are increasingly reaping the benefits of having a direct communication line to their customers. We, at DuPont Sustainable Solutions, use this forum to have a dialogue with you about the most pressing issues for a 21st Century business. Being able to interact and build more direct relationships with consumers provides innumerable opportunities for all companies including ours.

However, businesses using social media face inherently greater risks than traditional communications' channels, particularly to their reputation and financial success. Given DuPont's history and values in ethics, safety and risk management, DuPont understands the importance of engaging audiences through online social channels openly, honestly and transparently in order to gain their trust. Doing otherwise risks damage to the company's reputation.

So how do organizations find a way to balance the benefits of social media with the risks? The short answer is with corporate policies and proper training.

One risk businesses face when it comes to social media is the sharing of information. Either employees can share opinions in ways that reflect badly on the company or they can release confidential information or intellectual property. .Confidential information being leaked, data breaches, privacy violations, offensive tweets – all of these possibilities make organizations hesitant to adopt social media.

Aside from the risks stemming from improper employee use of social media, there is also the ability of consumers and competitors to use social media to attack a company's reputation.

However, not engaging in social media can become a risk in itself. Most every company and brand are being discussed through online social channels. It is important that the company provides its voice in these discussions and conversations. If the company does not provide a voice, others will regardless. Social media is developing into a standard practice for communications and rather than avoiding it, businesses should actively take steps to engage these online audiences.

Since social media is becoming a standard practice for many businesses or at the very least a good way to interact with customers and partners , it is important for employees to understand the best practices for engagement with particular emphasis on ethics and legal responsibilties. Having a good company policy on social media with an active governance plan is not longer optional. It is essential to mitigating the risks associated with social media communications.

At DuPont, we are mitigating the risks posed by social media through corporate policies and education and training for our employees. We have created a DuPont Web Based Social Networking and Digital Media Policy that provides guidelines and expectations related to our employees' use of social media on behalf of the company. The Policy addresses internal social networks, external social networks, codes of conduct and guidelines. We also have a Social Network Steering Committee that reviews and approves the creation and use of all social networks to ensure adherence to issued guidelines and standards. Education and training is equally important and DuPont is establishing a social media training program that will educate employees on the risks of social media, best practices for using social media and DuPont's policies.

Yes, there are risks to communicating through social media, but if a company communicates ethically and transparently, these risks are minimized. DuPont sees the value in using these new communicatoins' channels to reach our customers and prospects and we understand that these audiences expect us to use social media responsibly.

4.3 Social Networks and Cyber Forensics

Social networks like Facebook, Twitter, Foursquare and Google Buzz can be a treasure trove for forensics investigations. The expanding ocean of data in those networks is irresistible to investigators.

Marketers are already exploiting social data to analyze associations among consumers. A startup named 33Across looks at relationships among social media users to ascertain who, for example, would be a good prospect for viewing an ad on costume jewelry. If Jane is a good prospect, then some of her friends — or maybe just people who circulate in the same social group — might be too. 33Across uses tools like tracking cookies to follow relationships.

Just as this style of data gathering and analysis can help marketing, it can help law enforcement or dispute resolution. Police have already learned that drug dealers socialize online with other drug ring members, and street gangsters network with their cohorts.

A simple investigation might view just the publicly-available text and images posted on a suspect's social page. Deeper investigations may require the investigator to acquire special authority. In an internal corporate investigation, that authority might come in the form of consent from a company employee who has the right to access a page. Or, in a civil lawsuit or certain government investigations, the authority might come in the form of a subpoena. In a criminal investigation, it might be a search warrant.

A sophisticated investigation will examine more than just the data appearing on the face of social web page. It might, say, go for the cache of data collected at 33Across to ascertain who might be involved with a Medicare fraud scheme.

As an investigation team seeks authority such as a subpoena or search warrant, it will be prudent to address privacy concerns. Here are example steps to reduce privacy risks:

1. Deliberate in writing about the privacy risks, how they can be minimized and why they are justified taking in the case at hand.
2. Consult a third party expert (or panel of experts) on how to proceed with the investigation in a way that respects privacy.
3. Mask personally-identifying information from individual researchers.
4. Secure data against use or disclosure beyond the investigation.
5. Be transparent to the extent consistent with the mission of the investigation. Modern society rewards openness and transparency. Investigation teams do themselves a favor when they publicize their techniques and open them to scrutiny.
6. Document all efforts to protect privacy.

4.4 Using Social Media & Cyber Security Risks

Network World - Social media platforms such as Twitter, Facebook and LinkedIn increasingly are being used by enterprises to engage with customers, build their brands and communicate information to the rest of the world.

But social media for enterprises isn't all about "liking," "friending," "up-voting" or "digging." For organizations, there are real risks to using social media, ranging from damaging the brand to exposing proprietary information to inviting lawsuits.

Here are five of the biggest social media security threats:

1. Mobile apps

The rise of social media is inextricably linked with the revolution in mobile computing, which has spawned a huge industry in mobile application development. Naturally, whether using their own or company-issued mobile devices, employees typically download dozens of apps because, well, because they can.

Social networking security threats taken too lightly

But sometimes they download more than they bargained for. In early March, Google removed from its Android Market more than 60 applications carrying malicious software. Some of the malware was designed to reveal the user's private information to a third party, replicate itself on other devices, destroy user data or even impersonate the device owner[22].

And all because this new game is supposed to be even better than Angry Birds!

2. Social engineering

A favorite of smooth-talking scammers everywhere, social engineering has been around since before computer networks. But the rise of the Internet made it easier for grifters and flim-flam artists to find potential victims who may have a soft spot in their hearts for Nigerian royalty.

Social media has taken this threat to a new level for two reasons: 1) People are more willing than ever to share personal information about themselves online via Facebook, Twitter, Foursquare and Myspace, and 2) social media platforms encourage a dangerous level of assumed trust. From there it's a short step to telling your new friend about your company's secret project. Which your new friend really might be able to help with if you would only give him a password to gain access to a protected file on your corporate network. Just this once.

3. Social networking sites

Sometimes hackers go right to the source, injecting malicious code into a social networking site, including inside advertisements and via third-party apps. On Twitter, shortened URLs (popular due to the 140-character tweet limit) can be used to trick users into visiting malicious sites that can extract personal (and corporate) information if accessed through a work computer. Twitter is especially vulnerable to this method because it's easy to retweet a post so that it eventually could be seen by hundreds of thousands of people.[54]

4. Your employees

You knew this was coming, but even the most responsible employees have lapses in judgment, make mistakes or behave emotionally. Nobody's perfect all of the time.

But dealing with an indiscreet comment in the office is one thing; if the comment is made on a work-related social media account, then it's out there, and it can't be retrieved. Just ask Ketchum PR Vice President James Andrews, who two years ago fired off an infamous tweet trashing the city of Memphis, hometown of a little Ketchum client called FedEx, the day before he was to make a presentation to more than 150 FedEx employees (on digital media, no less!).

The tweet was discovered by Fedex employees, who emailed angry missives to Ketchum headquarters protesting the slight and wondering why FedEx was spending money on a snooty

New York PR firm while employees were dealing with a 5% salary cut during a severe recession. Andrews had to make a very public and humiliating apology.

Remember, this wasn't some low-level employee not tuned into the corporate mission. This was a high-level communications executive who damaged his company's brand and endangered an account. Imagine what a disgruntled low-level employee without as much invested in his job might be able to do with social media tools and a chip on his shoulder.

5. Lack of a social media policy
This one's totally on you. Without a social media policy for your enterprise, you are inviting disaster. You can't just turn employees loose on social networking platforms and urge them to "represent." You need to spell out the goals and parameters of your enterprise's social media initiative. Otherwise you'll get exactly what you're inviting - chaos and problems.

Who is allowed to use social media on behalf of the organization and what they're allowed to say are the two most obvious questions that must be addressed in a social media policy. You need to make all this clear or employees will make decisions on their own, on the fly. Does that sound like a good thing?[55].

Two more imperatives related to social media policy: 1) Organizations must conduct proper training for employees, if only to clear up issues regarding official social media policies, and 2) A social media initiative needs a coordinator and champion. And that means a social media manager.

4.5 Exploiting of Vulnerability to Secure User Privacy on a Social Networking Site

Attributes available for every user on a social network- ing site can be categorized into two major types: individual attributes and community attributes. Individual attributes characterize individual user information, including personal information such as gender, birth date, phone number, home address, etc. Community attributes characterize informa- tion about friends of a user, including friends that are trace- able from a user's pro_le (i.e., a user's friends list), tagged pictures, wall interactions, etc. Both types of attributes are always accessible to friends but may not to other users. Using privacy settings of a pro_le, a user can control the visibility of most individual attributes, but cannot control the visibility of most community attributes. For example, Facebook users can control traceable link information about friends but cannot control exposure of friends through photo tagging and wall interactions. On most social networking sites, privacy related e_orts have been concentrated on protecting individual attributes only. Thus, users are often left vulnerable through commu- nity attributes. We propose a mechanism that enables users to protect against vulnerability. The mechanism is tunable to accommodate individual preferences across the spectrum, from reclusive users to gregarious users. A novel way to de_ne a vulnerable friend from an indi- vidual user's perspective is dependent on whether or not the user's friends' privacy settings protect the friend and the individual's network of friends (which includes the user). An individual is vulnerable if any friend in the network of friends has insu_cient privacy settings to protect the entire network of friends. Thus, whether or not an individual re- ciprocates privacy settings with their friends within a friend network can impact the entire network. Before presenting a formal de_nition of vulnerable friends, we propose four indexes, I-index, C-index, P-index, and V-index, based on individual and community attributes for each user on a so- cial networking site. These indexes can be used to estimate user's privacy, quantify how well a user protects friends, spec- ify how public or private user pro_les are, and compute the vulnerability of individual users on a social network. For the _rst set of experiments, we compare V-index for each of user with two optimal algorithms and six intuitive strategies for unfriending to reduce vulnerability. For each graph in Figure 3, the X-axis and Y-axis indicate users and their V-index values, respectively. For simplicity, we sort all users in ascending order based on existing V-index, and then we plot their corresponding V-index before and after unfriending. Figure 3 indicates performance of all eight algorithms which will help us to decide whether unfriending makes users more or less vulnerable. The eight algorithms are[57]: _ Most vulnerable friend. For a user, the most vulnera- ble friend is the one whose removal lowers the V-index score the most.

For each user, we _rst _nd the most vulnerable friend and then estimate the new V-index value (M1-index) after unfriending him/her. As ex- pected, see Figure 3(a), V-index values for users de- crease in comparison with V-index values before un- friending the most vulnerable friend. Unfriending the most vulnerable friend makes all users more secure. _ Two most vulnerable friends. If we sort all of user's vulnerable friends in ascending order based on their new V-indexes (after unfriending), the top two in the list are the two most vulnerable friends. For each user, we _rst _nd two most vulnerable friends and then esti- mate the new V-index value (M2-index) after unfriend- ing them. As expected, see Figure 3(b), V-index values for all users decrease in comparison with V-index val- ues before unfriending the two most vulnerable friends. Unfriending the two most vulnerable friends also make all users more secure. _ Least V-friend. For each user, we choose to unfriend the friend whose V-index is the lowest among all friends. This friend is the least V-friend. V-index values in- crease for 65% of 100K users, and increase for 43% of the 2M+ users, in comparison with V-index values be- fore unfriending the least V-friend. See Figure 3(c), L-index refers to the new V-index value after unfriend- ing the least V-friend. V 0 u > Vu for some users because, Pl < Vu where Pl is the P-index of the least V-friend. Unfriending the least V-friend does not make all users insecure. _ Random friend. For each user, we randomly choose to unfriend a friend. V-index values increase for 24% of 100K users, and increase for 23.5% of the 2M+ users, in comparison with V-index values before unfriending a random friend. See Figure 3(d), R-index refers to the new V-index value after unfriending a random friend. V 0 u > Vu because Pr < Vu, where Pr is the P-index of the random friend. Unfriending a random friend does not make all users secure.

4. PRACTISES FOR PRIVICY ON SOCIAL NETWORKING SITES

5.1 Best Practices for privacy on Social Networking Sites Security and Privacy Issues

Below are some helpful tips regarding security and privacy while using social networking sites:

- Ensure that any computer you use to connect to a social media site has proper security measures in place. Use and maintain anti-virus software, anti-spyware software, and a firewall and keep these applications and operating system patched and up-to-date.
- Be cautious when clicking on links. If a link seems suspicious, or too good to be true, do not click on it....even if the link is on your most trusted friend's page. Your friend's account may have been hijacked or infected and now be spreading malware.
- If you are going to request that your account be deleted, first remove all of the data. Request that the account be deleted, rather than deactivated.
- Type the address of your social networking site directly into your browser or use your personal bookmarks. If you click a link to your site through email or another website, you might be entering your account name and password into a fake site where your personal information could be stolen.
- Be cautious about installing applications. Some social networking sites provide the ability to add or install third party applications, such as games. Keep in mind there is sometimes little or no quality control or review of these applications and they may have full access to your account and the data you share. Malicious applications can use this access to interact with your friends on your behalf and to steal and misuse personal data. Only install applications that come from trusted, well-known sites. If you are no longer using the app, remove it. Also, please note that installing some applications may modify your security and privacy settings.
- Use strong and unique passwords. Using the same password on all accounts increases the vulnerability of these accounts if one becomes compromised. Use different passwords for different accounts, and do not use a password you use to access your organizations network on any personal sites you access.
- Be careful whom you add as a “friend,” or what groups or pages you join. The more “friends” you have or groups/pages you join, the more people who have access to your information.
- Do not assume privacy on a social networking site. For both business and personal use, confidential information should not be shared. You should only post information you are comfortable disclosing to a complete stranger.

- Use discretion before posting information or comments. Once information is posted online, it can potentially be viewed by anyone and may not be able to be retracted afterwards. Keep in mind that content or communications on government-related social networking pages may be considered public records.
- When posting pictures, delete the meta data, which includes the date and time of the picture.
- Do not announce that you are on vacation or away for an extended period of time.
- Configure privacy settings to allow only those people you trust to have access to the information you post, and your profile. Also, restrict the ability for others to post information to your page. The default settings for some sites may allow anyone to see your information or post information to your page.
- Review a site's privacy policy. Some sites may share information, such as email addresses or user preferences, with other parties. If a site's privacy policy is vague or does not properly protect your information, do not use the site.

5.2 A Legal Guide to the Use of Social Media in the Workplace

An important tool in managing the legal risks associated with employees' use of technology and social networking sites is a wellcrafted technology and social media policy that balances company needs and concerns against employees' legal rights. Some of the business and legal risks that an employer should address in a technology and social media policy include:

• Covered technology and devices: Employers should consider whether the policy will extend only to employerpaid or provided devices or whether the employer may lawfully and should extend the policy to personallyowned devices used for work purposes. The law is still evolving in this area, and it is not clear that employers have the legal right in all jurisdictions to search an employee's personal device or personal email account on a company or personally-owned device. However, having a clearlyworded policy can improve an employer's legal position in arguing that it has the right to access any technology devices used by an employee for work purposes.
• Privacy considerations: Due to the privacy issues discussed above, a policy should include an express warning that the employer retains the right to monitor and review the use of and content on any technology and devices covered by the policy. As discussed above, however, there have been court decisions finding employers liable for improperly accessing or using online content, particularly where the content was on a website with restricted privacy settings, such as Facebook.com. As such, employers should take care to ensure they lawfully access online content, and they should consult with counsel as appropriate to ensure compliance.
• Permissible and impermissible uses: The policy should explain the permissible and impermissible uses of technology and social media. Items to address might include, for example, personal use of technology on work time, employees' obligation not to use technology to engage in unlawful behavior, the need to protect confidential or trade secret information, and the need to respect others' intellectual property rights. An employer may also want to prohibit employees from engaging in any company-related blogging, tweeting or the like without express written permission of the company to engage in such social networking activities on behalf of the business.
• Lawfully Protected Employee Activity: In setting out any prohibited conduct in a workplace policy, employers must take care to balance the employer's needs against employees' legal rights. As discussed above, a job applicant's or employee's use of technology and online content may be legally protected by discrimination, antiretaliation, lawful consumable products, lawful activity, labor law, or other laws. As such, an employer should be cautious in rejecting a job candidate or disciplining or terminating an employee for online activity to ensure that adverse action is not taken based on legally-protected activities by the individual.
• Wage and Hour issues: As discussed above, non-exempt employees generally must be paid at least minimum wage for all time worked and overtime pay, which can, depending on the circumstances, include time spent checking voice mails or e-mails away from work. In addition, wage and hour issues may arise for employees that use remote technology while telecommuting or while on a leave of absence. As such, an employer should consider addressing limits on the

use of technology by non-exempt employees outside of normal working hours or by employees on leave.

• Photography and Recording: Smartphones and other mobile devices make it far easier than in the past for employees to secretly record conversations at work or to take unauthorized photographs or videos that might be widely disseminated on the internet and go “viral.” Depending on the employer’s business and its unique risks, a technology policy might include language prohibiting the use of devices to make recordings or take photographs or videos. Again, however, an employer should consult with counsel to ensure that any such language does not run afoul of individuals’ Section 7 labor law rights or other employment law rights.

• Testimonials: As discussed above, the FTC has taken the position that false and misleading advertising laws apply to online postings. As such, employers should include language in any policy that advised employees of the need to comply with FTC requirements when making endorsements or testimonials about the company.

• Return of Company Data: An employer should make clear that all company data, including any electronic data stored on an employee’s personally-owned devices, such as a smartphone, tablet, or personal computer, must be returned to the company upon request or when an employee leaves employment. An employer that has a BYOD (bring your own device) approach to workplace technology should consider including language in a technology policy stating that employees agree to turn over their personal devices to the company to permit the company to wipe any company data from the device. In addition, many companies have the capability to remotely cut off access to company technology and to remotely wipe company-owned or employee-owned devices. An employer that has a BYOD approach, should consider including language in a policy that provides that an employee that is permitted to use a personal device for work agrees to permit the company to remotely wipe the device even if that may result in personal data on the device being deleted.

5.3 Protection against Online Spying, NSA PRISMFairview

Americans (and now other countries) are becoming increasingly concerned with the online protection of their personal information. There have been numerous stories in the press lately of hacked websites, identity theft, and eavesdropping, among many other alarming “criminal” acts. Recently, the Washington Post ran a story that included the publishing of four slides, all leaked from the United States National Security Agency (NSA). These slides clearly discussed, in extreme detail, how the domestic PRISM program collects its data.

Other countries are also being “spied” on with a different surveillance program called Fairview. A recent example is the surveillance on millions of citizens of Brazil.

Alarmingly enough, the collection process is incredibly simple. All that has to transpire in order for data collection to begin is for an NSA analyst to identify a potential target for surveillance and then ask a supervisor to approve the request. What if you happen to be one of these unfortunate individuals who is targeted? Is there a way that you can forgo the monitoring process and increase your online protection for your data and browsing activities?

Most experts agree that PRISM is capable of creating “wiretaps” on high capacity fiber optics and watches the data flows past. These wiretaps are easily placed at the Internet connections of such noteworthy conglomerates as Facebook, Yahoo, Apple, and Google, with most of their servers located in the United States.

Copies of the information traveling along these optic signals are diverted and re-routed to a location that is specifically operated by the NSA. Here it is categorized and indexed. Afterwards, it is sent back for analysis to the NSA. Most of this collected data is transferred by methods of plain text protocols. These are simply packets of information that contain a simple text header with a clear payload (the actual data).

However, what the Washington Post story also revealed is that when the payload is encrypted, the NSA is typically unable to crack the code. The below checklist includes ways to minimize the threat of potential NSA (and other hackers/spyware/trojans) monitoring for both digital voice and Internet traffic, leading to an increase in protection for your online activities[58].

1. Encryption of Internet traffic
A sure sign that communication between your web browser and a website is encrypted will be the use of the "https://" SSL protocol preceding the website address. By always typing this prior to entering your URL name into the browser, an encryption certificate will be downloaded from the website automatically (if it supports encryption). If the "https" is not clearly showing in your web browser, then the communication between your browser and the webserver is not encrypted. Check your most commonly visited websites for encryption. Some websites have better encryption than others. You can easily test all websites by visiting SSL Labs.

2. Disable website tracking
The two methods used to protect your site from online tracking include "black listing" (blocking) and "white listing" (approving). Programs of black listing will block spyware. PeerBlock is a good example of a free black list program. White listing programs turn off JavaScript, which is the very common language for programming that is often used in your search engine browsers. This is the language that is most often used for tracking techniques when you visit common websites with minimal online protection encryption. By turning off JavaScript, the spies find it infinitely more difficult to monitor your activities.

3. Search anonymously
Everyone is well aware that Google keeps track of all of our searches and Internet activity. This is how they create profiles on us and generate revenue. Ever notice that when 2 different people search for the same thing on Google, you each receive different results.

There are many other types of search engines which are significantly less concerned with our virtual actions. By switching to one of these other available options, such as Startpage or Duckduckgo, you'll see significant online protection against unwanted monitoring and tracking.

4. Never trust anyone for your confidential data
If you happen to be one of those online users who is using Dropbox, iCloud, SkyDrive, or Google Drive, your online protection may be in severe jeopardy. Make sure to encrypt all of your files before you transmit them online. Never rely on the encryption protection provided by the cloud provider. AxCrypt is considered a viable choice and complements online storage services.

5. File encryption
TrueCrypt (also free) is a type of encryption system used for scrambling the contents of files. By using this or a similar type of system with a strong algorithm and big "key", you will be better able to control the access of your files and documents by outside sources through the use of password protection methods.

6. Use traffic tunnels
The installation of a proxy server or VPN network will encrypt your online activity by sending your data to another outside location for decryption first, which will significantly increase your protection. Every time that you transmit data, emails, or web requests of any kind, the Internet tracks your activity through the use of headers that contain personal information of your address, the address of your destination, and the time and date of the activity. Proxy servers and VPN networks obscure your actions and effectively help make them anonymous. Here is an article with a listing of free anonymous web proxy servers.

7. Secure your operating system
Install an operating system that is "read-only". This will provide significant online protection for your PC against all forms of unwanted programs, applications, and software. Many Linux distributions have the ability to automatically detect PC hardware immediately upon booting of your system. Privatix and Puppy Linux are some good examples for increased security.

8. Use safe text
You might be surprised to learn that all Skype conversations are monitored by Microsoft. Additionally, texting features on digital phones is not at all secure. Most email services do not employ encryption protocols and therefore do not provide online protection either.

Some people try to secure their systems by installing PGP software (Pretty Good Privacy). This is a type of encryption software (supporting email) that attempts to make your conversations unreadable. However, PGP can be difficult to install and manage. There are instead some very reliable phone apps that can encrypt all types of text messages and conversations. Examples include iChat, CryptoCat, and Silent Circle apps.

9. Regulate your voice activity

Smartphones are a wonderful invention, but their lack of security and encryption features leaves us vulnerable to malicious apps, phishing scams, and numerous forms of malware. Microsoft even allows the weakening of Skype voice encryption services to accommodate lawful interception of our signals and voice conversations. Silent Circle launched recently to facilitate end-to-end encrypted communication (voice smartphone calls, conference calls, text messages, video, emails) where the encryption keys are kept by the subscribers and not on their servers.

Of course, even if you were to follow this entire checklist of options, there is no perfect method for protecting your systems from a very determined intruder. These recommendations will definitely make tracking your virtual activity increasingly more difficult. Ensure you research each option before implementing to understand what you are doing and don’t accidentally lock yourself out of your own files[58].

5.4 Solutions

In recent years, various commercial companies have expanded their traditional Internet security options, and now offer software solutions specifically for OSN users to better protect themselves against threats. In this section, we present mainstream software and application-protection solutions, which were developed by well-known security companies such as Symantec and Check Point, as well as solutions which were created by several startup companies such as United Parents. Internet Security Solutions. Many security companies, such as AVG, Avira, Kaspersky, Panda, McAfee, and Symantec [33], offer OSN users Internet security solutions. These software suites typically include anti-virus, firewall, and other Internet protection layers which assist OSN users in shielding their computers against threats such as malware, clickjacking, and phishing attacks. For example, McAfee Internet Security software [51] provides its users with protection against various threats, such as malware, botnet, and inappropriate sites. Norton Safe Web. Symantec’s Norton Safe Web [56] is a Facebook application with more than 500,000 users. It scans the Facebook user’s News Feed and warns the user about unsafe links and sites. Trend Micro Privacy Scanner. Trend Micro Privacy Scanner for Facebook [57] is an Android application which scans the user’s privacy settings and identifies risky settings which may lead to privacy concerns. It then assists the user in fixing the settings. Websense Defensio. Websense’s Defensio web service [56] helps in protecting social network users from threats like links to malware that can be posted on the user’s Facebook page. The Defensio service also assists in controlling the user’s published content, such as removing certain words from posts or filtering specific comments. ZoneAlarm Privacy Scan. Check Point’s ZoneAlarm Privacy Scan [55] is a Facebook application which scans recent activity in the user’s Facebook account to identify privacy concerns and to control what others can see. For instance, ZoneAlarm Privacy Scan can identify posts that expose the user’s private information. ContentWatch Net Nanny. ContentWatch’s Net Nanny [25] is software which assists parents in protecting their children from harmful content. Net Nanny lets parents monitor their children’s social media activity on different online social network websites, such as Facebook, Twitter, and Flickr [45]. Infoglide MinorMonitor. Infoglide’s MinorMonitor [56] is a parental control web service which gives parents a quick, dashboard view of their child’s Facebook activities and online friends. By using MinorMonitor, parents can be alerted about questionable content that may have been revealed to their child, or they can identify over-age friends in their child’s Facebook friends list. United Parents Child Protection Service. The United Parents Child Protection Service [55] assists parents in protecting their children by monitoring and tracking a child’s activity in online social networks and identifying suspicious patterns in their child’s online profile. The child’s privacy is maintained, but this service can help identify and report to the child’s parents, via email or text message, when the child is approached by online strangers or cyberbullied. Several recently published studies have proposed solutions to various online social networks threats. These solutions have primarily focused on identifying malicious users and applications. In this section, we present studies which provide solutions for detecting phishing attacks, spammers, cloned profiles, socware, and fake profiles.1 Phishing Detection. Many researchers have suggested anti-phishing methods to identify and prevent phishing attacks; most of these methods have been based on techniques which attempt to identify phishing websites and phishing URLs [56]. In recent years with the increasing number of phishing attacks on online social networks [55], several researchers have suggested dedicated solutions for identifying OSN phishing attacks. In 2012, Lee et al. [56] introduced the WarningBird, a suspicious URL detection system for Twitter, which can handle phishing attacks that conceal themselves by using conditional redirection URLs. Later in the same year, Aggarwal et al. [4] presented the PhishAri technique, which can detect if a tweet posted with a URL is phishing or not by utilizing specific Twitter features, such as the account age and the number of followers of the user who posted the suspicious tweet. Spammer Detection. Many researchers have recently proposed solutions for spammer detection in online social networks In 2009, Benevenuto et al. [12] offered algorithms for detecting video spammers, which succeeded in identifying spammers among YouTube users. In 2010, De-Barr and Wechsler used the graph centrality measure to predict whether a user is likely to send spam messages. Wang proposed a method to classify spammers on Twitter by using content and social network graph properties. Stringhini et al. created more than 300 fake profiles (also referred as “honey-profiles”) on three online social networks sites, Twitter, Facebook, and MySpace, and successfully identified spammers who sent spam messages to the fake profiles. Lee et al. also presented a method for detecting social spammers of different types by using honeypots combined with machine learning algorithms. In 2013, Aggarwal et al. [3] presented machine learning algorithms for detecting various type of spammers in Foursquare [46]. Recently, Bhat and Abulaish introduced a community-based framework to identify OSN spammers. Cloned Profile Detection. In 2011, Kontaxis et al. [57] proposed a methodology for detecting social network profile cloning. They designed and implemented a prototype which can be employed by users to investigate whether users have fallen victim to clone attacks. In 2013, Shan et al. presented the CloneSpotter which can be deployed into the OSN infrastructure and detect cloning attacks by using various users’ data records, such as a user’s login IP records, which are available to the OSN operator. Socware Detection. In the last few years a number of studies have tried to better understand and identify socware. In 2012, Rahman et al. [94] presented the MyPageKeeper Facebook application which aims to protect Facebook users from damaging posts on the user’s timeline. Rahman et al. also presented the Facebook’s Rigorous Application Evaluator (FRAppE) for detecting malicious applications on Facebook [57]. In 2013, Huang et al. [55] studied the socware ecosystem and discovered several insights about socware propagation characteristics that can assist in future research on the detection and prevention of socware propagation.

We advise OSN users who want to better protect themselves in these platforms to implement the following six recommendations in each of their OSN accounts:

1. Remove Unnecessary Personal Information. We advise OSN users to review the details they have inserted into their online social network accounts and remove extraneous information about themselves, their family, and their friends. It is also recommended that users hide their friends list if possible to prevent inference attacks.
2. Adjust Privacy and Security Settings. In many online social networks, like Facebook, the default privacy settings are insufficient. Moreover, a recent study showed that many Facebook users tend to stay with their default privacy settings [41]. In order for users to better protect themselves on Facebook and in other online social networks, we recommend modifying the privacy settings so that users’ personal data will be exposed only to themselves, or at most to their friends only (for example, see Figure 4). Additionally, if possible, we advise users to activate the secure browsing option and any other available additional authentication mechanisms, such as Twitter’s two-factor authentication [56].
3. Do Not Accept Friend Requests from Strangers. As we demonstrated in Section 3, fake profiles are quite common and often dangerous. Therefore, if a user receives a friend request from an unknown person, we recommend ignoring such a request. If the user is uncertain and is considering approving the friend request, we recommend performing a short background check on the new “friend” and, at a minimum, insert the friend’s profile image to Google Images Search [51] and submit the friend’s full name and other details to other search engines in order to validate the authenticity of the individual. In order to identify and remove strangers who are

already listed as friends with the user, we recommend OSN users examine their friends list, or use applications such as the Social Privacy Protector [40], and periodically remove friends whom they are not familiar with or friends, such as exes, who should not have access to personal information. Additionally, if parents want to better monitor their children's activity, it is also strongly recommended that they and their children scan the friends lists together in order to remove unwelcomed "friends."

4. Install Internet Security Software. We advise OSN users to install at least one of the many commercial Internet security software products; Facebook offers several free security downloads [33].

5. Remove Installed Third-Party Applications. Unbeknown to many users, third-party applications are frequently collecting online personal data. A recent study showed that 30% of an examined group of Facebook users had at least forty applications installed on their accounts [41]. It is recommended that a user not install new unnecessary applications on his or her account. Moreover, users are advised to go over their list of installed applications and remove any unnecessary applications. 6. Do Not Trust Your OSN Friends. As we described in Section 3.3, OSNs users tend to naively trust their friends in the social network. This trust has a negative side, as was demonstrated in the case of Abigail Pickett[53]. Therefore, we recommend OSN users take extra precautions when communicating with their online friends. We also recommend users to think twice before offering any personal and sensitive information about themselves, even when posting photos. Online social network users should definitely avoid revealing their home address, phone number, or credit cards numbers.

6. CONCLUSIONS

Online social networks have become part of our everyday life and, on average, most Internet users spend more time on social networks than in any other online activity. We enjoy using online social networks to interact with other people through the sharing of experiences, pictures, videos, and other types of information. Nevertheless, online social networks also have a dark side ripe with hackers, fraudsters, and online predators, all capable of using online social networks as a platform for procuring their next victim. In this paper, we have presented scenarios which threaten online social network users and can jeopardize their identities, privacy, and well-being both in the virtual world and in the real world. In addition, we have provided examples for many of the presented threats in order to demonstrate that these threats are real and can endanger every user. Moreover, we have emphasized certain threats which challenge the safety of children and teenagers inside the online social network cyberspace. There are remedies to these threats, and we have offered a range of solutions which aim to protect the online social network user's privacy and security (see Section 4). We have outlined some recommendations that are simple to implement for OSN users to better protect themselves. However, as demonstrated in Table 1, the presented solutions are not magical antidotes which provide full protection to a user's privacy and security. In order to be well protected against the various online threats, users need to employ more than one solution, and in many cases must count on the OSN provider's assistance in providing tools both to better protect their privacy and to identify the threats. Therefore, we recommend that online social network users not only adopt our recommendations, but also educate themselves and their loved ones about online threats which can entice users and potentially end disastrously. All social network users must very carefully consider what personal information is being revealed about themselves, about their friends, and about their workplace. Moreover, as parents, we are obligated to educate our children to be aware of potential threats and teach them not to engage with strangers either in the real world or in the cyber world. If a user's personal information falls into the wrong hands, it can cause a vast amount of damage, and in many cases there is no way to recapture what has been lost. The field of online social network security and privacy is a new and emerging one, filled with many directions to follow. Security researchers can continually offer better solutions to online threats; they can also discover new security threats to address. We believe that in order to improve the present solutions, the next step is to create synergy among the different security solutions. This will create more robust and effective security solutions for detecting fake profiles, spammers, phishing attacks, socware, and other threats. Another possible direction is to utilize various Natural

Language Processing (NLP) techniques and temporal analysis algorithms, combining them with existing solutions to provide better and more accurate protection against online social networks threats. For example, researchers can predict many users' private traits, such as age and gender, based on their Facebook likes [55]. Combining this algorithm with other topological-based fake profile detection methods can assist in spotting fake details, such as a false age, thus identifying fake profiles. An additional possible future research direction includes studying the emerging security threats due to the increasing popularity of geo-location tagging of social network users and offering solutions for threats unique to geosocial networking. All big social media security systems were analyzed in this research and existed and new ways of defense were suggested. Overall, researches can play a significant role by recognizing the value of solution synergies, by applying useful techniques and algorithms, and by understanding unique threat situations within online social networks.

7. REFERENCES

1. Key Facts - Facebook Newsroom. Internet: http://newsroom.fb.com/Key-Facts (Jan, 2014).

2. Joinson, A. “Looking at”, “looking up” or “keeping up with” people?: Motives and use of Facebook. In Proc. CHI 2008, ACM (2008), pp.1027-1036.

3. Spiliotopoulos T., Oakley I. Understanding Motivations for Facebook Use: Usage Metrics, Network Structure, and Privacy. Internet: http://tasos-spiliotopoulos.com/publications_assets/CHI13-Spiliotopoulos-FacebookMotivations.pdf (May, 2014).

4. Lampe, C., Vitak, J., Gray, R., and Ellison, N. Perceptions of facebook's value as an information source. In Proc. CHI 2012, ACM (2012).

5. Lampe, C., Wash, R., Velasquez, A., and Ozkaya, E. Motivations to participate in online communities. In Proc. CHI 2010, ACM (2010), pp.1927–1936.

6. Katz, E., Gurevitch, M., and Haas, H. On the use of the mass media for important things. American Sociological Review 38, (1973), pp.164-181.

7. Ikhalia E.J. A New Social Media Security Model (SMSM). International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 7, July 2013).

8. Imafidon, C.O, Ikhalia, E. The investigation and implementation of social media security‖. Proceedings of the 2nd global conference, London on communication information science and engineering (2013).

9. Devmane, M. A., Rana, N. K. Security Issues of Online Social Networks‖. Advances in Computing, Communication, and Control Communications in Computer and Information Science. 361 (1), pp. 740-746 (2013).

10. Lenkart, J. J. The Vulnerability of Social Networking Media and the Insider Threat: New Eyes for Bad Guys‖. MA Thesis. Monterey, California: naval postgraduate school (2011).

11. Kumar, A., Gupta, S.K., Rai, A. K. ,Sinha, S. Social Networking Sites and Their Security Issues‖. International Journal of Scientific and Research Publications, Volume 3, Issue 4, April 2013, pp.1-5.

12. Kontaxis, G., Polakis I., Ioannidis S., Markatos, E.P. Detecting Social Network Profile Cloning. 3rd International Workshop on Security and Social Networking, pp.295-300 (2011).

13. Alexa. Facebook.com—site info from Alexa. Internet: http://www.alexa.com/siteinfo/facebook.com (May, 2014).

14. Bailey, J. Life in the fishbowl: Feminist interrogations of webcamming. In Lessons from the identity trail: Anonymity, privacy and identity in a networked society. Oxford: OUP, pp. 283–301 (2009).

15. Benkler, Y. Through the looking glass: Alice and the constitutional foundations of the public domain. Law and Contemporary Problems, 66, pp. 173–224 (2003).

16. Hull G. Contextual gaps: privacy issues on Facebook. Ethics Inf Technol (2011), pp.289-302.

17. Lipford, H. R., Hull, G., Latulipe, C., Besmer, A., & Watson, J. Visual flows: Contextual integrity and the design of privacy mechanisms on social network sites. In Proceedings of the Workshop on Security and Privacy in Online Social Networking, IEEE International Conference on Social Computing (SocialCom) (Aug, 2009).

18. Keelty M. Social Media, Privacy and Security: Identifying and managing the public policy risks. Internet: https://crawford.anu.edu.au/public_policy_community/workshops/social_media_privacy_and_security/Social_Media_Privacy_and_Security_summary.pdf (May,2014).

19. Magid L. Facebook Groups Can Jeopardize Privacy. Internet: http://www.cubagreenscreen.com/forum/showthread.php?tid=10731, accessed (2010).

20. Yahoo Finance (2011). Apple slammed over iPhone, iPad location tracking (April, 2014).

21. World Facebook connections (2010). Internet: http://mashable.com/2010/12/13/facebook-membersvisualization/?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+Mashable+%28Mashable%29 (April, 2014).

22. Acquisti A., Gross R. Imagined Communities:Awareness, Information Sharing, and Privacy on the Facebook. Privacy Enhancing Technologies Workshop (PET), 2006.

23. Parker, R.: Alcohol policy violated. Kansan.com February 28 (2006)

24. Youngwood, S.: Networking by the ‘book’. The Times Argus February 26 (2006).

25. Kharif, O.: Big brother is reading your blog. BusinessWeek online February 28 (2006).

26. Gross, R., Acquisti, A.: Privacy and information revelation in online social networks. In: Proceedings of the ACM CCS Workshop on Privacy in the Electronic Society (WPES ’05). (2005).

27. D. boyd: Reflections on friendster, trust and intimacy. In: Intimate (Ubiquitous) Computing Workshop – Ubicomp 2003, October 12-15, Seattle, Washington, USA. (2003).

28. D. boyd: Friendster and publicly articulated social networking. In: Conference on Human Factors and Computing Systems (CHI 2004), April 24-29, Vienna, Austria. (2004).

29. Donath, J., D. boyd: Public displays of connection. BT Technology Journal 22 (2004) 71–82.

30. Liu, H., Maes, P.: Interestmap: Harvesting social network profiles for recommendations. In: Beyond Personalization - IUI 2005, January 9, San Diego, California, USA. (2005).

31. Social Media Compliance Policy. Internet: http://www.liv.ac.uk/media/livacuk/computingservices/regulations/socialmediapolicy.pdf (2014).

32. Omar Saeed Al Mushayt. Threats and Anti-threats Strategies for Social Networking Websites. International Journal of Computer Networks & Communications (IJCNC) Vol.5, No.4, July.

33. Social networking sites searchengine. Internet: /http://findasocialnetwork. com/search.php.

34. B. Stone. Is Facebook growing up too fast. The New York Times (March 29, 2009).

35. “Using Facebook to Social Engineer Your Way Around Security”. Internet: http://www.eweek.com/c/a/Security/Social-Engineering-Your-Way-Around-Security-With-Facebook-277803/ (2010).

36. Won Kim , Ok-Ran Jeong, Sang-Won Lee. On Social Websites. Information Systems 35 (2010), pp.215-236.

37. Facebook Delays New Privacy Policy. Internet: http://bits.blogs.nytimes.com/2013/09/05/facebook-delays-new-privacy-policy/?_php=true&_type=blogs&_php=true&_type=blogs&_r=1 (2014).

38. Daniele Quercia. Facebook and Privacy: The Balancing Act of Personality, Gender, and Relationship Currency. Internet: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/viewFile/4613/4997 (2014).

39. Social media and the law: A handbook for UK companies. Internet: http://www.linklaters.com/Publications/Social-media-law-A-handbook-UK-companies/Pages/Index.aspx (2014).

40. Lars Kaczmirek. Social Media Monitoring of The Campaigns for the 2013 German Bundestag Elections on Facebook and Twitter. Internet: http://arxiv.org/abs/1312.4476 (2014).

41. Glenn Greenwald. XKeyscore: NSA tool collects 'nearly everything a user does on the internet' Internet: http://www.theguardian.com/world/2013/jul/31/nsa-top-secret-program-online-data/print (2013).

42. Staff. XKeyscore Presentation from 2008 – Read in Full. The Guardian (July 31, 2013).

43. Norddeutscher Rundfunk. Snowden Interview Transcript. (27 January 2014).

44. Gallagher, Sean. NSA's Internet Taps Can Find Systems to Hack, Track VPNs and Word Docs – X-Keyscore Gives NSA the Ability to Find and Exploit Vulnerable Systems. Ars Technica. (August 4, 2013).

45. Mr Snowden interview. Internet: http://www.tagesschau.de/snowden-interview-englisch100.pdf (2014).

46. Dulles, Allen Welsh. The Craft of Intelligence, New York: Harper&Row (1963).

47. European Commission. Proposal for a General Data Protection Regulation (2011).

48. European Commission. Proposal for a General Data Protection Regulation, 25.1.2012, COM (2012).

49. European Commissioner - Reding, Viviane, Letter to the Attorney General, Ref. Ares (2013).

50. European Data Protection Supervisor - Hustinx, Peter. Data Protection and Cloud Computing Under EU Law, speech, Third European Cyber Security Awareness Day, BSA, European Parliament, 13 April 2010, Panel IV: Privacy and Cloud Computing.

51. The US surveillance programmes and their impact on EU citizens' fundamental rights. Internet: http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/dv/briefingnote_/briefingnote_en.pdf (2014).

52. White noise: An idea to defend privat sphere against XKeyscore. Internet: http://futuretechblog.com/?p=266 (2014).

53. Markus Ottela. TFC. Internet: http://www.cs.helsinki.fi/u/oottela/TFC.pdf (2014).

54. Report: NSA XKeyscore program allows access to emails, chats without warrant. Internet: http://thehill.com/policy/technology/314659-report-nsas-xkeyscore-program-allows-warrantless-internet-monitoring (2014).

55. Researcher argues for open hardware to defend against NSA spying. Internet: http://www.csoonline.com/article/2134047/network-security/researcher-argues-for-open-hardware-to-defend-against-nsa-spying.html (2014).

56. Michael Fire. Online Social Networks: Threats and Solutions (2014).

57. Pritam Gundecha. Exploiting Vulnerability to Secure User Privacy on a Social Networking Site. Internet: http://www.researchgate.net/publication/220272037_Exploiting_vulnerability_to_secure_user_privacy_on_a_social_networking_site/file/5046351a386dceb101.pdf (2014).

58. How to Protect Yourself against Online Spying, NSA PRISM/Fairview can spy on anyone. Internet: http://davidpapp.com/2013/09/05/how-to-protect-yourself-against-online-spying-nsa-prismfairview-can-spy-on-anyone/ (2014).

Comparative Analysis of Open-SSL Vulnerabilities & Heartbleed Exploit Detection

Eng. Mohanned Hassan Momani *m.hassan@securitywits.com*
Sr. Information Security & Technology Consultant
IT security trainer C|EI
Security Wits Technologies
Jordan

Adam Ali.Zare Hudaib *adamhudaib@gmail.com*
Information & Cyber Security Expert
Licensed Penetration Tester L|PT
Sweden

Abstract

Since its introduction in 1994 the Secure Socket Layer (SSL) protocol (later renamed to Transport Layer Security (TLS)) evolved to the de facto standard for securing the transport layer. SSL/TLS can be used for ensuring data confidentiality, integrity and authenticity during transport. A main feature of the protocol is its flexibility. Modes of operation and security aims can easily be configured through different cipher suites. During its evolutionary development process several flaws were found. However, the flexible architecture of SSL/TLS allowed efficient fixes in order to counter the issues. This paper presents an overview on theoretical and practical attacks of the last 20 years.

Keywords: SSL, TLS, BEAST Attack, CRIME Attack, Heartbleed Detection, RC4.

1. INTRODUCTION

In the last few years, we have witnessed a wide range of attacks on the SSL/TLS mechanism. In this article, we will try to cover various attacks that were prominent in the field of cryptography. Transport layer security (TLS) ensures integrity of data transmitted between two parties (server and client) and also provides strong authentication for both parties. The attacks launched in the last few years have exploited various features in the TLS mechanism. We are going to discuss these attacks one by one.

The Heartbleed bug is a serious vulnerability in the popular OpenSSL cryptographic software library. It allows an attacker to read the memory of systems using certain versions of OpenSSL, potentially allowing them to access user names, passwords, or even the secret cryptographic keys of the server used for SSL. Obtaining these keys would allow malicious users to observe all communications on that system, allowing further exploit. We will discuss this vulnerability too.

2. TLS, ATTACKS AND ANALYSIS

2.1 SSL Protocol and Types

Secure Sockets Layer, more commonly known as SSL, is a protocol that is used to maintain client and server authentication. A site is easily identified as using SSL if it has the small yellow padlock at the bottom of the browser.

In SSL, communication between the server and the client is encrypted using their certificates. This encryption creates virtual information that is not hackable by others. The steps of how SSL works is shown in the following diagram:

SSL Version 1, a test version, was quickly replaced by SSL version 2, which was the first version, released to the public and was shipped with the Netscape Navigator browser. Today version 2 is still supported despite having some security problems. Later, Microsoft came out with its own version of SSL called PCT. SSL Version 3 is a complete redesign of SSL and fixes the problems found in previous versions as well as having additional features [1].

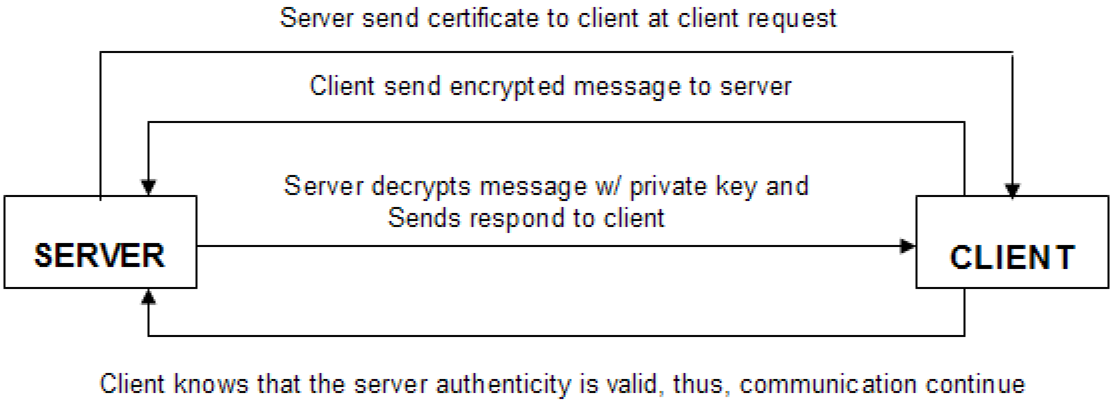

Server send certificate to client at client request
Client send encrypted message to server
Server decrypts message w/ private key and
Sends respond to client
SERVER
CLIENT
Client knows that the server authenticity is valid, thus, communication continue

FIGURE 1: SSL Algorithm.

SSL uses a handshake protocol. Suppose a client wants to make a purchase from a website server, but this server does not know anything about the client.

The first step is for the client to send a message to the server. After the server receives the message, it acknowledges it by sending the client a message in return. The server also sends the client its certificate and asks for the client's certificate. The client sends its certificate, a client key exchange message, and a certificate verification message. Both the client and server send change cipher spec messages and then send finished messages to end the handshake [2].

A website implements SSL by using HTTPS, which stands for Hypertext Transfer Protocol over Secure Socket Layer. This web protocol was developed by Netscape to encrypt and decrypt page requests as well as the pages that are returned by the web server. HTTPS uses port 443 instead of port 80, which is used for HTTP.

SSL uses a key size of 40-bits for the RC4 stream encryption algorithm. This is considered a sufficient degree of encryption for commercial exchange. Both HTTPS and SSL support the use of X.509 digital certificates from the server. This way, the user can authenticate the sender if needed [3].

One of SSL's strengths is its ability to help prevent some common attacks. SSL is strong against the brute force attack because it uses 128 bits. The dictionary attack which tends to be more efficient than a brute force attack is where an attack tries every word in a dictionary as a possible password for an encrypted message. This attack is also avoidable because SSL has very large key spaces. The replay attack which reruns messages that were sent earlier is prevented since SSL uses 128-bit nonce value to indicate a unique connection. And as mentioned earlier, the Man-In-the-Middle Attack is prevented by using signed certificates to authenticate the server's public key.

Despite the fact that SSL has the ability to prevent some common attacks, it still has some weaknesses. One of the weaknesses found in SSL is the brute force attack against weak ciphers. This weakness was forced by the US export on Netscape. This weakness still remains one of the most obvious weaknesses of the SSL protocol and it has broken many times [4].

Another weakness in SSL is the renegotiation of the master key. It is known that after a connection has been established, the same master key gets used all the way through the connection. This could be a serious security flaw if SSL are layered underneath a long running connection. One possible solution for this flaw is to force renegotiation of the master key at different times. This way, the difficulty and the cost of the any brute force attack will be multiplied by the number of times that the master key has changed [5].

The Transaction Layer Security protocol, commonly known as TLS, is based on SSL and became its successor. TLS has some changes in its MAC, has clearer and more precise specifications, cleaner handling because of not having a client certificate, and more flexibility.

TLS is an Internet Engineering Task Force (IETF) standards track protocol, first defined in 1999 and last updated in RFC 5246 (August 2008) and RFC 6176 (March 2011). It is based on the earlier SSL specifications (1994, 1995, 1996) developed by Netscape Communications[1] for adding the HTTPS protocol to their Navigator web browser.

Early research efforts towards transport layer security included the Secure Network Programming (SNP) application programming interface (API), which in 1993 explored the approach of having a secure transport layer API closely resembling Berkeley sockets, to facilitate retrofitting preexisting network applications with security measures [1].

The SSL protocol was originally developed by Netscape. Version 1.0 was never publicly released; version 2.0 was released in February 1995 but "contained a number of security flaws which ultimately led to the design of SSL version 3.0."[2] SSL version 3.0, released in 1996, was a complete redesign of the protocol produced by Paul Kocher working with Netscape engineers Phil Karlton and Alan Freier. Newer versions of SSL/TLS are based on SSL 3.0. The 1996 draft of SSL 3.0 was published by IETF as a historical document in RFC 6101.

The basic algorithm was written by Dr. Taher Elgamal. As the Chief Scientist of Netscape, Taher was recognized as the "father of SSL" [2].

TLS 1.0 was first defined in RFC 2246 in January 1999 as an upgrade of SSL Version 3.0. As stated in the RFC, "the differences between this protocol and SSL 3.0 are not dramatic, but they are significant to preclude interoperability between TLS 1.0 and SSL 3.0." TLS 1.0 does include a means by which a TLS implementation can downgrade the connection to SSL 3.0, thus weakening security.TLS 1.1 was defined in RFC 4346 in April 2006 [3]. It is an update from TLS version 1.0. Significant differences in this version include:

- Added protection against Cipher block chaining (CBC) attacks.
- The implicit Initialization Vector (IV) was replaced with an explicit IV.
- Change in handling of padding errors.
- Support for IANA registration of parameters.

TLS 1.2 was defined in RFC 5246 in August 2008. It is based on the earlier TLS 1.1 specification. Major differences include:

- The MD5-SHA-1 combination in the pseudorandom function (PRF) was replaced with SHA-256, with an option to use cipher suite specified PRFs.
- The MD5-SHA-1 combination in the Finished message hash was replaced with SHA-256, with an option to use cipher suite specific hash algorithms. However the size of the hash in the finished message is still truncated to 96-bits.
- The MD5-SHA-1 combination in the digitally signed element was replaced with a single hash negotiated during handshake, defaults to SHA-1.
- Enhancement in the client's and server's ability to specify which hash and signature algorithms they will accept.

- Expansion of support for authenticated encryption ciphers, used mainly for Galois/Counter Mode (GCM) and CCM mode of Advanced Encryption Standard encryption.
- TLS Extensions definition and Advanced Encryption Standard cipher suites were added [4].

All TLS versions were further refined in RFC 6176 in March 2011 removing their backward compatibility with SSL such that TLS sessions will never negotiate the use of Secure Sockets Layer (SSL) version 2.0.

2.2 Analysis of the Internet Protocol TLS

The TLS protocol allows client-server applications to communicate across a network in a way designed to prevent eavesdropping and tampering.

Since protocols can operate either with or without TLS (or SSL), it is necessary for the client to indicate to the server the setup of a TLS connection. There are two main ways of achieving this. One option is to use a different port number for TLS connections (for example port 443 for HTTPS). The other is for the client to request that the server switch the connection to TLS using a protocol-specific mechanism (for example STARTTLS for mail and news protocols).

Once the client and server have agreed to use TLS, they negotiate a stateful connection by using a handshaking procedure [5]. During this handshake, the client and server agree on various parameters used to establish the connection's security:

1. The client sends the server the client's SSL version number, cipher settings, session-specific data, and other information that the server needs to communicate with the client using SSL.
2. The server sends the client the server's SSL version number, cipher settings, session-specific data, and other information that the client needs to communicate with the server over SSL. The server also sends its own certificate, and if the client is requesting a server resource that requires client authentication, the server requests the client's certificate.
3. The client uses the information sent by the server to authenticate the server—e.g., in the case of a web browser connecting to a web server, the browser checks whether the received certificate's subject name actually matches the name of the server being contacted, whether the issuer of the certificate is a trusted certificate authority, whether the certificate has expired, and, ideally, whether the certificate has been revoked. If the server cannot be authenticated, the user is warned of the problem and informed that an encrypted and authenticated connection cannot be established. If the server can be successfully authenticated, the client proceeds to the next step.
4. Using all data generated in the handshake thus far, the client (with the cooperation of the server, depending on the cipher in use) creates the pre-master secret for the session, encrypts it with the server's public key (obtained from the server's certificate, sent in step 2), and then sends the encrypted pre-master secret to the server.
5. If the server has requested client authentication (an optional step in the handshake), the client also signs another piece of data that is unique to this handshake and known by both the client and server. In this case, the client sends both the signed data and the client's own certificate to the server along with the encrypted pre-master secret.
6. If the server has requested client authentication, the server attempts to authenticate the client. If the client cannot be authenticated, the session ends. If the client can be successfully authenticated, the server uses its private key to decrypt the pre-master secret, and then performs a series of steps (which the client also performs, starting from the same pre-master secret) to generate the master secret.
7. Both the client and the server use the master secret to generate the session keys, which are symmetric keys used to encrypt and decrypt information exchanged during the SSL session and to verify its integrity (that is, to detect any changes in the data between the time it was sent and the time it is received over the SSL connection).
8. The client sends a message to the server informing it that future messages from the client will be encrypted with the session key. It then sends a separate (encrypted) message indicating that the client portion of the handshake is finished.

9. The server sends a message to the client informing it that future messages from the server will be encrypted with the session key. It then sends a separate (encrypted) message indicating that the server portion of the handshake is finished.
10. The SSL handshake is now complete and the session begins. The client and the server use the session keys to encrypt and decrypt the data they send to each other and to validate its integrity [5].

This is the normal operation condition of the secure channel. At any time, due to internal or external stimulus (either automation or user intervention), either side may renegotiate the connection, in which case, the process repeats itself [2].

This concludes the handshake and begins the secured connection, which is encrypted and decrypted with the key material until the connection closes.

If any one of the above steps fails, the TLS handshake fails and the connection is not created.

In step 3, the client must check a chain of "signatures" from a "root of trust" built into, or added to, the client. The client must also check that none of these have been revoked; this is not often implemented correctly, but is a requirement of any public-key authentication system. If the particular signer beginning this server's chain is trusted, and all signatures in the chain remain trusted, then the Certificate (thus the server) is trusted.

The TLS protocol exchanges records—which encapsulate the data to be exchanged in a specific format. Each record can be compressed, padded, appended with a message authentication code (MAC), or encrypted, all depending on the state of the connection. Each record has a content type field that designates the type of data encapsulated, a length field and a TLS version field. The data encapsulated may be control or procedural messages of the TLS itself, or simply the application data needed to be transferred by TLS. The specifications (cipher suite, keys etc.) required to exchange application data by TLS, are agreed upon in the "TLS handshake" between the client requesting the data and the server responding to requests. The protocol therefore defines both the structure of payloads transferred in TLS and the procedure to establish and monitor the transfer.

When the connection starts, the record encapsulates a "control" protocol—the handshake messaging protocol (content type 22). This protocol is used to exchange all the information required by both sides for the exchange of the actual application data by TLS. It defines the messages formatting or containing this information and the order of their exchange. These may vary according to the demands of the client and server—i.e., there are several possible procedures to set up the connection. This initial exchange results in a successful TLS connection (both parties ready to transfer application data with TLS) or an alert message (as specified below).

A simple connection example follows, illustrating a handshake where the server (but not the client) is authenticated by its certificate:

1. Negotiation phase:
    - A client sends a ClientHello message specifying the highest TLS protocol version it supports, a random number, a list of suggested CipherSuites and suggested compression methods. If the client is attempting to perform a resumed handshake, it may send a session ID.
    - The server responds with a ServerHello message, containing the chosen protocol version, a random number, CipherSuite and compression method from the choices offered by the client. To confirm or allow resumed handshakes the server may send a session ID. The chosen protocol version should be the highest that both the client and server support. For example, if the client supports TLS1.1 and the server supports TLS1.2, TLS1.1 should be selected; SSL 3.0 should not be selected.

- The server sends its Certificate message (depending on the selected cipher suite, this may be omitted by the server).
- The server sends a ServerHelloDone message, indicating it is done with handshake negotiation.
- The client responds with a ClientKeyExchange message, which may contain a PreMasterSecret, public key, or nothing. (Again, this depends on the selected cipher.) This PreMasterSecret is encrypted using the public key of the server certificate.
- The client and server then use the random numbers and PreMasterSecret to compute a common secret, called the "master secret". All other key data for this connection is derived from this master secret (and the client- and server-generated random values), which is passed through a carefully designed pseudorandom function.

2. The client now sends a ChangeCipherSpec record, essentially telling the server, "Everything I tell you from now on will be authenticated (and encrypted if encryption parameters were present in the server certificate)." The ChangeCipherSpec is itself a record-level protocol with content type of 20.
   - Finally, the client sends an authenticated and encrypted Finished message, containing a hash and MAC over the previous handshake messages.
   - The server will attempt to decrypt the client's Finished message and verify the hash and MAC. If the decryption or verification fails, the handshake is considered to have failed and the connection should be torn down.
3. Finally, the server sends a ChangeCipherSpec, telling the client, "Everything I tell you from now on will be authenticated (and encrypted, if encryption was negotiated)."
   - The server sends its authenticated and encrypted Finished message.
   - The client performs the same decryption and verification.
4. Application phase: at this point, the "handshake" is complete and the application protocol is enabled, with content type of 23. Application messages exchanged between client and server will also be authenticated and optionally encrypted exactly like in their Finished message. Otherwise, the content type will return 25 and the client will not authenticate.

Public key operations (e.g., RSA) are relatively expensive in terms of computational power. TLS provides a secure shortcut in the handshake mechanism to avoid these operations: resumed sessions. Resumed sessions are implemented using session IDs or session tickets.

Apart from the performance benefit, resumed sessions can also be used for single sign-on as it is guaranteed that both the original session as well as any resumed session originate from the same client. This is of particular importance for the FTP over TLS/SSL protocol which would otherwise suffer from a man-in-the-middle attack in which an attacker could intercept the contents of the secondary data connections.

In an ordinary full handshake, the server sends a session id as part of the ServerHello message. The client associates this session id with the server's IP address and TCP port, so that when the client connects again to that server, it can use the session id to shortcut the handshake. In the server, the session id maps to the cryptographic parameters previously negotiated, specifically the "master secret". Both sides must have the same "master secret" or the resumed handshake will fail (this prevents an eavesdropper from using a session id). The random data in the ClientHello and ServerHello messages virtually guarantee that the generated connection keys will be different from in the previous connection. In the RFCs, this type of handshake is called an abbreviated handshake. It is also described in the literature as a restart handshake.

RFC 5077 extends TLS via use of session tickets, instead of session IDs. It defines a way to resume a TLS session without requiring that session-specific state is stored at the TLS server.

When using session tickets, the TLS server stores its session-specific state in a session ticket and sends the session ticket to the TLS client for storing. The client resumes a TLS session by sending the session ticket to the server, and the server resumes the TLS session according to the session-specific state in the ticket. The session ticket is encrypted and authenticated by the server, and the server verifies its validity before using its contents.

One particular weakness of this method is that it always limits encryption and authentication security of the transmitted TLS session ticket to AES128-CBC-SHA256, no matter what other TLS parameters were negotiated for the actual TLS session. This means that the state information (the TLS session ticket) is not as well protected as the TLS session itself. Of particular concern is OpenSSL's storage of the keys in an application-wide context (SSL_CTX), i.e. for the life of the application, and not allowing for re-keying of the AES128-CBC-SHA256 TLS session tickets without resetting the application-wide OpenSSL context (which is uncommon, error-prone and often requires manual administrative intervention) [5].

2.3 Transport Layer Security (TLS)

TLS has a variety of security measures:

- Protection against a downgrade of the protocol to a previous (less secure) version or a weaker cipher suite.
- Numbering subsequent Application records with a sequence number and using this sequence number in the message authentication codes (MACs).
- Using a message digest enhanced with a key (so only a key-holder can check the MAC). The HMAC construction used by most TLS cipher suites is specified in RFC 2104 (SSL 3.0 used a different hash-based MAC).
- The message that ends the handshake ("Finished") sends a hash of all the exchanged handshake messages seen by both parties.
- The pseudorandom function splits the input data in half and processes each one with a different hashing algorithm (MD5 and SHA-1), then XORs them together to create the MAC. This provides protection even if one of these algorithms is found to be vulnerable [5].

2.4 Heartbleed OpenSSL Vulnerability

The Heartbleed bug is a serious vulnerability in the popular OpenSSL cryptographic software library, affecting versions 1.0.1 to 1.0.1f. This weakness allows stealing the information protected, under normal conditions, by the SSL/TLS encryption used to secure the data payloads. SSL/TLS provides communication security and privacy over the Internet for applications such as web, email, instant messaging (IM) and some virtual private networks (VPNs).

The Heartbleed bug allows anyone on the Internet to read the memory of the systems protected by the vulnerable versions of the OpenSSL software. This compromises the secret keys used to identify the service providers and to encrypt the traffic, the names and passwords of the users and the actual content. This allows attackers to eavesdrop on communications, steal data directly from the services and users and to impersonate services and users.

On April 7 th 2014 OpenSSL and a team of security engineers published advisories regarding a severe vulnerability that "allows anyone on the Internet to read the memory of systems protected by vulnerable versions of the OpenSSL software". They have dubbed this vulnerability "Heartbleed" as it refers to a memory leak in a heartbeat function used by OpenSSL. SSL and TLS are cryptographic protocols designed to secure communications over the internet by way of certificates and asymmetric cryptography. This is implemented in conjunction with Certificate Authorities (CA) and Public Key Infrastructure (PKI). Collectively this forms the basis upon which trust is established on the Internet. For the non-technical person, these services are commonly associated with the acronym 'HTTPS' which enables secure online commerce and authentication. There are several different versions of SSL and unfortunately OpenSSL stands as one of the most commonly implemented versions on the Internet today. OpenSSL is bundled with many different operating systems, embedded systems, networked appliances, chat servers, VPNs, e-mail servers and client software making this vulnerability extremely wide reaching and dangerous. OpenSSL is deployed with open source web servers such as Apache and NGINX which account for over 50% of active sites on the internet [6].

The Heartbleed attack relies on what is considered a relatively simple programming error that when exploited allows attackers to read up to 64 KB of memory. Specifically the dtls1_process_heartbeat function contains values which are assigned in memory without performing correct error bounds checking. This allows an attacker to craft conversations with an OpenSSL Client or Server that reads outside of properly allocated memory. Due to the fact that OpenSSL handles account, certificate, and key information, reading in to memory can reveal extremely sensitive data. However, one of the discoverers of the vulnerability recently tweeted in an apparent attempt to allay concerns stating “heap allocation patterns make private key exposure unlikely for #heartbleed #dontpanic”. Although correct, this is highly dependent on how committed an attacker is in their efforts. Prolonged and recurring exploitation of this vulnerability against a system radically increases the likelihood of exposing private key information. The heartbeat function does serve a legitimate purpose in that it allows both parties in a communication channel to maintain a session while no longer actively exchanging data. There are multiple proof-of-concepts available in the wild demonstrating exploitation techniques against this vulnerability [7].

Researchers have classified the type of information being leaked into four categories:

1. Primary Key Material – encryption keys are leaked allowing attackers to inspect confidential traffic and impersonate the service.
2. Secondary Key Material – user account and password information can be stolen
3. Protected Content – actual confidential information contained within a previously assumed secure communication channel is exposed.
4. Collateral – incidental information gleaned during the attack with regards to OpenSSL implementation specifics and architecture [8].

Unfortunately, exploitation of this vulnerability does not record log evidence that can be used as an indicator of attack. However, IDS/IPS systems may be able to detect malicious heartbeat request/response communications based on the record type (and size) contained within the protocol. As described in the software vulnerability section above, detection is possible by comparing the size of a request against its reply.

To elaborate, systems with packet inspection capabilities (IDS/IPS, Analytics, Proxy) can look for request and response packets containing matches to specific hexadecimal values for different TLS versions. One must also factor in the size of the packet in order to reduce false positives and avoid simply identifying legitimate heartbeat communication [9].

Some operating system distributions that have shipped with potentially vulnerable OpenSSL version:

1. Debian Wheezy (stable), OpenSSL 1.0.1e-2+deb7u4.
2. Ubuntu 12.04.4 LTS, OpenSSL 1.0.1-4ubuntu5.11.
3. CentOS 6.5, OpenSSL 1.0.1e-15.
4. Fedora 18, OpenSSL 1.0.1e-4.
5. OpenBSD 5.3 (OpenSSL 1.0.1c 10 May 2012) and 5.4 (OpenSSL 1.0.1c 10 May 2012).
6. FreeBSD 10.0 – OpenSSL 1.0.1e 11 Feb 2013.
7. NetBSD 5.0.2 (OpenSSL 1.0.1e).
8. OpenSUSE 12.2 (OpenSSL 1.0.1c).

Operating system distribution with versions that are not vulnerable:

1. Debian Squeeze (oldstable), OpenSSL 0.9.8o-4squeeze14.
2. SUSE Linux Enterprise Server.
3. FreeBSD 8.4 – OpenSSL 0.9.8y 5 Feb 2013.
4. FreeBSD 9.2 – OpenSSL 0.9.8y 5 Feb 2013.
5. FreeBSD 10.0p1 – OpenSSL 1.0.1g (At 8 Apr 18:27:46 2014 UTC).
6. FreeBSD Ports – OpenSSL 1.0.1g (At 7 Apr 21:46:40 2014 UTC) [10].

That's a lot of system's that are vulnerable. We all thought Linux is secured and open source being the bearer of security flag, apparently not anymore!

There are several versions of the Heartbleed exploit actively in the wild, some are simply being used to test if systems are vulnerable, as well as more robust versions available in Metasploit and other frameworks. To watch potential exploits come through I have left a honeypot website purposely vulnerable to the Heartbleed bug, with a script that loads fake password and other random seemingly juicy data files into RAM [14].

To get a better picture of the Heartbleed vulnerability in our environment, we can use the full Tripwire suite. Tripwire IP360 provides reporting on the state of the vulnerability in your environment. Tripwire Log Center provides a guard dog on your network looking for indicators of Heartbleed exploits in real-time from IDS and other systems.

If we bring the two products together as well, when a Heartbleed exploit against a host is detected targeting a host, Tripwire Log Center can lookup vulnerability data on that host to better understand the risk. If the system attacked is vulnerable you can fire off alerts to your team, or activate scripts to automate remediation and counter measures in real-time (Fig.2).

Also there is another technique to detect Heartbleed vulnerability. This technique uses a BPF packet filter to automatically flag larger-than-typical TLS heartbeat responses from the server, and can be used with Wireshark and tcpdump as well as with the Riverbed AppResponse, Shark, and Pilot products. (AppResponse and Shark support many terabytes of stored packets, coupled with the ability to quickly analyze those packets; more Riverbed product specific hints will follow in a separate blog post) [15].

In addition, for the majority of published Heartbleed exploits so far (which are moving the compromised data in the clear on the wire), this technique also identifies what exact data was compromised (e.g., the set of user passwords exposed, vs. the "crown jewels" of a server's private keys).

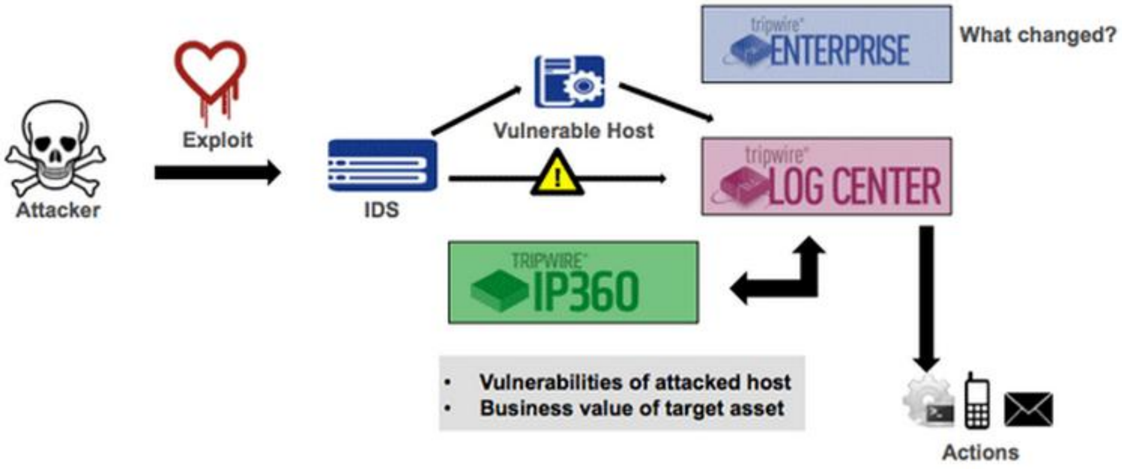


FIGURE 2: Heartbleed Vulnerability [14].

2.5 BEAST Attack

The ability to mount an adaptive chosen plaintext attack with predictable initialization vectors (IVs) against SSL/TLS using cipher block chaining (CBC) was known in 2004, but until late 2011 was thought to be largely theoretical.

Researchers Thai Doung and Juliano Rizzo found a way to exploit the vulnerability and demonstrated a live attack against Paypal at the Ekoparty security conference in September of 2011. Doung and Rizzo had notified and had been working with major software vendors including

Mozilla and Google to release a patch (CVE-2011-3389). Although the vulnerability is cryptographic in nature, it requires certain conditions to be successful. The proof of concept code presented at the conference also required a Java-based Same Origin Policy (SOP) bypass that they had found during their research and which has been patched by Oracle [16].

BEAST leverages a type of cryptographic attack called a chosen-plaintext attack. The attacker mounts the attack by choosing a guess for the plaintext that is associated with a known ciphertext. To check if a guess is correct, the attacker needs access to an encryption oracle [19] to see if the encryption of the plaintext guess matches the known ciphertext. To defeat a chosen-plaintext attack, popular configurations of TLS use two common mechanisms: an initialization vector (IV) and a cipher block chaining mode (CBC). An IV is a random string that is XORed with the plaintext message prior to encryption — even if you encrypt the same message twice, the ciphertext will be different, because the messages were each encrypted with a different random IV. The IV is not secret; it just adds randomness to messages, and is sent along with the message in the clear. It would be cumbersome to use and track a new IV for every encryption block (AES operates on 16-byte blocks), so for longer messages CBC mode simply uses the previous ciphertext block as the IV for the following plaintext block. The use of IVs and CBC is not perfect: a chosen-plaintext attack can occur if the attacker is able to predict the IV that will be used for encryption of a message under their control and the attacker knows the IV that was used for the relevant message they are trying to guess. 8 This new research demonstrated that the above attack can be mounted against TLS under certain conditions. When a SSL 3.0 or TLS 1.0 session uses multiple packets, subsequent packets use an IV that is the last ciphertext block of the previous packet, essentially treating the session as one long message. This allows an attacker who can see encrypted messages sent by the victim to see the IV used for the session cookie, the because cookie's location is predictable, and also know the IV that will be used at the beginning of the next message packet (the last ciphertext block from the current message packet). If the attacker can also ``choose'' a plaintext message sent on behalf of the victim, they can make a guess at the session cookie and see if the ciphertext matches.

A successful implementation of the attack requires browser or web technologies to meet the above criteria. For clarity, we walk through the example as follows: The network attacker (who we will call Mallory) has the ability to eavesdrop on the network (e.g., over a wireless network) and coerces Alice to visit http://mallory.com perhaps through phishing, advertising or another attack. The malicious website contains an attack script that forces Alice's browser to make a request to http://bob.com and Mallory records the encrypted cookie. Using a technology that allows Mallory to adapt the attack through a SOP bypass or technology that allows multi-origin communication, Mallory now tries to guess the session cookie as the first block of subsequent requests and sees if the resulting ciphertext matches the previously recorded session cookie ciphertext [20].

This class of attack is well known enough that it was mitigated in TLS version 1.1; however, due primarily to client compatibility reasons neither TLS 1.1 or 1.2 are widely supported on the web and most vendors still require support (i.e. fall back) for SSL v3.0 and TLS v1.0.

Browser vendors have attempted to implement a workaround to address the vulnerability at the implementation level while still remaining compatible with the SSL 3.0/TLS 1.0 protocol. These initially included inserting empty fragments into the message in order to randomize the IV as in the case of OpenSSL, and when that proved problematic to reliably implement, 1/n-1 record splitting where a single byte of the plaintext is injected in each record.

The resulting padding added to complete the block (16 or 15 bytes) is random, and its search space is too high for an attacker to guess.

2.6 SSL/TLS CRIME and BREACH Attacks

The authors of the BEAST attack are also the creators of the later CRIME attack, which can allow an attacker to recover the content of web cookies when data compression is used along with TLS

[21]. When used to recover the content of secret authentication cookies, it allows an attacker to perform session hijacking on an authenticated web session.

While the CRIME attack was presented as a general attack that could work effectively against a large number of protocols, including but not limited to TLS, and application-layer protocols such as SPDY or HTTP, only exploits against TLS and SPDY were demonstrated and largely mitigated in browsers and servers. The CRIME exploit against HTTP compression has not been mitigated at all, even though the authors of CRIME have warned that this vulnerability might be even more widespread than SPDY and TLS compression combined. In 2013 a new instance of the CRIME attack against HTTP compression, dubbed BREACH, was announced. Built based on the CRIME attack a BREACH attack can extract login tokens, email addresses or other sensitive information from TLS encrypted web traffic in as little as 30 seconds (depending on the number of bytes to be extracted), provided the attacker tricks the victim into visiting a malicious web link or is able to inject content into valid pages the user is visiting (ex: a wireless network under the control of the attacker). All versions of TLS and SSL are at risk from BREACH regardless of the encryption algorithm or cipher used.[20] Unlike previous instances of CRIME, which can be successfully defended against by turning off TLS compression or SPDY header compression, BREACH exploits HTTP compression which cannot realistically be turned off, as virtually all web servers rely upon it to improve data transmission speeds for users.[24] This is a known limitation of TLS as it is susceptible to chosen-plaintext attack against the application-layer data it was meant to protect.

Compression Ratio Info-leak Made Easy (CRIME) is an attack on SSL/TLS that was developed by researchers Juliano Rizzo and Thai Duong. CRIME is a side-channel attack that can be used to discover session tokens or other secret information based on the compressed size of HTTP requests. The technique exploits web sessions protected by SSL/TLS when they use one of two data-compression schemes (DEFLATE and gzip) which are built into the protocol and used for reducing network congestion or the loading time of web-pages. Rizzo and Doung demonstrated it at the Ekoparty security conference in September 2012 after notifying major affected software vendors, including Mozilla and Google (CVE-2012-4929 18). CRIME is known to work against SSL/TLS compression and SPDY, although other encrypted and compressed protocols are also likely vulnerable [24].

In a single session the same secret/cookie is sent with every request by the browser. TLS has an optional compression feature where data can be compressed before it is encrypted. Even though TLS encrypts the content in the TLS layer, an attacker can see the length of the encrypted request passing over the wire, and this length directly depends on the plaintext data which is being compressed. Finally, an attacker can make the client generate compressed requests that contain attacker-controlled data in the same stream with secret data. The CRIME attack exploits these properties of browser-based SSL. To leverage these properties and successfully implement the CRIME attack, the following conditions must be met:

- The attacker can intercept the victim's network traffic. (e.g. the attacker shares the victim's (W)LAN or compromises victim's router).
- Victim authenticates to a website over HTTPS and negotiates TLS compression with the server.
- Victim accesses a website that runs the attackers code [25].

This attack is feasible on all browsers and servers that support TLS compression. According to the Qualys SSL Lab's SSL Pulse test data showed 42% of servers and 45% of the browsers supported TLS compression when the attack was released. Internet Explorer, Safari, and Opera were not affected, as they did not support TLS compression.

Among the widely used web browsers, Google Chrome (NSS) and Mozilla Firefox, as well as Amazon Silk supported TLS compression as they implement DEFLATE. The attack also worked against several popular Web services that support TLS compression on the server side, such as

Gmail, Twitter, Dropbox and Yahoo Mail. This attack worked for all TLS versions and all cipher suites (AES and RC4) and even if HSTS is active and preloaded by the browser vendor.

CRIME is a the brute-force attack, so it requires O(W) requests where W is cookie charset, with the possibility to optimize to O(log(W)). The modified version of SSL Strip by Moxie Marlinspike can be used in a public network to launch a man-in-the-middle attack which will satisfy one requirement of the attack. This tool strips the ongoing SSL/TLS session and performs a man-in-the-middle attack by acting as a proxy. The proof of concept code by Krzysztof Kotowicz is also useful to simulate the attack. Duong and Rizzo's pseudo code works well in practice, but does not include a mechanism to sync the JavaScript with the program observing lengths on the network.
Browsers still support HTTP compression, and this attack is possible on HTTP compressed sessions. Timing Info-leak Made Easy (TIME) is an extension of this attack. Recently Browser Reconnaissance and Exfiltration via Adaptive Compression of Hypertext (BREACH) introduced a new targeted techniques to reliably retrieve encrypted secrets [26].

Like the CRIME attack, BREACH exploited the compression and encryption combination used to interact with users and web-servers. The working mechanism of BREACH is similar to CRIME, except CRIME targeted TLS compression, while BREACH targets HTTP compression. HTTP response compression compresses the body of responses but not header information. The algorithm used, DEFLATE, is comprised of two components. LZ77 replaces occurrences of three or more characters with ``pointer'' values to reduce space. Huffman coding replaces characters with symbols in order to optimize the description of the data to the smallest size possible. BREACH works by attacking the LZ77 compression while minimizing the effects of Huffman coding. If this isolation is not performed, too many false positives will result, reducing the effectiveness of the attack.

At a high level the attack works by injecting guesses in HTTP requests and measuring the sizes of the compressed and encrypted responses. The smallest response size indicates that the guess matches the secret value. This is then repeated on a character by character bases.

As BREACH focuses on the HTTP compression of the response body, it is possible to mount on all versions of SSL/TLS, and does not require TLS-layer compression. The cipher suite used during the session negotiation does not affect this attack. The number of requests required are proportional to the size of secret, but in general BREACH attack can be exploited with just a few thousand requests, and under a minute. In short, the scope of this attack includes a considerable portion of the HTTP traffic in the Internet as a large portion of enterprise applications and online websites use HTTP compression to optimize bandwidth.

The three main requirements for exploitation of the vulnerability to be effective are:

1. The application supports HTTP compression.
2. The response should reflect back user's input.
3. The response should have some sensitive/ secret information embedded in the body.

If the user's input is not reflected, there is no possible way to mount a chosen plaintext attack and measure the size of the responses. This attack targets the secret information in the response body (e.g. CSRF tokens), not the session cookie in the request header. So this is useful only if the the response of this attack contains sensitive information.

Like CRIME and TIME, the oracle needs to be aware of Huffman coding scheme and overcome the false positives generated due to the same. In their research paper, Gluck, Harris, and Prado gave a detailed explanation on methods to overcome the aberrations caused by the subtle inner working of the DEFLATE and how they were able to optimize the attack [27].

2.7 Crypt analysis of RC4 and its fails

In cryptography, RC4 (also known as ARC4 or ARCFOUR meaning Alleged RC4, see below) is the most widely used software stream cipher and is used in popular protocols such as Transport Layer Security (TLS) (to protect Internet traffic) and WEP (to secure wireless networks). While remarkable for its simplicity and speed in software, RC4 has weaknesses that argue against its use in new systems. It is especially vulnerable when the beginning of the output keystream is not discarded, or when nonrandom or related keys are used; some ways of using RC4 can lead to very insecure cryptosystems such as WEP.

As of 2013, there is speculation that some state cryptologic agencies may possess the capability to break RC4 even when used in the TLS protocol.[30] Microsoft recommends disabling RC4 where possible [28].

RC4 generates a pseudorandom stream of bits (a keystream). As with any stream cipher, these can be used for encryption by combining it with the plaintext using bit-wise exclusive-or; decryption is performed the same way (since exclusive-or with given data is an involution). (This is similar to the Vernam cipher except that generated pseudorandom bits, rather than a prepared stream, are used.) To generate the keystream, the cipher makes use of a secret internal state which consists of two parts:

- A permutation of all 256 possible bytes (denoted "S" below).
- Two 8-bit index-pointers (denoted "i" and "j").

The permutation is initialized with a variable length key, typically between 40 and 256 bits, using the key-scheduling algorithm (KSA). Once this has been completed, the stream of bits is generated using the pseudo-random generation algorithm (PRGA) [29].

Many stream ciphers are based on linear feedback shift registers (LFSRs), which, while efficient in hardware, are less so in software. The design of RC4 avoids the use of LFSRs, and is ideal for software implementation, as it requires only byte manipulations. It uses 256 bytes of memory for the state array, S[0] through S[255], k bytes of memory for the key, key[0] through key[k-1], and integer variables, i, j, and K. Performing a modular reduction of some value modulo 256 can be done with a bitwise AND with 255 (which is equivalent to taking the low-order byte of the value in question) [31].

Unlike a modern stream cipher (such as those in eSTREAM), RC4 does not take a separate nonce alongside the key. This means that if a single long-term key is to be used to securely encrypt multiple streams, the cryptosystem must specify how to combine the nonce and the long-term key to generate the stream key for RC4. One approach to addressing this is to generate a "fresh" RC4 key by hashing a long-term key with a nonce. However, many applications that use RC4 simply concatenate key and nonce; RC4's weak key schedule then gives rise to related key attacks, like the Fluhrer, Mantin and Shamir attack (which is famous for breaking the WEP standard) [33].

Because RC4 is a stream cipher, it is more malleable than common block ciphers. If not used together with a strong message authentication code (MAC), then encryption is vulnerable to a bit-flipping attack. The cipher is also vulnerable to a stream cipher attack if not implemented correctly [27]. Furthermore, inadvertent double encryption of a message with the same key may accidentally output plaintext rather than ciphertext because the involutary nature of the XOR function would result in the second operation reversing the first.

It is noteworthy, however, that RC4, being a stream cipher, was for a period of time the only common cipher that was immune[26] to the 2011 BEAST attack on TLS 1.0. The attack exploits a known weakness in the way cipher block chaining mode is used with all of the other ciphers supported by TLS 1.0, which are all block ciphers.

In 2013 there was a new attack scenario proposed by AlFardan, Bernstein, Paterson, Poettering and Schuldt that uses new statistical biases in RC4 key table[30] to recover plaintext with large number of TLS encryptions [32].

2.8 Attacks on RC4

In spite of existing attacks on RC4 that break it, the cipher suites based on RC4 in SSL and TLS were at one time considered secure because of the way the cipher was used in these protocols defeated the attacks that broke RC4 until new attacks disclosed in March 2013 allowed RC4 in TLS to be feasibly completely broken. In 2011 the RC4 suite was actually recommended as a work around for the BEAST attack. In 2013 a vulnerability was discovered in RC4 suggesting it was not a good workaround for BEAST. An attack scenario was proposed by AlFardan, Bernstein, Paterson, Poettering and Schuldt that used newly discovered statistical biases in the RC4 key table to recover parts of the plaintext with a large number of TLS encryptions. A double-byte bias attack on RC4 in TLS and SSL that requires 13 × 2^{20}encryptions to break RC4 was unveiled on 8 July 2013, and it was described as "feasible" in the accompanying presentation at the 22nd USENIX Security Symposium on August 15, 2013.

However, many modern browsers have been designed to defeat BEAST attacks (except Safari for Mac OS X 10.8 or earlier, for iOS 6 or earlier, and for Windows. As a result, RC4 is not the best choice for TLS 1.0 anymore. The CBC ciphers which were affected by the BEAST attack in the past are becoming a more popular choice for protection.

Microsoft recommends disabling RC4 where possible.

RC4 is an extremely simple and elegant algorithm. The first phase is the key scheduling algorithm (KSA). This algorithm takes an initial array and initializes it to values 0 to 255. For each index of the array a shuffling occurs that mixes in the key. Once this algorithm runs, the output of the KSA is input to the pseudo-random generation algorithm (PRGA) that continually shuffles the array. The output of the PRGA is exclusive-ored with the plaintext to produce a cipher text.

These recent attacks have found strong biases in the first 257 bytes of encryption which will allow recovery of roughly the first 200 bytes of plaintext in approximately 2 28 to 2 32 encryptions of the same plaintext under unique keys (referred to here as the broadcast attack). The attack recovers the plaintext at each position by gathering the set of observed ciphertexts (each encrypted with a different key) for the corresponding position. It tries each of the 256 candidate plaintexts and computes the PRGA output byte by exclusive-oring the candidate with each ciphertext. It then calculates which plaintext candidate resulted in PRGA outputs which most closely matches the known PRGA bias for that position. To take an extremely simplified example, consider when the PRGA output always outputs the value one. The correct plaintext will be the one which, when exclusive-ored with each of the ciphertexts, always results in one [34].

Additionally, using previously discovered long-term biases in RC4, 50% of a 16-byte secret value can be extracted after analysis of 6 ∗ 2 30 encryptions of the same plaintext message using a single key. This attack is different to the broadcast attack since it can recover plaintext from a one or more encryption streams in which the same plaintext is sent repeatedly. The attack recovers plaintext using transitional biases in the PRGA stream that recur at fixed positions in the stream. For example, if the PRGA output is zero at any offset that is a multiple of 256-bytes, then the next output is more likely to also be a zero. The attack works by starting with a known plaintext byte that repeats at a fixed position and finding a candidate for the plaintext byte that repeats at the next position. The candidate that most closely matches the transition bias of the PRGA is selected. This process is then repeated to find the next unknown plaintext byte.

In both attacks the request structure such as the URL, which may be known or the location and beginning of the cookie as well as the plaintext structure, such as the language or HTML can be used to further optimize the attack.

Additionally, many byte positions may quickly be probabilistically reduced to a limited number of candidate plain-texts. This includes positions that may have multiple biases. The candidate plaintexts can be used to attempt to authenticate to the system under attack in order to verify the correctness of the secret [35].

In the current and un-optimized state, these attacks do not represent a practical threat against the majority of implementations. However, optimization can be made and depending on the individual target, analysis of the structure, language can be made to reduce the attack time frame. A team from Royal Holloway, University of London, and the University of Illinois-Chicago has discovered that the small "biases" contained in RC4 can be manipulated in a way that reveals a limited amount of the plaintext in an encrypted data stream. It requires attackers to receive tens of millions of different encryptions of the same message. By statistically sampling them, the lack of randomness can be exploited to deduce parts of the encrypted message [36].

2.9 TLS Attacks in practice

Examples of attacks:

1. Prediction of random numbers
   In January 1996, Goldberg and Wagner published an article on the quality of random numbers used for SSL connections by the Netscape Browser. The authors gained access to the application's Source Code by decompiling it and identified striking weaknesses in the algorithm responsible for random number generation.
2. Limited entropy
   In 2008 Luciano Bello observed during code review that the PRNG of Debian-specific OpenSSL was predictable starting from version 0.9.8c-1, Sep 17 2006 until 0.9.8c-4, May 13 2008, due to an implementation bug. A Debian-specific patch removed two very important lines in the libssl source code responsible for providing adequate entropy.
3. Exception based DoS
   Zhao et al. provided an attack on the TLS handshake which leads to an immediate connection shutdown and can thus be used for a Denial of Service (DoS) attack. The authors exploited two previously discussed weaknesses to mount successful attacks.
   The first attack targets the Alert protocol of TLS and makes use of the fact that, due to yet missing completed cryptographic primitives negotiation during the handshake phase, all Alert messages remain strictly unauthenticated and thus spoof-able. This enables an obvious, but effective attack: Spoofing Fatal Alert messages which cause immediate connection shutdowns.
   The second attack simply confuses a communication partner by sending either misleading or replayed messages or responding with wrong messages according to the expected handshake flow
4. Renegotiation flaw
   Ray and Dispensa discovered a serious flaw induced by the renegotiation feature of TLS. The flaw enables an attacker to inject data into a running connection without destroying the session. A server would accept the data, believing its origin is the client. This could lead to abuse of established sessions - e.g., an attacker could impersonate a legitimate victim currently logged in to a web application.
5. SSL/TLS Stripping
   In February 2009, Moxie Marlinspike released the sslstrip tool which disables SSL/TLS at a higher layer. As a precondition it is necessary for an attacker to act as Mitm. To disable the security layer the tool sends HTTP 301 - permanent redirection responses and replaces any occurrence https:// with http:// (notice the missing s). This causes the client to move to the redirected page and communicate unencrypted and unauthenticated (when the stripping succeeds and the client does not notice that she is being fooled). Finally, the attacker opens a fresh session to the (requested) server and passes-through or alters any client and server data.
6. Computational DoS

In 2011, the German Hacker Group The Hackers Choice released a tool called THC-SSL-DoS, which creates huge load on servers by overwhelming the target with SSL/TLS handshake requests. Boosting system load is done by establishing new connections or using renegotiation. Assuming that the majority of computation during a handshake is done by the server the attack creates more system load on the server than on the own device - leading to a DoS. The server is forced to continuously recompute random numbers and keys [36].

3. CONCLUSIONS

We presented our analysis for SSL/TLS attacks. We found serious logic flaws in advanced attacks mechanisms. We discussed the weaknesses and ways of its protection.

SSL/TSL has been around for many years without any major modifications. This protocol was considered to be secure. The CRIME, BREACH,BEAST, Hartbleed attacks proved that in one very specific use case it can be compromised. While this use case can be avoided and SSL/TSL re-secured, will this have an effect on the thoughts of SSL/TSL security as a whole. People tend to lose faith in security protocols as soon as the simplest attack is successful. Will this be the end to SSL/TSL, or will users still have faith in the non-compressed version, that has yet to be broken, or will they run to a new protocol to be positive that they are secure? This will only be answered in time.

We believe that our study takes some steps in the security problem space that SSL protocols have brought. Also our study suggests and analyses Heartbleed exploit detection. We believe that our study brings some new chain of trust between the client and the protocol security. In future work we are considering the security challenges that come with other advanced SSL attacks. Fundamentally, we believe that vulnerabilities of SSL/TSL demands new research efforts on ensuring the security quality of the protocols.

4. REFERENCES

[1] “The Secure Sockets Layer Protocol”. Internet: http://www.cs.bris.ac.uk/~bradley/publish/SSLP/chapter4.html [Nov. 22, 2013].

[2] “SSL: Intercepted today, decrypted tomorrow”. Netcraft, pp. 10-12, May 25, 2013.

[3] “SSL/TLS in Detail“. Microsoft TechNet, July 31, 2003.

[4] “Description of the Secure Sockets Layer (SSL) Handshake“. Internet: http://www.support.microsoft.com [Dec. 1, 2013].

[5] “Secure electronic transaction”. Internet: http://en.wikipedia.org/wiki/Secure_Electronic_Transaction [Dec. 12, 2013].

[6] OpenSSL TLS/DTLS Heartbeat Read Overrun Vulnerability. Internet: http://herjavecgroup.com/admin/pdf/THG_TAB_Heartbleed.pdf [May, 2014].

[7] “The Heartbleed Bug”. Codenomicon, Internet: http://heartbleed.com/ [Apr, 2014].

[8] “April 2014 Web Server Survey”. Netcraft. April 2 204. Internet: http://news.netcraft.com/archives/2014/04/02/april-2014-web-server-survey.html [Apr, 2014].

[9] “OpenSSL Security Advisory”. OpenSSL. Internet: https://www.openssl.org/news/secadv_20140407.txt [Apr, 2014].

[10] “Wild at Heart: Were Intelligence Agencies Using Heartbleed in November 2013?”. Electronic Frontier Foundation. Internet: https://www.eff.org/deeplinks/2014/04/wild-heart-were-intelligence-agencies-using-heartbleed-november-2013 [Apr, 2014].

[11] “Daily Ruleset Update Summary”. Emerging Threats Snort Ruleset. Internet: http://www.emergingthreats.net/2014/04/09/daily-ruleset-update-summary-04092014/ [Apr, 2014].

[12] “Detecting OpenSSL Heartbleed with Suricata”. Inliniac. Internet: http://blog.inliniac.net/2014/04/08/detecting-openssl-heartbleed-with-suricata/ [Apr, 2014].

[13] „Detect Exploit openSSL Heartbleed vulnerability using Nmap and Metasploit on Kali Linux“. Internet: http://www.blackmoreops.com/2014/05/03/detect-exploit-openssl-heartbleed-vulnerability-using-nmap-metasploit-kali-linux/ [June, 2014].

[14] „Heart attack: detecting heartbleed exploits in real-time“. Internet: http://www.tripwire.com/state-of-security/incident-detection/heart-attack-detect-heartbleed-exploits-in-real-time-with-active-defense/ [June, 2014].

[15] „How to Detect a Prior Heartbleed Exploit“. Internet: http://www.riverbed.com/blogs/Retroactively-detecting-a-prior-Heartbleed-exploitation-from-stored-packets-using-a-BPF-expression.html [June, 2014].

[16] Ivan Ristić. „SSL/TLS Deployment Best Practices“. Internet: https://www.ssllabs.com/downloads/SSL_TLS_Deployment_Best_Practices_1.3.pdf [June, 2014].

[17] Nadhem AlFardan, Dan Bernstein, Kenny Paterson, Bertram Poettering, Jacob Schuldt. „On the Security of RC4 in TLS“. Internet: http://www.isg.rhul.ac.uk/tls/ [June, 2014].

[18] Lars Nybom, Alexander Wall. „SSL/TLS and MITM attacks“. Internet: http://www.it.uu.se/edu/course/homepage/distrinfo/ht09/presentations/Group7.pdf [June, 2014].

[19] Pratik Guha Sarkar. „Attacks on ssl a comprehensive study of beast, crime, time, breach, lucky 13 & rc4 biases“. Internet: https://www.isecpartners.com/media/106031/ssl_attacks_survey.pdf [June, 2014].

[20] „List of browsers support for different TLS version“. Internet: https://en.wikipedia.org/wiki/Transport_Layer_Security#Web_browsers [June, 2014].

[21] Hong lei Zhang. „Three attacks in SSL protocol and their solutions“. Internet: https://www.cs.auckland.ac.nz/courses/compsci725s2c/archive/termpapers/725zhang.pdf [June, 2014].

[22] „Vulnerability Summary for CVE-2012-4929“. Internet: http://web.nvd.nist.gov/view/vuln/detail?vulnId=CVE-2012-4929 [June, 2014].

[23] „Software >> sslstrip“. Internet: http://www.thoughtcrime.org/software/sslstrip/ [June, 2014].

[24] „SSL, GONE IN 30 SECONDS“. Internet: http://breachattack.com/ [June, 2014].

[25] Christopher Meyer. „SoK: Lessons Learned From SSL/TLS Attacks“. Internet: http://www.nds.rub.de/media/nds/veroeffentlichungen/2013/08/19/paper.pdf [June, 2014].

[26] Scott C. Johnson. „CRIME Attack on SSL/TSL“. Internet: http://www.cs.rit.edu/~sxj4236/crypto2_paper2.pdf [June, 2014].

[27] Be'ery, Tal and Amichai Shulman. "TIME Prefect CRIME." Internet: https://media.blackhat.com/eu-13/briefings/Beery/bh-eu-13-a-perfect-crime-beery-wp.pdf [June, 2014].

[28] Constantin Lucian. „Researchers resurrect and improve CRIME attack against SSL“. Internet: http://www.networkworld.com/news/2013/031413-researchers-resurrect-and-

improve-crime-267698.html?page=1 [June, 2014].

[29] Kelsey John. "3091." 2002. IACR.org. 9 4 2013. Internet; http://www.iacr.org/cryptodb/archive/2002/FSE/3091/3091.pdf [June, 2014].

[30] „zOompf. Explaining the Crime weakness in SPDY and SSL“. Internet:

http://zoompf.com/2012/09/explaining-the-crime-weakness-in-spdy-and-ssl [June, 2014].

[31] Vlastimil Klíma. „Attacking RSA-based Sessions in SSL/TLS“. Internet:

http://eprint.iacr.org/2003/052.pdf [June, 2014].

[32] Canvel, B.“Password Interception in a SSL/TLS Channel“. Internet: http://lasecwww.epfl.ch/memo_ssl.shtml [February, 2003].

[33] Jonsson J. „On the Security of RSA Encryption in TLS“. In Proc. of CRYPTO '02, pp. 127 -142, 2002.

[34] Kurt Seifried. „As with marriage, SSL security success is in the details Attacks Against SSL“. Internet: http://www.linux-magazine.com/Issues/2010/112/Security-Lessons-Secure-Programming [June, 2014].

[35] Dan Goodin. „Two new attacks on SSL decrypt authentication cookies“. Internet: http://arstechnica.com/security/2013/03/new-attacks-on-ssl-decrypt-authentication-cookies/ [June, 2014].

[36] „Another crypto-attack on SSL/TLS encryption“. Internet: http://www.h-online.com/security/news/item/Another-crypto-attack-on-SSL-TLS-encryption-1823227.html [June, 2014].

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Computer Science and Security (IJCSS)* is a refereed online journal which is a forum for publication of current research in computer science and computer security technologies. It considers any material dealing primarily with the technological aspects of computer science and computer security. The journal is targeted to be read by academics, scholars, advanced students, practitioners, and those seeking an update on current experience and future prospects in relation to all aspects computer science in general but specific to computer security themes. Subjects covered include: access control, computer security, cryptography, communications and data security, databases, electronic commerce, multimedia, bioinformatics, signal processing and image processing etc.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJCSS.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 8, 2014, IJCSS is appearing with more focused issues. Besides normal publications, IJCSS intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

IJCSS LIST OF TOPICS

The realm of International Journal of Computer Science and Security (IJCSS) extends, but not limited, to the following:

- Authentication and authorization models
- Computer Engineering
- Computer Networks
- Cryptography
- Databases
- Image processing
- Operating systems
- Programming languages
- Signal processing
- Theory
- Communications and data security
- Bioinformatics
- Computer graphics
- Computer security
- Data mining
- Electronic commerce
- Object Orientation
- Parallel and distributed processing
- Robotics
- Software engineering

CALL FOR PAPERS

Volume: 9 - Issue: 2

i. Submission Deadline : January 31, 2015 **ii. Author Notification:** February 28, 2015

iii. Issue Publication: March/April 2015

CONTACT INFORMATION

Computer Science Journals Sdn BhD

B-5-8 Plaza Mont Kiara, Mont Kiara

50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6204 5627

Fax: 006 03 6204 5628

Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2014
COMPUTER SCIENCE JOURNALS SDN BHD
B-5-8 PLAZA MONT KIARA
MONT KIARA
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6204 5627
FAX: 006 03 6204 5628
EMAIL: cscpress@cscjournals.org