NATIONAL DEFENCE UNIVERSITY "CAROL I"

REGIONAL DEPARTMENT OF DEFENSE RESOURCES MANAGEMENT STUDIES

NEW CHALLENGES IN INFORMATION SECURITY MANAGEMENT

Workshop unfolded during the postgraduate course in Information Security Management

21- 22.02.2011, Brasov

Coordinator:
LTC Prof. eng. Daniel Sora, PhD

National Defense University „Carol I" Publishing House
BUCHAREST 2011

Scientific board:

LTC Professor eng. Daniel Sora, PhD
LTC Senior Lecturer Cezar Vasilescu, PhD
Junior lecturer Aura Codreanu, PhD

ISBN: 978-973-663-951-7

CONTENT

REMARKS ON THE NEW EMERGING IT TECHNOLOGIES TO SUPPORT MILITARY ENTERPRISES, AND THEIR IMPACT ON THE INFORMATION SECURITY

COL. Cristian MOSORESCU

I. INTRODUCTION

Information technology is very often seen as a holly grail to support the continuous changing in the Military Business and demanding operations. Since the new technologies have broken trough, revolutionary concepts have been developed and updated to reflect the operational needs and the new technological hype. However, just a few of them, finally succeed in providing the expected benefits. How many of them have actually delivered any good to the *boots on the ground*, and how many of them have been just buzz, and have fizzle out. It is worthwhile mentioning some of modern military initiatives, as follows: the Network Centric Warfare (NCE), Network Enable Capability (NEC) or NATO Network Enable Capabilities (NNEC). However, the new security challenges facing modern countries are so demanding that the governments are making great efforts to get into the technological cycle hoping to successfully address the military gaps and to achieve the military supremacy. As a consequence, the NATO countries, have taken steps to alleviate on both national and communitarian areas the identified issues related to acquisition, program management, and implementing the new emerging technologies, leveraging information technologies to create a more efficient and effective support to the military on the ground, at a more affordable cost and within expected time frame.

But despite spending billions on information technology over the past decade, the military has achieved little of the productivity improvements that private industry has realized from IT. Too often, IT projects run over budget, behind schedule, or fail to deliver promised functionality. Many projects use "grand design" approaches that aim to deliver functionality every few years, rather than breaking projects into more manageable chunks and demanding new functionality every few quarters. In addition, the governments too often rely on large, custom, proprietary systems when "light technologies" or shared services exist[1] .

[1] 25 Point Implementation Plan to Reform Federal Information Technology Management, Vivek Kudra, U.S. Chief Information Officer, 2010

II. ACHIEVING OPERATIONAL EFFICIENCY

The military business innovation and transformation of the military capabilities is crucial to achieve the military supremacy on the battlefield. Often to improve your capabilities, military planners and the industry have to determine which technology is enough mature to be incorporated in the new capability. Is a key decision when, what, and when a specific technology military should implement to avoid capability gaps, on one hand and to deliver an efficient and effective tool to support military actions, on the other hand.

In today's environment, where the threat landscape changes daily and the cyber defense of (military) networks is constantly being tested, finding ways to simplify network topologies and provide for a more effective event aggregation and correlation is crucial[2].

Gartner's Hype Cycle provides an instrument to graphically present the typical progression of an emerging technology, from over enthusiasm through a period of disillusionment to an eventual understanding of the technology's relevance and role in a market or domain (see Figure II.1).

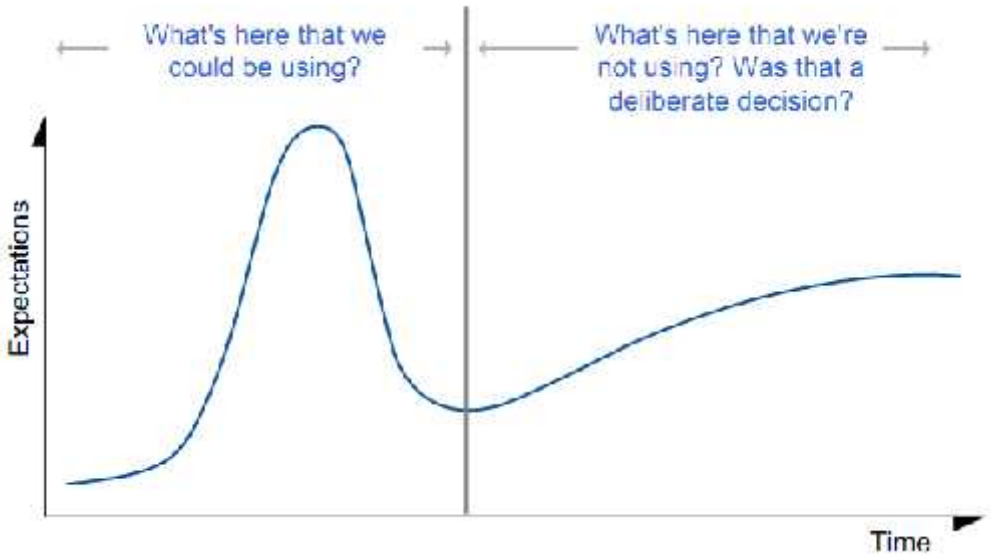


Figure II-1 - The Hype Cycle of Innovation: Key Questions (Source: Gartner)

The 2009 and 2010 Gartner Hype Cycle Special Report evaluates the maturity of over 1,650 technologies and trends in 79 technology, topic and industry areas. New Hype Cycles in 2010 include cloud computing, virtualization, data center power and cooling technologies, and mobile device technologies (see Figure II.2).

According to Gartner, cloud computing is the most hyped technology of years 2009 and 2010.

[2] Virtualization Arsenal, Jeff Lake, Fortinet, http://www.military-information-technology.com/mit-archives/190-mit-2008-volume-13-issue-6/1807-virtualization-arsenal.html

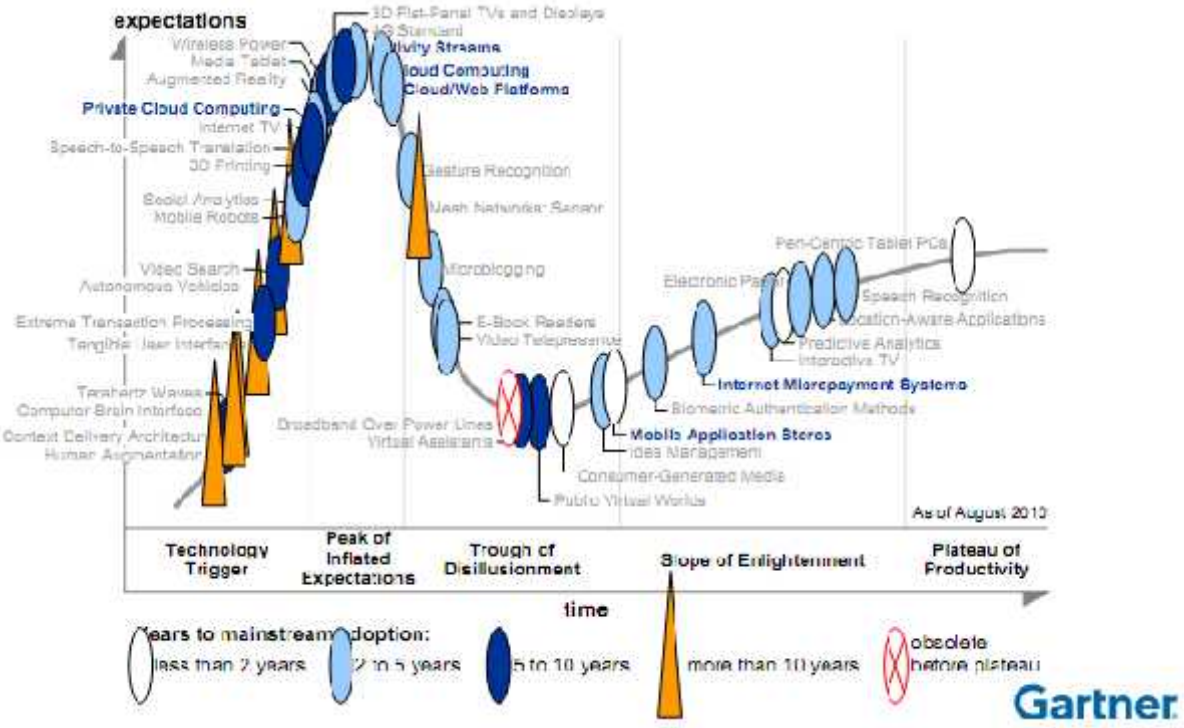


Figure II.2 - Cloud and platforms; Source: Gartner

Cloud computing is the latest super-hyped concept in IT. Although cloud computing is about a very simple idea — consuming and/or delivering services from "the cloud," there are many issues regarding types of cloud computing and scope of deployment that make the details not nearly so simple. Improved security and reduced costs are among the attractions for the military of this increasingly popular approach to utilizing computing resources.

Massive scalability is another crucial military requirement to support more and more demanding operations. Besides, the tactical needs to decrease the physical footprint of IT infrastructure on troops' deployability and agility, such as garrison data centers, or deployed tactical environment, can be addressed using virtual networking technology.

Implementing virtual networking technologies allows a single network device to transparently host multiple networks or echelons on a common infrastructure. Virtual local area networks (VLANs) allow network links to be shared by virtualized servers to help improve network performance, reduce management complexity and enable more granular usage policies.

Two important areas to review further in the virtual world are virtual domains (VDOMs) and VLANs. VDOMs enable the capability to use a common infrastructure to provide routing and network protection for several organizations or echelons.

The primary reasons for implementing VDOMs and VLANs are to improve network manageability, scalability and security. Security solutions for virtual networks must allow management on a per customer or per-application basis, while ensuring availability of the control itself and the systems it protects. Also required is a high-performance security platform that is capable of scaling to support

thousands of virtual networks with management, logging and reporting customized for each customer or application. [3]

a. Minding the Gap

The current paradigm to develop, field and support military systems seems to rich their technological limits. Military looks for agility, flexibility, high integration and security of their applications, everything fully packaged at a low cost. As a response to these operational requirements, the industry provides just cumbersome, complex stove pipes at an unaffordable cost. Thus, it is crystal clear that industry needs to dramatically changing their paradigm to deliver solutions to meet military expectation.

As part of a broader IT transformation, the military needs to fundamentally shift its mindset from building custom systems to adopting light technologies and shared solutions. Too often, military units build large standalone systems from scratch, segregated from other systems. *These systems often duplicate others already within the military, wasting taxpayer dollars. The growth in data centers from 432 in 1998 to 2,094 in 2010 highlights this problem*[4].

Allied Command Transformation (ACT) has been tasked by NATO to tackle this problem.

ACT considers that a cloud based system may facilitate the technical consolidation of hardware, and data, and will expand interoperability.

"We're investigating how command-and-control can be used and what benefits it would bring," said Johan Goossens, the head of ACT's Technology Branch in Norfolk, Virginia, in the US.

It is considered that the new trends in commercial developments in 'cloud computing', including service oriented architectures and virtualisation, are being widely promoted as a means of making more efficient use of IT infrastructure.

The main gains are cost savings and increased interoperability: 'cloud computing' offers opportunities to reduce operating costs for information systems, together with increased efficiency and flexibility in the way information is stored, managed and used. This is achieved through shared network-delivered services, both public and private, in which each user sees only the service, as the implementation or infrastructure is managed elsewhere. Most computer experts both in and outside of the Alliance cite cloud computing's ability to perform complex computing tasks cheaply to be the main draw.

This is especially attractive with 28 different member states – and their 28 different computer networks plus that of NATO's itself – which need at least some degree of interoperability on a daily

3 Virtualization Arsenal, Jeff Lake, Fortinet, http://www.military-information-technology.com/mit-archives/190-mit-2008-volume-13-issue-6/1807-virtualization-arsenal.html

[4] 25 Point Implementation Plan to Reform Federal Information Technology Management, Vivek Kudra, U.S. Chief Information Officer, 2010

basis, and in the battlefield of Afghanistan, for example, the number of interacting countries rises to more than 40.[5]

Some technical experts recommend[6] that the military should take the lead to the cloud computing and *show other large organizations how it should be done*. They claim that military *could run more effectively using cloud technologies*. In support of these statements, this technology exposes some features that could support a couple of military strategic requirements like: data centers' consolidation and agility to quickly align to mission changes.

III. NEW EMERGING TECHNOLOGIES

a. Virtualization, open standards, and Service-oriented architecture

Data integration is a real concern for the military organizations. Old fashion design paradigm has little capacity to provide military with timely tools and capabilities properly aligned with the new operational requirements of the military business and to keep paces with the fluid tactical environment.

Common Operational Picture (COP), Data integration, common formats and standards, system development agility and collaboration are the most important requirements to support operations.

Open standards, enterprise architecture, service-oriented architecture and virtualization are the industry new proponents to enhance interoperability, and lower the operational cost of IT infrastructure.

Virtualization, in computing, is the creation of a virtual (rather than actual) version of something, such as a hardware platform, operating system, a storage device or network resources[7]. The usual goal of virtualization is to centralize administrative tasks while improving scalability and work loads.

Virtualization includes software, hardware, memory, storage, data and network virtualization.

Information assurance or IT professionals concerned with network security in the Department of Defense are confronted by a constantly evolving array of threats and increasing compliance requirements. They must balance the ability to manage this dynamic "threatscape" against many other imperatives, including capital and operating costs, limited data center space, manageability and, increasingly, environmental concerns. In the DoD world, the other factor of great consideration is the balance of deployable network security assets between tactical and garrison environments.

Driven by space, power, budget and other constraints, consolidation has become both a tactical and strategic imperative for DoD IT and network defense professionals at all levels. The benefits of consolidation, whether physical or virtual, are well-known, including lower equipment and

[5] NATO embraces cloud computing, http://www.dworld.de/dw/article/0,,14824382,00.html, 08 February 2011

[6] How and why the military should adopt the cloud, http://www.infoworld.com/d/cloud-computing/how-and-why-the-military-should-adopt-the-cloud-484, 09 February 2011

[7] WIKIPEDIA, The free Encyclopedia, Virtualization, 07 February 2011

operations costs, less power consumption, improved manageability, and a better environmental footprint[8].

b. Cloud computing – integrating the new technologies

The new fancy technology, cloud computing, has been developed based on the electricity grid model. It is a natural evolution of some other mature technologies, to include virtualization, service-oriented architecture and utility computing. The concept is a new layer of abstractization in which computing resources (servers, software, data and other devices) are offered to a consumer by a third-party which controls the cloud, on demand and byproduct.

The cloud term identifies the Internet network and it is an abstraction of underlying infrastructure it represents.

The intent is to deliver online applications that can be consumed by another Web services or clients (Web browser), while software and data are located on virtualized servers and storage.

The operational requirements and the quality of service (QoS) are commercially managed under a contract, typically including a service level agreement (SLA).

IV. CLOUD COMPUTING TECHNOLOGY

In his book The Big Switch (W.W. Norton & Co., 2008), Nicholas Carr proposes an information revolution very similar to an important change within the industrial era. Specifically, Carr equates the rise of cloud computing in the information age to electrification in the industrial age. It used to be that organizations had to provide their own power (water wheels, windmills).

With electrification, however, organizations no longer provide their own power; they just plug in to the electrical grid. Carr argues that cloud computing is really the beginning of the same change for information technology. Now organizations provide their own computing resources (power). The emerging future, however, is one in which organizations will simply plug in to the cloud (computing grid) for the computing resources they need.

Cloud technologies and Infrastructure-as-a-Service enable IT services to efficiently share demand across infrastructure assets, reducing the overall reserve capacity across the enterprise. Additionally, leveraging shared services of "commodity" applications such as e-mail across functional organizations allows organizations to redirect management attention and resources towards value-added activities[9] .

Controversially, security experts argue that security, interoperability, and portability issues may hamper the success of migration to the cloud architecture.

---

[8] Virtualization Arsenal, Jeff Lake, Fortinet, http://www.military-information-technology.com/mit-archives/190-mit-2008-volume-13-issue-6/1807-virtualization-arsenal.html

[9] 25 Point Implementation Plan to Reform Federal Information Technology Management, Vivek Kudra, U.S. Chief Information Officer, 2010

The expectation is that the industry will shorten the adoption cycle, enabling cost savings and an increased ability to quickly create and deploy enterprise applications.

Since there are some experts considering that the cloud computing model is similar to the development of IT domain itself, the others see it as an Internet development.

Figure IV-1 illustrates the evolution of cloud computing as a natural extension of the Internet service provider (ISP) model.

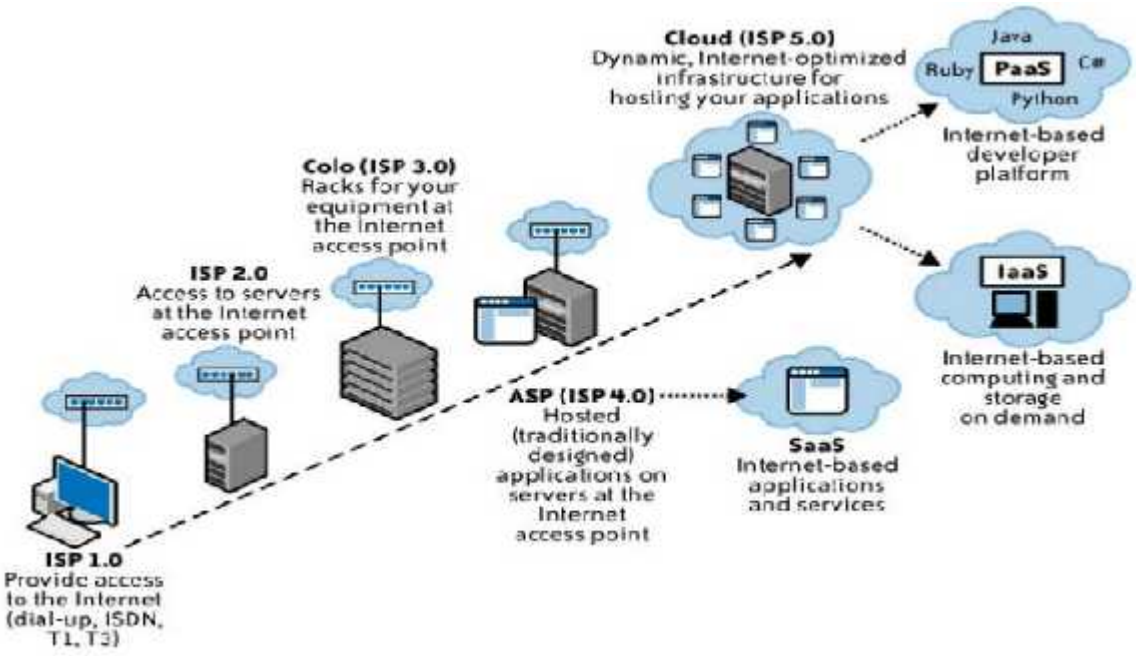


Figure IV-1. Evolution of cloud computing (Source: [6])

There are some authors that consider cloud computing as a natural development of IT domain. Starting with mainframe computers and getting through minicomputers, personal computers, cloud computing is just around the corner (see the figure IV-2).

However, technology and its impact on the word economy did not take place overnight, but though waves of changes. Many of them put a great stamp on the human development. Cloud computing has a great potential to be the next disruptive wave.

In his book The Big Switch (W.W. Norton&Co., 2008), Nicholas Carr, sees the new coming information revolution, especially the rise of cloud computing very similar to the revolutionary change within the industrial era.

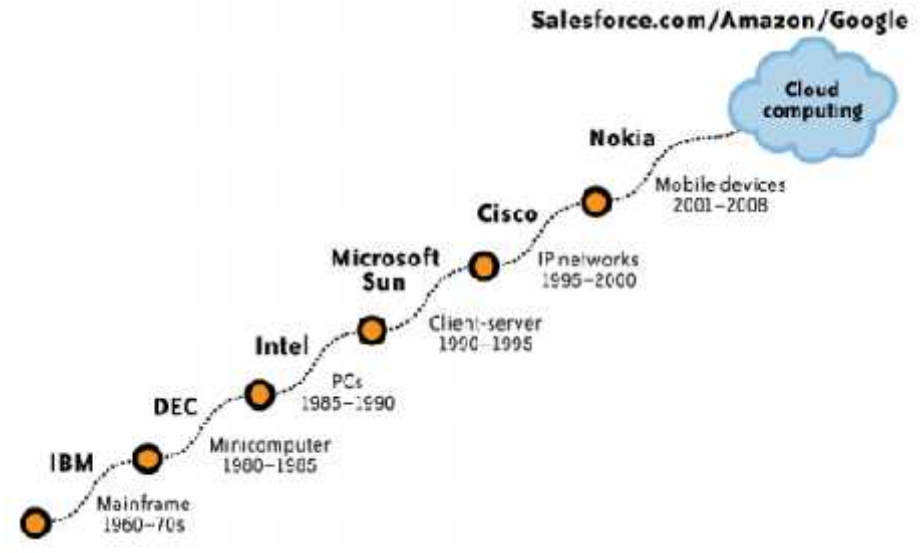


Figure IV-2. Subwaves within the information age (Source: [6])

a. Cloud computing – setting up the scene

Cloud computing is still an evolving technology having a lot of uncertainty and buzzwords around. However, cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential **characteristics,** three **service models**, and four **deployment models**[10].

b. Cloud computing characteristics

Cloud computing is described by five **characteristics**, namely:

multitenancy (shared resources), massive scalability, elasticity, pay as you go, and self-provisioning of resources[11].

A brief description of these characteristics is provided hereafter.

1. Multitenancy (shared resources)

Unlike previous computing models, which assumed dedicated resources (i.e., computing facilities dedicated to a single user or owner), cloud computing is based on business model in which resources are shared (i.e., multiple users use the same resource) at the network level, host level, and application level.

2. Massive scalability

[10] The NIST Definition of Cloud Computing, Peter Mell and Tim Grance,Version 15, 10-7-09, National Institute of Standards and Technology, Information Technology Laboratory

[11] Cloud Security and Privacy, Tim Mather, Subra Kumaraswamy, and Shahed Latif, 2009, O'Reilly Media, p. 26

Although organizations might have hundreds or thousands of systems, cloud computing provides the ability to scale to tens of thousands of systems, as well as the ability to massively scale bandwidth and storage space.

3. Elasticity

Users can rapidly increase and decrease their computing resources as needed, as well as release resources for other uses when they are no longer required.

The elasticity concept is briefly illustrated in the Figure IV-3.

4. Pay as you go

Users pay for only the resources they actually use and for only the time they require them.

5. Self-provisioning of resources

Users self-provision resources, such as additional systems (processing capability, software, storage) and network resources.

These **characteristics** are very promising to overcome the current identified military gaps and systems' flaws and limitations.

Elasticity of resources will allow a great flexibility to support with IT resources various operations when demands are unknown, and constantly changing, addressing spikes in usage.

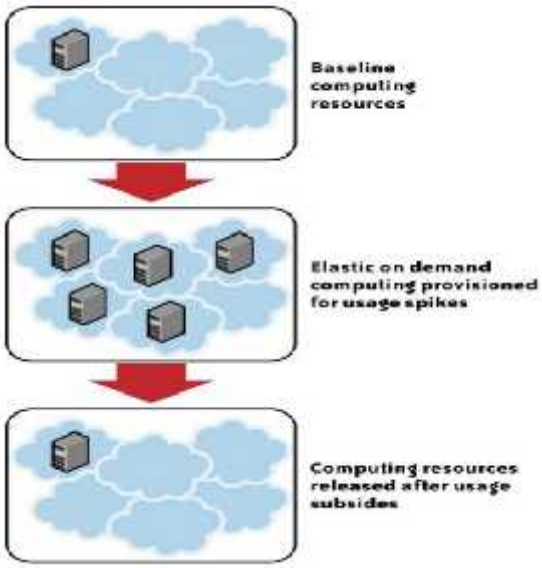


Figure IV-3. Attribute of elasticity (Source: [6])

Since the traditional Software application paradigm is based on very cumbersome model entailing upfront licensing costs and annual support cost. Increasing demand of resources (e.g. number of users) can raise the base cost of solution due to the need for additional hardware capacity and IT manpower. Licensing scheme costs are very often defined by metrics that are not align with the business usage (number of CPUs, machines types etc.).

As a consequence, security architecture to protect the critical assets is highly customizable and expensive in terms of money and manpower.

Besides, the capacity of cloud to provide all the necessary capabilities on demand as a service to the military consumers, using Internet technologies, appears very attractive.
All these benefits comes at very good cost and alleviate the impact of lack of that very sophisticate support which might be necessary at every unit level, now easily provided by the cloud itself.

c. Service Models[12]

1. *Cloud Software as a Service (SaaS).*

The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

2. *Cloud Platform as a Service (PaaS).*

The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.

3. *Cloud Infrastructure as a Service (IaaS).*

The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

d. Deployment Models[13]

1. *Private cloud*

The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premise or off premise.

2. *Community cloud*

[12] "NIST.gov - Computer Security Division - Computer Security Resource Center". Csrc.nist.gov. http://csrc.nist.gov/groups/SNS/cloud-computing/. Retrieved 2010-08-22
[13] "NIST.gov - Computer Security Division - Computer Security Resource Center". Csrc.nist.gov. http://csrc.nist.gov/groups/SNS/cloud-computing/. Retrieved 2010-08-22

The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by the organizations or a third party and may exist on premise or off premise.

3. *Public cloud*

The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

4. *Hybrid cloud*

The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).

The figure IV-4 provides a general view on the cloud computing technology.

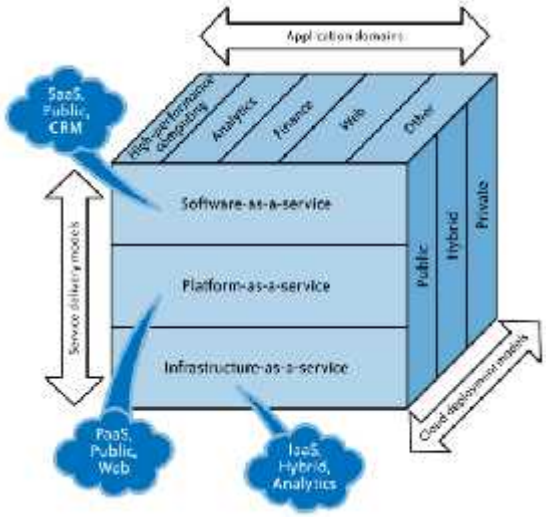


Figure IV-4. SPI service model (Source: [6])

Figure IV-5 describes the more relevant technologies that are exposed by the cloud computing architecture.

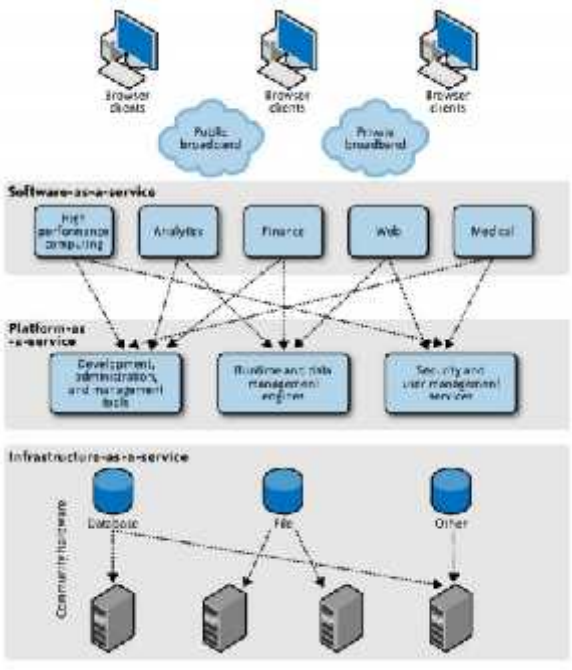


Figure IV-5. Architecture for relevant technology (Source:[6])

e. Strategy to migrate to cloud computing

David Linthicum practically recommends a three-step strategy to move the military to the cloud[14], namely:

1. ***Focus on componentization before moving to cloud computing.*** The idea is not to force-fit IT onto cloud computing platforms; it's about rebuilding the core IT infrastructure as many components parts: data, services, processes, images perhaps bundled into virtual appliances that could be portable among cloud platforms. The idea is to treat these components as items that can be moved to any platform easily, allowing the DoD to run its systems on the platforms that are most efficient and effective, and to quickly align to the requirements of the missions.
2. ***Leverage private cloud first.*** *While I often don't mention this as general advice, the security issues around DoD systems are so sensitive that they can't live on public clouds -- at least for now. The DoD needs to get good at private clouds and move to better and more effective platforms when it can. If it componentized well, the use of private cloud technology should not be much of a challenge.*
3. ***Bring in new people.*** It is important that new people with a holistic, meaningful strategy around cloud computing to participate in the new approach.

U.S. DoD has already announced a program to get the cloud capabilities.

"The (U.S.) Army has 160-plus installations and over 400-plus NIPR access points, Cloud computing is becoming one of the tools of getting control of those assets. APC2 and the containerized data centers will provide both fixed and tactical approaches to cloud computing." [15]

V. INFORMATION SECURITY SIDE EFFECTS AND IMPACT ON THE MILITARY BUSINESS

Without denying the real benefits which the new technology may bring up, the collateral effects of implementing it is still difficult to gauge and should be properly considered.

Some military officials recognize that the cloud computing balance good features with some bad ones, such as security, bandwidth and last but not the least a *culture that is averse to sharing*, *which the Achilles' heel for cloud were computing.*

"*It's an illusion to think data is less safe because there aren't two Army guys sitting there with it. We have to prove successes so that people saying, 'You can't do this,' can understand and get on board and no longer be barriers*".[16]

---

[14] How and why the military should adopt the cloud, http://www.infoworld.com/d/cloud-computing/how-and-why-the-military-should-adopt-the-cloud-484, 09 February 2011

[15] DOD tackles information security in the cloud, Amber Corrin, Jan 20, 2011, http://www.defensesystems.com/Articles/2011/01/24/Defense-IT-1-DOD-cloud-computing-security-issues.aspx?Page=2

Due to the inherent security problems, military propose that a cloud computing in-house, with services managed internally or through commercial companies.

The envisaged solution includes one of them or any combination of them.

For instance, the US Defense Information Systems Agency is consider a private cloud, which could eventually be available across DOD, that contracted services will support

"In order to secure not only our classified data but also our official-business sensitive but not classified data, we are implementing a private cloud to support these requirements. This private cloud is under positive DOD control, hosted in our secure Defense Enterprise Computing Centers, managed by appropriately cleared and certified personnel, directly connected to the DOD's enterprise networks and securely configured to meet DOD's Security Technical Implementation Guides." [17]

Each of the three cloud computing models has some specific security issues, which are briefly exposed here after:

• With SaaS, military need to rely heavily on their cloud providers for security. The provider retains all the responsibilities to protect sensitive information. It's very difficult for military side to get details to help provide assurance that the right things are being done. Besides, it's tough to get assurance that the application will be available enough.

• With PaaS, the provider may give some control to the people building applications on top of its platform. For instance, developers might be able to do their own authentication systems and data encryption, but any security below the application level is still going to be completely under the provider's responsibility.

• With IaaS, the developer has much better control over the security environment, primarily because applications run on virtual machines that are separate from other virtual machines running on the same physical machine. As a consequence, it is easier to ensure that security and compliance concerns are properly addressed[18].

Generally, the main security concerns are referring to four main categories, as follows:

a. Infrastructure security concerns

This includes network-, host-, and application-level security and the issues related to the surrounding each level with specific regard to cloud computing.

b. Data Security and Storage concerns

---

[16] DOD tackles information security in the cloud, Amber Corrin, Jan 20, 2011, http://www.defensesystems.com/Articles/2011/01/24/Defense-IT-1-DOD-cloud-computing-security-issues.aspx?Page=2

[17] DOD tackles information security in the cloud, Jan 20, 2011, http://www.defensesystems.com/Articles/2011/01/24/Defense-IT-1-DOD-cloud-computing-security-issues.aspx?Page=2

[18] The Myths of Security: What the Computer Security Industry Doesn't Want You to Know, John Viega, O'Reilly Media

In the new cloud computing infrastructure, data security becomes more important when and should be addressed at all levels: infrastructure-as-a-service (IaaS), platform-as-a-service (PaaS), and software-as-a-service(SaaS).

The new security aspects should carefully consider, but not limited to:

- Data-in-transit
- Data-at-rest
- Processing of data, including multitenancy
- Data lineage
- Data provenance
- Data remanence.

c. Identity and Access Management concerns

Military organizations considering cloud services (IaaS, PaaS, or SaaS) should consider their organization's operational, security, privacy, and compliance requirements to provision and manage the user identity life cycle, in order to manage user account provisioning, authentication, and authorization in the cloud.

d. A shift in Security Management

With the adoption of cloud services, a large part of your network, system, applications, and data will move to a third-party provider's control. The cloud services delivery model brings new challenges to the IT operations and management staff in the area of availability, access control, vulnerability, and security patch and configuration management. As a first step, cloud customers will have to understand the service delivery model (SPI) and the layers they own, touch, or interface with—network, host, application, database, storage, and web services, including identity services. To tackle these challenges, you will need to understand the scope of IT system management responsibilities, including your system management responsibilities for access, change, configuration, patch, and vulnerability management.

e. **Attacks Against the Cloud**

A cloud-based system does not necessary bring more security than its traditional counterpart. The truth is that the cloud computing can make a system even less secure.

The traditional unsecured applications are just ported to the cloud from standalone or dedicated servers.

Thus, current existing security issues such as buffer overflows, SQL injection, cross-site scripting (XSS), command injection, and other common application-level vulnerabilities do still stand.

More over, the cloud architecture brings up a new set of security classes.

Since some of the new threats have already been theorized, studied, and accepted as potential avenues of attack, due to the complexity and novelty of this technology, many others will be discovered and dealt with hardening application[19].

VI. A NOTIONAL STRATEGY TO THE CLOUD COMPUTING

a. National military authorities shall identify the military proponent services, and the necessary levels of protection;
b. Appropriate community of interest shall be identified, and established a Cloud Governance committee to elaborate a strategy, a roadmap, and a CONOPS for coordinating the migration to the cloud;
c. All common services shall be prioritized and get migrated to the cloud, as per their maturity and levels of protection;
d. A security committee shall considerate and up-front address the security issues to support a smooth transition to the cloud;
e. Military shall encourage industry to cooperate and support the migration efforts;
f. Military shall develop three cloud-architecture, namely:
    i. Public and FOUO cloud– to accommodate unclassified and For Official Use Only services (e.g. INTERNET Access, email and web services)
    ii. Private cloud – to provide support for High classified services
    iii. Hybrid cloud – assure the services for deployed units and for interoperability and interagency support.

VII. CONCLUSION

The cloud computing is a very promising technology which is luring the military organizations into a new technological wave.

It is pretty clear that the new fashion brigs up not only efficiency and address some current systemic flaws, but will also reshape the military business.

Security is just a domain that needs properly addressing, just upfront getting to cloud, but is definitely not the single one.

"The main problem is a lack of a clear legal framework, a lack of transparency in the market, a lack of understanding between parties", said Daniele Catteddu, a NATO official[20]

[19] Hacking: The Next Generation, Nitesh Djanjani, Billy Rios, and Brett Hardin, O'Reilly Media
[20] NATO embraces cloud computing, http://www.dworld.de/dw/article/0,,14824382,00.html, 08 February 2011

References

[1]-25 Point Implementation Plan to Reform Federal Information Technology Management, Vivek Kudra, U.S. Chief Information Officer, 2010

[2]-Handbook of Research on Information Security and Assurance, Jatier N.D. Gupta and Sushil K. Sharma, 2009

[3]-Hacking: The Next Generation, Nitesh Djanjani, Billy Rios, and Brett Hardin, O'Reilly Media

[4]-Inside Cyber Warfare, Jeffrey Carr, 2010, O'Reilly Media

[5]-The Myths of Security: What the Computer Security Industry Doesn't Want You to Know, John Viega, O'Reilly Media

[6]-Cloud Security and Privacy, Tim Mather, Subra Kumaraswamy, and Shahed Latif, 2009, O'Reilly Media

[7]-Junos Security, Rob Cameron, Brad Woodberg, Patricio Giecco, Tim Eberhard, and James Quinn

[8]-WIKIPEDIA, The free Encyclopedia, Cloud computing, 07 February 2011

[9]-"NIST.gov - Computer Security Division - Computer Security Resource Center". Csrc.nist.gov. http://csrc.nist.gov/groups/SNS/cloud-computing/. Retrieved 2010-08-22

[10]-WIKIPEDIA, The free Encyclopedia, Virtualization, 07 February 2011

[11]-The Open Group to look at EA, SOA and cloud computing, By James Denman, 03 Feb 2011, SearchSOA.com

[12]-NATO embraces cloud computing, http://www.dworld.de/dw/article/0,,14824382,00.html, 08 February 2011

[13]-How and why the military should adopt the cloud, http://www.infoworld.com/d/cloud-computing/how-and-why-the-military-should-adopt-the-cloud-484, 09 February 2011

[14]-DOD tackles information security in the cloud, Jan 20, 2011, http://www.defensesystems.com/Articles/2011/01/24/Defense-IT-1-DOD-cloud-computing-security-issues.aspx?Page=2

[15]-Gartnet, Jackie Fenn, Emerging Technology Hype Cycle 2010: What's Hot and What's Not

[16]-Gartnet, Gartner's Hype Cycle Special Report for 2009

[17]-Virtualization Arsenal, Jeff Lake, Fortinet, http://www.military-information-technology.com/mit-archives/190-mit-2008-volume-13-issue-6/1807-virtualization-arsenal.html.

SHARE TO WIN - NATO NETWORK-ENABLED CAPABILITIES

LTC Ştefan GROSU

Introduction

The Information Age carries implications for virtually all human endeavors, including the military profession. It's likely that these implications have or will produce revolutionary
changes in warfare, but that issue remains unresolved among academics and military specialists alike. The search for answers, however, has generated a new intellectual excitement about military theory. It also has uncovered some preliminary notions about national security that require attention now.

Unfortunately, the technical precision which characterizes information warfare techniques is insufficient for answering two other fundamental questions in international politics: Who are the players? and What are their intentions regarding one another? While it is clear that information warfare techniques are available to empower a far broader spectrum of both nation and non-nation state actors, the extent to which this has occurred remains ambiguous. We simply don't know with precision who the information warfare players are or will be. In like manner, it is not yet clear how enthusiastic the new players will be about using their new-found weapon.

The battlespace associated with Information Warfare (IW) has been a constantly expanding one, moving far beyond traditional military situations. In some quarters, IW has even been associated with the leveraging of information technologies to achieve greater effectiveness and efficiency. This has stretched the meaning of IW to the breaking point and has sowed more confusion than enlightenment[21].

The scope of information warfare and strategy (IWS) can be defined by the players and three dimensionsùthe nature of their interactions, the level of their interactions, and the arena of their interactions.

Nation states or combinations of nation states are not the only players. Non-state actors (including political, ethnic, and religious groups; organized crime; international and transnational organizations; and even individuals empowered by information technology) are able to engage in information attacks and to develop information strategies to achieve their desired ends.

[21] Alberts, David S. - *Defensive Information Warfare*, Washington, D.C.: Department of Defense Command and Control Research Program, 1996, p. 1

I. NETWORK CENTRIC WARFARE

I.1 Information superiority

Information superiority in military operations is a state that is achieved when competitive advantage (e.g., full-spectrum dominance) is derived from the ability to exploit a superior information position. In military operations this superior information position is, in part, gained from information operations that protect our ability to collect, process, and disseminate an uninterrupted flow of information while exploiting and/or denying an adversary's ability to do the same[22]. Achieving information superiority increases the speed of command preempting adversary options, creates new options, and improves the effectiveness of selected options. This promises to bring operations to a successful conclusion more rapidly at a lower cost. The result is an ability to increase the tempo of operations and to preempt or blunt adversary initiatives and options.

The bottom line for value creation in military operations involves the detection, identification, and disposal of the most important targets at any given time.

In order to successfully engage a target, all of the following must

be accomplished within a certain amount of time. First, the target must be detected. Second, it must be identified. Third, the decision to engage the target

must be made. Fourth, the decision must be conveyed to a weapon. Fifth, the weapon must be aimed and fired.

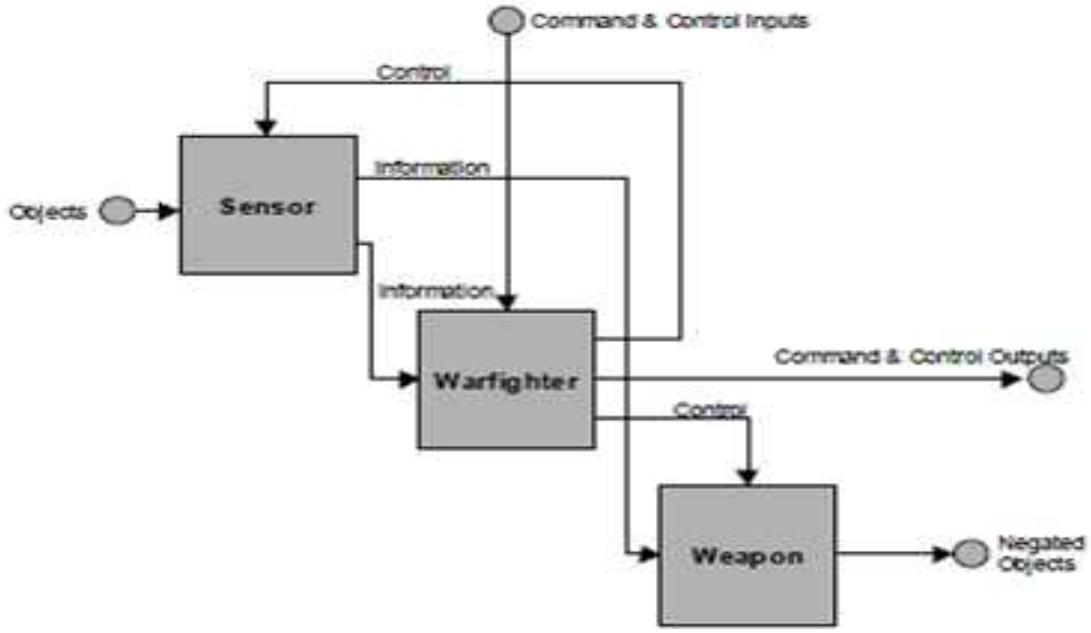


Fig.1 Platform Centric Shooter

To illustrate the power derived from sharing information, take the problem of assigning targets to actors. This problem can be formulated either as a centralized (unconstrained) or a decentralized

[22] Alberts, David S., Garstka, John J, Stein, Frederick P. - *NETWORK CENTRIC WARFARE: Developing and Leveraging Information Superiority*, Washington, D.C.: Department of Defense Command and Control Research Program, 1999, p. 54

(constrained) problem. That is, either there is: 1) one decision maker with no constraints on the information or processing power available to this decision maker, or on the decision maker's ability to communicate; or

2) there are several decision makers, each with limited vision and limited processing power (the sum of which may actually exceed that of the single decision maker).

The function of the network in this case is to bring together partial pictures, assemble them into a unified whole, and then convey the product of the decision making process to each actor.

I.2 Interoperability

Interoperability is "the ability of different forces to exchange services so as to operate effectively together." The key terms are "to exchange" and "operate effectively together." Both are vital to the progress and success of nations and organizations working together, or interoperating. The term interoperability is also used in technical systems engineering, which takes into account cultural, social, political and organizational factors that impact system-to-system performance. Together, these combined interoperability definitions, when put into practice, can pave the way for unwavering mission success.

It seems appropriate that any discussion of transformation should start with Network Centric Warfare (NCW), the concept of linking all aspects of warfighting into a shared situation awareness and shared understanding of command intent so as to achieve a unity and synchronicity of effects that multiplies the power of military forces.

1.3 Network Centric Warfare

NCW is about human and organizational behavior. NCW is based on a new way of thinking, network-centric thinking, and applying it to military operations. NCW focuses on the combat power that can be generated from the effective linking or networking of the warfighting enterprise. It is characterized by the ability of geographically dispersed forces to create a high level of shared battle space awareness that can be exploited via self-synchronization and other network-centric operations to achieve commander's intent. NCW supports speed of command, the conversion of a superior information position to action. NCW is transparent to mission, force size, and geography. Furthermore, NCW has the potential to contribute to the coalescence of the tactical, operational, and strategic levels of war. In brief, NCW is not narrowly about technology, but broadly about an emerging military response to the Information Age[23].

[23] Alberts, David S., Garstka, John J, Stein, Frederick P. - *NETWORK CENTRIC WARFARE: Developing and Leveraging Information Superiority*, Washington, D.C.: Department of Defense Command and Control Research Program, 1999, p. 88

Network-centric operations, then, are the application of the concepts and principles of Network Centric Warfare to military operations across the spectrum of conflict from peace, to crisis, to war.

As this description suggests, Network Centric Warfare and network-centric operations are closely aligned with the emerging new technologies of the so-called Information Age. But, the description does more than that. It implies that the new technologies by themselves are not enough and that the real potential of network-centric operations stems from some innovative thinking as to how to use these technologies. Thus, the new technologies must be accompanied by changes in organization, doctrine, and tactics.

The common thread that runs through the definition of Network Centric Warfare, the introduction of new technologies, and the exploration of a concept of effects-based warfare is the search for greater combat efficiency. That is, the purpose of each technology and concept is a reduction in the relative amount of military or other power needed to undertake a given mission, to fulfill a given task, or to create a specific outcome. The attraction of Network Centric Warfare and effects-based warfare is the prospect that they can yield improved combat efficiency. The challenge is to understand how they might do this and what combination of technologies used in support of which concepts would yield the greatest combat efficiency.

NCW recognizes the centrality of information and its potential as a source of power. This potential is realized as a direct result of the new relationships among individuals, organizations, and processes that are developed. These new relationships create new behaviors and modes of operation.

NCW focuses on the combat power that can be generated from the effective linking or networking of the warfighting enterprise.

NCW supports speed of command—the conversion of superior information position to action.

As the ranges of our sensors and weapons increase and as our ability to move information rapidly improves, we are no longer geographically constrained. Hence, in order to generate a concentrated effect, it is no longer necessary to concentrate forces.

NCW, with the significantly improved capabilities, has the potential to significantly impact the outcome of military operations and enable commanders to change their operational and strategic calculus. For example, by increasing battlespace awareness, creating shared awareness, and helping to ensure that the most accurate information is made available to those who need it, situations like those that arose in Mogadishu, Belgrade, and the Gulf War can be avoided in the future, or have more favorable outcomes[24].

Empowered by knowledge, derived from a shared awareness of the battlespace and a shared understanding of commanders' intent, our forces will be able to self-synchronize, operate with a

[24] Alberts, David S., Garstka, John J, Stein, Frederick P. - *NETWORK CENTRIC WARFARE: Developing and Leveraging Information Superiority*, Washington, D.C.: Department of Defense Command and Control Research Program, 1999, p. 84

small footprint, and be more effective when operating autonomously. Effective linking requires the establishment of a robust, high-performance information infrastructure, or infostructure, that provides all elements of the warfighting enterprise with access to high-quality information services. NCW is built around the concept of sharing information and assets. Networking enables this.

There is significant need to harmonize the technical and operational aspects of net-centric warfare and net-centric operations among multiple nations, in order to support coalition activities and joint operations. The general theme of any network- centric platform is to achieve mission success by delivering information to the warfighter in a timely, efficient manner. However, network-enabled capabilities (NEC) will never be developed without proper operational concepts, policy and doctrine.

While there may never be complete interoperability throughout the various commands, net-centric capabilities and operations, it is possible for systems and capabilities to have a common link without creating a new system to function with a legacy system.

II. Network-Enabled Capabilities

II.1 NEC Themes

While NCW suggested a new way of looking at how to accomplish the functions associated with command and control, many chose to focus on providing the information infrastructure to support network-centric operations, thereby neglecting the need to explore new approaches to command and control. The network-enabled capabilities (NEC), is aimed at emphasising capability rather than the infrastructure[25].

NEC aspires to enable platforms and C2 capabilities to exploit shared awareness and collaborative planning, to support the understanding of command intent, and to enable seamless battlespace management. To accomplish this will require a wide range of network enablers from across the Lines of Development.

These enablers have been brigaded into nine NEC Themes (eight of the themes cover equipment capability and one the acquisition process.):

a. Agile Mission Groups

The formation of network-centric forces must not be constrained by 'hardwired' equipment configurations based on organisational structures. The network-centric force will be composed of capability components brought together to form Agile Mission Groups to undertake specific operational tasks. The tasks may be long lasting. Once complete, the elements of the Agile Mission Group will return to their host, functionally or environmentally oriented organisation. Shared awareness within an Agile Mission Group will need to be very high in order to understand and

[25] Alberts, David S., Huber, Reiner K., Moffat, James - *NATO NEC C2 maturity model*, Washington, D.C.: Department of Defense Command and Control Research Program, 2010, p. 21

achieve their common goal, but lower between it and other Agile Mission Groups, where a general understanding of the intent or position may be all that is required. The high level of Shared Awareness will require elements within an Agile Mission Group to have a corresponding high level of "application interworking" to ensure the synchronisation of planning, control and effects. The concept of "asymmetric collaborative working" makes attainment of a high level of application interworking complex. This concept recognises that capability components within an Agile Mission Group may have very different levels of IT support (or indeed training and expertise). For example, An Agile Mission Group composed of dismounted infantrymen and HQ based users. This requires an interworking regime that can cope with differing levels of capability.

b. Fully Networked Support

The membership of operational forces, including Agile Mission Groups, should not be restricted to the in-theatre forces but will include non-frontline government bodies, industry and public services. Hence, a dynamic resourcing mechanism is required that makes use of non-frontline government bodies, industry, academia and public service capabilities to support the in-theatre capability, e.g., logistics, data/image analysis and medical.

c. Flexible Working

Agile Mission Groups will be how network-centric forces exert effect within the battlespace. Ideally, Agile Mission Groups will always be made up from elements suited to the role they have been tasked with. However, this will not always be possible.

Elements will need the flexibility to:

- Undertake tasks not supported by their primary roles.
- Work with elements that it was never intended to work with.
- Work in multiple Agile Mission Groups simultaneously, maintaining coherent "situational representation" between the Agile Mission Groups and not compromising their role in any of the Groups.
- Be able to change rapidly from one Agile Mission Group to another without disrupting the operation of either Group.

d. Synchronised Effects

An efficient, effective dynamic planning and C2 system is a key element of NEC, and is vital to co-ordinating the multiple and diverse strands of operations to achieve synchronicity. Without it, the complexity of planning and managing a number of simultaneous tasks with different tempos, and of making dynamic use of resources, will not be feasible. This will require breaking down the barriers within command and control and exercising it as a single process; the hard distinction between planning and execution must be broken down and replaced by a single dynamic planning, tasking and execution process, thereby increasing tempo and responsiveness. The battlespace will contain many separate such planning teams, who themselves could be distributed, and their planning

processes must be synchronised; thereby creating a more synchronised force. The co-ordination between the planning groups will include the co-ordinated use of the battlespace environment, which encompasses such diverse elements as airspace, waterspace and RF spectrum. This co-ordination is done as part of "command management".

e. Effects Based Planning

Network centric forces will have access to many other effectors within the battlespace above and beyond the traditional effectors. In particular, Information Operations, considered currently as a separate, stand-alone capability, should be brought into the mainstream of military planning and execution; thereby treating Information Operations as just another battlespace effector and hence providing more operational scope to the battlespace commander. To fully utilise all these effectors operational planning will have to change from an attrition based process to an effects-based one.

The following are required to allow the full scope of Effects Based Planning (EBP) and Effects Based Operations (EBO):

- A fully capable EBP capability, operational through all levels in the Ministry of Defence and in all other Government Departments that has an impact on political/military/ economic aims (including Foreign Office, Home Office, Treasury). Within the MoD (and potentially elsewhere) this capability will have to be able to operate with the dynamic, distributed planning systems required for the delivery of synchronised effects and the management of Agile Mission Groups;
- Modelling tools that can allow prediction and "what-if"ing across the whole domain of Effects Based Operations (EBO), including predicting the interaction between military, diplomatic and financial effects;
- Tools to assess the effects of operations across all domains - for assessing military effect, capitalising on the greater sharing of information to allow more rapid/simultaneous assessment, and add assessment of the effect in non-military domains (political and economic).

f. Shared Awareness

Shared awareness is a central facet of NEC and underpins many of the other themes, including Agile Mission Groups, Synchronised Effects and Effects Based Planning. Achieving awareness is a cognitive activity that results in a gaining of a personal understanding of what is happening, why and what could happen in the battlespace. Gaining
understanding requires appropriate processes and training as well as supporting equipment.

Shared awareness, in the context of NEC, is the ability to communicate an individual's
understanding to others in order that, as a group, there is some level of shared understanding.

Shared Awareness has two principal elements:

- The gathering, maintenance and presentation of relevant information. This will include

extracting information from all relevant, available sources, seeking specific additional information and/or clarification, and combining all this information to produce a local representation or "picture" that meets an individual user's needs. However, Shared Awareness will only be supported if separate local "pictures" are consistent with each other. The goal for NEC is a set of consistent pictures across the battlespace rather than a common one.

- Developing a shared *understanding* of the situation. Understanding exists not in the underlying information gathered from across the battlespace, but in the mind of the user. To achieve a common shared awareness, the understanding must be communicated to others. If the users are co-located then verbal and non-verbal (body language) means can and are used. If the users are distributed equipment must be used to support the "encoding of understanding" and the transmission of it. The equipment could attempt to replicate co-location, for example video teleconferencing, or could encode understanding for presentation on standard IT equipment, for example using text and graphics.

g. Full Information Availability

The future battlespace will be teeming with information. NEC will make much of this information available to users. This will include access to the widest range of information sources, including military sources (ISTAR, intelligence sensors, weapons sensors etc), civil sources (news feeds, environmental information, etc), encyclopaedic information, archived information, information available from sensors of opportunity and information collected but not fully exploited. However, this does not mean that all this information will be pushed to the user; overwhelming him with irrelevant information. On the contrary, only a very small part of this information "pool" will be presented to any user (for example orders, plans and pre-defined information needs). The rest of the information the user, or application, will have to actively search for from across and beyond the battlespace. To enable this, the user, software application or system, will be provided with the capability, tools and mechanisms, to proactively, rather than reactively, search for, manipulate and exchange information. The capability must allow the searching and exchanging to take place not only within national systems but also those of coalition partners and the Internet. This will require the tools and mechanisms to handle data of different classifications securely. In summary, this proactive searching mechanism must be an adjunct to, not a replacement of, other information management mechanisms such as selected information push and broadcast, providing the user with a rich set of information access and retrieval mechanisms.

h. Resilient Information Infrastructure

A Resilient Information Infrastructure is required to provide a secure and assured environment to meet the requirements of a dynamic battlespace equipment capability, and in particular the demanding, dynamic requirements of Agile Mission Groups.

The requirements of the Resilient Information Infrastructure include:

- The capability to share information across the battlespace, and allow all users (human or machine) access to the information that they require for planning, execution and monitoring of operations. This capability should allow information to flow transparently across domains, be robust in the face of communications limitations and ECM, and should support the information user (human or machine).
- Efficient management of information sharing, as demanded by the operational situation, and the requirements for information access.
- The provision of an assured end to end performance based upon the business need.

i. Inclusive Flexible Acquisition

The equipment acquisition process must be enabled to allow it to realise the aspirations of NEC. These requirements range from a more coordinated approach to equipment capability definition through to a holistic view of the equipment programme: the relationship between individual acquisitions and the delivery of coherent packages of military capability. Of prime importance in a domain where the fundamental technology is evolving rapidly is the ability to take advantage of new technology. Without this agility, exploitation of leading edge technology will be impossible.

III. NATO Network-Enabled Capabilities

III.1 Concept

By being a leader in collaboration between various nations and having the ability to take joint action in addressing security and other issues of concern, NATO strives to work among the alliance and the Partnership for Peace countries to promote interoperability. Streamlining this process, identifying and implementing warfighter capabilities continues to be a challenge, but its success is maintained by the alliance's and industry's support.

The optimal aim for NATO and the nations is to advance NATO's operational capability by improving the way it shares and uses information available throughout the alliance to achieve the desired operational outcomes. Additional benefits of achieving NNEC include a faster and seamless flow of information and a higher operational tempo. It also serves as a key enabler for moving from de-confliction through coordination, to integration and to coherence in reference to the armed forces, and encourages communication and discussion across the alliance.

The NATO NEC C2 Maturity Model (N2C2M2) provides a framework that can be used to assess appropriateness of the C2 approaches and related capabilities possessed by a collection of entities (both military and non-military). The model consists of five C2 maturity levels that are associated with the degree to which an entity or a collective is able to effectively conduct network centric operations.

Operating at a higher level of C2 maturity provides collections of entities (or an entity) with a larger set of C2 approach options from among which to employ. Having options is of little value unless one understands which of the available options is appropriate for the situation at hand. Thus, a maturity level not only involves being able to select from a particular set of C2 approaches but also the ability to recognise the appropriate C2 approach and the ability to transition from one approach to another, as appropriate. This dynamic applies not only to preparing for an endeavour but also during an endeavour as required.

Since increasing command and control capability is not an end unto itself, progress towards NEC requires that links be made between C2 maturity levels and NNEC capability levels. The maturity model establishes these performance-related links. Knowing where you are is not sufficient for the journey at hand. One also needs a roadmap that shows how to get to the next step along the way. The N2C2M2 helps in this regard by identifying what is needed to move an entity (a nation, or a coalition) from one maturity level to the next.

Thus, the N2C2M2 provides a set of milestones that can be used by NATO as well as nations for C2 and NEC planning (strategic planning for an expected set of mission contexts or planning for a particular mission). It also provides a set of metrics to measure progress toward the achievement of a desired level of C2 maturity .

III.2 Evolution

At the Prague Summit in November 2002, NATO recognized that transformation of the military based upon Information Age principles was essential, and pursued a course of transformation denoted as NATO Network-Enabled Capabilities (NNEC).

In November 2003, nine NATO nations (Canada, France, Germany, Italy, The Netherlands, Norway, Spain, The United Kingdom and The United States) signed an arrangement to join in funding a feasibility study on NATO Network Enabled Capability (NNEC) as an important step towards NATO transformation. The study was carried out by the NATO C3 Agency (NC3A).

In June 2009, NATO defense ministers approved an action plan to improve interoperability, including the development of revised policies and a new strategy to integrate interoperability into the alliance's defense planning process. In December, the North Atlantic Council agreed to the new interoperability policy and the associated strategy for enhancing interoperability. In the agreed policy, interoperability is re-defined in practical terms and a set of principles has been established to guide the application of this new policy: Interoperability is the ability to act together coherently, effectively and efficiently to achieve allied tactical, operational and strategic objectives.

According to the new policy, interoperability has three main dimensions:

- technical (hardware and systems),
- procedural (doctrine and procedures) and

- human (language, terminology and training)—that complement each other.

Strengths in one dimension can mitigate weaknesses in another.

In keeping with enhancing military effectiveness of existing and emerging programs, NATO has recognized the importance and relevance of working towards a network-centric approach to operations throughout the nations and, when applicable, with partner nations. Achieving network enabled capabilities within NATO will take dedication to the process and adoption of a series of best practices. However, the outcome will enable operational planning, joint deployment and sustainment of forces with relevant and timely information, which will increase combat power and mission effectiveness.

III.3 Definition

According to the official definition The NATO Network Enabled Capability (NNEC) programme is the Alliance's cognitive and technical ability to federate various capabilities at all levels, military (strategic to tactical) and civilian, through an information infrastructure.

NNEC is NATO's commitment to standardize and harmonize NATO and national Network Enabled Capability (NEC) programmes. Common cognitive and operational aims will revolutionize the way NATO forces fight in future conflicts. By improving collaboration in an open and dynamic information environement.

NNEC enhances the efficiency and effectiveness of the Alliance through a networking and information infrastructure.

But the main objective of the NNEC programme, illustrated by the slogan "Share to Win", is to initiate a culture change that begins with people. Interacting with each other and sharing information will lead to better situational awareness and faster decision making, which ultimately saves lives, resources and improves collaboration between nations.

III.4 Mechanisms

To create the framework in which NNEC capabilities can evolve, three coherence areas have been defined:

- Operational Concept Requirement Implications: maintains a focus on military and civilian operations requirements.
- Architectures and Services Definition and Standardization: maintains a focus on the service-oriented approach and architectures for specification and definition of the use and reuse of capabilities.
- Implementation: maintains a focus on developing the roadmap for NNEC's success.

In addition to the three coherence areas, a steering group monitors the progress and reports to the political level.

The NNEC is a complex and multifaceted concept that promises to significantly improve military operations by federating the constituent capabilities used to conduct them. NNEC implies several dimensions, components and principles that need to be applied cross functionally to all capabilities used to conduct a mission. Only when such understanding is homogenous across stakeholders and users can we achieve the real benefits of providing the required information when and where needed, and of having the necessary decision support tools.

In line with the above, the NNEC Integrated Capability Team (ICT) at NATO HQ SACT has been established with a dual purpose in mind: to develop the NNEC concept (focused on the transformational elements), and ensure homogeneous understanding and coordinated implementation.

By following these common rules and guidance, the nations will have the opportunity to converge to the NNEC compliancy for all their future projects and programs while drastically increasing their level of interoperability. This task started late 2009 and should be completed by 2014.

Finally, to promote NNEC to NATO and the nations, NATO/ ACT is conducting a NNEC awareness campaign that includes the NNEC conference, a Website (http://nnec.act.nato.int), a quarterly newsletter and a multimedia DVD development featuring NNEC programs and projects.

III.5 Components of the policy

The networking and information infrastructure (NII) is the supporting structure that enables collaboration and information sharing amongst users and reduces the decision-cycle time. This infrastructure enables the connection of existing networks in an agile manner.

As such, the NII is the "physical reality" of NNEC. The NII, the technical standards, associated policies and procedures are the components which are necessary to realising the promises of NNEC and providing NATO's leaders and commanders with the "information and decision superiority" discussed above.

Although the benefits of NNEC are only achieved when the cognitive capabilities of the leaders are engaged, the need for the environmental conditions and tools necessary to implement them do not "happen by accident". The NII is the very intentional and deliberate framework necessary for NATO and Nations to begin the transition to a knowledge enterprise necessary to not only increase the effectiveness of the cognitive process but enable the implementation and execution of guidance with synergy of components and responsive alacrity.

Bringing NNEC capabilities to existing projects and programs that will affect and improve warfighters' missions is a complex process, but it can be and has been done.

NNEC has an important technical component, but the capability is mainly a question of people. If people in NATO and in the NATO nations are not strongly involved, nothing will happen to realize NNEC. It comes down to being a cultural change: changing the way we share information instead

of protecting it, and preparing the operator to be able to make use of all available technical opportunities.

Nations' and NATO's future operational environments will demand substantial transformation at all levels—strategic, operational and tactical. Integrated planning and execution, information sharing and responsive support are necessary to achieve success with fewer personnel and fewer resources. Ideally, this would also mean fewer casualties, but an overall more effective alliance.

In today's dynamic environment, NNEC will enable NATO and the nations to conduct more complex operations and—despite the deployment of fewer forces—conduct these operations with more efficiency and greater overall effectiveness.

Doing more with less, particularly in today's economic climate, is vital to nearly everyone. The alliance is no exception. Although its resources and budgets have tightened, it must deliver the same, or greater, levels of service on the battlefield, in peacekeeping missions and to internal programs. Interoperability will increase warfighters' probability of mission success, even though they rely on current levels of resources and capabilities.

III.6 Human factor

The ultimate goal of NNEC is to provide an operational advantage to the warfighter. One aspect of this challenge is to apply our knowledge of human behavior, organizational dynamics and technological innovation in ways that optimize NNEC's benefits to warfighters.

The human factor, consisting of the complex and largely intangible web of human behaviors and abilities, is more likely than the technology factor to determine ultimate system effectiveness. In fact, while a certain amount of automation may enhance human performance, too much automation can actually degrade the human's ability to assimilate and process information in ways that result in effective performance. Achieving a proper balance will allow humans to exploit more effectively the capabilities that modern technology offers.

Interoperability must be achieved in both the human and technical arenas. Human interoperability includes complex factors such as language, ethics and social beliefs. While shared data and information services form a technological foundation for NNEC, people and processes are necessary to transform the technical capability into knowledge and action. To fully achieve NNE, the alliance must address human cognitive processes such as situational awareness, sense making and decision-making.

Maximizing human performance requires a clear understanding of factors that impact information sharing, processing and decision-making. Those factors include policy barriers, information flow, communication gaps, standards, procedures and protocols for social information sharing and flow, and trust building across different cultures under stressed conditions. New concepts for doctrine, organization, strategy and tactics must be developed to effectively deliver net-centric capabilities to

both conventional teams and distributed, virtual teams that may need to assemble at a moment's notice to support a joint endeavor.

III.7 Other essential players

Industry and academics are essential players in the realization of NNEC. For this reason, NATO constantly reaches out to these groups. As a part of that effort, NATO permanently liaises with the Network Centric Operations Industry Consortium (NCOIC) in order to support their vision of "facilitating global realization of the benefit inherent in network-centric operations."

With this relationship, NATO aims to facilitate industry's working together, as a global organization with membership open to all interested parties. The aim is to apply the potential of NNEC to NATO operational challenges in a coordinated, interoperable and efficient manner, by pursuing adoption of the same NNEC tenets across the variety of military and civilian capabilities, and using existing and emerging open standards and processes. The ultimate benefit for warfighters will include maximizing information age capabilities. Industry adoption of common principles and standards will improve operational resilience, reduce the cost of developing new capabilities, and lower integration and administrative costs for the procuring agencies, resulting in capabilities that can cost effectively and securely interoperate.

Transformation in the context of the alliance is a continuous and proactive process of developing and integrating innovative concepts, doctrines and capabilities in order to improve the effectiveness and interoperability of NATO and partner forces. Allied Command Transformation will deepen the transformational effort by clearly identifying the specific military problems that need to be solved to increase alliance mission effectiveness, and will then focus maximum staff effort on achieving the necessary changes.

NEC is not limited to NATO or the nations; it truly affects everyone at every level. To this end, the NNEC ICT organizes an annual NNEC conference aimed at disseminating the latest information on the topic and to foster an environment for discussion. The conference is NATO's primary forum to exchange information and views on a wide array of net-centric, NEC and NNEC related topics between all stakeholders.

III.8 Share to win

It is clear that NNEC and the human aspects go hand in hand. However, there is a technical aspect that must not be overlooked. NNEC will play into a new revolution in how information is handled and delivered to the commander and soldier on the battlefield. This revolution will bring changes to warfare, which can be compared to those changes caused by the first and second industrial revolutions.

The same revolutionary developments can be expected from today's information age. It will be seen that transmitting information from different sources through military hierarchies while supporting information exchange will allow for unprecedented military forms of organization, agility and changes of structures on the fly. Just as both revolutions happened, so will NNEC.

Complete, seamless interoperability will likely never be achieved. Gateways, interfaces and more will always exist due to the nations' and military services' different approaches. The reliance on similar material, such as common standards, similar doctrine and culture, shape the best conditions for interoperability. Experience has proved that the cultural aspects have always been underestimated, causing unaffordable delays in achieving interoperability.

There is just one last piece that will assist in ensuring NNEC realization: You share to win!

The NNEC programme provides various benefits to all levels, military and civilian.

Some of these benefits are:

- Improved efficiency
- Drastic increase in interoperability between nations
- Improved and secure way of sharing information
- Better information quality
- Faster decisions and speed of command.

The NNEC can be used in various ways by a number of different communities, including:

- Strategic planners can use the model to determine what C2-related capabilities are needed to face current and future challenges in a variety of different contexts;
- Programmers and budgeters can use the model to support a variety of investment decisions and doctrine development;
- Educators and trainers can use the model to help individuals and organisations better understand the nature of collective command and control and its implications;
- Researchers can use the model to help design experiments, campaigns of experimentation, and exercises;
- Professionals, schools, and colleges/universities can use the model to structure lessons learned and analyses;
- Researchers can use the model to formulate hypotheses and as a framework for conceptual C2 models.

NATO's future operational environment will demand delivery on the promise of NNEC but these will be focused on providing substantive improvement to the commander at all levels: strategic, operational and tactical. Information sharing, data deconfliction, integrated planning and execution and responsive support are necessary to achieve success in the complex environment. NATO no longer has the comparative "luxury" of assuming its dominant role in the international environment. More and more, military assets are finding themselves as partners with NonGovernmental and

International Organizations (NGO/IO) a community inherently proscriptive in their relationships with government organizations, particularly the military. In the future, NATO must be able to share information openly with these partners and receive information from the as well to accomplish an evolving mission set in an environment increasingly dynamic, increasingly volatile and increasingly lethal.

NNEC will enable NATO's accomplishment of more complex operations with smaller, yet more effective forces.

The goal is to get the right people to the right place with right equipment in a timely manner, a manner that will achieve the greatest effect with the greatest efficiency. Once engaged, whether in support of humanitarian missions or combat, commanders will have greater awareness, supported by a responsive network which supports critical decisions at the lowest levels.

CONCLUSIONS

Evolving in the future, the first step is to have NNEC be adopted by the people of NATO. Then NNEC will transform everyday business from the onset of a new project to its completion. This evolution will meld previous "hard" delineations between strategic, operational and tactical into a more holistic vision of the military environment. This singularly critical evolution will maintain NATO's significant role as a force majeure.

NNEC will evolve slowly as does every change but that change will be informed by the operational requirements we are seeing not only in ISAF but through NATO's aggressing training and exercise programme as well. Initially, people will look for ways of keeping what they are used to today and only "build" some sort of "translation" capability that will allow them to share information with others. However, once they begin to see the power which shared information and situational awareness creates they will quickly look for ways to enhance the capability to the point that procedures or processes become complimentary and there is seamless information flow. Then it will be a short step to enabling all of NATO and the Nations by empowering the commanders with the tools to make rapid, quality decisions.

Nations will realize a better sense of military situations and a decreased level of casualties caused by lack of interoperability.

NNEC will continue to be a cultural change for all. What NNEC requires is a reliable concerted effort from all stakeholders, since everyone will benefit from a realized NNEC environment, which will allow a new way of doing business and an enhanced way of obtaining and sharing important data among the NATO nations by fusion of information from different assets.

REFERENCES

1. Alberts, David S., Garstka, John J, Stein, Frederick P. - *NETWORK CENTRIC WARFARE: Developing and Leveraging Information Superiority*, Washington, D.C.: Department of Defense Command and Control Research Program, 1999

2. Alberts, David S., Huber, Reiner K., Moffat, James - *NATO NEC C2 maturity model*, Washington, D.C.: Department of Defense Command and Control Research Program, 2010

3. Moffat, James - *Complexity Theory and Network Centric Warfare*, Washington, D.C.: Department of Defense Command and Control Research Program, 2003

4. Alberts, David S. - *Defensive Information Warfare*, Washington, D.C.: Department of Defense Command and Control Research Program, 1996

5. Smith, Edward A., Jr. -*Effects Based Operations: Applying Network-centric Warfare in Peace, Crisis, and War*, Washington, D.C.: Department of Defense Command and Control Research Program, 2002

6. Wheatley, Gary F., Hayes, Richard E. - *Information Warfare and Deterrence*, Washington, D.C.: Institute for National Strategic Studies, 1996

ANONYMITY ON THE INTERNET WITH TOR: THE SECOND GENERATION ONION ROUTER

Cpt. Eng. Florian ŞOICA

INTRODUCTION

21st century modern life cannot be imagined outside of the information technology revolution. For most of us the daily routine includes without a doubt a lot of Internet related activities. Even if it's about online banking, searching documentation for some project on Google, finding leisure related information, social networking, online shopping or only the simple e-mails, any Internet user became to some degree dependable upon the global network.

Staying in front of your home personal computer or laptop offers to most of the internauts a false sense of security. Most of us registered at least once to some website using a peculiar username, not at all related to our life or daily existence. Others are using mIRC or another online chatting technology using nicknames inspired by their personal idols, mythological characters, movie heroes or Hollywood stars and hall of famers. But are we in fact hiding our identities only by not revealing it to the persons we come in contact with in the virtual environment? Does the Internet provide some layer of security that can really hide our identity from the general public only by sitting in front of the home computer? Or there are methods of identifying an Internet user's identity? And what online identity really means? Could an Internet user be identified by its service provider and by all the Internet websites and services that he connects with? And if so, what can we do if we want online anonymity? Can this virtual identity be concealed by any means and, if so, to what extent these methods provide true anonymity for those interested in achieving it?

The concept of online anonymity refers, at least for the purpose of this paper, to the means necessary to provide the Internet users a browsing platform which does not give anyone the ability to trace or link web activity or personal information back to the user. Internet anonymity is about using the Internet without revealing your true identity. And identity includes, of course, personal information, but extends also to computer information and geographic location.

There are more software solutions providing various degrees of anonymity for the Internet users, starting with anonymous web proxy services, VPN services and continuing with solutions such as Jondonym service (mixed cascades) and the Tor project, also known as the second generation onion routing network. The purpose of this paper is to familiarize the reader with the inner workings of the Tor network. The history of the Tor project will be the subject of one of the first chapters of the paper. We will try to explain why someone would want to become anonymous online and we will try to present the legal issues related with the usage of such a technology. We will address the

limitations and weaknesses of the solution and, also, we will try to identify some of the future Tor-based online anonymity projects.

Tor – free, open-source anonymity platform

I. WHAT IS ONLINE ANONYMITY AND WHO NEEDS IT?

Let's begin by trying to analyse in depth the concept of online anonymity and by presenting some of the reasons why the virtual space is not at all a private environment. As mentioned during the introduction, when we talk about online anonymity systems, we usually refer to software platforms for the Internet users which will not give anyone the ability to trace or link web activity or personal information back to them.

Why is this sort of tracing and tracking possible? The main reason is that, in order to connect to the Internet, a user needs to receive from his internet service provider (ISP) an IP address, a sort of virtual ID which uniquely identifies him on the Internet. The IP address is the main obstacle regarding one's online anonymity because it enables the ISP and the websites and services that the users are accessing to log all sorts of information based on this virtual identifier.

Also, for the outer world, the IP address reveals the internet service provider of the regular internet user, many times his geographic location and in the case of a company or computer centre, even what terminal the user is working on. In many cases, an IP address can be related directly to one person. The Internet user must always bear in mind that any online communication leaves all sorts of digital traces which can be acquired, saved and analysed. There are companies specialized in creating individual user profiles based on surfing related data, a part of a larger process, called data accumulation, of building databases of high economic value.

Anonymizing services such as Tor or Jondonym address the issue of IP tracking. Using those systems makes IP-based tracing of the internet traffic more difficult. Visits to websites, online posts, instant messages and other online communication means cannot be linked back to the original source, if the source's internet activity has been anonymized through Tor or similar technologies.

Anonymity techniques should not me confused with cryptography and encryption (although they embed such techniques), which are usually used to hide only the contents of the messages that is being sent online and not the source and the destination of the communication channel. Also,

anonymity is not steganography. Steganography is the art and science of writing hidden messages in such a way that nobody, apart from the sender and intended recipient, can suspect the existence of the message. This sort of approach is also called security through obscurity. In the case of Tor or similar anonymity systems, the attackers can tell when the intended target is talking through the anonymity network, but are not able to discern the destination of the target's traffic. The same is true at the other end of the communication channel, with a potential attacker being able to detect that the traffic reaching the destination is coming from the Tor network, without knowing anything about the original source.

Anonymous communications systems can serve different interests for different user groups, being useful in a plethora of activities and institutions, from government agencies, corporations, journalists and, of course, the regular internet user interested in achieving online anonymity. According to the Tor Project website, businesses can use Tor to research competition. Regular citizens don't want to be watched and tracked for reasons such as freedom of speech, child protection, personal medical and financial data privacy. Activists can use Tor to report anonymously about abuses, journalists can protect their information and sources using it and even militaries and law enforcement agencies could use Tor in order to protect their communications, investigations, and intelligence gathering online. With Tor, a blocked website can be reached, so Tor provides the means for circumventing Internet restrictions imposed by some countries, ISP or even websites themselves.

II. HISTORY OF THE TOR PROJECT

In a Technology Review article, an MIT independent media initiative, David Talbot mentions the fact that, as in the case of the Internet itself, Tor's origins can be traced back to a military research project. In the mid 1990s, in the Research Laboratory of the U.S. Navy based in Washington, an internet traffic anonymizer software prototype was built. The utility of such a software solution ranges from anonymizing a cover agent's Internet traffic, to hiding a wi-fi home user's browsing habits from the neighbour „sniffers" and to hiding relevant search queries and internet traffic from the prying eyes of data miners.

The original navy project ran exclusively on public military machines and caught the attention of Roger Dingledine, a cryptographer concerned about all Internet privacy related issues, such as the data collected about the internet users, their browsing activity and search history at ISP and website level. During a conference in 2000 he met a Naval Research Lab mathematician, Paul Syverson and they revived the project using money from DARPA and the Navy.

According to Wikipedia, an alpha version of the Tor software, with the onion routing network "functional and deployed", was announced on 20 September 2002. Roger Dingledine, Nick

Mathewson, and Paul Syverson presented "Tor: The Second-Generation Onion Router" at the 13th USENIX Security Symposium on Friday, August 13, 2004.

Tor was financially supported by the Electronic Frontier Foundation from 2004 to 2005 and nowadays the Tor software is developed by the Tor Project, a research and education non-profit organization based in the United States of America, through a diverse base of financial support.

III. WHAT IS TOR AND HOW IT WORKS?

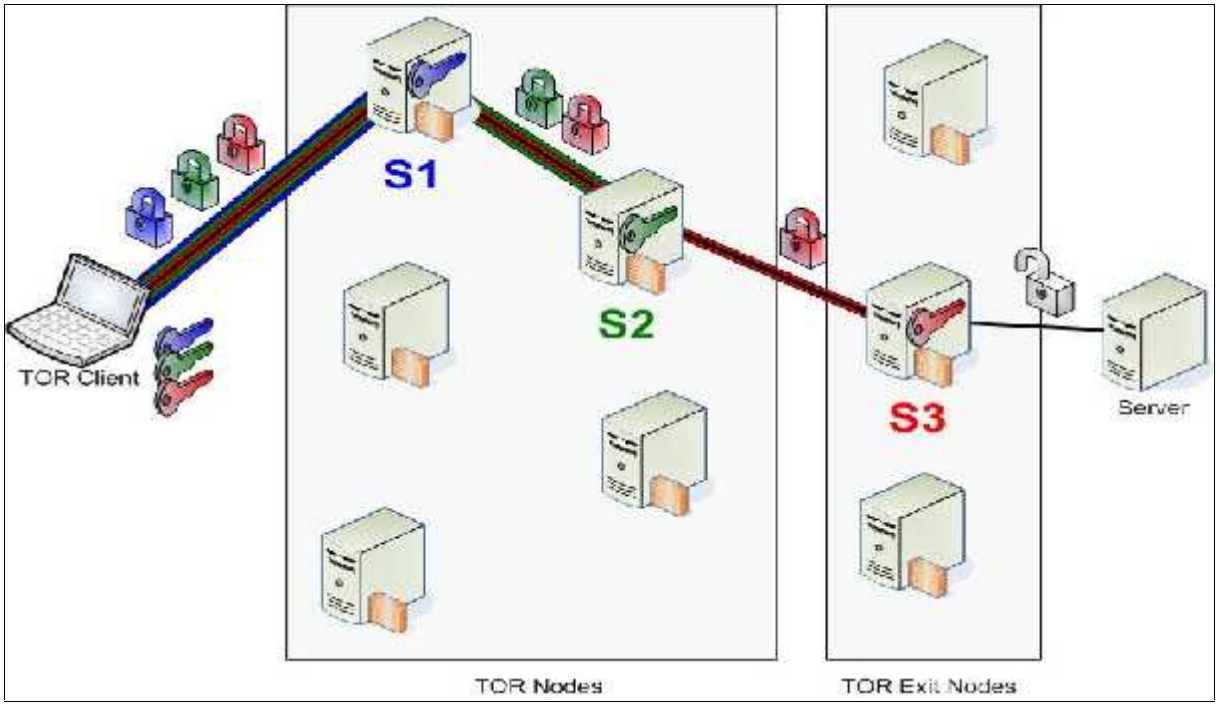
In this chapter we will try to take a more detailed look at the Tor's inner workings and following, in the next one, a very thorough technical overview of this online anonymity system. The main issue addressed by anonymization software such as Tor is, as previously mentioned, concealing its users' identities and their network activity from surveillance and traffic analysis. We've seen that encryption, as sometimes used with web browsing (SSL for HTTP secure connections), only hides message content, and not the traffic data: source, destination, size and timing. Traffic analysis is the study of such additional data to discover the behaviour and interests of groups and individuals. Traffic analysis based techniques are widely used to track people, for instance for marketing purposes, and the role of the anonymity systems is to protect the privacy of Internet users from traffic analysis.

Regarding from the Internet protocol suite point of view (link layer which manages ethernet addresses, internet layer – IP addresses, transport layer – TCP/UDP protocols, application layer which consists of the applications themselves), protocols such as TLS and his predecessor SSL are functioning at the application level. TLS and SSL encrypt the segments of network connections above the Transport Layer, using symmetric cryptography for privacy and a keyed message authentication code for message reliability. One must bear in mind that the encryption of the messages can prevent the ISP or any eavesdropper from discovering what is being transmitted through the network, but not the source and the destination of the information that is being sent. This sort of information can be very important, as previously mentioned. If a data mining company or ISP can sell information about the browsing of ISP's clients to marketing companies or about the medical conditions of those internet users to some job recruiting company, this could have a big impact on their lives.

Apart from the contents of the messages sent online (the data), which are of interest and therefore are usually encrypted through protocols such as SSL and TLS, one may want to also protect the metadata, so the data found in the packet headers: source, destination, timing etc. Normally, any TCP connection made by some software run on an Internet user's computer reveals his IP address. Tor is a privacy enhancing technology which works at the transport layer and allows an Internet user to make a TCP connection without revealing its IP address. It's most commonly used for HTTP connections, so for the usual online website browsing activities.

Tor works by passing encrypted messages from server to server (in this case the servers are called onion routers) within a distributed network. Each such node within the Tor network receives the encrypted message and decrypts the addressing information for the next server. The rest of the message remains encrypted with a different key and is then sent to the next server in the path. Each server can decrypt only the layer intended for it. This layering of encryption and routes ensures that no single server knows at the same time where the message being sent originally came from and its final destination. This technique, along with frequently changing the network path used for messages, prevents detection by traffic pattern analysis.

In the figure below the Tor client sens its data packets through a communication channel consisting of himself and 3 nodes (S1, S2, S3). The data packets from the client are then encrypted with the encryption key negotiated with the last node (S3, see the figure below), which is also called the exit node, afterwards encrypted with the encryption key negotiated with the middle node (S2) and finally encrypted with the key exchanged with the first node (S1). When the data packet is sent, it is decrypted once on each TOR server and forwarded to the next hop until it reaches the exit node (S3) which sends the decrypted packet to the destination server. Packets in the other direction are encrypted and decrypted the opposite way. TOR does not provide end-to-end encryption. Traffic from the exit node to the destination server is not encrypted by TOR. It "only" provides anonymity, nothing else. The exit node is able to view all the original traffic bits and bytes.



Credit: http://www.csnc.ch/misc/files/publications/the_onion_router_v1.1.pdf

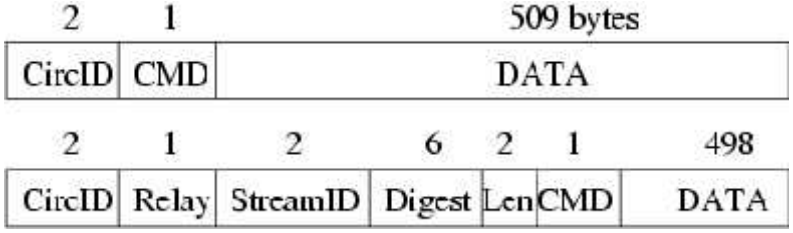
IV. TECHNICAL DETAILS

The Tor network is an overlay network; each onion router (OR, there are more than 1500 such nodes around the world) runs as a normal user-level process without any special privileges. Each onion router maintains a TLS connection to every other onion router. Each user runs local software called an onion proxy (OP) to fetch directories, establish circuits across the network, and handle connections from user applications. These onion proxies accept TCP streams and multiplex them across the circuits. The onion router on the other side of the circuit connects to the requested destinations and relays data.

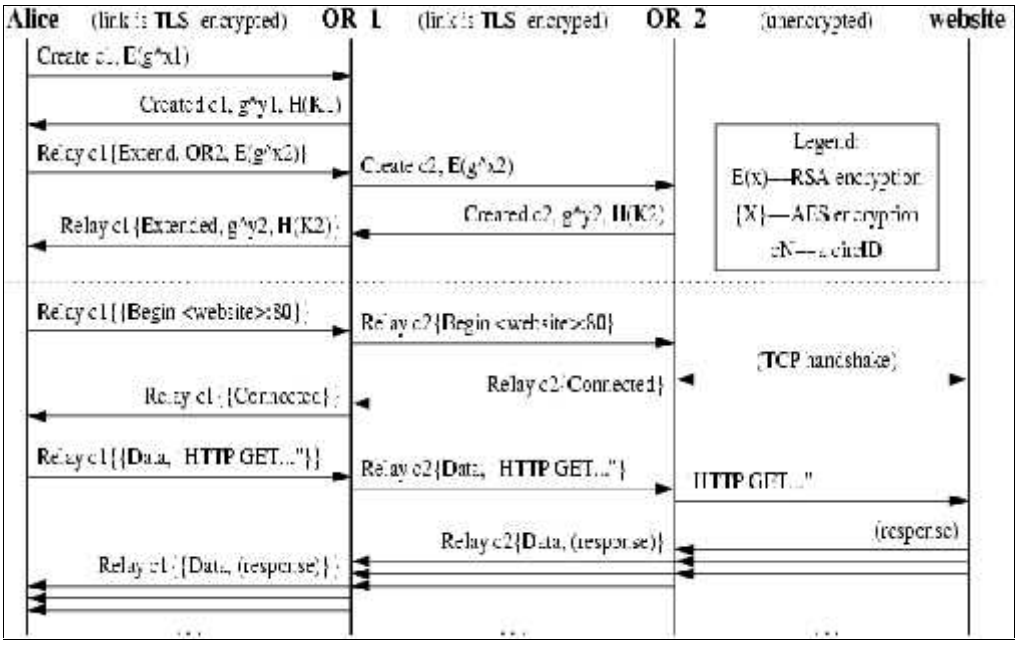
To create a network pathway (a circuit), the Tor client installed on the internet user's computer (OP) first obtains a list of nodes (ORs) from a directory server. The Tor clients know what the directory authorities are because the Tor software comes with a built-in list of locations and public keys for each directory authority (directory server). This also means that tricking the users into entering a fake Tor network could be accomplished only by offering them a modified version of the software. The original packages are digitally signed using GNU Privacy Guard, and a signature checking procedure is available.

After obtaining the list of nodes, the OP extends its path to the destination through encrypted connections, one hop at a time (a technique called telescoping path building). This constructed routing path changes at regular intervals. Each node in the path has information only about its predecessor and its successor, and not about other nodes.

Traffic flows down the circuit in fixed-size cells. Each cell is 512 bytes, and consists of a header and a payload. The header includes a circuit identifier (circID) that specifies which circuit the cell refers to (many circuits can be multiplexed over the single TLS connection), and a command to describe what to do with the cell's payload. Based on their command, cells are either control cells, always interpreted by the node that receives them, or relay cells, which carry end-to-end stream data. The control cell commands are: padding (currently used for keepalive, but also usable for link padding); create or created (used to set up a new circuit); and destroy (to tear down a circuit). The relay commands are: relay data (for data flowing down the stream), relay begin (to open a stream), relay end (to close a stream cleanly), relay teardown (to close a broken stream), relay connected (to notify the OP that a relay begin has succeeded), relay extend and relay extended (to extend the circuit by a hop, and to acknowledge), relay truncate and relay truncated (to tear down only part of the circuit, and to acknowledge), relay sendme (used for congestion control), and relay drop (used to implement long-range dummies). Below, a visual overview of the cell structure:



The circuit is constructed by the user's OP incrementally by sending an encrypted message and make key exchange with each OR in the circuit, one hop at a time. To extend the circuit to more ORs, the OP sends special extend-circuit packages through the circuit in order to inform the last reached node to extend it to one more hop (one more OR). The OP-OR connections are encrypted with the TLS protocol and the connection between the last OR in any circuit and the destination in not encrypted. After the Tor client finishes building the circuit and he shares keys with every OR in the circuit, the communication over that particular circuit can begin. As previously mentioned, the clients encrypts the message with each ORs key in layers and along the path every router can decrypt one layer and then sends the package to the next node in the circuit. Below you can see a 2-hop circuit construction phase and the beginning of data relaying through the constructed circuit, which in this case in the fetching of a webpage.



The hand-shake with every new OR in the circuit is made with packets encrypted using public-key cryptography (RSA) and, once the circuit built, the packets exchanged between the nodes are encrypted using AES, a symmetric cryptography protocol.

For an even more detailed technical description of the onion router network design, the original article signed by the projects' authors (Roger Dingledine, Nick Mathewson and Paul Syverson) is available at https://svn.torproject.org/svn/projects/design-paper/tor-design.pdf.

V. SOFTWARE NEEDED TO BECOME A TOR USER. HIDDEN SERVICES.

The main software projects developed by the Tor team are the Tor browser bundle - a Firefox-based browser package which includes Tor and is tweaked properly in order to provide safe browsing capabilities to the users, Torbutton - a Firefox add-on providing a 1-click way for the Firefox users to enable or disable Tor for Firefox, Vidalia - a graphical interface which provides a way to view and control Tor's settings and connections and Check - a webpage used for verifying the proper functionality of Tor on your home computer.

The Tor browser bundle can run on Windows, Linux or Mac OS X and also can run off a USB flash drive as a self contained and preconfigured solution. A variant of it it's the Tor IM browser bundle, which includes the facility to run over the Tor network an instant messaging client and chat named Pidgin.

Torbutton provides a plethora of configuration options such as disabling other plugins while using Tor, isolate dynamic content while Tor is running, hooking dangerous javascript, cookie clearing, browser cache management, browser history management, user–agent and time zone spoofing and others.

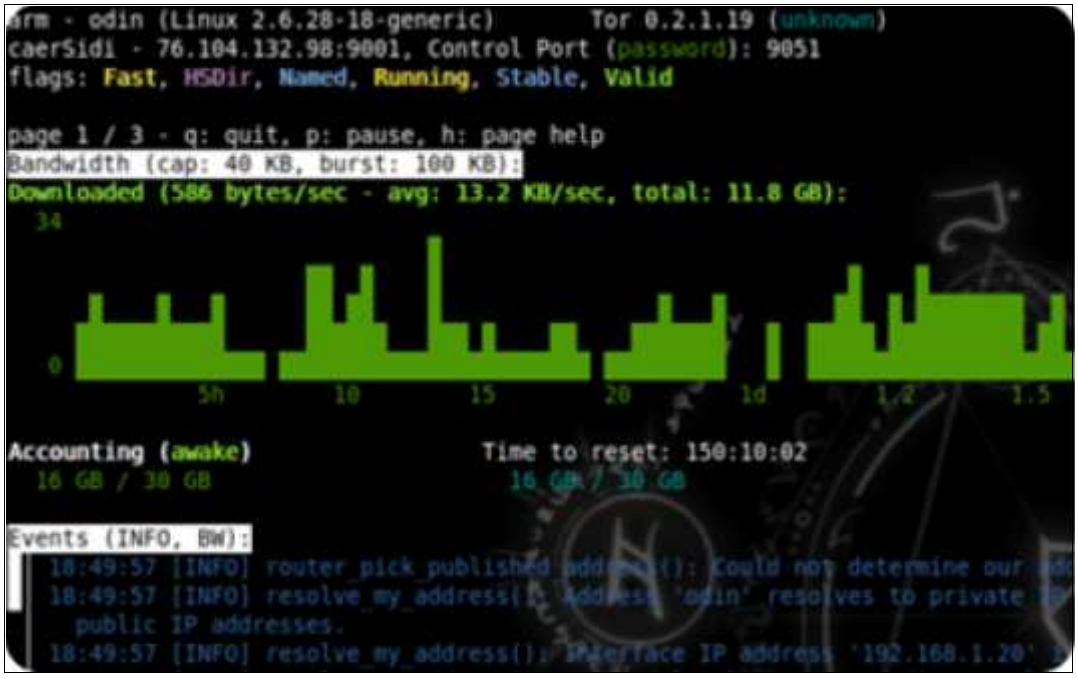
Vidalia is a cross-platform graphical controller for the Tor software and runs on Microsoft Windows, Apple OS X, Linux and Unix variants using the X11 window system. Vidalia, which requires Tor to be installed, offers a various range of control functionalities such as starting and stopping Tor, bandwidth consumption visualization, active circuits visualization and mapping, configuration of the client as a Tor bridge or relay.

Vidalia snapshot

Check is basically an webpage which runs at https://check.torproject.org/ and tells the person who accesses it if Tor is enabled or not on their home computer.

Check snapshot

Another nice software package is arm (the anonymizing relay monitor) - a command line interface which provides the functionality to monitor Tor's parameters such as resource usage (bandwidth, cpu, memory), general relaying information etc.

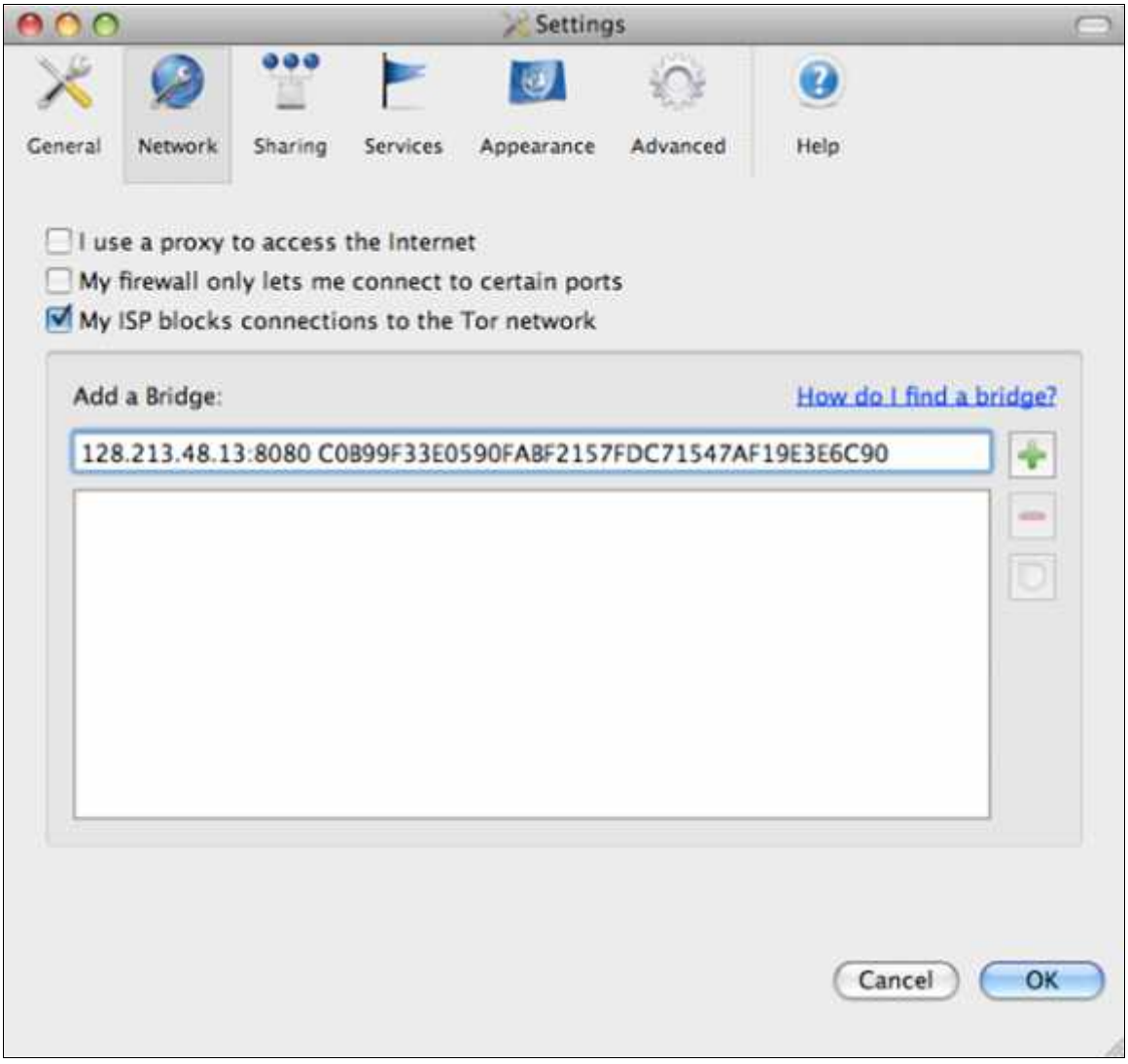


Arm snapshot

Tor can also provide anonymity to servers (web, ssh servers) in the form of location-hidden services, which are Tor clients or relays running specially configured server software. The server's IP address, so also its location, will not be revealed and the server is accessed through Tor-specific .onion pseudo top-level domain (TLD). The Tor network understands this pseudo-TLD and routes data anonymously both to and from the hidden service. Since this type of architecture does not rely on a public IP address, such a hidden service may be hosted behind personal firewalls or NAT gateways. One must keep in mind that in order to access a hidden service one needs to run a Tor

client. A detailed explanation of the hidden service protocol is available at this page: http://www.torproject.org/docs/hidden-services.html.en.

Since some ISPs and countries as a whole may choose to block access to the Tor network by filtering all connections to all known Tor relays (which are listed in the public Tor directory servers), the Tor project designers came up with the idea of Tor bridges, a special type of Tor relay not listed in the main Tor directory. Vidalia lets the user adapt to the situation when his ISP blocks access to the Tor network and configure bridge routing. A guide for finding and running Tor bridges is also offered on the project's website.



Vidalia bridge configuration snapshot

VI. WHAT TOR DOES AND WHAT IT DOESN'T? TOR'S LIMITATIONS.

Tor maintains the Internet users' anonymity. There are websites, such as WIMIA service (http://www.whatismyipaddress.com), which can be used in order to check the difference between the actual IP address and the Tor (or other anonymity technology) provided IP address. Since Tor nodes

are public, such services (e.g. https://check.torproject.org) can also identify the fact that someone uses Tor. So Tor does not provide secrecy in this regard. A possible solution would be to use a Tor bridge relay rather than connecting directly to the public Tor network. The future will probably provide another type of solution. If Tor's usage will grow and with it the number of relays, it will be more and more difficult to map and monitor the entire network of nodes.
Below are some other things that a Tor user should have in mind for a good understanding of the services provided by this onion routing solution:

Tor should be used along with Torbutton extension for anonymizing browsing activity. Don't forget the fact that the mere installation of the Tor client is not enough. The actual applications should be configured properly in order to be relayed through Tor. For Firefox, this is accomplished with the help of the Torbutton extension.
The Torbutton user should keep in mind that this extension blocks the execution of all sort of browser related technologies, such as Java, Flash, ActiveX, RealPlayer, Quicktime, Adobe's PDF plugin, and others. That's because any of these technologies can be used in order to reveal the actual IP address of the Tor user. All those security tweaks come with associated downsides, such as disabling Youtube content. The Torbutton can be reconfigured, but one needs to keep in mind that any tweak mean an opportunity for the potential attacker. One solution would be to use 2 browsers, one for anonymized browsing, and the other for normal browsing.
Cookies also represent means for identifying internet users. Since the Torbutton offers the option of enabling and disabling anonymous browsing at will, one should bear in mind that if the browser accepts a cookie during a non-anonymized working session, that cookie is stored on the user's home computer and can be used to identify him even during a later Tor-enabled browsing session. Tor project official website recommends CookieCuller, a Mozilla Firefox add-on permitting the user to keep the cookies he wants and automatically delete the rest of them.
Although Tor anonymizes the origin of the user's traffic, and it encrypts everything between you and the Tor network and everything inside the Tor network, it cannot encrypt the traffic between the Tor network (the exit node) and its final destination. There are means to encrypt end-to-end communication, such as https, and the Tor project recommends HTTPS Everywhere, a Firefox extension which encrypts communications with a number of major websites.
A compromised or misconfigured Tor exit node can deliver wrong pages or even embedded Java applets disguised as domains you trust, so the user must be careful when opening documents or applications downloaded through Tor. Their integrity should be at first verified through appropriate means.

VII. TOR VS. VPN/ONLINE ANONYMITY SERVICES

One can wonder why to use Tor instead of a free (or very cheap) VPN service or another anonymity solution. The main advantage when using a VPN is the speed of the connection. The actual design and size of the distributed Tor network limits drastically the connection speed for its users.

But, as an advantage, one may think that a VPN server or an anonymity service represents a single point of failure. Also, the providers of such services can find out the identity (IP address) and browsing habits of its clients, although most of them probably don't do such things. But the truth is that it's technically possible. In the case of the onion routing network, a compromised exit node cannot find out the virtual identity (IP address) of the senders, but only the IP address of the previous node. A certain level of trust is necessary in all cases. The user has to answer himself the question: who do I trust more – a public anonymity/VPN service provider or an anonymous Tor node owner?

One must also keep in mind that an overtaken exit node has access to the data sent in clear, which may contain access usernames and passwords, but this is the case also for a VPN service provider or an alternative online anonymity solution. The best way of avoiding such schemes would be to visit only secured websites, such as banking or e-commerce portals, and the aforementioned HTTPS Everywhere is a good point to start with.

When the common internet user browses through Tor or using a VPN connection, the traffic between him and the exit node, the VPN server respectively, is encrypted. But there are a number of online anonymity solutions, usually simple proxy services providers, who don't even use SSL to secure your connection to them.

VIII. TOR'S WEAKNESSES. POSSIBLE ATTACKS ON TOR

There are several security aspects involving the functionality and inner workings of the Tor network, some of which will be mentioned below.

First of all, as mentioned before, the last node through which traffic passes in the network (the so-called exit node) has to decrypt the communication before delivering it to its final destination. Someone operating that node can see the communication passing through this server. Because of this limitation in the Tor network design, in September 2007, Dan Egestard, a Swedish security consultant, was able to intercept a large number of usernames and passwords for e-mail accounts by operating and monitoring five Tor exit nodes. Each Tor user must bear in mind this limitation and access only websites with secure login procedures. Tor users must not mistake anonymity for end-to-end encryption.

Secondly, it is possible for an observer who can view both ends of the communication channel (let's say Alice and the destination website or Alice and the Tor exit node) to correlate the timings of

Alice's traffic as it enters and exits the Tor network. Tor does not defend against such a threat model called traffic confirmation, also known as end-to-end correlation.

Nonetheless, a study published in 2009 established that anonymity systems such as Tor and JonDonym are more resilient to size and timing of the encrypted data streams analysis (local traffic analysis attacks) than alternatives such as VPNs.

IX. TOR-BASED PROJECTS FOR THE FUTURE

Since October 2009, Android mobile phone users are able to anonymize their Internet browsing activity, instant messaging and e-mail traffic using a Tor-based application named Orbot, developed by a team from the Guardian Project lead by Nathan Freitas, which worked very closely with Tor project members. According to the "Tor on Android" project page, Orbot provides a local HTTP proxy and the standard SOCKS4A/SOCKS5 proxy interfaces into the Tor network. Orbot has the ability to transparently torify all of the TCP traffic on an Android device when it has the correct permissions. In an article published in March 2010 on the Tor official blog, caution was advised with this beta-stage release. "The Android web browser is not yet protected by Torbutton and a fully anonymous browser for the Android platform is yet to come", mentions the blog entry's author. Although there are some others Tor implementations on Android, the Tor official blog mentiones that this Android package, Orbot, the official Tor-on-Android release, which is using the C reference implementation of Tor, should be a lot safer than other similar packages.

Very recently, on the 22nd of December 2010, Technology Review, an MIT publication, published an article about a low-cost home router prototype with the Tor security built-in. According to Tom Simonite, the author of the above mentioned article, a number of volunteers are already testing a small number of Tor-adapted modified routers, using a popular low-cost wireless router available on the market from Buffalo Technology. The router's modified software was developed by a Tor team lead by Jacob Appelbaum, a Tor project developer, and is based on OpenWrt project which is a Linux distribution for embedded devices, an open-source distribution often used for networking equipments. The idea is that the Tor-enabled-routers will offer the possibility for the users to pass their entire Internet traffic, or only some of their applications through Tor, without the need for networking technical knowledge and without any additional software installation or configuration on the home computers. The nice thing about it is that anyone with minimal technical knowledge will be able to install the software on personal routers. Also, in addition to behaving like a Tor client, each router will be able to act as a Tor node, helping in the general effort of anonymizing the Internet traffic of other users and improving the overall Tor infrastructure and performances, speed and security-wise.

X. LEGAL ISSUES RELATED TO ONLINE ANONYMITY

There are a lot of possible legal issues regarding the Tor network, especially since there is no legal precedent related to its functionalities. Each Tor user should be aware of the risks of using the software, especially in the countries where access to some websites and services is restricted. Circumventing such restrictions through the use of Tor could be in some cases equivalent to breaking a law.

Internet users who downloaded and installed the Tor software can easily act as a relay for others by activating an option in the application menus, but they should be aware that this process could make them a link in the distribution chain of some child pornography and other illegal materials. Which brings us to the next question: could a Tor node owner be held responsible for the relayed contents? The logical answer seems to be not, but each country legal system has its own particularities.

For instance, according to a Cnet.com September 2007 article, in Germany a Tor node owner was arrested by the German police during a bomb threat investigation. The threat was posted online and the police traced one of the online messages posted on a forum to the IP address of that Tor node. The article mentions that the respective Tor node carried daily more than 40GB of random strangers' Internet traffic. After some hours of interrogation, the owner of the Tor relay was released and the police admitted making a mistake. In the summer of 2006 the German authorities conducted a simultaneous raid of seven different data centres, seizing 10 Tor servers in the process. Agents took the servers believing them to be related to a child pornography investigation.

Tor project webpage contains a legal guide for relay operators intended only for information purposes and not as legal advice.

CONCLUSIONS

Tor is a technique used to protect the privacy of the internet users, to prevent attackers from determining who is talking to whom on the internet by hiding the ends of the conversations through the use of a distributed overlay network. It hides information required for traffic analysis, but the information is concealed only until the so-called exit node, which forwards it to the destination in clear text. So every Tor user should bare in mind the necessity for additional encryption such as encrypted e-mails and https browsing. The need for a more complex network is also an important concern regarding the Tor network. Since the number of nodes in the network is constantly increasing, the overall speed and reliability of the second generation onion routing network will become better and better.

REFERENCES

- Tor project's official webpage – anonymity online: http://www.torproject.org/
- Roger Dingledine, Nick Mathewson, Paul Syverson – Tor, the second generation onion router: https://svn.torproject.org/svn/projects/design-paper/tor-design.pdf
- Wikipedia Tor webpage - http://en.wikipedia.org/wiki/Tor_%28anonymity_network%29
- Staying anonymous online - http://whatismyipaddress.com/internet-anonymity/
- Steven J. Murdoch - Tor: Anonymous internet communication: http://www.cl.cam.ac.uk/research/security/posters/sjm217-tor.pdf
- Martin Suess - Breaking Tor anonymity: http://www.csnc.ch/misc/files/publications/the_onion_router_v1.1.pdf
- Jonas Petterson, Yuxiang Zhu, Jiying Tian – Tor: http://www.it.uu.se/edu/course/homepage/sakdat/ht05/assignments/pm/programme/tor.pdf
- David Talbot – Dissent made safer, Technology Review, May/June 2009 issue: http://www.technologyreview.com/printer_friendly_article.aspx?id=22427
- Tom Simonite – Home internet with anonymity built-in, Technology Review, 22nd of Dec. 2010: http://www.technologyreview.com/web/26981/
- CERIAS Security: Tor: Anonymous communications for government agencies, corporations and you (parts 1 through 6): http://www.youtube.com/watch?v=dv2dnlVNrss
- Anna Lysyanskaya – Cryptography: How to you're your secrets safe, Scientific American, 20th of August 2008: http://www.scientificamerican.com/article.cfm?id=cryptography-how-to-keep-your-secrets-safe
- Various online references:
  - JAP anonymity and privacy – The world wide web and your privacy: http://anon.inf.tu-dresden.de/help/jap_help/en/help/wwwprivacy.html
  - http://arstechnica.com/software/news/2006/09/7709.ars
  - http://news.cnet.com/8301-13739_3-9779225-46.html
  - http://www.drdobbs.com/security/197002414;jsessionid=ISQCEE4PCMYAJQE1GHRSKHWATMY32JVN?pgno=1
  - http://www.wired.com/politics/security/news/2007/09/embassy_hacks?currentPage=1
  - http://www.iusmentis.com/society/privacy/remailers/onionrouting/

THE PENETRATION TEST MANAGEMENT

Maj. Eng. Ion GRUJDIN

1. INTRODUCTION

The CIS security assessment issue is no one trivial. In this area nothing can be white or black and, furthermore, the grey is sometimes so soft insomuch that be easily confused with the false or true. False positive and false negative terminology derives from here. But what can be the reason? It is quite simple. The communication and information systems are actually expressions of human being creativity, willingness and performance. The threats (excepting the natural and contextual ones) are also imagine of this side of human being. Can be human mind quantified and put it in a paper? Can we actually thinking in the same manner like our "enemy"? Why have I to consider the hacker my enemy? Actually is my friend. We have the some problems, we have to study a lot of arid things, and our nights are the single "white" and black things.

Jack Koziol posted on http://www.infosecinstitute.com/ : "Much of the confusion surrounding penetration testing stems from the fact it is a relatively recent and rapidly evolving field. Additionally, many organisations will have their own internal terminology (one man's penetration test is another's vulnerability audit or technical risk assessment)."

But not the terminology seems to be "lapis philosoforum". The same publisher said: "Well the goal behind penetration testing is to try to find as much serious vulnerability as possible. In order to do this, you must develop the "mindset" of your attacker. You should look at your assessed system or application in all of the possible ways you think it could be misused, abused and exploited. You should then take a break, drink some well-deserved coffee, and then think of entirely new "misuse cases" for the system under review. Using a cut and dry methodology runs counter to the basic and essential premise of penetration testing; that a penetration test is an exercise in system abuse and cannot be readily scripted."

To conclude, it seems that following a methodology during the pen tests is usefulness. Indeed, there are a lot of methodologies and no criteria to make decision which of them to be applied at a peculiar case. Furthermore, the reports are brushy (hundreds of pages) and extremely confuse. Many false positives and false negatives in the report, no real vulnerabilities are ever acted on.

However, a methodology represents the rationale and the philosophical assumptions that underlie a particular study relative to the scientific method. Following a methodology it means to introduce an order in your test. Based on experience, you can skip, complete or redesign some steps in order to adapt at your concrete situation. In order to minimize this effort, I consider that firstly is necessary to try harmonizing the provisions of existent methodologies in order to retain the appropriate elements in accordance with our peculiar purpose. In other words, it is necessary to synchronize and

stabilize the methodologies between them in accordance with a peculiar purpose. In this moment two of these methodologies seems to be the most widely accepted and used: Open Source Testing Methodology Manual (OSSTMM) issued by the ISECOM (Institute for Security and Open Methodologies) and NIST 800-53A completed with NIST Special Publication 800-115 "Technical Guide to Information Security Testing and Assessment" released by National Institute of Standards and Technology. Obviously, the both are valuable and are the same purpose, the both are pragmatic but, strange, and the approaches are different.

This paper tries to find out a superposition of these references' provisions in the peculiar area of penetration testing.

1.1. Penetration testing. Definition, scope, purpose, risks, methodologies

There are a lot of definitions of penetration testing, depending of source and the evolution of concept in time:

"Penetration testing (also called pen testing) is the practice of testing a computer system, network or Web application to find vulnerabilities that an attacker could exploit." (Jonathan Gershater, Puneet Mehta, 2003)

"Penetration testing is security testing in which assessors mimic real-world attacks to identify methods for circumventing the security features of an application, system, or network."

(NIST Special Publication 800-115 "Technical Guide to Information Security Testing and Assessment")

"A test of a network's vulnerabilities by having an authorized individual actually attempt to break into the network. The tester may undertake several methods, workarounds and "hacks" to gain entry, often initially getting through to one seemingly harmless section, and from there, attacking more sensitive areas of the network." (Computer Desktop Encyclopedia)

"The legal intrusion into a computer system by hackers in order to test the security mechanisms in the system".(High Beam Research, Inc.)

"Double Blind - The Analyst engages the target with no prior knowledge of its defenses, assets, or channels. The target is not notified in advance of the scope of the audit, the channels tested, or the test vectors. A double blind audit tests the skills of the Analyst and the preparedness of the target to unknown variables of agitation. The breadth and depth of any blind audit can only be as vast as the Analyst's applicable knowledge and efficiency allows. This is also known as a Black Box test or Penetration test." (Open Source Testing Methodology Manual, version 3)

But the most intuitive and complete, even if, at the first sight, has no connection with the scope addressed by the previous definitions, seems to be the engineering definition of penetration test **found in McGraw-Hill Science & Technology Dictionary - "**A test to determine the relative

values of density of no cohesive sand or silt at the bottom of boreholes." **I let you to think about this definition while you read this paper…**

Herein after, the expression "penetration test" will refer only Information Technology security area and, more specific, to Data Networks Channel as described in OSSTMM 3.

The purpose is described very well in NIST Special Publication 800-115 "Technical Guide to Information Security Testing and Assessment":

"Most penetration tests involve looking for combinations of vulnerabilities on one or more systems that can be used to gain more access than could be achieved through a single vulnerability. Penetration testing can also be useful for determining:

- How well the system tolerates real world-style attack patterns
- The likely level of sophistication an attacker needs to successfully compromise the system
- Additional countermeasures that could mitigate threats against the system
- Defenders' ability to detect attacks and respond appropriately."

Some words about the methodologies advantages and disadvantages. Penetration testing can be invaluable, but it is labor-intensive and requires great expertise to minimize the risk to targeted systems. Systems may be damaged or otherwise rendered inoperable during the course of penetration testing, even though the organization benefits in knowing how a system could be rendered inoperable by an intruder. Although experienced penetration testers can mitigate this risk, it can never be fully eliminated. Penetration testing should be performed only after careful consideration, notification, and planning. This is one of the methodology purposes.

They can be the powerful base of your assessment or might be the main reason to fail in a lamentable way …. The methodologies are in intuitive descriptions, algorithms which we'll follow in order to obtain the result of a specific problem, based on some input data. But in this case, the input data are "provided" by the human brain and at this level the single rule seems to be "no rules". But a good methodology will let more time to human tester to think.

1.2 How and why have to synchronize the methodologies?

In this time two methodologies that refer pen test seems to be the most widely accepted and used: Open Source Testing Methodology Manual (OSSTMM) issued by the ISECOM (Institute for Security and Open methodologies) and NIST 800-53A "Guide for Assessing the Security Controls in Federal Information System" completed with NIST Special Publication 800-115 "Technical Guide to Information Security Testing and Assessment" released by National Institute of Standards and Technology. Herein after, the term OSSTMM will refers the OSSTM 3 and NIST the NIST SP 800-115.

We are beginning from a simple assumption. Two methodologies that follow the same purpose should have common parts and different parts as is shown in Figure 1.

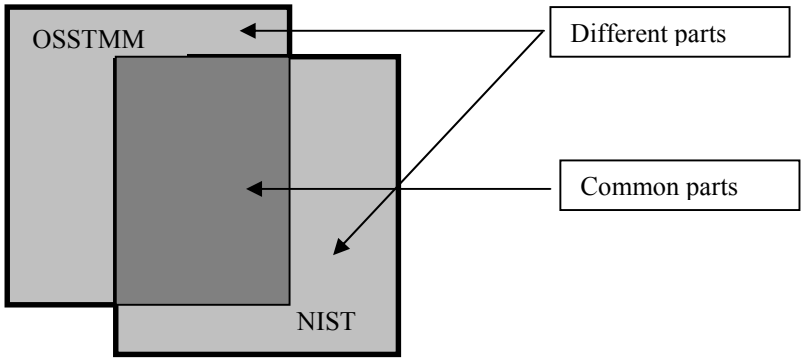


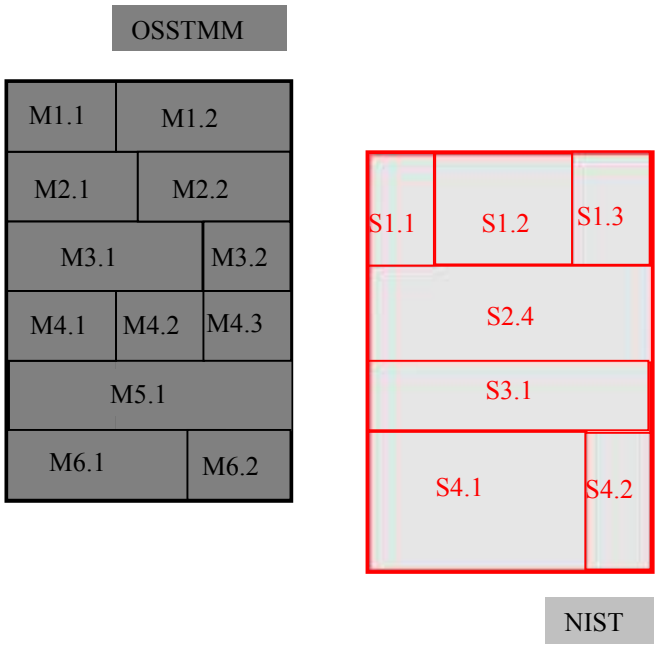
Figure 1. The ideal common and different parts of methodologies (pen test section)

The next question is "Which are the similar high-level description terms?" The answer is quite simple:

- NIST **phases (P)** – OSSTMM **phases (F)**;
- NIST **stages (S)** – OSSTMM **modules (M)**
- NIST **activities (A)** – OSSTMM **tasks (A)**

In order to obtain a mapping image as above is necessary to find the appropriate granularity of these terms. Actually, it means "How deep the methodologies terms can be parsing in order to have an ideal one-to-one items' correspondence? And, which of them have to be parsed? "

a)



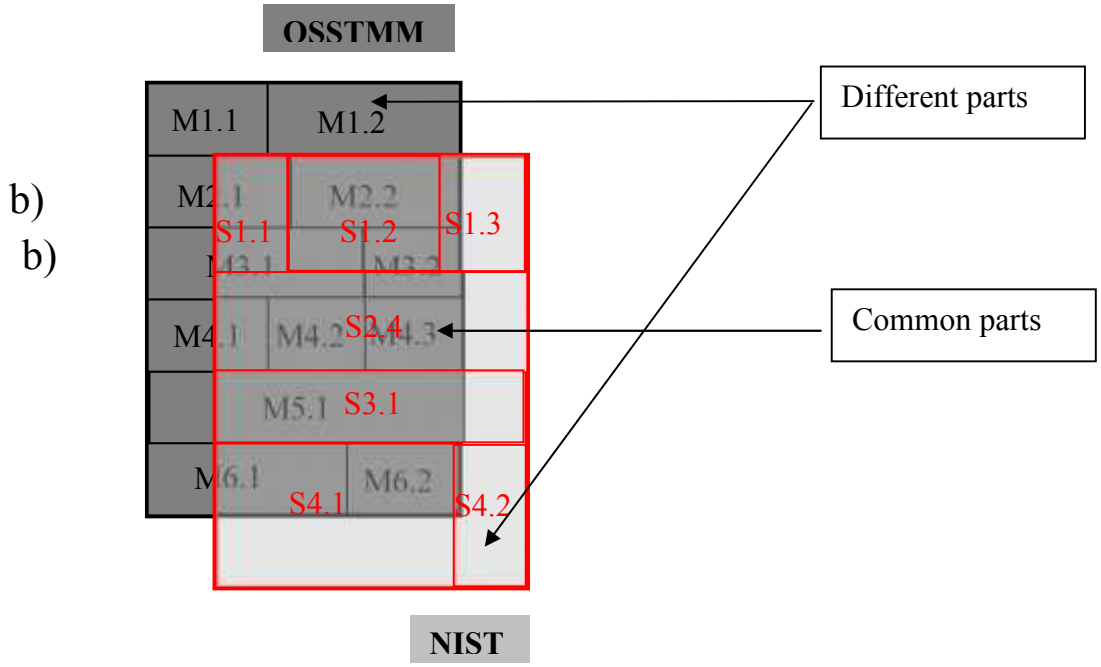


Figure 2. (a) OSSTMM modules and NIST stages

(b) Superposition of OSSTMM modules and NIST stages

(no real approach is presented here, only for demonstrative purpose)

Normally, the first image that we obtain trying to sync the methodologies looks like in the bellow figure. Quite confuse, isn't it? It is very easy to observe that the correspondence is not at all perfect even I used a simplified imagine of the objects. There are stages and modules that have one-to-one correspondence but majority represents one-to- multiple or multiple-to-one cases. The parse has to be made in order to obtain the most possible one-to-one elements. So, the parse cannot be symmetric. I mean the both methodologies' steps have to be "broken" in order to meet the maximum superposition requirement. Furthermore, other modules have to be added in order to meet the specific pen test requirements.

Why to synchronize the methodologies? There are several reasons:

- the common parts reinforce each other, becoming the "spine" of applicable methodology; this will not be a new one but will be more powerful than each separate methodology
- if this technique is applied aiming an concrete goal, than will be selected the optimal number of elements necessary to accomplish the task; we can say that the "new" methodology is stabilized, being eliminated the unnecessary elements that can alter the results

2. PENETRATION TESTING. MAPPING NIST AND OSSTMM PHASES

2.1. Introduction

In the Figure 1 and 2 are presented the four phases of penetration testing in accordance with **NIST Special Publication 800-115 "Technical Guide to Information Security Testing and**

Assessment", with details regarding the attack phase and Figure 3 shows the methodology flow in accordance with **OSSTMM 3**.

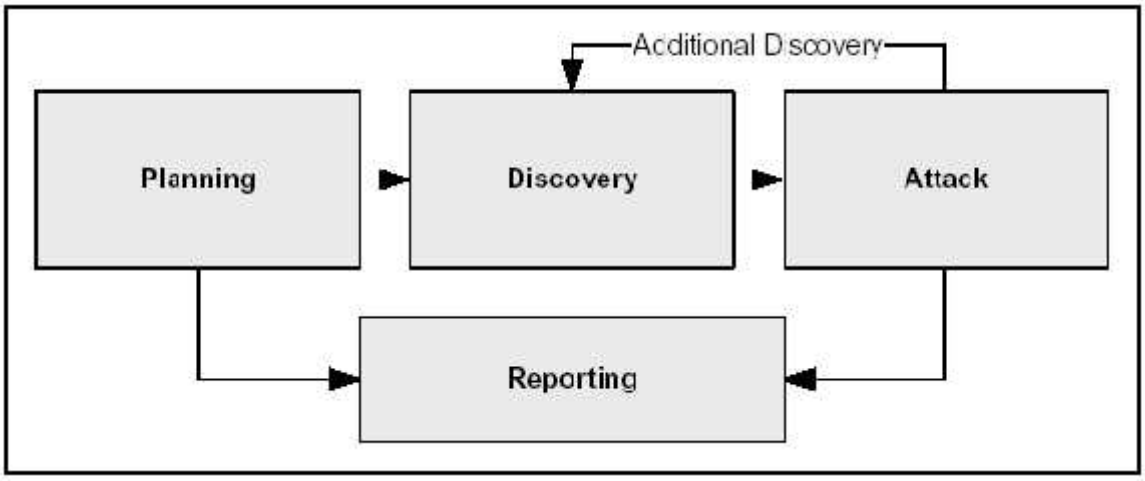


Figure 3. NIST Penetration Testing Phases

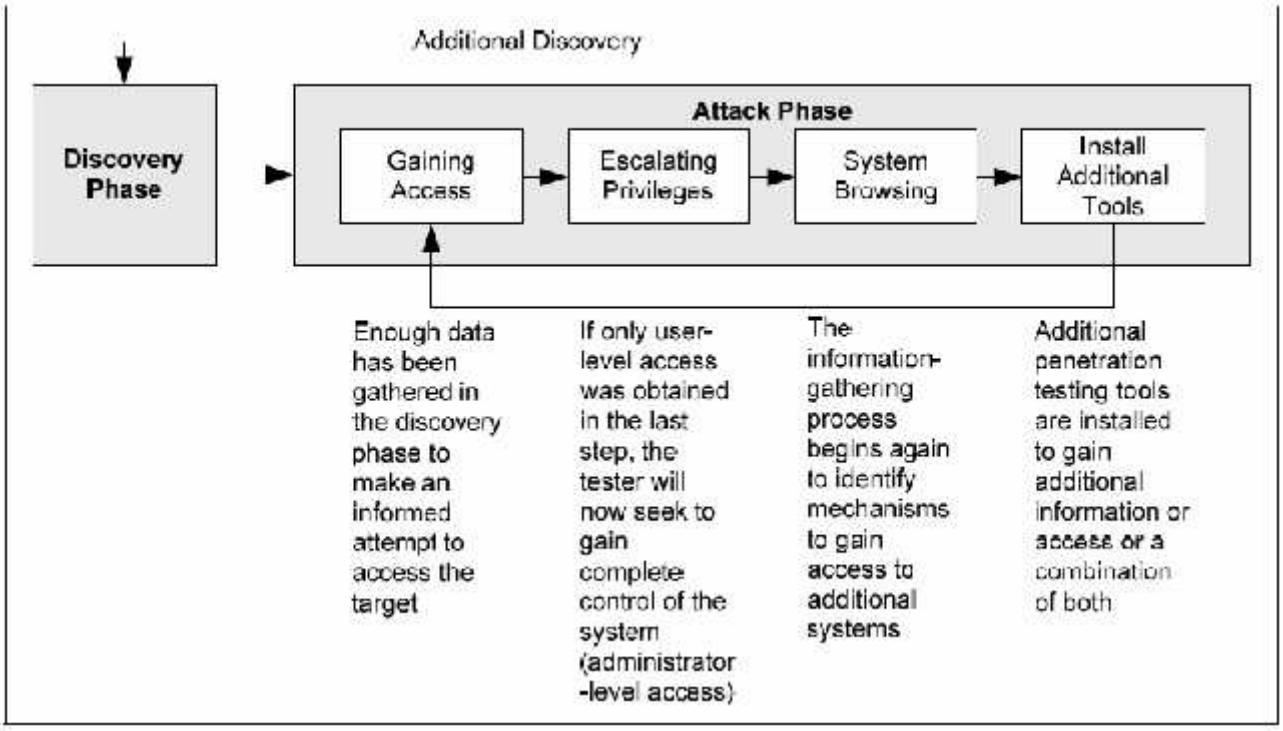


Figure 4. NIST Penetration Testing – Attack Phase detail

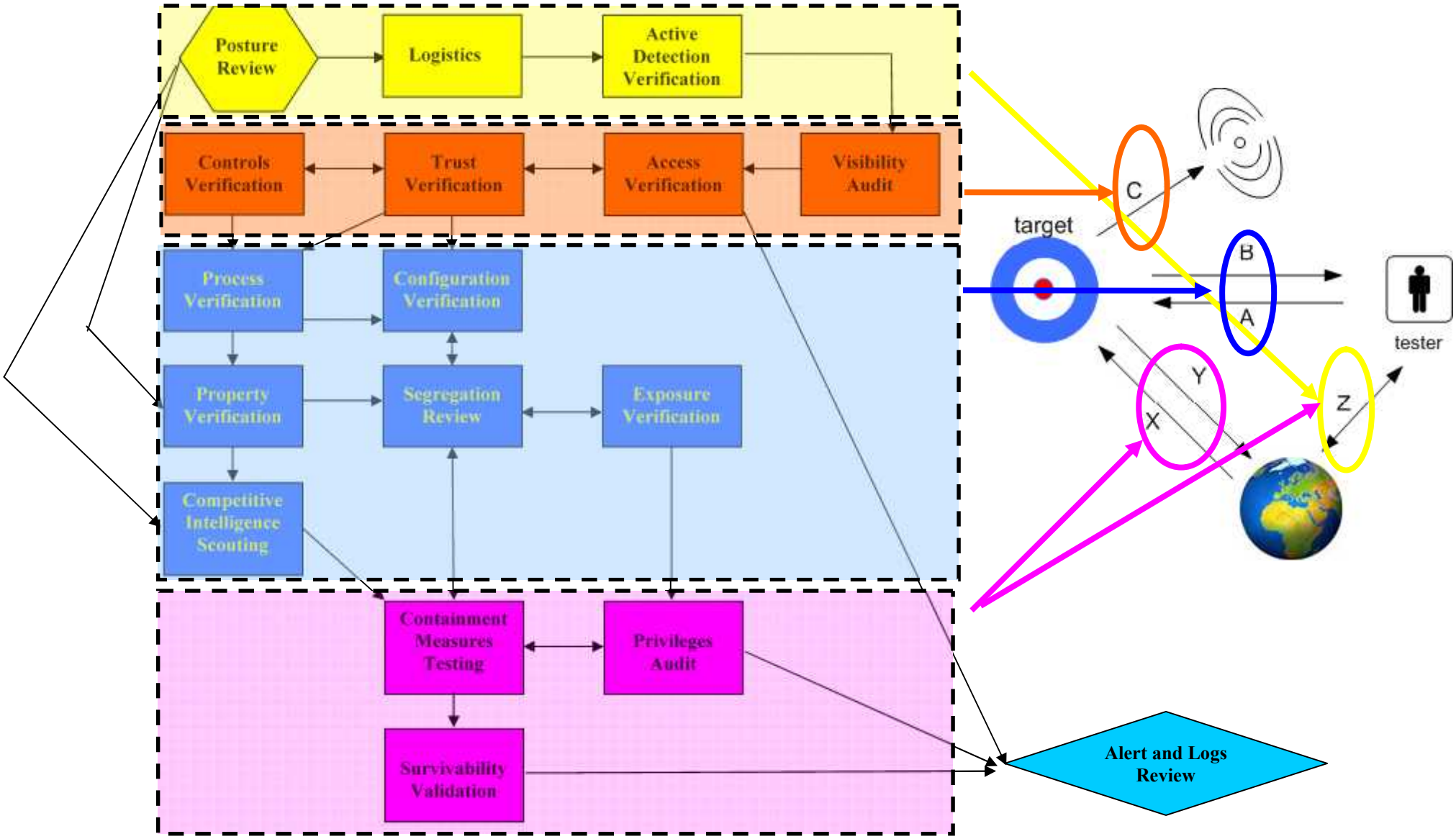


Figure 5. OSSTMM Modules mapped to Four Point Access method

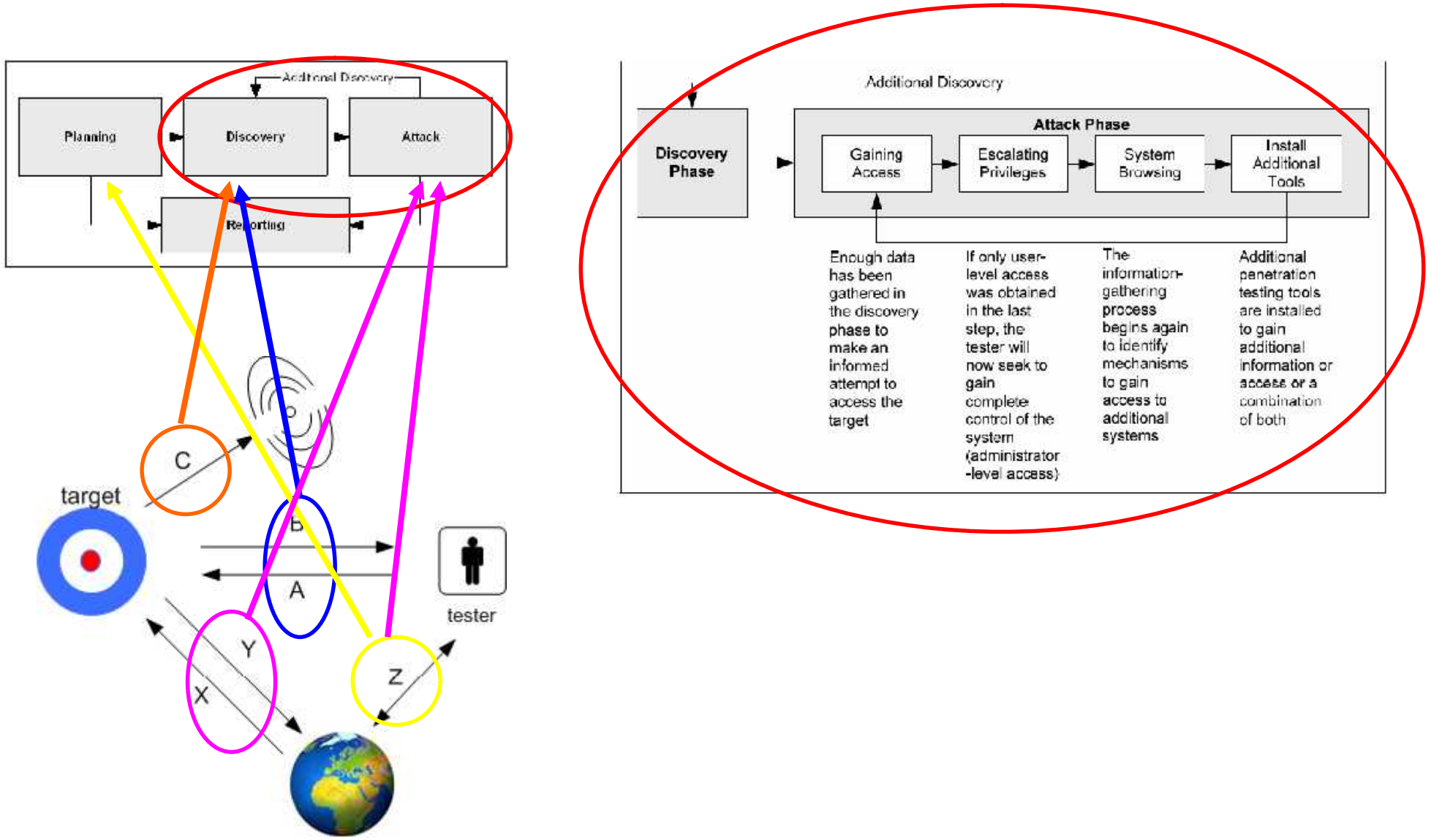


Figure 6. High-level mapping between NIST and OSSTMM

Even if the flow methodology seems to be different, a mapping between these two methodologies can be established in order to obtain a better penetration test methodology tailored for a peculiar purpose. I use the NIST methodology as baseline due to the fact that is easy, "naturally", to follow, and provides natural breaking points for staff transitions.

2.2. Planning phase

2.2.1 Prerequisites

Usually, in this phase the following actions should be taken:

- Sets the objectives of the penetration test and attackers profiles for the tests;
- Determine the test types;
- Make a decision what are the success criteria with which organization can measure results against predetermined criteria, for both external and internal attacks;
- Plane whether exploits will be performed and to what extent;
- Coordinate the plan with the appropriate IT team in order not to cause any damages to the network;
- Obtain management approval for the pen test;
- Define a time scale.

2.2.2. OSSTMM point of view

The OSSTMM phase is Induction – "Establishing principle truths about the target from environmental laws and facts. The Analyst determines factual principles regarding the target from the environment where the target resides. As the target will be influenced by its environment, its behavior will be determinable within this influence. Where the target is not influenced by its environment, there exists an anomaly to be understood.

The first module for this phase is **Posture Review**. Initial studies of the posture include the laws, ethics, policies, industry regulations, and political culture which influence the security and privacy requirements for the scope. This review forms a matrix against which the testing should be mapped but not constrained due to the ubiquity of the channel endpoints. Therefore, it is important to consider, as some legislation requires, the target market or end users of this channel which must also be added to the scope for this module. The Posture Review comprises:

- **Policy**. Review and document appropriate organizational policy regarding security, integrity, and privacy requirements of the scope. Review and document contracts and Service Level Agreements (SLAs) with service providers and other involved third parties.
- **Legislation and Regulations**. Review and document appropriate regional and national legislation, and industry regulations regarding the security and privacy requirements of the

organization in the scope as well as that which includes the appropriate customers, partners, organizational branches, or resellers outside the scope.

- **Culture**. Review and document appropriate organizational culture in the scope towards security and privacy awareness, required and available personnel training, organizational hierarchy, help desk use, and requirements for reporting security issues.
- **Age**. Review and document the age of systems, software, and service applications required for operations.
- **Fragile Artifacts**. Review and document any systems, software, and service applications which require special care due to high use, instabilities, or a high rate of change.

Other modules of OSSTMM that can be assimilated with the planning phase are **Logistics** and **Active Detection Verification.**

Logistics represents the preparation of the channel test environment needed to prevent false positives and false negatives which lead to inaccurate test results. It comprises:

- ***Framework***

(1) Verify the scope and the owner of the targets outlined for the audit.
(2) Determine the property location and the owner of the property housing the targets.
(3) Verify the owner of the targets from network registration information.
(4) Verify the owner of the target domains from domain registration information.
(5) Verify the ISP(s) providing network access or redundancy.
(6) Search for other IP blocks and targets related to the same owner(s).
(7) Search for similar domain names or mistyped domain names which can be confused with the target.
(8) Verify which target domain names resolve to systems outside of the owner's control such as caching devices.
(9) Verify which target IP addresses trace back to locations different from the owner's location.
(10) Verify that reverse name look-ups of target system addresses correspond with the scope and the scope owner.
(11) Find and verify the paths of network services which interact outside of target for the paths they follow into and out of the scope.
(12) Prepare local name resolution to map domain names only to the specific systems to be tested and not any devices outside the target or target ownership.
(13) Use reverse name look-ups as an additional information source towards determining the existence of all the machines in a network.

- Network Quality

(1) Measure the rate of speed and packet loss to the scope for a requested service in TCP, UDP, and ICMP both as a whole service request and as a request/response pair. Repeat each request in succession at least 100 times and record the average for both whole service requests and packet responses for each of the three protocols.

(2) Determine sending and receiving packet rates for a total of 6 averages (per protocol) as requests per second per network segment in the scope.

(3) Record packet loss percentages for the determined packet sending and receiving rates. The planning phase sets the groundwork for a successful penetration test. No actual testing occurs in this phase. The scenarios have to be put in accordance with the real world threats. Nobody is trust. Some persons are supposed to be trusted but many times the vulnerabilities reside around them.

- Time

(1) Verify time zone, holidays, and work schedules for the various systems within the scope including partners, resellers, and influential customers interacting with the scope.

(2) Identify the Time To Live (TTL) distance to the gateway and the targets.

(3) Assure the Analyst's clock is in sync with the time of the targets.

Active Detection Verification is the determination of active and passive controls to detect intrusion to filter or deny test attempts must be made prior to testing to mitigate the risk of corrupting the test result data as well as changing the alarm status of monitoring personnel or agents. It may be necessary to coordinate these tests with the appropriate persons within the scope.

- Filtering

(1) Test whether INCOMING network data or communications over web, instant messaging, chat, web-based forums, or e-mail, are monitored or filtered by an authoritative party for relay of improper materials, code injections, malicious content, and improper conduct and record responses and response time.

(2) Test whether OUTGOING network data or communications over web, instant messaging, chat, web-based forums, or e-mail, are monitored or filtered by an authoritative party for relay of improper materials, code injections, malicious content, and improper conduct and record responses and response time.

- Active Detection

(1) Verify active responses to probes from systems and services. This could be human or machine readable notifications, packet responses, silent alarm trips, or the like.

(2) Map any applications, systems, or network segments within the scope which produce logs, alarms, or notifications. This could include Network or Host based Intrusion Detection or Prevention Systems, system logs, Security Information Management tools (SIMs), application logs, and the like.

2.2.3. NIST point of view

In the planning phase, rules are identified, management approval is finalized and documented, and testing goals are set. The planning phase sets the groundwork for a successful penetration test. No actual testing occurs in this phase.

It is very interesting that some activities described in other sections can be included in this phase, without to have an apparent connection with pen test:

- Documentation Review (**NIST Special Publication 800-115, paragraph 3.1**.)
- Log review (**NIST Special Publication 800-115, paragraph 3.2**.)
- Rule set View (**NIST Special Publication 800-115, paragraph 3.3**.)
- System Configuration Review (**NIST Special Publication 800-115, paragraph 3.4**.)

The findings from these activities have to be logged and used at final report but the problems have to be fixed in order to obtain a more accurate imagine about other vulnerabilities that otherwise can be hidden by false positives and false negatives provided by these discrepancies. The pen test purpose is not to "discover" the breaches that are the result of inappropriate application of rules, policies, documentation, etc. but those breaches that are outside of this scope.

2.2.4. Conclusions

The entire pen test planning operations seems to be identical for the both methodologies with the mention that is more detailed in OSSTMM 3.

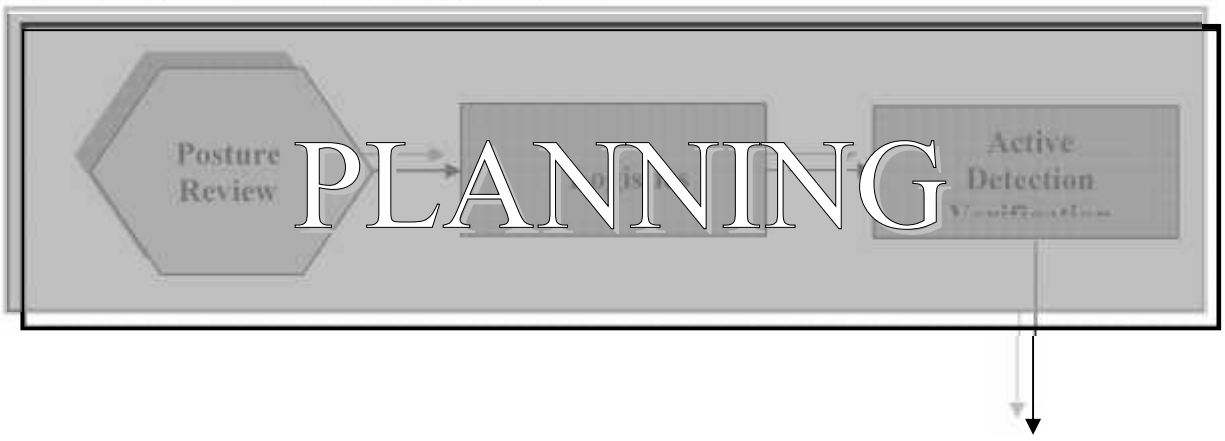


Figure 7. The superposition of NIST and OSSTMM planning phase

As I already mentioned in 2.2.3, a feed-back using other elements that are not specially designated for pen test can be useful in order to put the system in security state that **has to be**, the major gain being reduction of false positives and false negatives that can hide other vulnerabilities that hadn't take into account so far. The reason is that I don't want to discover once again the vulnerabilities that have been already identified but to create an appropriate environment to discover others.

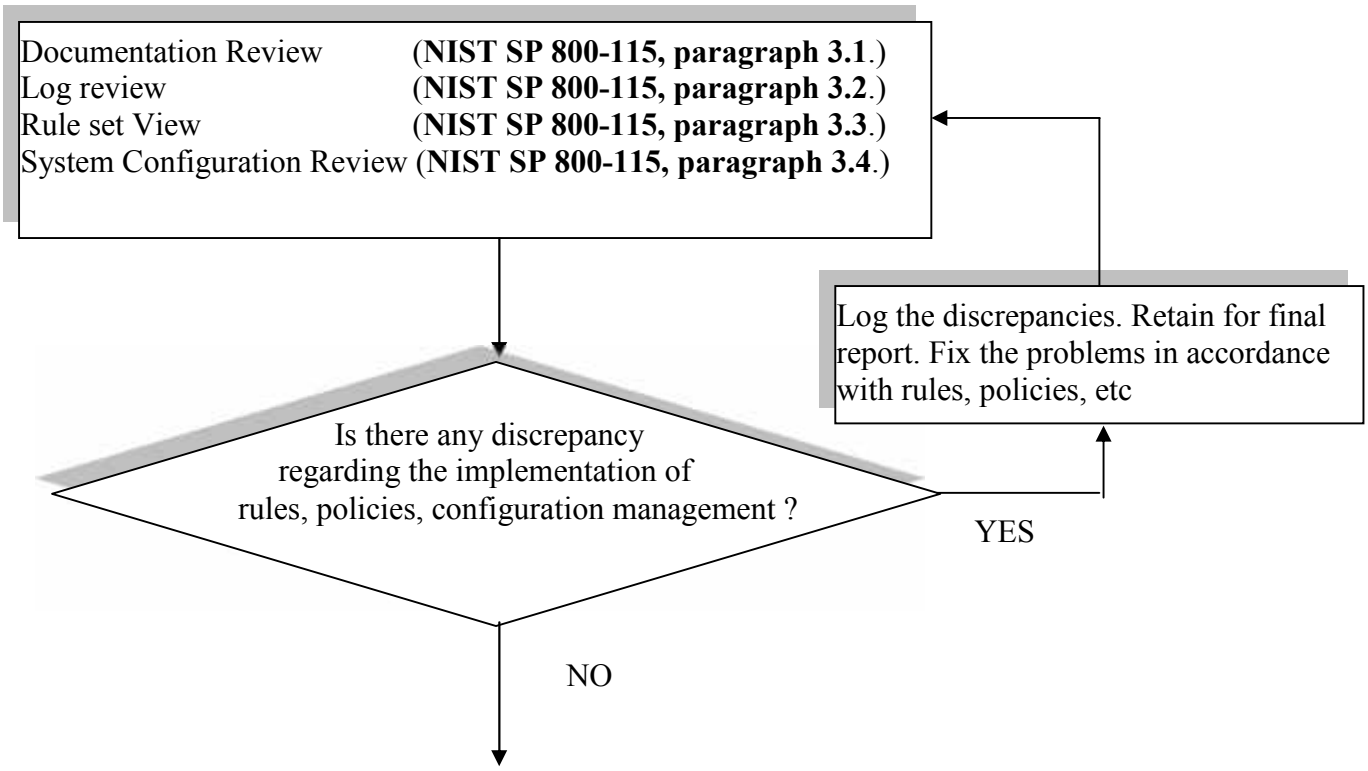


Figure 8. A suggestion to reduce the false positives and false negatives in planning phase

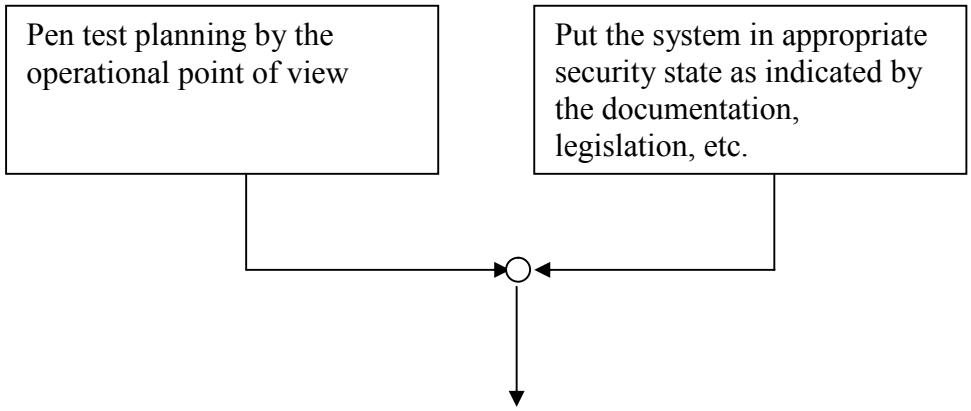


Figure 9. A suggestion regarding the planning phase management

2.3. Discovery phase

2.3.1 Prerequisites.

Discovery phase is strongly dependent of the type of test. **Open Source Testing Methodology Manual (OSSTMM) 3** refers to six common types, dependent on the amount of information the tester knows about the targets and what the target knows about the tester or expects from test. In its assumption the **penetration test** (**Black Box test**) is considered to be only the "**double blind**" type, where the Analyst engages the target with no prior knowledge of its defenses, assets, or channels. The target is not notified in advance of the scope of the audit, the channels tested, or the test vectors. A double blind audit tests the skills of the Analyst and the preparedness of the target to unknown variables of agitation. The breadth and depth of any blind audit can only be as vast as the Analyst's applicable knowledge and efficiency allows. If we take into account the fact that most successful attacks are coming from connections that are inside your perimeter security, being performed by the persons that know very well the asset, then the "**reversal**" type represents in this case an appropriate approach for penetration testing. The Analyst engages the target with full knowledge of its processes and operational security, but the target knows nothing of what, how, or when the Analyst will be testing. The true nature of this test is to audit the preparedness of the target to unknown variables and vectors of agitation. The breadth and depth depends upon the quality of the information provided to the Analyst and the Analyst's applicable knowledge and creativity. This is also often called a **Red Team exercise**.

Also, an insider or an operator, with more or less knowledge about the asset can try to exploit the system using different network access points, and include each logical and physical segment of its. As following, in my opinion, the "grey" area also represents the penetration test types simulating real situations.

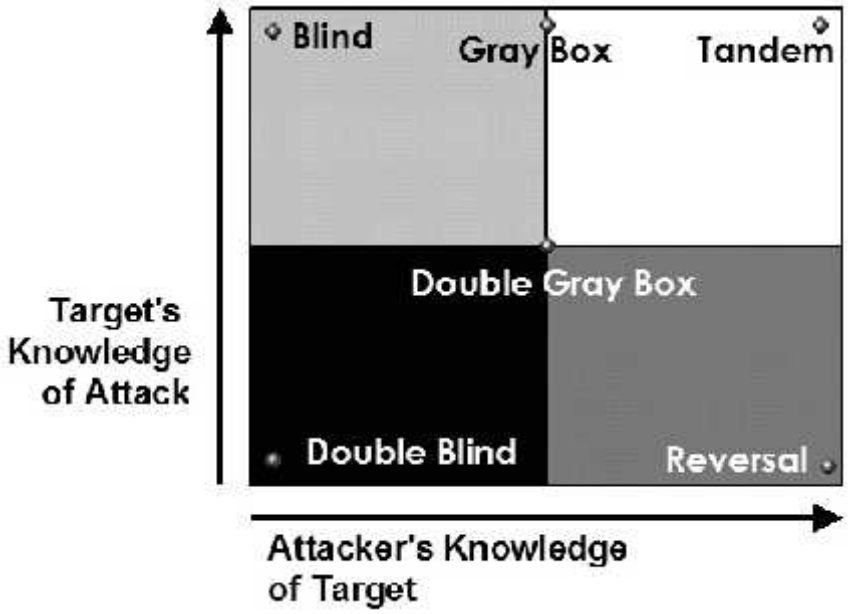


Figure 10. The OSSTMM common types of tests

2.3.1 The discovery phases (NIST 800-115 and OSSTMM 3 mapping)

In accordance with **NIST Special Publication 800-115 "Technical Guide to Information Security Testing and Assessment"**, the discovery usually comprises two parts:

2.3.2.1. Information gathering and scanning: hostnames and IPs, the used ports, employee names and contact information, system information (names and shares), application and service information.

Host name and IP address information can be gathered in many ways: DNS interrogation, Inter NIC (WHOIS) queries, network sniffing

Employee names and contact information can be obtained by searching the organization's Web servers or directory servers.

System information, such as names and shares can be found through methods such as NetBIOS enumeration (generally only during internal tests) and Network Information System (NIS) (generally only during internal tests)

*Application and service information***,** such as version numbers, can be recorded through banner grabbing.

I think that an important point omitted in this phase is the *network mapping.* This has to be established during all the penetration test types but for the "double-blind" type is mandatory, being the main means to identify network access points. The information can be gathered initially from Internet and can be completed following social engineering specific tasks. No penetration test would be complete without addressing this non technical approach to exploitation. Social engineering preys on human interaction to obtain or compromise information about an organization and its computer systems. In a social engineering scheme, the attacker relies on human nature to gain access to unauthorized network resources. This could be in the form of eavesdropping or "shoulder surfing" (i.e., direct observation practices) to obtain access. It can also include data aggregation through "dumpster diving" (e.g., looking for passwords written on sticky notes) or talking to multiple sources and building on data from each source until the attacker has enough information to commence an attack.

The correspondent OSSTM phase is **Interaction –** "like echo tests, standard and non-standard interactions with the target to trigger responses. The Analyst will inquire or agitate the target to trigger responses for analysis."

The correspondent modules are: **Visibility Audit**, **Access Verification**, **Trust Verification**, and **Control Verification**.

Visibility Audit represents the enumeration and indexing of the targets in the scope through direct and indirect interaction with/or between live systems. This comprises the following tasks:

- Network Surveying

(1) Identify the perimeter of the target network segment(s) and the vector from which they will be tested.

(2) Use network sniffing to identify emanating protocols from network service responses or requests where applicable. For example: Netbios, ARP, SAP, NFS, BGP, OSPF, MPLS, RIPv2, etc.

(3) Query all name servers and the name servers of the ISP or hosting provider, if available, for corresponding A, AAAA, and PTR records as well as ability to perform zone transfers to determine the existence of all targets in the network and any related redundancies, load balancing, caching, proxying, and virtual hosting.

(4) Verify broadcast requests and responses from all targets.

(5) Verify and examine the use of traffic and routing protocols for all targets.

(6) Verify ICMP responses for ICMP types 0-255 and ICMP codes 0-2 from all targets.

(7) Verify default and likely SNMP community names in use are according to practical deployments of all SNMP versions.

(8) Verify responses from targets to select ports with TTL expiration set to less than 1 and 2 hops from the targets. For example:

TCP 8, 22, 23, 25, 80, 443, 445, 1433

UDP 0, 53, 139, 161

ICMP T00:C00, T13:C00, T15:C00, T17:C00

(9) Trace the route of ICMP packets to all targets.

(10) Trace the route of TCP packets to all targets for ports SSH, SMTP, HTTP, and HTTPS ports.

(11) Trace the route of UDP packets to all targets for DNS and SNMP ports.

(12) Identify TCP ISN sequence number predictability for all targets.

(13) Verify IPID increments from responses for all targets.

(14) Verify the use of Loose Source Routing to the target gateway and outer perimeter systems to route packets to all targets.

- Enumeration

(1) Search newsgroups, forums, IRC, IM, P2P, VoIP, and web-based communications for connecting information of the target to determine outgoing gateway systems and internal addressing.

(2) Examine e-mail headers, bounced mails, read receipts, mail failures, and malware rejections to determine outgoing gateway systems and internal addressing.
(3) Examine target web-based application source code and scripts to determine the existence of additional targets in the network.
(4) Examine service and application emanations. Manipulate and replay captured traffic to invoke new requests or responses, gain depth, or expose additional information. For example, SQL, Citrix, HTTP, SAP, DNS, ARP, etc.
(5) Search web logs and intrusion logs for system trails from the target network.
(6) Verify all responses from UDP packet requests to ports 0-65535.
(7) Verify responses to UDP packet requests FROM SOURCE ports 0, 53, 139, and 161 to 0, 53, 69, 131, and 161.
(8) Verify responses to UDP packet requests with BAD CHECKSUMS to all discovered ports and for 0, 53, 69, 131, and 161.
(9) Verify service request responses to common and contemporary UDP remote access malware ports.
(10) Verify responses from TCP SYN packet requests to ports 0-65535.
(11) Verify responses from TCP service requests to ports 0, 21, 22, 23, 25, 53, 80, and 443.
(12) Verify responses from a TCP ACK with a SOURCE port of 80 to ports 3100-3150, 10001-10050, 33500-33550, and 50 random ports above 35000.
(13) Verify responses from TCP SYN fragments to ports 0, 21, 22, 23, 25, 53, 80, and 443.
(14) Verify responses from all combinations of TCP flags to ports 0, 21, 22, 23, 25, 53, 80, and 443.
(15) Verify the use of all targets with HTTP or HTTPS based VPNs, proxies, and URL redirectors to redirect requests for targets within the scope.
(16) Verify the use of all targets with sequential IPIDs to enumerate systems within the network.
(17) (q) Map and verify for consistency visible systems and responding ports by TTLs.

- Identification

Identify targets' TTL response, system uptime, services, applications, application faults, and correlate this with the responses from system and service fingerprinting tools.

Access Verification refers to tests for the enumeration of access points leading within the scope.

- Network

(1) Request known, common services which utilize UDP for connections from all addresses.
(2) Request known, common VPN services including those which utilize IPSEC and IKE for connections from all addresses.
(3) Manipulate network service and routing to access past restrictions within the scope.
(4) Request known, common Trojan services which utilize UDP for connections from all addresses.
(5) Request known, common Trojan services which utilize ICMP for connections from all addresses.
(6) Request known, common Trojan services which utilize TCP for connections from all addresses and unfiltered ports which have sent no response to a TCP SYN.

- Services

(1) Request all service banners (flags) for discovered TCP ports.
(2) Verify service banners (flags) through interactions with the service comprising of both valid and invalid requests.
(3) Match each open port to a daemon (service), application (specific code or product which uses the service), and protocol (the means for interacting with that service or application).
(4) Verify system uptime compared to the latest vulnerabilities and patch releases.
(5) Verify the application to the system and the version.
(6) Identify the components of the listening service.
(7) Verify service uptime compared to the latest vulnerabilities and patch releases.
(8) Verify service and application against TTL and OS fingerprint results for all addresses.
(9) Verify HTTP and HTTPS for virtual hosting.
(10) Verify VoIP services.
(11) Manipulate application and service requests outside of standard boundaries to include special characters or special terminology of that service or application to gain access.

- Authentication

(1) Enumerate accesses requiring authentication and document all privileges discovered which can be used to provide access.
(2) Verify the method of obtaining the proper Authorization for the authentication.

(3) Verify the method of being properly Identified for being provided the authentication.
(4) Verify the logic method of authentication.
(5) Verify the strength of the authentication through password cracking and re-applying discovered passwords to all access points requiring authentication.
(6) Verify the process for receiving authentication.
(7) Test for logic errors in the application of the authentication.

Trust Verification refers to tests for trusts between systems within the scope where trust refers to access to information or physical property without the need for identification or authentication.

- Spoofing

(1) Test measures to access property within the scope by spoofing your network address as one of the trusted hosts.
(2) Verify if available caching mechanisms can be poisoned.

- Phishing

(1) Verify that URLs for submissions and queries on the target are concise, within the same domain, use only the POST method, and use consistent branding.
(2) Verify that target content images/records/data do not exist on sites outside of the target to create a duplicate of the target.
(3) Examine top level domain records for domains similar to those identified within the scope.
(4) Verify that the target uses personalization in websites and mail when interacting with authenticated users.
(5) Verify the control and response of the target to mail bounces where the FROM is spoofed in the header field to be that of the target domain.

- Resource Abuse

(1) Test the depth of access to business or confidential information available on web servers without any established, required credentials.
(2) Test if information is sent to the outside of the scope as padding to network packets such as that which has occurred previously as “Ether leak”.
(3) Verify that continuity measures, specifically load balancing, are seamless outside the scope to prevent users from using, referring, linking, bookmaking, or abusing just one of the resources.

Control Verification refers to tests to enumerate and verify the operational functionality of safety measures for assets and services.

- Non-repudiation

(1) Enumerate and test for use or inadequacies of daemons and systems to properly identify and log access or interactions to property for specific evidence to challenge repudiation.

(2) Document the depth of the recorded interaction and the process of identification.

(3) Verify that all methods of interactions are properly recorded with proper identification.

(4) Identify methods of identification which defeat repudiation.

- Confidentiality

(1) Enumerate all interactions with services within the scope for communications or assets transported over the channel using secured lines, encryption, "quieted" or "closed" interactions to protect the confidentiality of the information property between the involved parties.

(2) Verify the acceptable methods used for confidentiality.

(3) Test the strength and design of the encryption or obfuscation method.

(4) Verify the outer limits of communication which can be protected via the applied methods of confidentiality.

- Privacy

(1) Enumerate services within the scope for communications or assets transported using specific, individual signatures, personal identification, "quieted" or "closed room" personal interactions to protect the privacy of the interaction and the process of providing assets only to those within the proper security clearance for that process, communication, or asset.

(2) Correlate information with non-responsive TCP and UDP ports to determine if availability is dependent upon a private type of contact or protocol.

- Integrity. Enumerate and test for inadequacies of integrity where using a documented process, signatures, encryption, hash, or markings to assure that the asset cannot be changed, redirected, or reversed without it being known to the parties involved.

2.3.2.2. Vulnerability analysis, which involves comparing the services, applications, and operating systems of scanned hosts against vulnerability databases (a process that is automatic for vulnerability scanners) and the testers' own knowledge of vulnerabilities. The team conducts the authorized attacks using public, custom, and professional tools to search for vulnerabilities in the targets, which will allow access permission. Manual processes can identify new or obscure vulnerabilities that automated scanners may miss, but are much slower than an automated

scanner. These tests will expose compromised hosts that will be used as escalating points during the next stages. Next, the team collates information gathered during the previous stage in order to plan a series of subsequent actions. These will include planning of the overall approach for the pen test in question, as well as formalizing which targets require further research. While vulnerability scanners check only for the possible existence of vulnerability, the attack phase of a penetration test exploits the vulnerability to confirm its existence. Most vulnerability exploited by penetration testing fall into the following categories:

- **Misconfigurations.** Misconfigured security settings, particularly insecure default settings, are usually easily exploitable.
- **Kernel Flaws.** Kernel code is the core of an OS, and enforces the overall security model for the system—so any security flaw in the kernel puts the entire system in danger.
- **Buffer Overflows.** A buffer overflow occurs when programs do not adequately check input for appropriate length. When this occurs, arbitrary code can be introduced into the system and executed with the privileges—often at the administrative level—of the running program.
- **Insufficient Input Validation.** Many applications fail to fully validate the input they receive from users. An example is a Web application that embeds a value from a user in a database query. If the user enters SQL commands instead of or in addition to the requested value, and the Web application does not filter the SQL commands, the query may be run with malicious changes that the user requested—causing what is known as a SQL injection attack.
- **Symbolic Links.** A symbolic link is a file that points to another file. Operating systems include programs that can change the permissions granted to a file. If these programs run with privileged permissions, a user could strategically create symbolic links to trick these programs into modifying or listing critical system files.
- **File Descriptor Attacks.** File descriptors are numbers used by the system to keep track of files in lieu of filenames. Specific types of file descriptors have implied uses. When a privileged program assigns an inappropriate file descriptor, it exposes that file to compromise.
- **Race Conditions.** Race conditions can occur during the time a program or process has entered into a privileged mode. A user can time an attack to take advantage of elevated privileges while the program or process is still in the privileged mode.
- **Incorrect File and Directory Permissions.** File and directory permissions control the access assigned to users and processes. Poor permissions could allow many types of

attacks, including the reading or writing of password files or additions to the list of trusted remote hosts.

The correspondent OSSTM phase is **Inquest – "**Much of security auditing is about the information that the Analyst uncovers. In this phase, the various types of value or the detriment from misplaced and mismanaged information as an asset is brought to light.**"** The correspondent modules are: **Process Verification**, **Configuration Verification**, **Property Validation**, **Segregation Review, Exposure Verification,** and **Competitive Intelligence Scouting.**

Process Verification represents tests to examine the maintenance of functional security in established processes and due diligence as defined in the Posture Review.

- Maintenance

(1) Examine and document the timeliness, appropriateness, access to, and extent of processes for notification and security response in regards to network and security monitoring.

(2) Verify the appropriateness and functionality of incident response and forensics capabilities for all types of systems.

(3) Verify the level of incident or compromise which the support channels can detect and the length of response time.

- Misinformation

Determine the extent to which security notifications and alarms can be expanded or altered with misinformation.

- Due Diligence

Map and verify any gaps between practice and requirements as determined in the Posture Review through all channels.

- Indemnification

(1) Document and enumerate targets and services which are protected from abuse or circumvention of employee policy, are insured for theft or damages, or use liability and permission disclaimers.

(2) Verify the legality and appropriateness of the language in the disclaimers.

(3) Verify the affect of the disclaimers upon security or safety measures.

(4) Examine the language of the insurance policy for limitations on types of damages or assets.

Configuration Verification represents tests to gather all information, technical and non-technical, on how assets are intended to work, and to examine the ability to circumvent or disrupt functional security in assets, exploiting improper configuration of access controls, loss controls, and applications.

- Configuration Controls

(1) Examine controls to verify the configurations and baselines of systems, equipment and applications meet the intent of the organization and reflect a business justification.

(2) Examine Access Control Lists and business roles configured on networks, systems, services, and applications within the scope to ensure they meet the intent of the organization and reflect a business justification.

- Common Configuration Errors

(1) Verify services available are not unnecessarily redundant and that they match the systems' intended business role.

(2) Verify default settings have been changed. Some devices or applications ship with a default or hidden administrative account. These accounts should be changed, or if possible, disabled or deleted and replaced with a new administrative account.

(3) Verify that Administration is done locally or with controls to limit who or what can access the remote administration interfaces of the equipment.

- Limitations Mapping

(1) Check for unnecessary or unused services/features available.

(2) Check for default credentials.

(3) Identify if any known vulnerabilities are residing on the systems.

Property Validation represents tests to examine information and data available within the scope or provided by personnel who may be illegal or unethical.

- Sharing

Verify the extent to which individually licensed, private, faked, reproduced, non-free, or non-open property is shared either intentionally through sharing processes and programs, libraries, and personal caches or unintentionally through mismanagement of licenses and resources, or negligence.

- Black Market

Verify the extent to which individually licensed, private, faked, reproduced, non-free, or non-open property is promoted, marketed, or sold between personnel or by the organization.

- Sales Channels

Verify whether any public, out of scope businesses, auctions, or property sales provide contact information from targets within the scope.

Segregation Review represents **t**ests for appropriate separation of private or personal information property from business information. Like a privacy review, it is the focal point of the

legal and ethical storage, transmission, and control of personnel, partner, and customer private information property.

- Privacy Containment Mapping

Map key locations of private information property within the scope, what information is stored, how and where the information is stored, and over which channels the information is communicated.

- Disclosure

(1) Examine and document types of disclosures of private information property for segregation according to policy and regulations as determined in the Posture Review.

(2) Verify that private information and confidential intellectual property, such as documents, service contracts, OS/Software keys, etc. are not available to anyone without proper privileges.

- Limitations

(1) Verify that design considerations or channel alternatives exist for people with physical limitations to interact with the target.

(2) Identify any parts of the infrastructure designed to interact with children legally identified as minors and verify what and how identifying information is provided from that child.

-Discrimination

Verify information requested and privileges granted from gatekeepers in cases where age (specifically minors), sex, race, custom/culture and religion are factors which may be discriminated against in accordance to the Posture Review.

Exposure Verification represents tests for uncovering information which provides for or leads to access or allows for access to multiple locations with the same authentication.

- Exposure Enumeration

(1) Enumerate information regarding the organization such as organization charts, key personnel titles, job descriptions, personal and work telephone numbers, mobile phone numbers, business cards, shared documents, resumes, and organizational affiliations, private and public e-mail addresses, log-ins, log-in schemes, passwords, back-up methods, insurers, or any particular organizational information stated implicitly as confidential in regulations and policy.

(2) Enumerate system, service and application exposures detailing the design, type, version, or state on the targets or from resources outside the scope such as from postings or leaks.

Competitive Intelligence Scouting represents tests for scavenging information that can be analyzed as business intelligence. While competitive intelligence as a field is related to marketing, the process here includes any form of competitive intelligence gathering, including but not limited to economic and industrial espionage. Business information includes but is not limited to business relationships like employees, partners, or resellers, contacts, finances, strategy, and plans.

-Business Grinding

Enumerate and evaluate access points (gateways) to business property within the scope: what business information is stored, how it is stored, and where the information is stored.

- Profiling

(1) Profile employee skill requirement types, pay scales, channel and gateway information, technologies, and organizational direction from sources outside the scope.

(2) Profile data network set-ups and configurations from job databases and newspapers hiring ads for data networking positions within the organization relating to hardware and software engineering or administration within the target's default business language(s).

- Business Environment

(1) Explore and document from individual gateway personnel business details such as alliances, partners, major customers, vendors, distributors, investors, business relations, production, development, product information, planning, stocks and trading, and any particular business information or property stated implicitly as confidential in regulations and policy.

(2) Review third party web notes, annotations, and social bookmark site content made for the web presence of the scope.

- Organizational Environment

Examine and document types of disclosures of business property from gatekeepers on operations, processes, hierarchy, financial reporting, investment opportunities, mergers, acquisitions, channel investments, channel maintenance, internal social politics, personnel dissatisfaction and turn-over rate, primary vacation times, hiring, firings, and any particular organizational property stated implicitly as confidential in regulations and policy.

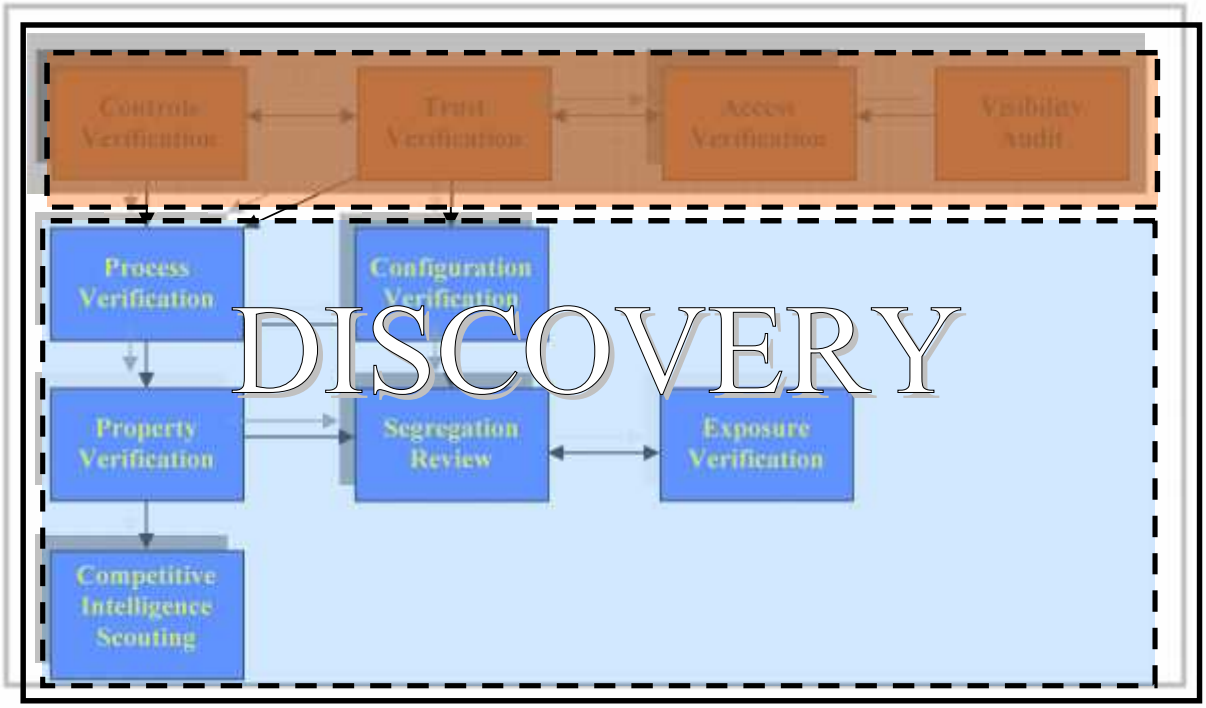


Figure 11. The superposition of NIST and OSSTMM discovery phase

2.3.2.3 Conclusions

It is quite difficult in this moment to make an assertion like in previous section. Intuitively (but it can be a fatal error) we can do the same thing. I mean, to find a feed-back in order to reduce the false reports. More effort and study is necessary in this phase before to make a statement of this kind.

2.3.2 The attack phases (NIST 800-115 and OSSTMM 3 mapping)

The attack is at the heart of any penetration test. If an attack is successful, the vulnerability is verified and safeguards are identified to mitigate the associated security exposure. In many cases, exploits provide exploit instructions or code for many identified vulnerabilities.) that are executed do not grant the maximum level of potential access to an attacker. They may instead result in the testers learning more about the targeted network and its potential vulnerabilities, or induce a change in the state of the targeted network's security. Some exploits enable testers to escalate their privileges on the system or network to gain access to additional resources. If this occurs, additional analysis and testing are required to determine the true level of risk for the network, such as identifying the types of information that can be gleaned, changed, or removed from the system. In the event an attack on a specific vulnerability proves impossible, the tester should attempt to exploit another discovered vulnerability. If testers are able to exploit vulnerability, they can install more tools on the target system or network to facilitate the testing process. These tools are used to gain access to additional systems or resources on the network,

and obtain access to information about the network or organization. Testing and analysis on multiple systems should be conducted during a penetration test to determine the level of access an adversary could gain. This process is represented in the feedback loop in Figure 5-1 between the attack and discovery phase of a penetration test.

The correspondent OSSTM phase is **Intervention – "**These tests are focused on the resources the targets require in the scope. Those resources can be switched, changed, overloaded, or starved to cause penetration or disruption. This is often the final phase of a security test to assure disruptions do not affect responses of less invasive tests and because the information for making these tests may not be known until other phases have been carried out. The final module of Alert and Log Review is required to verify prior tests which provided no interactivity back to the Analyst. Most security tests that do not include this phase may still need to run an end review from the perspective of the targets and assets to clarify any anomalies.**"**

The correspondent modules are: **Containment Process Identification Process, Privileges Audit Verification, Survivability Validation** and **Alert and Log Review.**

Containment Process Identification identifies and examines quarantine methods for aggressive and hostile contacts such as malware, rogue access points, unauthorized storage devices, etc.

- Containment Levels

(1) Measure the minimum resources that need to be available to this subsystem in order for it to perform its task.

(2) Verify any resources available to this subsystem that it does not need to perform its tasks and what resources are shielded from use by this subsystem.

(3) Verify the detection measures present for the detection of attempted access to the shielded resources.

(4) Verify the features of the containment system.

(5) Verify detection measures are present for detection of 'unusual' access to the needed resources

(6) Measure the response and process against encoded, packaged, condensed, renamed, or masqueraded inputs.

(7) Verify the state of containment and length of time for quarantine methods both into and out of the scope. Ensure the completeness and thoroughness of the methods and that they are within legal context and boundaries.

Privileges Audit represents tests where credentials are supplied to the user and permission is granted for testing with those credentials.

-Identification

Examine and document the authorization process for obtaining identification from users through both legitimate and fraudulent means on all channels.

-Authorization

(1) Examine and verify any means for gaining fraudulent authorization to gain privileges similar to that of other personnel.

(2) Enumerate the use of default accounts on targets.

(3) Test access to authenticated access points through the most appropriate and available cracking techniques. Password cracking via dictionary or brute-force may be limited by the time frame of the audit and therefore not a valid test of the protection from that authentication schema however any successful discoveries do attest to its weakness.

-Escalation

(1) Collect information on persons with high privileges. Look for trusted roles or positions, access gateways for trusted persons, and any required physical access media such as tokens or smart cards.

(2) Verify the boundaries of privileges on the target or across multiple targets and if the means exists to escalate those privileges.

Survivability Validation determining and measuring the resilience of the targets within the scope to excessive or hostile changes designed to cause failure or degradation of service. Denial of Service (DoS) is a situation where a circumstance, either intentionally or accidentally, prevents the system from functioning as intended. In certain cases, the system may be functioning exactly as designed however it was never intended to handle the load, scope, or parameters being imposed upon it. Survivability tests must be closely monitored as the intent is to cause failure and this may be unacceptable to the target's owner.

- Resilience

(1) Verify single points of failure (choke points) in the infrastructure where change or failure can cause a service outage.

(2) Verify the impact to target access which a system or service failure will cause.

(3) Verify the privileges available from the failure-induced access.

(4) Verify the operational functionality of controls to prevent access or permissions above lowest possible privileges upon failure.

- Continuity

(1) Enumerate and test for inadequacies from all targets with regard to access delays and service response times through back-up systems or the switch to alternate channels.

(2) Verify intruder lock-out schemes cannot be used against valid users.

- Safety

Map and document the process of gatekeepers shutting down target systems due to evacuation or safety concerns as a gap analysis with regulation and security policy.

Alert and Log Review represents a gap analysis between activities performed with the test and the true depth of those activities as recorded or from third-party perceptions both human and mechanical.

- Alarm

Verify and enumerate the use of a localized or scope-wide warning system, log, or message for each access gateway over each channel where a suspect situation is noted by personnel upon suspicion of circumvention attempts, social engineering, or fraudulent activity.

- Storage and Retrieval

(1) Document and verify unprivileged access to alarm, log, and notification storage locations and property.

(2) Verify the quality and the length of time of the document storage to assure the data will maintain integrity on that storage medium for the required duration.

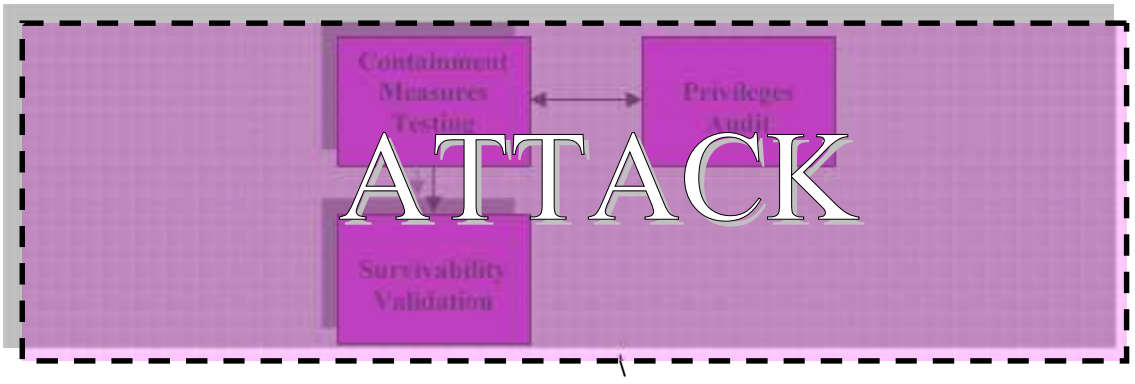


Figure 12. The superposition of NIST and OSSTMM attack phase

2.3.3 The clean-up and reporting phases

Cleanup phase is not specified in the NIST nor OSSTMM methodology but represents a very important phase. This stage concerns cleaning up log files and making sure whatever settings or parameters were changed during the Pen Test are set back to their original condition. The team cleans up all traces of the pen test by removing all testing traces of compromised systems, returning the system and any compromised hosts to the exact configurations that they had prior to the penetration test.

The reporting phase occurs simultaneously with the other three phases of the penetration test (see Figure 1). In the planning phase, the assessment plan—or ROE—is developed. In the discovery and attack phases, written logs are usually kept and periodic reports are made to system administrators and/or management. The final report must map the findings (vulnerabilities found, exploits performed) to the risk the company may has been exposed to if the threats were realized. At this point the team is ready to report high-risk vulnerabilities to the IT decision-makers so that the IT organization is better informed and better prepared to conduct their own penetration testing or to direct additional consulting services.

The report will also review:

- The objectives and scope of the penetration test
- Conclusions from each test phase regarding remediation required and the relative priority of these recommendations

Details gathered on every system, including the high-risk systems found vulnerable to attack, and detailed lists of vulnerabilities.

3. CONCLUSIONS

Sometimes, it seems that following a methodology during the penetration tests is usefulness. Indeed, there are a lot of methodologies and no criteria to make decision which of them to be applied at a peculiar case. Furthermore, the reports are brushy (hundreds of pages) and extremely confuse. Many false positives and false negatives in the report, no real vulnerabilities are ever acted on.

However, a methodology represents the rationale and the philosophical assumptions that underlie a particular study relative to the scientific method. Following a methodology it means to introduce an order in your test. Based on experience, you can skip, complete or redesign some steps in order to adapt at your concrete situation. In order to minimize this effort, I consider that firstly is necessary to try harmonizing the provisions of existent methodologies in order to retain the appropriate elements in accordance with our peculiar purpose. In other words, it is necessary to synchronize and stabilize the methodologies between them in accordance with a peculiar purpose and this paper demonstrated that such approach can reveal other mechanisms that are useful in pen test organizing and conduct.

REFERENCES

1. **Institute for Security and Open Methodologies (ISECOM) -** Open Source Testing Methodology Manual (OSSTMM), version 3
2. **National Institute of Standard and Technology - NIST 800-53A** – Guide for Assessing the Security Controls in Federal Information Systems
3. **National Institute of Standard and Technology - NIST 800-115 -** Technical Guide to Information Security Testing and Assessment

THREATS AND VULNERABILITIES

MAJ. Marius ȚÎRDOIU

1. INTRODUCTION

For much of the last decade, the world has been striving to adapt to changes brought by the dramatic improvements in information and communication technology, particularly as offered by the Internet. These changes have had a significant effect on the economics of information and have created new business models that have resulted in a new information economy. Companies are exchanging goods, services and information in new ways that are more efficient and that are blurring geographic and geopolitical boundaries.

The technologies and resultant environment are evolving so rapidly that is difficult, if not impossible, to fully digest, adapt and incorporate changes before newer and better capabilities are developed. Incorporating these changes is especially difficult in the context of the military, where operational concepts and doctrine are prone to change very slowly. To illustrate the concept of Risk Management in Information Security, I like to use a popular diagram from Common Criteria, shown below:

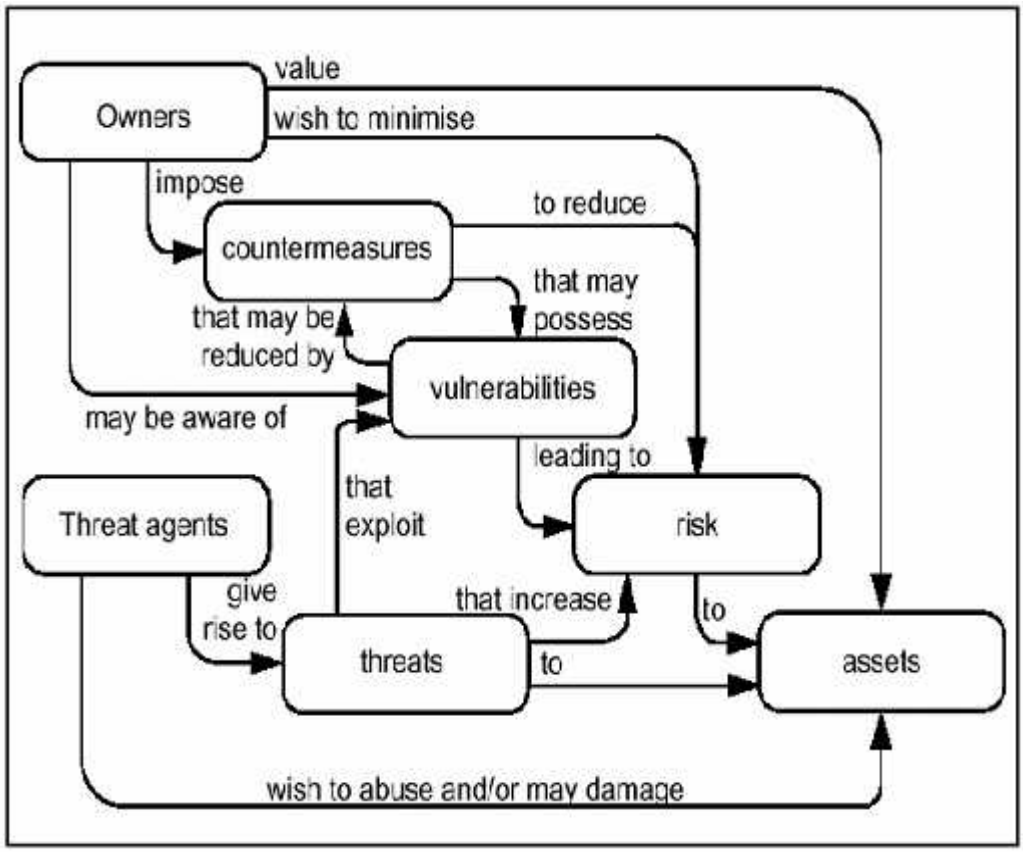


Fig. 1

In the center of this diagram you'll find the term vulnerabilities. ***Vulnerabilities*** are any weaknesses of a system. A system *always* contains vulnerabilities. You cannot build a 100% perfect system with no vulnerabilities, even if you have unlimited power, money, and time to build such a system. All systems contain imperfect components, and the integration of imperfect components produces an imperfect system that always possesses certain vulnerabilities.

Threats are elements from various sources that can exploit vulnerabilities and that increase risk. Threats can be initiated by ***threat agents***. A common threat agent for IT systems is people. They can accidentally or intentionally exploit vulnerabilities of a system to impact an IT system.

Risk is the probability that the system's asset will be damaged/abused by the threats that exploit the vulnerabilities. Assets can be tangible (such as hardware/software) or intangible (such as good will and customers' confidence).

In order to manage risk, we deploy ***countermeasures*** (controls) to a system to reduce the vulnerabilities. The decision to deploy certain countermeasures to reduce the vulnerabilities and hence reduce risk lies solely on the information owner, who bears all consequences arising from the risk.

In a formal risk management exercise, an organization should undergo an intense brainstorming session to discover all possible threats that can exploit the vulnerabilities of a system. The difficult part of this step is not determining whether a certain threat will cause risk to a system, but the effort required to locate all possible threats to a system. Anything overlooked could lead to possible serious exposure to risks that have not been identified.

It is of the utmost importance for the owner (the "Owners" in the diagram) of an organization to identify all possible threats to its information system to the very best of his/her effort and knowledge, in order to fulfill fiduciary duties to customers and other stakeholders. Without knowing what the risks are, it's impossible to implement suitable countermeasures to contain and mitigate those risks.

2. THREATS

The term "*threat*" refers to the source and means of a particular type of attack. A threat assessment is performed to determine the best approaches to securing a system against a particular threat, or class of threat. Penetration testing exercises are substantially focused on assessing threat profiles, to help one develop effective countermeasures against the types of attacks represented by a given threat. Where risk assessments focus more on analysing the potential and tendency of one's resources to fall prey to various attacks, threat assessments focus more on analysing the attacker's resources. Analysing threats can help one develop specific

security policies to implement in line with policy priorities and understand the specific implementation needs for securing one's resources.

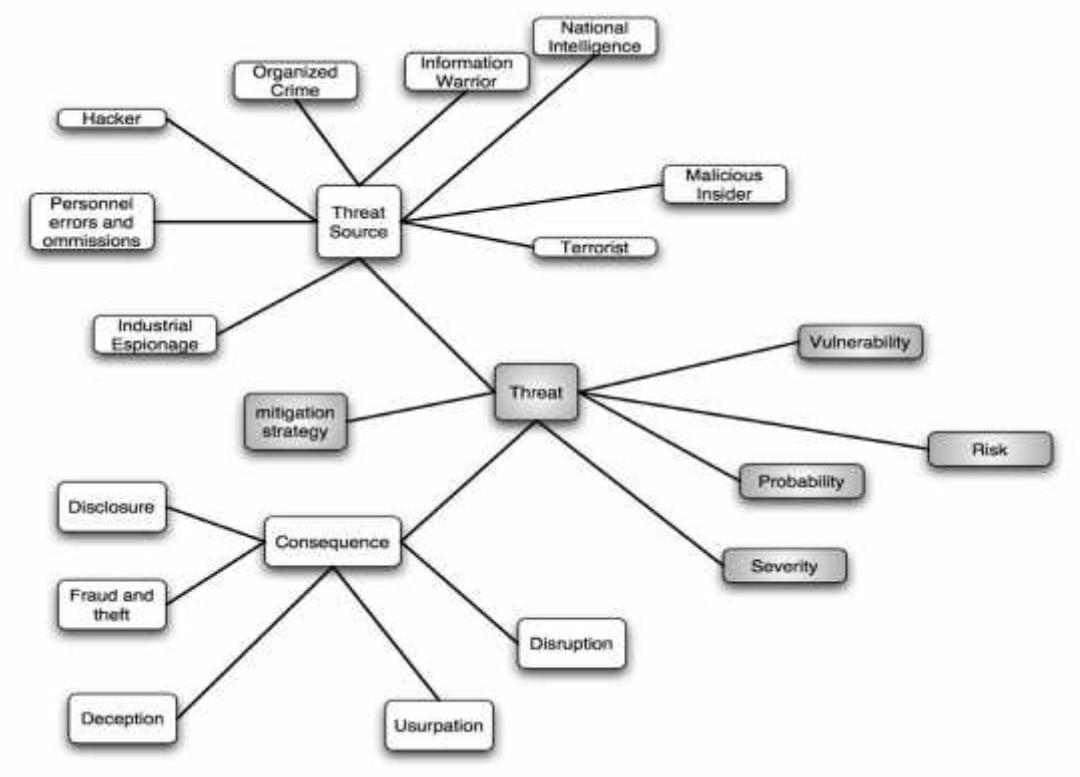


Fig. 2

2.1. Definitions

In Computer security a **threat** is a potential for violation of security, which exists when there is a circumstance, capability, action, or event that could breach security and cause harm. That is, a **threat** is a possible danger that might exploit vulnerability. A threat can be either "intentional" (i.e., intelligent; e.g., an individual cracker or a criminal organization) or "accidental" (e.g., the possibility of a computer malfunctioning, or the possibility of an "act of God" such as an earthquake, a fire, or a tornado). The definition is as IETF RFC 2828.

ISO 27005 defines **threat** as:

[*A potential cause of an incident that may result in harm of systems and organization.*]

A more comprehensive definition, tied to an Information assurance point of view, can be found in "*Federal Information Processing Standards (FIPS) 200, Minimum Security Requirements for Federal Information and Information Systems*" by NIST of United States of America.

[*Any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, or individuals through an information system via unauthorized access, destruction, disclosure, modification of information, and/or denial of service. Also, the potential for a threat-source to successfully exploit a particular information system vulnerability*.]

National Information Assurance Glossary defines **threat** as:

[*Any circumstance or event with the potential to adversely impact an IS through unauthorized access, destruction, disclosure, modification of data, and/or denial of service.*]

ENISA gives a similar definition:

[*Any circumstance or event with the potential to adversely impact an asset [G.3] through unauthorized access, destruction, disclosure, modification of data, and/or denial of service.*]

The Open Group defines **threat** in as:

[*Anything that is capable of acting in a manner resulting in harm to an asset and/or organization; for example, acts of God (weather, geological events, etc.); malicious actors; errors; failures*.]

Factor Analysis of Information Risk defines **threat** as:

[*Threats are anything (e.g., object, substance, human, etc.) that are capable of acting against an asset in a manner that can result in harm. A tornado is a threat, as is a flood, as is a hacker. The key consideration is that threats apply the force (water, wind, exploit code, etc.) against an asset that can cause a loss event to occur.*]

The widespread of computer dependencies and the consequent raising of the consequence of a successful attack, led to a new term cyberwarfare.

It should be noted that nowadays the many real attacks exploit Psychology at least as much as technology. Phishing and Pretexting and other methods are called social engineering techniques. The Web 2.0 applications, specifically Social network services, can be a mean to get in touch with people in charge of system administration or even system security, inducing them to reveal sensitive information.

The most widespread documentation on Computer insecurity is about technical threats such computer virus, trojan and other malware, but a serious study to apply cost effective countermeasures can only be conducted following a rigorous IT risk analysis in the framework of an ISMS: a pure technical approach will let out the psychological attacks, that are increasing threats.

2.2. Threat modeling

Security threat modeling, or threat modeling, is a process of assessing and documenting a system's security risks. Security threat modeling enables you to understand a system's threat profile by examining it through the eyes of your potential foes. With techniques such as entry point identification, privilege boundaries and threat trees, you can identify strategies to mitigate potential threats to your system. Your security threat modeling efforts also enable your team to justify security features within a system, or security practices for using the system, to protect your corporate assets.

There are five aspects to security threat modeling:

a. Identify threats. The first thing to do is to identify assets of interest, you first model the system either with data flow diagrams (DFDs) or UML deployment diagrams. You can identify entry points to your system such as data sources, application programming interfaces (APIs), Web services and the user interface itself. Because an adversary gains access to your system via entry points, they are your starting points for understanding potential threats.

b. Understand the threat(s). To understand the potential threats at an entry point, you must identify any security-critical activities that occur and imagine what an adversary might do to attack or misuse your system. Ask yourself questions such as "How could the adversary use an asset to modify control of the system, retrieve restricted information, manipulate information within the system, cause the system to fail or be unusable, or gain additional rights. In this way, you can determine the chances of the adversary accessing the asset without being audited, skipping any access control checks, or appearing to be another user. To understand the threat posed by the interface between the order and payment processing modules, you would identify and then work through potential security scenarios.

c. Categorize the threats. To categorize security threats, consider the STRIDE (Spoofing, Tampering, Repudiation, Information disclosure, Denial of Service, and Elevation of privilege) approach. Classifying a threat is the first step toward effective mitigation.

d. Identify mitigation strategies. To determine how to mitigate a threat, you can create a diagram called a threat tree. At the root of the tree is the threat itself, and its children (or leaves) are the conditions that must be true for the adversary to realize that threat. Conditions may in turn have subconditions.

e. Test. Your threat model becomes a plan for penetration testing. Penetration testing investigates threats by directly attacking a system, in an informed or uninformed manner. Informed penetration tests are effectively white-box tests that reflect knowledge of the system's internal design, whereas uninformed tests are black box in nature.

2.3. Threat classification

Microsoft has proposed a threat classification called STRIDE, from the initial of threat categories:

- **S**poofing of user identity
- **T**ampering
- **R**epudiation
- **I**nformation disclosure (privacy breach or Data leak)
- **D**enial of Service (D.o.S.)
- **E**levation of privilege

Microsoft used to risk rating security threats using five categories in a classification called DREAD: Risk assessment model. The model is considered obsolete by Microsoft. The categories were:

- **D**amage - how bad would an attack be?
- **R**eproducibility - how easy it is to reproduce the attack?
- **E**xploitability - how much work is it to launch the attack?
- **A**ffected users - how many people will be impacted?
- **D**iscoverability - how easy it is to discover the threat?

The DREAD name comes from the initials of the five categories listed.

The spread over a network of threats can led to dangerous situations. In military and civil fields, threat level as been defined: for example INFOCOM is a threat level used by USA. Leading antivirus software vendors publish global threat level on their websites.

2.4. Associated terms

2.4.1. Threat Agents

Threat Agents

[*Individuals within a threat population; Practically anyone and anything can, under the right circumstances, be a threat agent – the well-intentioned, but inept, computer operator who trashes a daily batch job by typing the wrong command, the regulator performing an audit, or the squirrel that chews through a data cable.*]

Threat agents can take one or more of the following actions against an asset:

- Access – simple unauthorized access
- Misuse – unauthorized use of assets (e.g., identity theft, setting up a porn distribution service on a compromised server, etc.)
- Disclose – the threat agent illicitly discloses sensitive information
- Modify – unauthorized changes to an asset
- Deny access – includes destruction, theft of a non-data asset, etc.

It's important to recognize that each of these actions affects different assets differently, which drives the degree and nature of loss. For example, the potential for productivity loss resulting from a destroyed or stolen asset depends upon how critical that asset is to the organization's productivity. If a critical asset is simply illicitly accessed, there is no direct productivity loss. Similarly, the destruction of a highly sensitive asset that doesn't play a critical role in productivity won't directly result in a significant productivity loss. Yet that same asset, if disclosed, can result in significant loss of competitive advantage or reputation, and generate legal costs. The point is that it's the combination of the asset and type of action against the asset that

determines the fundamental nature and degree of loss. Which action(s) a threat agent takes will be driven primarily by that agent's motive (e.g., financial gain, revenge, recreation, etc.) and the nature of the asset. For example, a threat agent bent on financial gain is less likely to destroy a critical server than they are to steal an easily pawned asset like a laptop.

It is important to separate the the concept of the event that a threat agent get in contact with the asset (even virtually, i.e. through the network) and the event that a threat agent act against the asset.

OWASP collects a list of potential threat agents in order to prevent system designers and programmers insert vulnerabilities in the software.

The term *Threat Agent* is used to indicate an individual or group that can manifest a threat. It is fundamental to identify who would want to exploit the assets of a company, and how they might use them against the company.

Threat Agent = Capabilities + Intentions + Past Activities

These individuals and groups can be classified as follows:

- Non-Target Specific Threat Agents are computer viruses, worms, trojans and logic bombs.
- Employees: Staff, contractors, operational/maintenance personnel, or security guards who are annoyed with the company.
- Organized Crime and Criminals: Criminals target information that is of value to them, such as bank accounts, credit cards or intellectual property that can be converted into money. Criminals will often make use of insiders to help them.
- Corporations: Corporations are engaged in offensive information warfare or competitive intelligence. Partners and competitors come under this category.
- Human, Unintentional: Accidents, carelessness.
- Human, Intentional: Insider, outsider.
- Natural: Flood, fire, lightning, meteor, earthquakes.

2.4.2. Threat Communities

The following threat communities are examples of the human malicious threat landscape many organizations face:

- Internal
    - Employees
    - Contractors (and vendors)
    - Partners
- External
    - Cyber-criminals (professional hackers)

- Spies
- Non-professional hackers
- Activists
- Nation-state intelligence services
- Malware (virus/worm/etc.) authors

2.4.3. Threat management

Threats should be managed by operating an ISMS (information security management system), performing all the IT risk management activities foreseen by laws, standards and methodologies. Very large organizations tend to adopt business continuity management plans in order to protect, maintain and recover business-critical processes and systems. Some of these plans foreseen to set up **computer security incident response team** (CSIRT) or **computer emergency response team** (CERT).

There are some kinds of verification of the threat management process:

- Information security audit.
- Penetration test.

Most organizations perform a subset of these steps, adopting countermeasures based on a non systematic approach: Computer insecurity studies the battlefield of computer security exploits and defences that results.

Countermeasures may include tools such as firewalls, intrusion detection system and anti-virus software, Physical Security measures, policies and procedures such as regular backups and configuration hardening, training such as security awareness education.

3. VULNERABILITY

The term “vulnerability” refers to the security flaws in a system that allows an attack to be successful. Vulnerability testing should be performed on an ongoing basis by the parties responsible for resolving such vulnerabilities, and helps to provide data used to identify unexpected dangers to security that need to be addressed. Such vulnerabilities are not particular to technology — they can also apply to social factors such as individual authentication and authorization policies.

Testing for vulnerabilities is useful for maintaining ongoing security, allowing the people responsible for the security of one’s resources to respond effectively to new dangers as they arise. It is also invaluable for policy and technology development, and as part of a technology selection process; selecting the right technology early on can ensure significant savings in time, money, and other business costs further down the line.

In computer security, **vulnerability** is a weakness which allows an attacker to reduce a system's information assurance.

Vulnerability is the intersection of three elements: a system susceptibility or flaw, attacker access to the flaw, and attacker capability to exploit the flaw. To be vulnerable, an attacker must have at least one applicable tool or technique that can connect to a system weakness. In this frame, vulnerability is also known as the attack surface.

Vulnerability management is the cyclical practice of identifying, classifying, remediating, and mitigating vulnerabilities. This practice generally refers to software vulnerabilities in computing systems.

A security risk may be classified as a vulnerability. The usage of vulnerability with the same meaning of risk can lead to confusion. The risk is tied to the potential of a significant loss. Then there are vulnerabilities without risk: for example when the affected asset has no value. Vulnerability with one or more known instances of working and fully-implemented attacks is classified as an exploitable vulnerability - a vulnerability for which an exploit exists. The **window of vulnerability** is the time from when the security hole was introduced or manifested in deployed software, to when access was removed, a security fix was available / deployed, or the attacker was disabled.

Security bug is a narrower concept: there are vulnerabilities that are not related to software: hardware, site, personnel vulnerabilities are examples of vulnerabilities that are not security software bugs.

Constructs in programming languages that are difficult to use properly can be a large source of vulnerabilities.

3.1. Definitions

ISO 27005 defines **vulnerability** as:

[*A weakness of an asset or group of assets that can be exploited by one or more threats.*]

where an *asset is anything that can has value to the organization, its business operations and their continuity, including information resources that support the organization's mission*

IETF RFC 2828 defines **vulnerability** as:

[*A flaw or weakness in a system's design, implementation, or operation and management that could be exploited to violate the system's security policy.*]

The Committee on National Security Systems of United States of America defined **vulnerability** in CNSS Instruction No. 4009 dated 26 April 2010 National Information Assurance Glossary:

[*Vulnerability - Weakness in an IS, system security procedures, internal controls, or implementation that could be exploited.*]

Many NIST publications define **vulnerability** in IT contest in different publications: FISMApedia term provide a list. Between them SP 800-30, give a broader one:

[*A flaw or weakness in system security procedures, design, implementation, or internal controls that could be exercised (accidentally triggered or intentionally exploited) and result in a security breach or a violation of the system's security policy.*]

ENISA defines **vulnerability** as:

[*The existence of a weakness, design, or implementation error that can lead to an unexpected, undesirable event [G.11] compromising the security of the computer system, network, application, or protocol involved.(ITSEC)*]

The Open Group defines **vulnerability** as:

[*Anything that is capable of acting in a manner resulting in harm to an asset and/or organization; for example, acts of God (weather, geological events,etc.); malicious actors; errors; failures*.]

Factor Analysis of Information Risk (FAIR) defines **vulnerability** as:

[*The probability that an asset will be unable to resist the actions of a threat agent.*]

According FAIR vulnerability is related to Control Strength, i.e. the strength of a control as compared to a standard measure of force and the threat Capabilities, i.e. the probable level of force that a threat agent is capable of applying against an asset.

ISACA defines **vulnerability** in Risk It framework as:

[*A weakness in design, implementation, operation or internal control.*]

3.2. Classification

Vulnerabilities are classified according to the asset class they related to:

- hardware
    - susceptibility to humidity
    - susceptibility to dust
    - susceptibility to soiling
    - susceptibility to unprotected storage
- software
    - insufficient testing
    - lack of audit trail
- network
    - unprotected communication lines
    - insecure network architecture
- personnel

    - inadequate recruiting process
    - inadequate security awareness
- site
    - area subject to flood
    - unreliable power source
- organizational
    - lack of regular audits
    - lack of continuity plans

3.3. Causes

- Complexity: Large, complex systems increase the probability of flaws and unintended access points
- Familiarity: Using common, well-known code, software, operating systems, and/or hardware increases the probability an attacker has or can find the knowledge and tools to exploit the flaw
- Connectivity: More physical connections, privileges, ports, protocols, and services and time each of those are accessible increase vulnerability
- Password management flaws: The computer user uses weak passwords that could be discovered by brute force. The computer user stores the password on the computer where a program can access it. Users re-use passwords between many programs and websites.
- Fundamental operating system design flaws: The operating system designer chooses to enforce sub optimal policies on user/program management. For example operating systems with policies such as default permit grant every program and every user full access to the entire computer. This operating system flaw allows viruses and malware to execute commands on behalf of the administrator.
- Internet Website Browsing: Some internet websites may contain harmful Spyware or Adware that can be installed automatically on the computer systems. After visiting those websites, the computer systems become infected and personal information will be collected and passed on to third party individuals.
- Software bugs: The programmer leaves an exploitable bug in a software program. The software bug may allow an attacker to misuse an application.
- Unchecked user input: The program assumes that all user input is safe. Programs that do not check user input can allow unintended direct execution of commands or SQL statements (known as Buffer overflows, SQL injection or other non-validated inputs).

- Too feeble learning system from occurred accidents: for example most vulnerabilities discovered in IPV4 protocol software where discovered in the new IPV6 implementations

The research has shown that the most vulnerable point in most information systems is the human user, operator, designer, or other human: so humans should be considered in their different roles as asset, threat, information resources.

3.4. Vulnerability consequences

The impact of a security breach can be very high. The fact that IT managers, or upper management, can (easily) know that IT systems and applications have vulnerabilities and do not perform any action to manage the IT risk is seen as a misconduct in most legislations. Privacy law forces managers to act to reduce the impact or likelihood that security risk. Information technology security audit is a way to let other independent people certify that the IT environment is managed properly and lessen the responsibilities, at least having demonstrated the good faith. Penetration test is a form of verification of the weakness and countermeasures adopted by an organization: a White hat hacker tries to attack organization information technology assets, to find out how is easy or difficult to compromise the IT security. The proper way to professionally manage the IT risk is to adopt an Information Security Management System, such as ISO/IEC 27002 or Risk IT and follow them, according to the security strategy set forth by the upper management.

One of the key concept of information security is the principle of defence in depth: i.e. to set up a multilayer defence system that can:

- prevent the exploit
- detect and intercept the attack
- find out the threat agents and persecute them

Intrusion detection system is an example of a class of systems used to detect attacks.

Physical security is a set of measures to protect physically the information asset: if somebody can get physical access to the information asset is quite easy to made resources unavailable to its legitimate users.

Some set of criteria to be satisfied by a computer, its operating system and applications in order to meet a good security level have been developed: ITSEC and Common criteria are two examples.

3.5. Vulnerability disclosure date

The time of disclosure of vulnerability is defined differently in the security community and industry. It is most commonly referred to as "a kind of public disclosure of security information

by a certain party". Usually, vulnerability information is discussed on a mailing list or published on a security web site and results in a security advisory afterward.

The **time of disclosure** is the first date security vulnerability is described on a channel where the disclosed information on the vulnerability has to fulfill the following requirement:

- The information is freely available to the public
- The vulnerability information is published by a trusted and independent channel/source
- The vulnerability has undergone analysis by experts such that risk rating information is included upon disclosure

3.6. Identifying and removing vulnerabilities

Many software tools exist that can aid in the discovery (and sometimes removal) of vulnerabilities in a computer system. Though these tools can provide an auditor with a good overview of possible vulnerabilities present, they can not replace human judgment. Relying solely on scanners will yield false positives and a limited-scope view of the problems present in the system.

Vulnerabilities have been found in every major operating system including Windows, Mac OS, various forms of UNIX and Linux, and others. The only way to reduce the chance of a vulnerability being used against a system is through constant vigilance, including careful system maintenance (e.g. applying software patches), best practices in deployment (e.g. the use of firewalls and access controls) and auditing (both during development and throughout the deployment lifecycle).

3.6.1. Examples of vulnerabilities

Vulnerabilities are related to:

- physical environment of the system
- the personnel
- management
- administration procedures and security measures within the organization
- business operation and service delivery
- hardware
- software
- communication equipment and facilities
- and their combinations.

It is evident that a pure technical approach cannot even protect physical assets: you should have administrative procedure to let maintenance personnel to enter the facilities and people with adequate knowledge of the procedures, motivated to follow it with proper care.

Four examples of vulnerability exploit:

- an attacker finds and uses an overflow weakness to install malware to export sensitive data;
- an attacker convinces a user to open an email message with attached malware;
- an insider copies a hardened, encrypted program onto a thumb drive and cracks it at home;
- a flood damage your computer systems installed at ground floor.

3.6.2. Software vulnerabilities

Common types of software flaws that lead to vulnerabilities include:

- Memory safety violations, such as:
  - Buffer overflows
  - Dangling pointers
- Input validation errors, such as:
  - Format string bugs
  - Improperly handling shell met characters so they are interpreted
  - SQL injection
  - Code injection
  - E-mail injection
  - Directory traversal
  - Cross-site scripting in web applications
  - HTTP header injection
  - HTTP response splitting
- Race conditions, such as:
  - Time-of-check-to-time-of-use bugs
  - Symlink races
- Privilege-confusion bugs, such as:
  - Cross-site request forgery in web applications
  - Clickjacking
  - FTP bounce attack
- Privilege escalation
- User interface failures, such as:
  - Warning fatigue or user conditioning
  - Blaming the Victim Prompting a user to make a security decision without giving the user enough information to answer it
  - Race Conditions

Some set of coding guidelines have been developed and a large number of static code analysers has been used to verify that the code follows the guidelines.

3.6.3. CAULDRON

CAULDRON (Combinatorial Analysis Utilizing Logical Dependencies Residing on Networks) is a tool that was developed to automate vulnerability analysis, the task of examining network security to identify deficiencies and predict the effectiveness of proposed improvements. Vulnerability analysis is performed manually today. To perform this analysis, engineers must find the vulnerabilities that an attacker could exploit and the many paths that an attack could take in order to traverse a network and reach the attacker's target. This has become an intractable task, as systems and networks have grown more complex and as exploits have become more numerous. Given thousands of exploits, vulnerabilities and possible network configurations, vulnerability analysis needs to be automated.

An attack may penetrate a network at one node and then hop from that node to reach a target at a remote node in the network. A multistage attack may employ different exploits along the way, as different nodes may have different vulnerabilities. It may also traverse the network via many possible attack paths. A vulnerability analysis should ideally identify all possible attack paths, and the exploits and vulnerabilities used to traverse them.

Once the attack paths and exploits are known, developers may add security mechanisms or reconfigure the network in order to "harden" the network. Proposed changes can then be analyzed to predict their effectiveness before they are implemented. Multiple solutions can be explored at minimal cost if the process is automated.

Vulnerability analysis needs to be a continuing activity. Networks are dynamic places: they expand and are upgraded; new vulnerabilities are discovered, and so are new exploits. Each of these changes can affect the security posture of a network. By automating vulnerability analysis, CAULDRON makes it practical to periodically perform thorough vulnerability analyses, and find and eliminate new vulnerabilities before an attacker finds and exploits them.

This figure shows CAULDRON's inputs. Commercial off-the-shelf tools provide information about network topology, known threats and intrusions. The user provides CAULDRON with attack scenarios that identify an attacker's potential network entry point(s) and target(s). CAULDRON then finds all of the paths and exploits that an attacker could use to reach those targets.

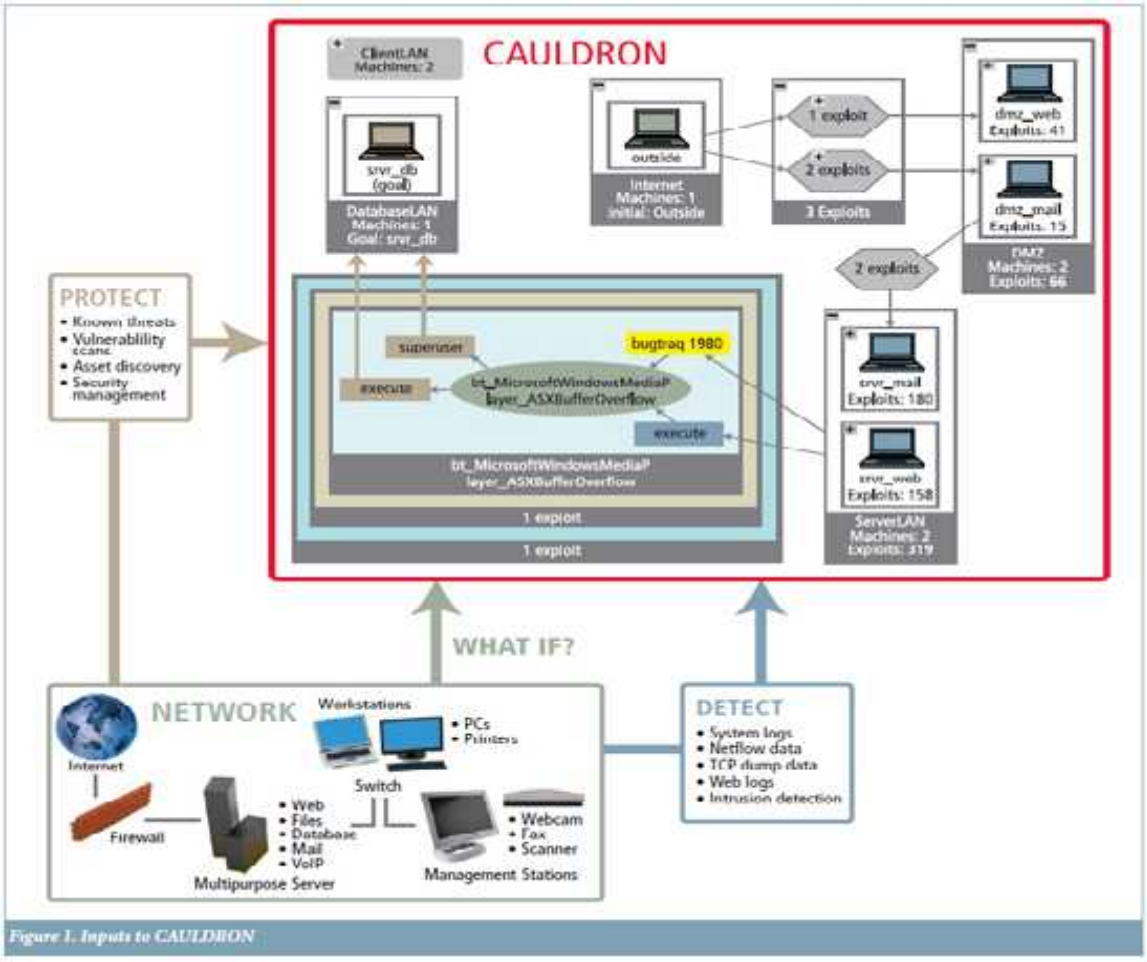


Fig. 3

CAULDRON provides the user with visualizations of its analysis results, as shown in next figure. This gives the user information about attack paths, vulnerabilities, and exploits used, as well as recommendations for how network security can be effectively improved with minimal addition of security mechanisms.

On one of these programs, an 81-host system with more than 2,300 open Internet ports was analyzed for vulnerabilities. Current practice would have required engineers to manually interpret vulnerability scan data, find critical attack paths and eliminate critical vulnerabilities. This would have taken weeks to do. CAULDRON found the attack paths, identified the critical exploits, recommended solutions, and helped eliminate 75 percent of the vulnerabilities in a few hours.

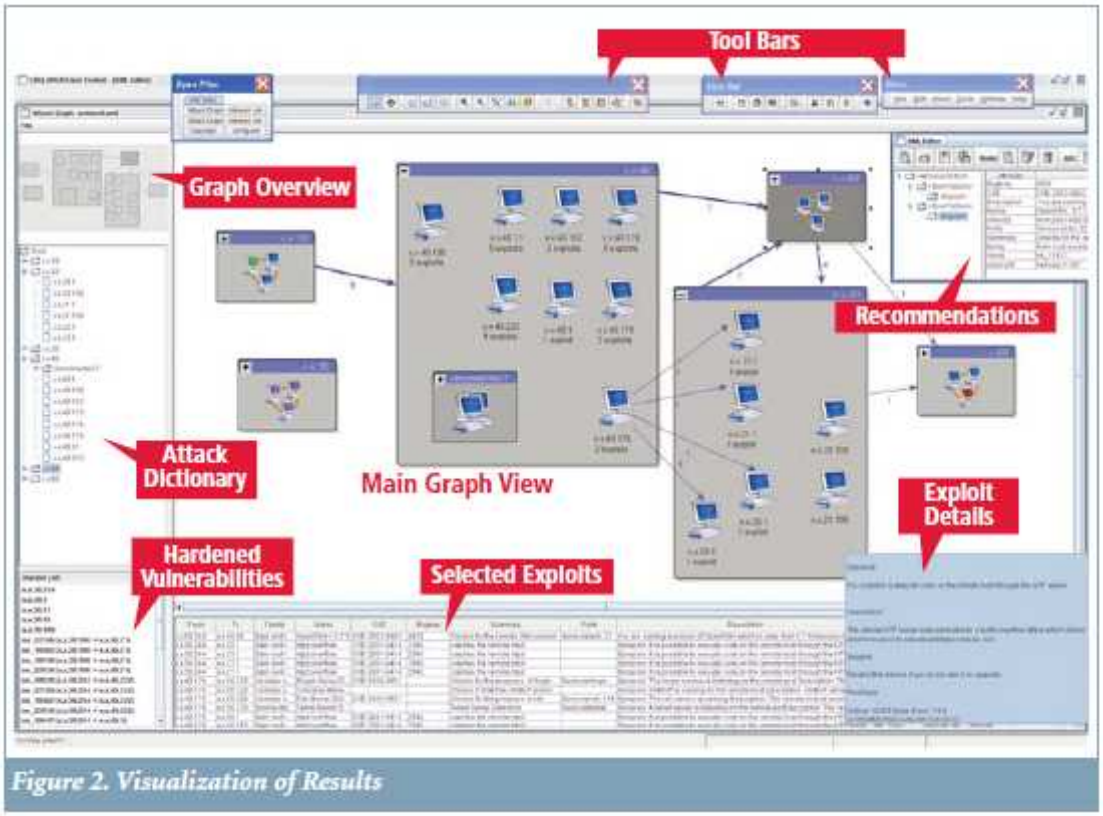


Fig. 4

4. Conclusions

In short, the information developed in the era of the Internet, each computer is running behind there computer system vulnerability every moment threatens the safe operation of the computer system. As a network user, when you browse the web, videos, pictures, and send documents at the same time must take the security of computer systems in advance of factors to consider: installing antivirus software, updated virus database, system vulnerability scanning, to install the patch software so you may say, take preventive measures.

Information security is the ongoing process of exercising due care and due diligence to protect information, and information systems, from unauthorized access, use, disclosure, destruction, modification, or disruption or distribution. The never ending process of information security involves ongoing training, assessment, protection, monitoring & detection, incident response & repair, documentation, and review.

References

1. Federal Information Processing Standards (FIPS) 200, Minimum Security Requirements for Federal Information and Information Systems
2. http://www.enisa.europa.eu/act/rm/cr/risk-management-inventory/glossary#G51 ENISA Glossary threat

3. Wright, Joe; Jim Harmening (2009) "15" *Computer and Information Security Handbook* Morgan Kaufmann Pubblications Elsevier Inc ISBN 978-0-12-374354-1
4. Security engineering:a guide to building dependable distributed systems, second edition, Ross Anderson, Wiley, 2008 ISBN 978-0-470.06852-6
5. Eweek Using Facebook to Social Engineer Your Way Around Security
6. Networkworld Social engineering via Social networking
7. Uncover Security Design Flaws Using The STRIDE Approach
8. OWASP Threat agents categorization
9. FIPS PUB 31 FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION 1974 JUNE
10. "The Three Tenents of Cyber Security". U.S. Air Force Software Protection Initiative. http://www.spi.dod.mil/tenets.htm. Retrieved 2009-12-15.
11. ISO/IEC, "Information technology -- Security tecniques-Information security risk management" ISO/IEC FIDIS 27005:2008
12. British Standard Institute, Information technology -- Security techniques -- Management of information and communications technology security -- Part 1: Concepts and models for information and communications technology security management BS ISO/IEC 13335-1-2004
13. Internet Engineering Task Force RFC 2828 Internet Security Glossary
14. CNSS Instruction No. 4009 dated 26 April 2010
15. NIST SP 800-30 Risk Management Guide for Information Technology Systems
16. Risk Management Glossary Vulnerability
17. Technical Standard Risk Taxonomy ISBN 1-931624-77-1 Document Number: C081 Published by The Open Group, January 2009.
18. "An Introduction to Factor Analysis of Information Risk (FAIR)", Risk Management Insight LLC, November 2006;
19. NIATEC Glossary
20. Technical Report CSD-TR-97-026 Ivan Krsul The COAST Laboratory Department of Computer Sciences, Purdue University, April 15, 1997
21. The Web Application Security Consortium Project, Web Application Security Statistics 2009
22. The Tech Herald: The new era of vulnerability disclosure - a brief chat with HD Moore

SOCIAL NETWORKS

1st LT Loreta GAVRILĂ

Introduction

In both professional and personal life, human beings naturally form groups based on affinities and expertise. We gravitate to others with whom we share interests. Most of us belong to real world networks that formed organically. Not surprisingly, these networks rapidly migrated to the online world. Online social networking has been around in various forms for nearly a decade, and has begun to achieve wide notice in the past few years[26].

Online social networks take many forms, and are created for many reasons, like: to network with new contacts, reconnect with former friends, maintain current relationships, build or promote a business or project, participate in discussions about a certain topic, or just have fun meeting and interacting with other users.

Since their introduction, social network sites (SNSs) such as MySpace, Facebook, Cyworld, and Bebo have attracted millions of users, many of whom have integrated these sites into their daily practices. As of this writing, there are hundreds of SNSs, with various technological affordances, supporting a wide range of interests and practices. While their key technological features are fairly consistent, the cultures that emerge around SNSs are varied. Some sites cater to diverse audiences, while others attract people based on common language or shared racial, sexual, religious, or nationality-based identities. Sites also vary in the extent to which they incorporate new information and communication tools, such as mobile connectivity, blogging, and photo/video-sharing.

Social network tools have changed the way we interact in our personal lives and are in the process of transforming our professional lives. Increasingly, they play a significant role in how business gets done. But they're also high risk. With hundreds of millions of users, these tools have attracted attackers more than any other target in recent years.

The information revolution has given birth to new economies structured around flows of data, information, and knowledge. In parallel, social networks have grown stronger as forms of organization of human activity[27].

[26] www.cerado.com
[27] Social network analysis, Oliver Serrat

I. CONCEPTUAL DELIMITATIONS

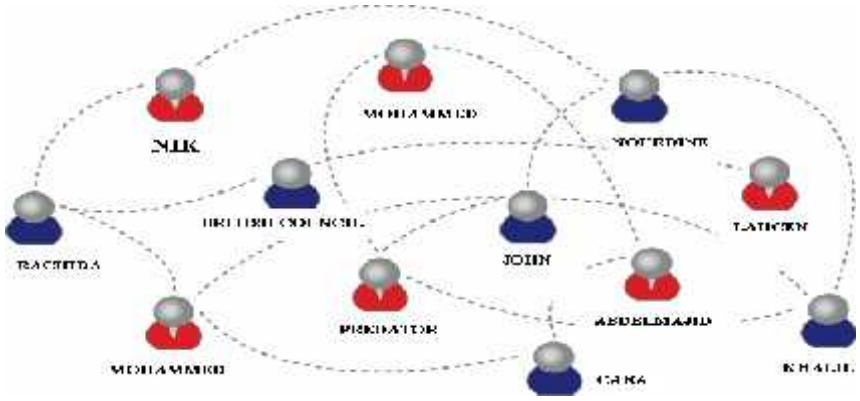
A **network** is a set of nodes, points, or locations connected by means of data, voice, and video communications for the purpose of exchange.

Social refers to the interaction of people and other organisms with each other, and to their collective co-existence[28].

A **social network** is a description of the social structure between actors, mostly individuals or organizations. It indicates the ways in which they are connected through various social familiarities ranging from casual acquaintance to close familiar bonds.

In its simplest form, a social network is a map of specified ties, such as friendship, between the nodes being studied. The network can also be used to measure social capital – the value that an individual gets from the social network. These concepts are often displayed in a social network diagram, where nodes are the points and ties are the lines[29].

We define **social network sites** [SNS] as web-based services that allow individuals to construct a public or semi-public profile within a bounded system, articulate a list of other users with whom they share a connection, and view and traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from site to site[30].



Social network analysis views social relationships in terms of network theory consisting of nodes and ties (also called edges, links, or connections). Nodes are the individual actors within the networks, and ties are the relationships between the actors. The resulting graph-based structures are often very complex. There can be many kinds of ties between the nodes. Research in a number of academic fields has shown that social networks operate on many levels, from families up to the level of nations, and play a critical role in determining the way problems are

[28] http://en.wikipedia.org/wiki/Social_(disambiguation)
[29] http://en.wikipedia.org/wiki/Social_network
[30] Social Network Sites: Definition, History, and Scholarship, Nicole Ellison

solved, organizations are run, and the degree to which individuals succeed in achieving their goals[31].

Examples of social networks include “Facebook”, “You Tube”, “Linkedin”, “Yahoo!Groups”, “Wikipedia”, “Myspace” and hundreds of other sites all focused on empowering individuals to:

a) connect with friends, colleagues or strangers;

b) create, contribute and publish content;

c) comment on, rank or embellish that content;

d) communicate freely and creatively using multiple formats including: email, instant messaging, mobile devices, voice and video and all for free or next to free in terms of real costs.

II. HISTORY OF SOCIAL NETWORK SITES

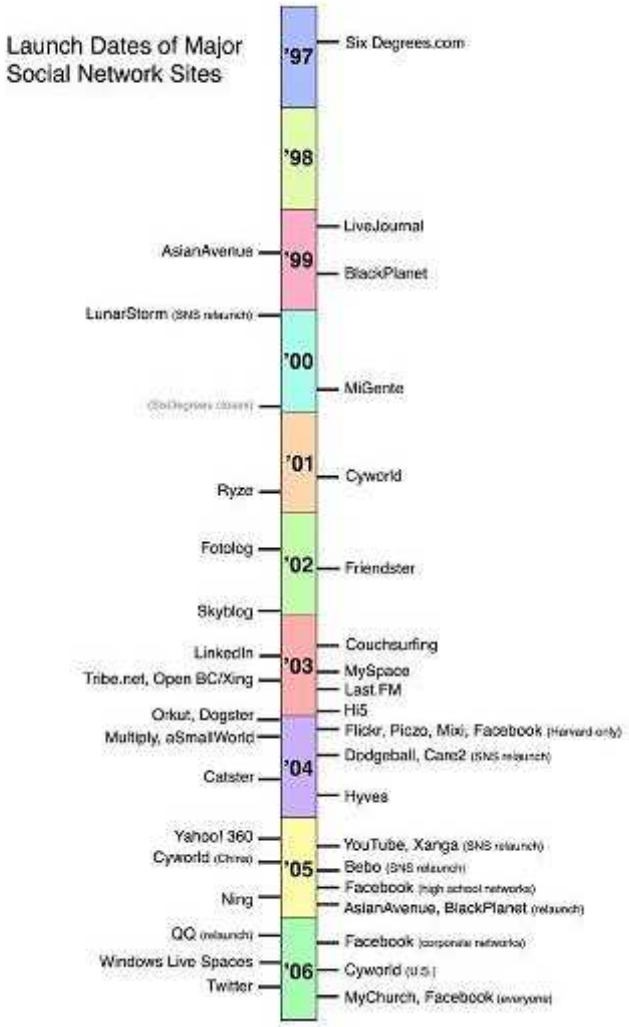
II.1. Timeline of social networks sites

Sporting a name based on the theory somehow associated with actor Kevin Bacon that no person is separated by more than six degrees from another, the site sprung up in 1997 and was one of the very first to allow its users to create profiles, invite friends, organize groups, and surf other user profiles. Its founders worked the six degrees angle hard by encouraging members to bring more people into the fold. Unfortunately, this "encouragement" ultimately became a bit too pushy for many, and the site slowly de-evolved into a loose association of computer users and numerous complaints of spam-filled membership drives. SixDegrees.com folded completely just after the turn of the millennium.

From 1997 to 2001, a number of community tools began supporting various combinations of profiles and publicly articulated Friends. AsianAvenue, BlackPlanet, and MiGente allowed users to create personal, professional, and dating profiles—users could identify Friends on their personal profiles without seeking approval for those connections. Likewise, shortly after its launch in 1999, LiveJournal listed one-directional connections on user pages. LiveJournal's creator suspects that he fashioned these Friends after instant messaging buddy lists—on LiveJournal, people mark others as Friends to follow their journals and manage privacy settings. The Korean virtual worlds site Cyworld was started in 1999 and added SNS features in 2001, independent of these other sites. Likewise, when the Swedish web community LunarStorm refashioned itself as an SNS in 2000, it contained Friends lists, guestbooks, and diary pages.

[31] http://en.wikipedia.org/wiki/Social_network

The next wave of SNSs began when Ryze.com was launched in 2001 to help people leverage their business networks. Ryze's founder reports that he first introduced the site to his friends—primarily members of the San Francisco business and technology community, including the entrepreneurs and investors behind many future SNSs. In particular, the people behind Ryze, Tribe.net, LinkedIn, and Friendster were tightly entwined personally and professionally. They believed that they could support each other without competing. In the end, Ryze never acquired mass popularity, Tribe.net grew to attract a passionate niche user base, LinkedIn became a powerful business service, and Friendster became the most significant, if only as "one of the biggest disappointments in Internet history"[32].



II.2. Most important social networks

In the following section we discuss about the SNSs that shaped the business, cultural, and research landscape.

Friendster—In 2003, Friendster hit the Internet and blew up. It quickly gained worldwide media attention and was featured in magazines such as Spin and Time.

Livejournal—Although Livejournal was created before Friendster, it started gaining popularity around the same time. Kids everywhere got to journal their lives and deepest emotions for everyone to see. But at this point, the whole social networking thing remained pretty much underground.

MySpace—Friendster and Livejournal didn't enjoy success for long. Enter MySpace. MySpace took Friendster's formula, combined it with the blogging of Livejournal, and quickly dominated the market. By 2007, MySpace was the undisputed champion. MySpace also became the go-to

[32] Social Network Sites: Definition, History, and Scholarship, Nicole Ellison

method for bands to get their music out to the masses. This tool became useful to small and big bands alike.

Facebook– Facebook started out as a social networking website for college kids. In fact, you had to enter the name of the college you attended to even sign up. So at the onset, MySpace was killing Facebook. However, FaceBook eventually decided to go public and make the site available to everyone, while adding new features. And by doing so, Facebook has successfully thrust social networking into the mainstream. Kids, adults, seniors, corporations—everyone has a Facebook account.

Twitter– Twitter started getting big around the time Facebook took over. And while Facebook is in the lead, Twitter fulfills a different niche. Twitter creators capitalized on the same notion as fast food providers. People want something quick and easy. So they limited "tweets" to a small word count and now it's one of the best ways to share news with the world. Athletes tweet from games. Reporters tweet breaking news before TV and newspapers can pick it up[33].

III. SOCIAL NETWORK ANALYSIS

III.1 Overview

Social network analysis [SNA] is the mapping and measuring of relationships and flows between people, groups, organizations, computers, URLs, and other connected information/knowledge entities. The nodes in the network are the people and groups while the links show relationships or flows between the nodes. SNA provides both a visual and a mathematical analysis of human relationships.

Social Network Analysis is an approach to analysing organizations focusing on a network-based view of the relationships between people and/or groups as the most important aspect. Going back to the 1950's, it is characterised by adopting mathematical techniques especially from graph theory. It has applications in organizational psychology, sociology and anthropology. Social Network Analysis provides an avenue for analysing and comparing formal and informal information flows in an organization, as well as comparing information flows with officially defined work processes.

The first goal of Social Network Analysis is to visualise relationships between people and/or groups by means of diagrams. The second goal is to study the factors which influence relationships (for example the age, cultural background, and previous training of the people involved) and also to study the correlations between relationships. The third goal is to draw out implications of the relational data, including bottlenecks where multiple information flows funnel through one person or section (slowing down work processes), situations where

[33] http://www.techvert.com/history-social-networking-sites

information flows does not match formal group structure, and individuals who carry out key roles that may not be formally recognised by the organization. The fourth and most important goal of Social Network Analysis is to make recommendations to improve communication and workflow in an organization[34].

III.2 Social network analysis software

Social network analysis software is used to identify, represent, analyze, visualize, or simulate nodes (e.g. agents, organizations, or knowledge) and edges (relationships) from various types of input data (relational and non-relational), including mathematical models of social networks. The output data can be saved in external files. Various input and output file formats exist.

Network analysis tools allow researchers to investigate representations of networks of different size - from small (families, project teams) to very large (the Internet, disease transmission). The various tools provide mathematical and statistical routines that can be applied to the network model.

Visual representations of social networks are important to understand network data and convey the result of the analysis. Visualization is often used as an additional or standalone data analysis method.

Social network tools are:

- for business oriented social network tools: iPoint, NetMiner, InFlow, Keyhubs, Sentinel Visualizer, KXEN Social Network, NodeXL.;
- For large networks with millions of nodes: Sonamine or ORA;
- For mobile telecoms Idiro SNA Plus is recommended;
- An open source package with GUI for Linux, Windows and Mac, is Social Networks Visualizer or SocNetV, developed in Qt/C++;
- Another generic open source package for Windows, Linux and OS X with interfaces to Python and R is "igraph", "Tulip";
- Another generic open source package with [GUI] for Windows, Linux and OS X is RapidNet is a generic freely available open source solution for network analysis and interactive visual network exploration and drill-down;
- For Mac OS X a related package installer of SocNetV is available[35].

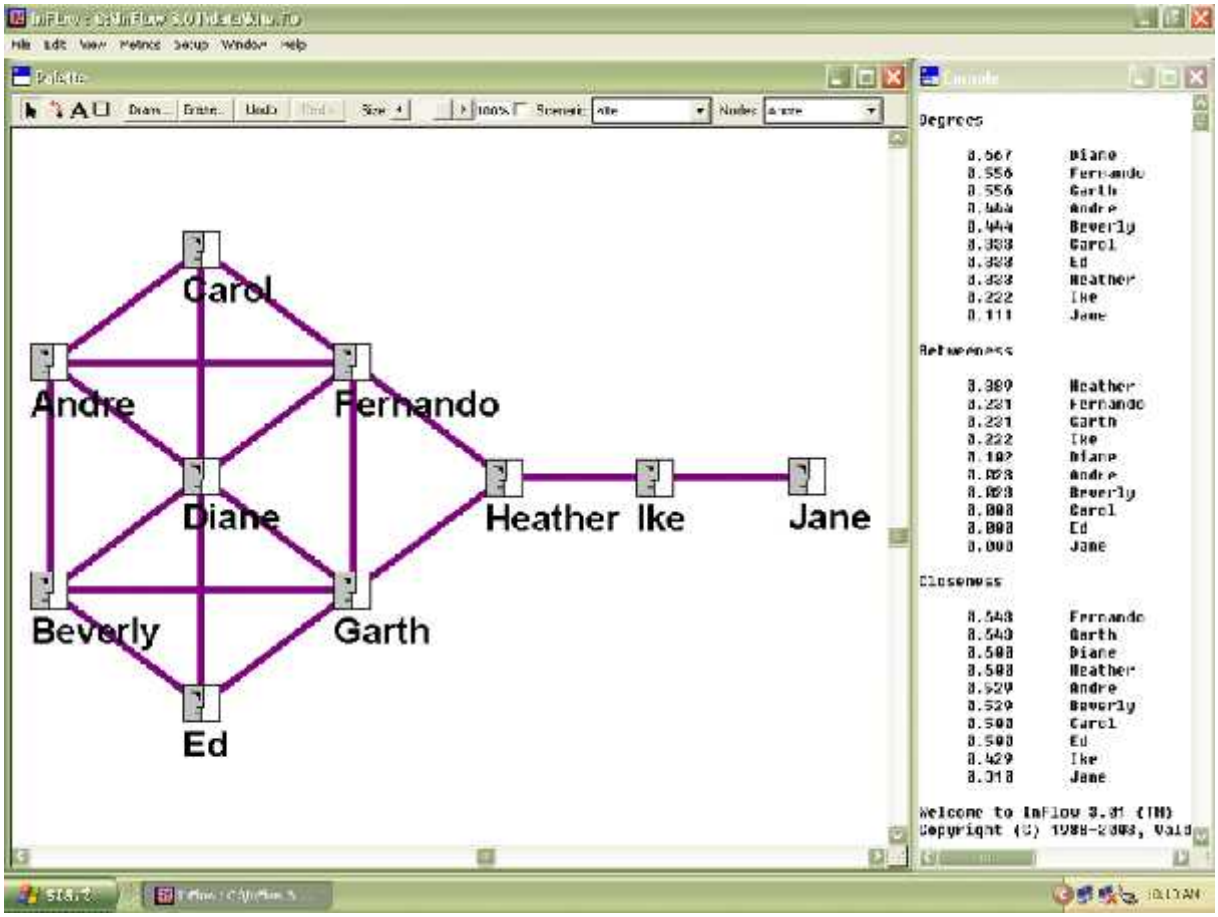
To understand networks and their participants, we evaluate the location of actors in the network. Measuring the network location is finding the centrality of a node. These measures give us insight into the various roles and groupings in a network -- who are the connectors, mavens,

[34] Applying Social Network Analysis Concepts to Military C4ISR Architectures, Anthony Dekker

[35] http://en.wikipedia.org/wiki/Social_network_analysis_software

leaders, bridges, isolates, where are the clusters and who is in them, who is in the core of the network, and who is on the periphery?

Next, we present how to use such software and its benefits, using InFlow:



We look at a social network - the "Kite Network" above - developed by David Krackhardt, a leading researcher in social networks. Two nodes are connected if they regularly talk to each other, or interact in some way. Andre regularly interacts with Carol, but not with Ike. Therefore Andre and Carol are connected, but there is no link drawn between Andre and Ike. This network effectively shows the distinction between the three most popular individual centrality measures: degree centrality, betweenness centrality, and closeness centrality.

Degree centrality - Social network researchers measure network activity for a node by using the concept of degrees - the number of direct connections a node has. In the kite network above, Diane has the most direct connections in the network, making hers the most active node in the network. She is a 'connector' or 'hub' in this network. Common wisdom in personal networks is "the more connections, the better." This is not always so. What really matters is where those connections lead to - and how they connect the otherwise unconnected! Here Diane has connections only to others in her immediate cluster - her clique. She connects only those who are already connected to each other.

Betweenness centrality - While Diane has many direct ties, Heather has few direct connections -- fewer than the average in the network. Yet, in may ways, she has one of the best locations in the network -- she is between two important constituencies. She plays a 'broker' role in the network. The good news is that she plays a powerful role in the network, the bad news is that she is a single point of failure. Without her, Ike and Jane would be cut off from information and knowledge in Diane's cluster. A node with high betweenness has great influence over what flows - and does not - in the network. Heather may control the outcomes in a network.

Closeness centrality - Fernando and Garth have fewer connections than Diane, yet the pattern of their direct and indirect ties allow them to access all the nodes in the network more quickly than anyone else. They have the shortest paths to all others - they are close to everyone else. They are in an excellent position to monitor the information flow in the network - they have the best visibility into what is happening in the network.

Network centralization - Individual network centralities provide insight into the individual's location in the network. The relationship between the centralities of all nodes can reveal much about the overall network structure.

A very centralized network is dominated by one or a few very central nodes. If these nodes are removed or damaged, the network quickly fragments into unconnected sub-networks. A highly central node can become a single point of failure. A network centralized around a well connected hub can fail abruptly if that hub is disabled or removed. Hubs are nodes with high degree and betweeness centrality.

A less centralized network has no single points of failure. It is resilient in the face of many intentional attacks or random failures - many nodes or links can fail while allowing the remaining nodes to still reach each other over other network paths. Networks of low centralization fail gracefully.

Network reach - Not all network paths are created equal. More and more research shows that the shorter paths in the network are more important. Noah Friedkin, Ron Burt and other researchers have shown that networks have horizons over which we cannot see, nor influence. They propose that the key paths in networks are 1 and 2 steps and on rare occasions, three steps. The "small world" in which we live is not one of "six degrees of separation" but of direct and indirect connections < 3 steps away. Therefore, it is important to know: who is in your network neighborhood? Who are you aware of, and who can you reach?

In the network above, who is the only person that can reach everyone else in two steps or less?

Boundary spanners - Nodes that connect their group to others usually end up with high network metrics. Boundary spanners such as Fernando, Garth, and Heather are more central in the overall network than their immediate neighbors whose connections are only local, within their

immediate cluster. You can be a boundary spanner via your bridging connections to other clusters or via your concurrent membership in overlappping groups.

Boundary spanners are well-positioned to be innovators, since they have access to ideas and information flowing in other clusters. They are in a position to combine different ideas and knowledge, found in various places, into new products and services.

Peripheral players - Most people would view the nodes on the periphery of a network as not being very important. In fact, Ike and Jane receive very low centrality scores for this network. Since individuals' networks overlap, peripheral nodes are connected to networks that are not currently mapped. Ike and Jane may be contractors or vendors that have their own network outside of the company - making them very important resources for fresh information not available inside the company.[36]

IV. Consequences of using social networks

IV.1. In the field of information security

When you share information online, you need to understand the potential risks, and you need to be wary of what you share and with whom. Attackers may use social networking services to spread malicious code, compromise users' computers, or access personal information about a user's identity, location, contact information, and personal or professional relationships. You may also unintentionally reveal information to unauthorized individuals by performing certain actions.

IV.1.1. Social network threats

The following are some common threats to social networking services:

- Viruses – The popularity of social networking services makes them ideal targets for attackers who want to have the most impact with the least effort. By creating a virus and embedding it in a website or a third-party application, an attacker can potentially infect millions of computers just by relying on users to share the malicious links with their contacts.
- Tools – Attackers may use tools that allow them to take control of a user's account. The attacker could then access the user's private data and the data for any contacts that share their information with that user. An attacker with access to an account could also pose as that user and post malicious content.
- Social engineering attacks – Attackers may send an email or post a comment that appears to originate from a trusted social networking service or user. The message may contain a malicious URL or a request for personal information. If you follow the instructions, you may disclose sensitive information or compromise the security of your system.

[36] http://www.orgnet.com/sna.html

- Identity theft – Attackers may be able to gather enough personal information from social networking services to assume your identity or the identity of one of your contacts. Even a few personal details may provide attackers with enough information to guess answers to security or password reminder questions for email, credit card, or bank accounts.
- Third-party applications – Some social networking services may allow you to add third-party applications, including games and quizzes, that provide additional functionality. Be careful using these applications—even if an application does not contain malicious code, it might access information in your profile without your knowledge. This information could then be used in a variety of ways, such as tailoring advertisements, performing market research, sending spam email, or accessing your contacts[37].
- Data leaks - Social networks are all about sharing. Unfortunately, many users share a bit too much about the organization -- projects, products, financials, organizational changes, scandals, or other sensitive information. Even spouses sometimes over-share how much their significant other is working late on top-secret project, and a few too many of the details associated with said project. The resulting issues include the embarrassing, the damaging and the legal.

You may risk professional opportunities, personal relationships, and safety by posting certain types of information on social networking services.

IV.1.2 Solutions

Social networking services are useful and enjoyable, but it is important to take proactive steps to protect your computer, your personal information, and your company data. By protecting yourself, you also help to protect the people you are connected to on these services.

Taking general security precautions will reduce the risk of compromise.

1. Use strong passwords8, and use a unique password for each service.
2. Keep anti-virus software9 up to date.
3. Install software updates10 in a timely manner, particularly updates that affect web browsers.

Social networking services offer unique risks, and you can minimize these risks by adopting good security practices.

1. **Use strong privacy and security settings** – Take advantage of the security options provided by social networking services. When choosing appropriate options, erron the side of privacy to better protect your information. These services may change their options periodically, so regularly evaluate your security and privacy settings, looking for changes

[37] http://www.us-cert.gov/reading_room/safe_social_networking.pdf

and ensuring that your selections are still appropriate. Also periodically review the services' privacy policies to see if there are any changes.

2. **Avoid suspicious third-party applications** – Choose third-party applications wisely. Look for applications developed by vendors you trust, and avoid applications that seem suspicious. Limit the amount of information third-party applications can access.

3. **Treat everything as public** – The best way to protect yourself is to limit the amount of personal information you post to these services. This recommendation applies not only to information in your user profile, but also to any comments or photos you post. It is important that you consider information that you post about yourself and about others, particularly children.

4. **Share only with people you know** – Although many users seek to establish as many contacts on these services as possible, consider sharing personal information only with people you know. If you expand your contacts beyond people you are sure you can trust, check the service's settings to see if you can group your contacts and assign different levels of access based on your comfort level. Attackers may adopt different identities to try to convince users to add them as contacts, so try to confirm that contacts are who they claim to be before giving them access to your information.

Regardless of how restrictive you make your security settings, they may not offer complete privacy. An attacker or application may take advantage of software vulnerabilities, or another user may repost your information. When using social networking services, be responsible and always consider the risks. Operate as if all of the content is public, and only post information you would be comfortable sharing with other people.

IV.2. In the field of business

Social and business networking sites are changing the way people communicate with each other, both for business and pleasure. Some might think it makes sense for organizations to simply block employees' access to these sites while at work, citing cyber-slacking as the reason, but it isn't that straight forward. These sites provide employees and the organizations they work for with a very real business advantage. Some of the benefits of allowing employees to access social and business networking sites while at work are outlined below.

IV.2.1 Benefits of social networks in business

a) Networking, Collaboration and Information Sharing

Social networking sites can be very effective for business networking. Almost like an informal CRM system, people can use social networking sites such as LinkedIn to maintain business contacts and to introduce colleagues or contacts to one another in an informal manner. There are

also some less well known social networking sites that have been set up specifically to encourage information sharing and collaboration between professionals operating in a particular industry. Sermo, for example, is a site exclusively for US physicians. It has teamed up with the pharmaceutical giant Pfizer to allow doctors direct online access to employees from the drug company, which encourages feedback and ongoing communication.

These specialized social networking sites offer much more than just an unmoderated free-for-all. Given a few restrictions in terms of membership and content, a social network can provide a valuable, easy-to-use forum for academic debate or business discussion. The use of restricted groups in Facebook is another good example of this: companies can set up a private area and use it to share ideas in an informal environment that encourages creativity.

b) Marketing

Social networking sites have also opened up new marketing and promotional opportunities for businesses. Companies can pay for banner ads on the sites themselves and can also create their own home pages. Appealing to the tech-savvy, less formal, Web 2.0 generation who have become used to hearing about the latest bands on MySpace, social networking sites have become a valuable, low cost marketing tool, particularly for consumer-facing organizations. Publishing corporate blogs on social networking sites can also be a very effective way of sharing information and strengthening brand image.

c) The MySpace Generation

But it goes further than that. They also play a very important part in the lifestyle of anyone under 30: accessing social networking sites is as important to these younger employees as using their mobile phone. Preventing these employees from using all the technology tools they take for granted will only lead to disgruntled, unhappy workers. By contrast, giving them the freedom – albeit regulated - to use these social networking sites in the workplace can help both employees, and the organizations they work for, to flourish.

It is therefore important to get the balance right: allow employees to use these sites, but ensure that they do so without subjecting themselves or the organization to undue risk. Most employees will have the common sense to use these networks to socialize and do business without compromising security, but it only needs one employee to use a social networking site unwisely for the repercussions to be significant[38].

IV.2.2 Threats of social networks in business

From a purely technical perspective, social networking is simply another example of employees accessing websites while at work. However, social networks do present specific challenges for

[38] www.zdnet.co.uk/i/s/ads/.../WhitePaper_SocialNetworking.pdf

employers due to the type of content published on these sites. Some of the key threats that organizations need to guard against are discussed below.

a) Viruses/Malware

Criminal gangs target social networking sites because they offer an effective way of propagating malware to a wide, unsuspecting audience. The MySpace trojan (2006), the Orkut worm (2007) and the Secret Crush Facebook widget (2008) are examples of how criminal gangs can use social networking sites to their advantage. For the 'bot-herders', who can charge based on the size of their botnets, social networking sites provide an easy way to play a percentage game: with so many users, they know they can rely on some of them to become victims. And if the user is accessing the social networking site from a work PC, then the organization's whole network risks being compromised. This is especially the case when people believe they are receiving something from a friend and hence their defences are automatically lowered.

b) Privacy

It is very easy for people to get carried away and post too much information about themselves on social networking sites. This can lead to identity theft or phishing attacks and helps to promote cybercrime. There have also been several instances where employers or prospective employers have used information posted on these sites in evaluating employees. Many sites, such as Facebook, recommend that users do not post sensitive information on these sites and that they apply the necessary security measures to prevent their personal home pages from being viewed illicitly. That said, there have also been some concerns over what social networks do with the information that they are privy to. Towards the end of 2007, social networking site Quechup came under a lot of fire for using its members' address books to send out spam to try and swell its ranks.

c) Cyberbullying/Cyberstalking

Similarly, employees using these sites are putting themselves at risk of becoming victims of cyberbullying or cyberstalking. A survey carried out by the trade union Amicus, reported that one fifth of employees in the UK were being bullied electronically. Whilst cyberbullying includes emails, it also extends to social networking sites; the overall effect can be seriously detrimental to morale within an organisation. Amicus estimates that bullying costs the UK economy over £2 billion per annum in sick pay, staff turnover and productivity. Often, a cybervictim's only recourse is to secure or remove his profile from the offending site.

d) Data Leakage

It is very easy for an employee to post confidential information about their company – be it unwittingly or deliberately – in a blog or on a social networking site. Whether it is the product road map, confidential financial information or even just derogatory comments about

management, data leakage can lead to internal reprimands or worse: litigation, fines or even imprisonment of company officials may occur as the result of poor data control.

There are other considerations too: whilst LinkedIn and other such sites can be used advantageously as a cheap and simple CRM system, they are usually attached to an individual rather than a company, so the data becomes very portable. It would not be difficult for an employee to take the entire sales database with him to a rival after having built up an extended network of friends/business colleagues through a social networking site.

e) Brand Credibility

Warren Buffett said that it takes twenty years to build a reputation and five minutes to ruin it. When an organisation tries to use a social networking site to its advantage, it needs to be careful. Six major companies seeking to benefit from advertising through Facebook found their banner ads appearing on the neo-fascist British National Party‘s pages. They all pulled their advertisements, one of them publicly declaring that it was doing so to “protect its brand.”

f) Lost Productivity

Social networking sites can become addictive, so much so that it is relatively easy to spend two or three hours of the working day socialising online instead of working. Recent surveys indicate that 43 percent of organizations in the UK have banned the use of social networking sites at work completely, for productivity and security reasons. Indeed, in August 2007, Kent County Council banned all of its 32,000 employees from using Facebook, citing ‘time-wasting’ as the principle reason. This was shortly after the ‘I have dossed around on Facebook all day and consequently have done no work’ group had been set up.

So where does all this leave organizations that are concerned about the use of social networking sites in the workplace? The answer is that it doesn’t have to be that black and white. The technology is available – in the form of secure Web gateways – to allow employees to use social networking sites safely and securely. A secure Web gateway, such as WebMarshal, combines advanced Web access controls, data leakage prevention and inbound threat controls in one centrally managed solution or service that makes accessing social networking sites a low-risk, high-reward option for organizations.

There are both technical challenges and personal use issues that to be addressed: organizations have to determine their own modus operandi, identify and deploy the appropriate underlying technology solution and then communicate to employees how social networks can be used in accordance with their Acceptable Use Policy[39].

[39] www.zdnet.co.uk/i/s/ads/.../WhitePaper_SocialNetworking.pdf

CONCLUSIONS

Social networking services are useful and enjoyable, but it is important to take proactive steps to protect your computer, your personal information, and your company data. By protecting yourself, you also help to protect the people you are connected to on these services.

Taking general security precautions will reduce the risk of compromise.

Social networking services offer both, unique and specific risks of any computer network and you can minimize these risks by adopting good security practices.

In my opinion, we can mitigate these risks through a rigorous education of users, because the users are the ones who determine how they use these online social networks. They decide what to do with the information that they have access; they use it in order to build term relationships or not.

Also, we can say that online social networks are relatively new to business, the MySpace and Facebook generation has grown up with them. For these individuals entering the workforce, online social networking is simply be a part of the fabric of business. Accordingly, the organizations that have determined how to best integrate social networking into their operations will be the ones that are most successful.

REFERENCES

1. Social network analysis, Oliver Serrat
2. Social Network Sites: Definition, History, and Scholarship, Nicole Ellison
3. Applying Social Network Analysis Concepts to Military C4ISR Architectures, Anthony Dekker
4. www.cerado.com
5. http://en.wikipedia.org/wiki/Social_(disambiguation)
6. http://en.wikipedia.org/wiki/Social_network
7. http://en.wikipedia.org/wiki/Social_network
8. http://www.techvert.com/history-social-networking-sites
9. http://en.wikipedia.org/wiki/Social_network_analysis_software
10. http://www.orgnet.com/sna.html
11. http://www.us-cert.gov/reading_room/safe_social_networking.pdf
12. www.zdnet.co.uk/i/s/ads/.../WhitePaper_SocialNetworking.pdf
13. www.zdnet.co.uk/i/s/ads/.../WhitePaper_SocialNetworking.pdf

CYBER THREATS TO MOBILE DEVICE

LTC Iulian CIAUŞU

ABSTRACT

Today's advanced mobile devices are well integrated with the Internet and have far more functionality than mobile phones of the past. They are increasingly used in the same way as personal computers (PCs), potentially making them susceptible to similar threats affecting PCs connected to the Internet. Since mobile devices can contain vast amounts of sensitive and personal information, they are attractive targets that provide unique opportunities for criminals intent on exploiting them. Both individuals and society as a whole can suffer serious consequences if these devices are compromised. This paper introduces emerging threats likely to have a significant impact on mobile devices and their users.

INTRODUCTION

Mobile devices are becoming more and more similar to desktop computers although their computational and storage capacities remain smaller. Mobile devices normally stay connected online all the time because of their default characteristics and user behavior. As a consequence of the integration of mobile networks into the Internet, security treats on one network will affect the other network.

As mobile device technology evolves, consumers are using it at unprecedented levels. Mobile cellular technology has been the most rapidly adopted technology in history, with an estimated 5 billion mobile cellular subscriptions globally at the end of 2010. Furthermore, technological advances have fueled an unprecedented portable computing capability, increasing user dependence on mobile devices and skyrocketing mobile broadband subscriptions. Mobile devices have become an integral part of society and, for some, an essential tool. However, the complex design and enhanced functionality of these devices introduce additional vulnerabilities. These vulnerabilities, coupled with the expanding market share, make mobile technology an attractive, viable, and rewarding target for those interested in exploiting it.

In the past, malicious activity targeting mobile phones was relatively limited compared to that of PCs. The proprietary nature and limited functionality of the hardware and software architectures previously used by individual mobile phone manufacturers made this market a less than ideal target for mass exploitation. Current mobile devices have much greater functionality and more accessible architectures, resulting in an increase in malicious activity affecting them. These

smartphones include the Apple iPhone, Google Android, Research in Motion (RIM) Blackberry, Symbian, and Windows Mobile-based devices.

Due to the similar functionality of mobile devices and PCs, the distinction between the two has blurred. Mobile devices have become equally susceptible to malicious cyber activity and will likely be affected by many of the same threats that exist for PCs on the Internet. The variety of sensitive information available from a mobile device is also potentially greater and more enticing than that of a traditional mobile phone or computer. Users are more likely to take advantage of the portability and convenience of mobile devices for activities such as banking, social networking, emailing, and maintaining calendars and contacts. The features of mobile devices also introduce additional types of information not typically available from a PC, such as information related to global positioning system (GPS) functionality and text messaging.

A multitude of threats exist for mobile devices, and the list will continue to grow as new vulnerabilities draw the attention of malicious actors. This paper provides a brief overview of mobile device malware and provides information on the following threats to mobile devices: social engineering, exploitation of social networking, mobile botnets, exploitation of mobile applications and exploitation of m-commerce.

I. COMING CYBER THREATS TO TARGET MOBILE DEVICES

The biggest cyber threats in 2011 are expected to include, among other new risks, malicious applications on mobile devices and attacks aimed at stealing government secrets and sabotaging business operations, according to McAfee.

The computer security firm annually issues a list predicting what will be the biggest cyber scares during the coming year. New for 2011 is the projection that perpetrators will target social media communications on mobile devices - a means of interaction that businesses, including agencies, increasingly depend on for work.

The societal shift from desk-based e-mail communications to mobile instant messaging and Twitter insta-blogging has transformed the threat landscape, according to the report.

The specialists employed by McAfee Labs, the firm's research arm, expect to see apps - online tools for mobile devices - expose privacy and identity data. These tools have historically weak coding and security practices, and will allow cybercriminals to manipulate a variety of physical devices through compromised or controlled apps.

McAfee Labs anticipates that attackers will hide malicious software in programs that look like legitimate applications, including federal data apps, Dmitri Alperovitch, McAfee's vice president

for threat research, said in an interview.[40] According to the threat list, "friendly fire" malware, which appears to come from contacts on social networks, will grow.

"Social media connections will eventually replace e-mail as the primary vector for distributing malicious code and links. The massive amount of personal information online coupled with the lack of user knowledge of how to secure this data will make it far easier for cybercriminals to engage in identity theft and user profiling than ever before."

For example, phishing - traditionally scam e-mails that appear to come from your bank or from Nigerians - will move to Twitter because e-mail is no longer vulnerable, Alperovitch said. "E-mail is a fairly well-protected channel these days, and people are starting to finally get the message that if that they get an e-mail that looks too good to be true ... it potentially needs to be reported."

The transition to mobile communications also creates an easy opportunity for fraud purveyors to pinpoint the location of potential victims. More Internet users are logging on to the Web via portable devices with Global Positioning System satellite technology. Many GPS tools essentially broadcast people's coordinates to friends and colleagues so they can see where they are.

"You can easily search, track and plot the whereabouts of friends and strangers," the report stated. "In just a few clicks cybercriminals can see in real time who is Tweeting and where, what they are saying, what their interests are, and the operating systems and applications they are using."

In 2011, shortened Web addresses -- ideal for inserting website locations in word-constrained mobile messages and Tweets -- will become ideal for masking fake websites, the researchers noted. "The trouble -- and abuse -- follows because users do not know where these shortened links actually lead until they click them."

Alperovitch said malware distributors and phishers will start using these abbreviated Web addresses, or short URLs, to bypass the Web-filtering tools in offices.

But information technology managers cannot prohibit federal officials from conducting business via mobile devices, as President Obama demonstrated when he refused to part with his BlackBerry upon taking office. "The real answer is not to fight these things because they will get in," Alperovitch said. "The key is to make sure they are secure."

The motivation of attackers also is changing, according to the study. Instead of carrying out attacks to steal money or to send a political message, some groups, including nation-states and corporations, increasingly are interested in stealing intelligence.

[40] Aliya Sternstein - http://www.nextgov.com/nextgov/ng_20101228_6846.php

McAfee defines these new so-called advanced persistent threats as government or organization-sponsored attempts at cyberespionage or cybersabotage for something other than political protest, or financial gain.

Work mobile devices will become breeding grounds for APTs, Alperovitch said. "Those are essentially full-blown computers now -- and they are connected to the network," he added. "Companies of all sizes that have any involvement in national security or major global economic activities should expect to come under pervasive and continuous APT attacks that go after e-mail archives, document stores, intellectual property repositories and other databases."

Other 2011 predictions detailed in the report: Cybercriminals will target more Apple-manufactured technologies; botnets -- compromised computers that hackers hijack all at once to send viruses -- will filch data from breached computers instead of sending spam; and "hacktivism" attacks, intended to discredit political opponents, will intensify.

"The popularity of iPads and iPhones in business environments and the easy portability of malicious code between them could put many users and businesses at risk next year and beyond" adding botnets will be a common occurrence on Apple platforms in 2011.

More activists will mimic the WikiLeaks model of harming companies and individuals by manipulating their online operations, as sympathizers of the document-leaking site did by knocking MasterCard services offline. The company stopped processing payments for WikiLeaks because the site violated MasterCard's acceptable use policies.

"Hacktivism as a diversion could be the first step in cyberwarfare", where governments secretly arm grass-roots groups with sophisticated cyber weapons. That hacktivism initiated by nongovernmental organizations serves as a good cover for government-sponsored cyberwar. It grants nation-states plausible deniability.

Everyone within information security will have to be vigilant to recognize the difference between hacktivism and the beginning of a cyberwar. As in the physical world, we expect that hacktivist attacks will inspire and foment riots and other real-world demonstrations.

II. MOBILE MALWARE

1. MOBILE CYBER THREATS ARE GETTING WORSE

Malware of all kinds keeps spreading on computing platforms. But mobile malware grew at a particularly fast clip in 2010, according to McAfee.

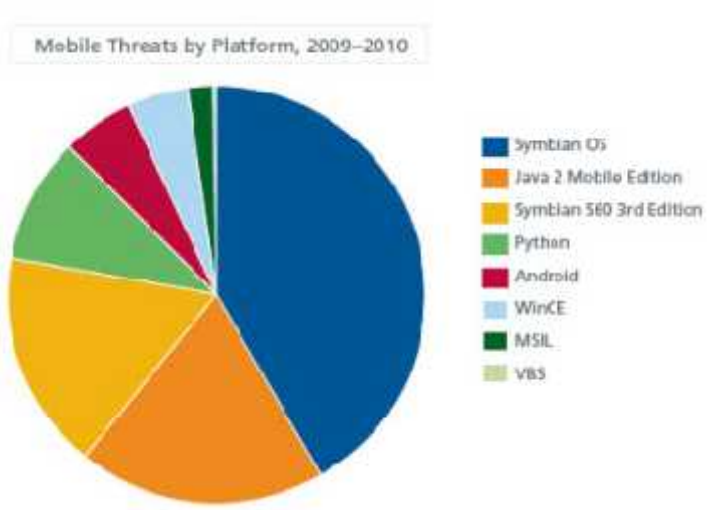
Mobile malware was up 46 percent in 2010 to 967 threats, compared to 704 in 2009, according to the McAfee Threats Report for the fourth quarter.

That’s small compared to PC threats, but the trend is clear. The renewed interest in attacking mobile platforms comes as smartphones and tablets become a primary computing tool for

millions of users. If this trend continues, mobile security may begin to consume as much resources as PC security, which accounts for billions of dollars in investment.

Altogether McAfee said there were 20 million new pieces of malware in 2010, equating to nearly 55,000 new malware threats each day. That's because cyber criminals are able to automate the creation of new variants of malware. To date, McAfee has identified 55 million pieces of malware, and 360 percent of those were created in 2010.

Spam saw a surprising decline recently. Spam accounted for 80 percent of total email traffic in the fourth quarter. That was the lowest level since the first quarter of 2009. But McAfee said the decline is only because several large spam botnets (or herds of compromised computers that are controlled by cybercriminals) were taken down and spammers are moving to new botnets.



Vincent Weafer, senior vice president of McAfee Labs, said there is a direct correlation between the popularity of a device and attacks against the device. One of the most high-profile threats was SymbOS/Zitmo.A, which attacked phones with the Symbian operating system, which is still the most popular mobile platform despite Nokia's significant loss of market share. Another prominent threat was Android/Geinimi, which hid a Trojan in legitimate Android mobile apps.

McAfee said that attacks against mobile platforms were also successful because so many mobile users aren't aware of mobile security threats. People who believe in putting antivirus software on a PC don't think about doing that for mobile phones.

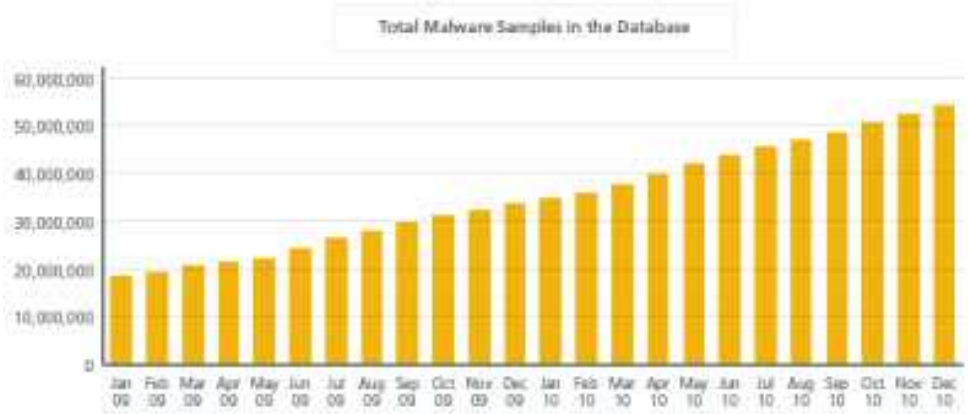


Figure 4: Total count of unique malware (including variants) in the McAfee Labs database.

Overall, malware keeps spreading. McAfee found that, within the top 100 results of the top daily search terms, 51 percent of the results led to malicious sites. McAfee said that Adobe product vulnerabilities have turned software such as Adobe's PDFs into prime carriers of malware. McAfee said that trend would continue, as mobile devices support various Adobe technologies.

2. MALICIOUS ACTIVITY

Malicious actors have created and used malware targeted to mobile devices since at least 2000. The total number of malware variants significantly increased in 2004 with the public release of Cabir source code[41]. Cabir is a Bluetooth worm and the first widespread sample of mobile malware. It runs on mobile phones using the Symbian Series 60 platform and spreads among Bluetooth-enabled devices that are in discoverable mode. The worm causes a phone to constantly attempt to make a Bluetooth connection, subsequently draining the battery. While this worm was an inconvenience to device users, today's mobile malware is more insidious and often has more severe effects on devices and their users.

A recent and more nefarious example of mobile malware is the Ikee.B, the first iPhone worm created with distinct financial motivation. It searches for and forwards financially sensitive information stored on iPhones and attempts to coordinate the infected iPhones via a botnet command and control server.[42] This worm only infects iPhones that have a secure shell (SSH) application installed to allow remote access to the device, have the root password configured as "alpine"—the factory default—and are "jailbroken." A jailbroken iPhone is one that has been configured to allow users to install applications that are not officially distributed by Apple. Although Ikee.B has limited growth potential, it provides a proof of concept that hackers can migrate the functionality typical to PC-based botnets to mobile devices. For example, a victim iPhone in Australia can be hacked from another iPhone located in Hungary and forced to exfiltrate its user's private data to a Lithuanian command and control server.

Spy software also exists for mobile devices, including some programs being sold as legitimate consumer products. FlexiSpy is commercial spyware sold for up to $349.00 per year. Versions are available that work on most of the major smartphones, including Blackberry, Windows Mobile, iPhone, and Symbian-based devices. The following are some of the capabilities provided by the software[43]:

- Listen to actual phone calls as they happen;
- Secretly read Short Message Service (SMS) texts, call logs, and emails;

[41] Ken Dunham, et al. Mobile Malware Attacks and Defense. 2009. Burlington, MA: Syngress Publishing, Inc.
[42] F-Secure. Worm:iPhoneOS/Ikee.B. 2009. Retrieved February 16, 2010 from http://www.f-secure.com/v-descs/worm_iphoneos_ikee_b.shtml.
[43] Flexispy Ltd. FlexiSpy Homepage. 2010. Retrieved February 17, 2010 from http://flexispy.com/.

- Listen to the phone surroundings (use as remote bugging device);
- View phone GPS location;
- Forward all email events to another inbox;
- Remotely control all phone functions via SMS;
- Accept or reject communication based on predetermined lists; and
- Evade detection during operation.

FlexiSpy claims to help protect children and catch cheating spouses, but the implications of this type of software are far more serious. Imagine a stranger listening to every conversation, viewing every email and text message sent and received, or tracking an individual's every movement without his or her knowledge. FlexiSpy requires physical access to a target phone for installation; however, these same capabilities could be maliciously exploited by malware unknowingly installed by a mobile user.

Cross-platform mobile malware further complicates the issue. The Cardtrp worm infects mobile devices running the Symbian 60 operating system and spreads via Bluetooth and Multimedia Messaging Service (MMS) messages. If the phone has a memory card, Cardtrp drops the Win32 PC virus known as Wukill onto the card.[44] Two proof-of-concept Trojans, Crossover and Redbrowser, further show how widespread attacks could simultaneously hit desktops and mobile devices.[45] Both Trojans can infect certain mobile devices from PCs.

SMS, MMS, Bluetooth, and the synchronization between computers and mobile devices are all examples of potential attack vectors that extend the capabilities of malicious actors. Inherent vulnerabilities exist in modern mobile device operating systems that are similar to those of PCs and may provide additional exploitation opportunities. For example, the most recent Apple security update for iPhone OS 3.1.3 provided fixes for scenarios where playing a maliciously crafted mp4 audio file, viewing a maliciously crafted Tagged Image File Format (TIFF) image, or accessing a maliciously crafted File Transfer Protocol (FTP) server could result in arbitrary code execution. To help mitigate malicious activity affecting known vulnerabilities, users should install security patches and software updates as they become available.

3. MOBILE MALWARES - A BIG THREAT FOR SMARTPHONE USERS

Carrying a smart cell phone is more or less like carrying a powerful handy computer into your pocket. A rapid increase in the number of phones selling everyday across the globe gave the boost to the fascinating handy gadgets and more over these handsets include not only a good

[44] Cyrus Peikari. Analyzing the Crossover Virus: The First PC to Windows Handheld Cross-infector. 2006. Retrieved February 17, 2010 from http://www.informit.com/articles/article.aspx?p=458169&seqNum=3.
[45] Bill Brenner. Proof-of-concepts heighten mobile malware fears. 2006. Retrieved February 17, 2010 from http://searchexchange.techtarget.com/news/article/0,,sid43_gci1171168,00.html.

camera but also they have extensive online access, qwerty keyboards, and several other typical computer functions. Although with power and expediency comes with a cost. Just like our conventional desktops and laptops these smart phones face high security threats. The irony of this advancement is that the greater the functionality it gains, the more they become vulnerable to the unavoidable threats which corrupts our laptops and desktops.[46]

What so far, Mobile Malware has been one of the most awful nightmares of millions of people worldwide these days. Now the year is about to get over, another more year is on its threshold. Is 2011 going to be a large piece of cake for the Mobile Malware developers? It is a big question of concern whether the mobile software developers will be able to redeem this issue to a significant degree or not. It is not really possible to predict anything based on any view because it is hard to specify the exact source of certain significant Mobile Malware.

The standard modus operandi of these malwares has been in the state of drafted applications and also the traditional email structure. Often you can even log into your email account through your desktop, and surprisingly the email that would redirect you to some external file download link! And here it comes! The next thing is your system is into the trap of a slew of viruses or the background key logger program is sending your important information to some nomad hacker.

It is known to everyone that the Mobile Malwares are not at all the treat issue for only smartphones because there are laptops, notebooks and handheld gadgets (tablets and PDAs). Notebooks are as much vulnerable as the other ones. Instead of all these treat issues smartphones are the instant way of getting in touch with the internet and also a dynamic way of being an easy target for the Mobile Malware app developers in recent times.

Microsoft by default applies a certificate system to protect the Windows Mobile API. Only the program threads with signed certificates are able to call mobile APIs. This system works good till a user is willing to add an unsigned new program. On the other hand Rogue apps work like a charm. They are uploaded in different smartphone application stores. Thousands of the users end

[46] Kreaty Ferguson - http://www.shaswatpatel.com/mobile-malwares-a-big-threat-for-smartphone-users-in-2011/

up downloading them. Even very recently a Vietnamese hacker uploaded **iTunes** related applications and hacked a lot of bucks from several credit card numbers which were later on utilized for buying legit and costly applications.

4. CROSSOVER VIRUS

Recently, we have seen a rapid evolution of "blended" mobile malware. Much of this activity has been seen on the Symbian Smartphone platform. For example, "Skulls" was the second trojan to infect Symbian Series 60 smart phones (the first was Mosquito). When launched, the application claims to be an "Extended Theme Manager by Tee-222." However, it then disables all other applications on the phone and replaces their icons with a skull and crossbones. Worse, it was more recently merged with Caribe to form the first "crossover" malware for Smartphones. [47]

Skulls and Caribe also merged to form Metal Gear, a trojan that masquerades as the game with the same name. Metal Gear uses Skulls to deactivate the devices' antivirus. Thus, it was the first anti-AV malware for Symbian phones. The malware also drops SEXXXY.sis to the device, an installer that adds code to disable the handset menu button. The Trojan then uses Caribe to transmit itself to new devices

Another example of blending is the Gavno.a Trojan, which is spread via a file called patch.sis (it masquerades as phone patch). Gavno uses a malformed file to crash an internal Symbian process, thus disabling the phone. The effect is to disable all handset buttons and to completely prevent the user from making calls. It may also cause a continual rebooting loop. It is only 2kb in size, and it has already seen variants merged with Caribe to spread to other phones.

Other examples of viral evolution include the following:

- Dampig trojan: Notable in that it corrupts the system uninstallation settings, making it more difficult to remove
- Mabir virus: Similar to Cabir, but instead of Bluetooth it uses SMS to spread
- Commwarrior: also tries to disable the onboard antivirus software
- Frontal virus: causes a total system crash of the phone until it is removed

Lastly, a new Symbian Trojan called Doomboot-A that now loads a Commwarrior variant when it infects Smartphones. Doomboot-A destroys the boot process so that the phone is not useable.

Cross-platform mobile malware

A newer development, and one that may be the most troubling, is the new breed of "cross-platform" mobile infectors. For example, the first mobile phone virus capable of infecting a PC was the Cardtrp worm. Cardtrp infects handsets running the Symbian 60 operating system and

[47] Cyrus Peikari. Analyzing the Crossover Virus: The First PC to Windows Handheld Cross-infector. Retrieved February 17, 2010

spreads via Bluetooth and MMS. If the phone has a memory card, it will drop the Win32 PC virus known as Wukill onto the card.

Conversely, the most recent type of malware does the opposite: it now cross-infects mobile devices from a PC. The first example of such malware, and the subject of this article, is a Trojan dubbed "crossover," which spreads from a Win32 desktop machine to a Windows Mobile Pocket PC handheld.

When executed from Win32, the Trojan checks what version the current OS is; if it is not Windows CE or Windows Mobile, the virus makes a copy of itself and puts a startup command in the registry key of local-machine-current-version-run. The trojan then quietly waits for an ActiveSync connection to be detected; it can wait indefinitely. When an ActiveSync connection is detected, the trojan automatically copies itself to the handheld device and remotely executes the trojan. The handheld device is now infected. The Trojan will then begin to delete documents on the handheld.

III. SOCIAL ENGINEERING

One of the more common methods of spreading malware on the Internet is through social engineering. Most malicious activity is often successful because users are deceived into believing it is legitimate. Exploitation by social engineering is extremely lucrative and will likely significantly increase in the mobile market.

Phishing is the criminal act of attempting to manipulate a victim into providing sensitive information by masquerading as a trustworthy entity. This technique is a well-established, significant cyber threat, and mobile devices provide unique opportunities for phishing, including variants such as vishing and smishing.

Vishing is the social engineering approach that leverages voice communication. This technique can be combined with other forms of social engineering that entice a victim to call a certain number and divulge sensitive information. Advanced vishing attacks can take place completely over voice communications by exploiting Voice over Internet Protocol (VoIP) solutions and broadcasting services. VoIP easily allows caller identity (ID) to be spoofed, which can take advantage of the public's misplaced trust in the security of phone services, especially landline services. Landline communication cannot be intercepted without physical access to the line; however, this trait is not beneficial when communicating directly with a malicious actor.

Smishing is a form of social engineering that exploits SMS, or text, messages. Text messages can contain links to such things as webpages, email addresses or phone numbers that when clicked may automatically open a browser window or email message or dial a number. This

integration of email, voice, text message, and web browser functionality increases the likelihood that users will fall victim to engineered malicious activity.

Regardless of the communication medium, users must ensure that any exchange of information occurs between their intended parties. Links contained in suspicious or unsolicited emails and text messages should be avoided, and to help prevent disclosing sensitive information to an unintended party via voice communication, users can initiate the phone call to a known, trusted number.

IV. MOBILE BOTNETS

Botnets represent a serious security threat on the Internet. Current security mechanism are typically inadequate for protecting against the latest breed of botnets, as botherders constantly develop new techniques and methods to frustrate investigators. Until recently, mobile networks have been relatively isolated from the Internet, so there has been little need for protecting against botnets. However, this situation is rapidly changing. Mobile networks are now well integrated with the Internet, so that threats on the Internet most likely will migrate over to the mobile networks and vice versa. Botnets of malware injected into mobile devices will probably appear very soon, and there are already signs of this happening. This chapter analyses the potential threat of botnets based on mobile networks.

1. BOTNET ON THE INTERNET

A botnet is a set of computers that are infected by a specific bot virus which gives an attacker (aka. Botherder) the ability to remotely control those computers.

Most botnets are developed for organized crime where doing targeted attack to gain money. Example of attacks are sending spam, denial of service attack (DOS) or collecting and sale of information that can be exploited for illegal purposes.

Researchers who fight against botnets are in general one step behind the attackers. Once they have found a solution to discover and take down a botnet the botherders change their botnet to fight back.

Botnet starts their lifecycle when a vulnerability in an operating system or software are exploited or users have been fooled to run unwanted software on their computer. Malware is often distributed as spam within a malicious attachment, spam linked to infected websites, open file shares, through instant messaging (IM) or by scanning after vulnerabilities. Compared to older malware they spread much faster on the Internet than through floppy disk or Bluetooth.

After exploiting the desktop computer a secondary infection that installs and updates the botclient appears.

Evolution of botnets has made them more difficult to discover and take down because they hide their communication between the legitimate traffic on the Internet by using TCP port 80 and they use peer to peer (P2P) technology that make them more resilient.

Web based C&C uses pull rather than push technology. Web servers can passively wait for the botclients to make contact. This lowers the network traffic between the botclients and the servers making it harder to spot the botnet on the network. To evade shutdown efforts the C&C needs to be constantly moving by using techniques like multihoming, fast flux and distributed C&C (Superbotnet).

Waledac is a botnet on the Internet that is web based and communicates using XML messages as payload. MMS messages have a body field where the XML message from Waledac can be hidden. Can Waledac infect mobile devices and use MMS and SMS to communicate with its C&C?

2. THREAT OF BOTNETS SPREADING TO MOBILE NETWORKS

Mobile devices are nowadays capable of using Internet connection through High-Speed Downlink Packet Access (HSDPA), Evolution-Data Optimized (EVDO), Universal Mobile Telecommunication System (UMTS), Enhanced Data Rates for GSM Evolution (EDGE) and General Packet Radio Service (GPRS) which are different IP based technologies evolved within the mobile network and wireless network (WLAN).

The next generation of network technology will be mobile broadband where the mobile device will stay online and connected to the Internet all the time. Mobile terminals become more and more like desktop computers. Several studies on botnets and mobile devices predict communicating on the Internet.

The connection between the traditional Internet and the mobile network may act as a gateway for malware to move freely between these networks. Infection vectors used to spread Internet malware are extended using MMS, SMS, Bluetooth and synchronizing between the computer and the mobile device. Malware on mobile devices can move using infections vectors on the Internet as email, web pages and social engineering.

What benefits can be gained using mobile device as a botclient? Are there any economic gains, is there any malicious action to do through the mobile network? Can mobile botnet become as difficult to track and close down as botnets on the Internet?

By having a botclients on mobile devices a botherder will be able to exploit services in the mobile network. By their nature mobile devices will not be available on the Internet all the time. They switch between available communications channels on the Internet or on the mobile network.

Mobile devices using the Internet may stay behind a firewall and normally they get dynamic private IP addresses that are not available on the Internet. The botherder will thus not be able to contact and send commands to the botclient directly through Internet connection. This challenge has already been solved with web based botnets where the botclients connect to web servers and poll updates.

Waledac is a web based botnet on the Internet, very similar to Storm known as the biggest mass emailing botnet on the Internet. Waledac seems to be an updated version of Storm and is based on a three-tiered architecture existing of C&C, proxy servers and botclients. Botclients that respond to HTTP requests act as proxies to the backend C&C servers. None of the botclients communicate directly with the backend C&C.

The payload of HTTP requests falls in two parts. The first part of the payload and contains the XML elements. The second part of the payload contains IP addresses if the sender is a proxy and AAAAAA (36 0-bits that are Base 64 encoded) if the sender is a botclient. All requests and responses have a payload forming at least <lm> XML element in plaintext. Sub XML elements differ depending on the type of message that is being send within the botnet.

Mobile devices mostly communicate on the mobile network and thus are unavailable on the Internet. Sending SMS and MMS messages containing payload like, XML elements those used in Waledac, will give the botclient on the mobile device the opportunity to maintain connection to its C&C through the proxy servers. The proxy server may be a device that understand and read the SMS and MMS message and have access to the Internet. By monitoring incoming messages the botclient can delete the message before it is added to the inbox to hide its existence.

Waledac uses P2P technology, meaning that infected mobile devices would have to communicate with each other to exchange list of active proxy servers. This can be done through MMS messages communicated between the infected devices on the mobile network.

Waledac on the Internet update their list with IP-addresses. Mobile devices do not use IP-addresses when sending SMS or MMS. Every device on the mobile network has a International Mobile Subscriber Identity (IMSI) and MSISDN.

IMSI can be used to identify, authenticate and register the device on the mobile network. MSISDN is the number you use when calling another mobile device or sending a message. Then the list of proxy servers must contain MSISDN numbers together with IP-addresses in case the botclients would connect to the Internet thus communicate directly with the proxy servers on the Internet.

Domain name does not exist on the mobile network. This makes it impossible to use techniques like fast flux and multihoming on devices in the mobile network. The consequence is that mobile devices cannot act as proxy servers as they very soon will be detected.

Symantec reported on 13 July 2009 that the first botclient on Symbian OS may have been developed. SymbOS.Exy.C, (aka Sexy Space) is a worm similar to other worms made for Symbian OS but the difference is the client that tries to contact a malicious server reporting the mobile device‘s phone type, International Mobile Equipment Identity (IMEI) and IMSI. Symantec mentions that this may be the first occurrence of a true botnet client on a mobile device.

Waledac has not been ported to an operating system for mobile devices and therefore does not pose a threat yet, but this may only be a matter of time.

Possible attacks on mobile devices

Waledac is known for sending email. A infected mobile device can send MMS or SMS to other mobile devices or to service numbers. Victims can be chosen by the botherder or they can randomly be chosen from the address book or contact list on the infected mobile device.

In Norway there are some contests where you can vote for your favorite song, person, TV program and so on. A botnet consist of many infected computers and the voting system will not be able to detect if a botherder uses his botnet to send a short message to the voting service. But will anyone pay a botherder to vote up his or her favorite song or person? Maybe someone would in context of political elections.

Instead of using the voting application, a DDOS attack against core of the mobile network can be done to stop people voting by making the voting system unavailable during the voting period.

There exist service phones where you can give money to charity. If you call a specific service number the mobile device subscriber pays a preset amount. Example are Nationalforeningen for Folkehelse or Kirkens Nødhjelp in Norway where you can call their respective service numbers to donate NOK 100, or other service numbers where you would donate twice as much.What if a botherder creates his own service number and programs all his botclients to call that number.

The price should be low so the subscribers would not notice and be suspicious about the extra charges.

Mobile devices are being more common to our daily life. They are small and you can carry them everywhere you go and you will probably lose some of them too. Mobile devices can be used to communicate between people both through voice and messaging, play games alone or with others on the Internet, make payments, check the status on your bank accounts, store private information like contact information (name, phone number, email addresses), personal information (social security number, PIN codes, account numbers, private pictures or business related data) and other informations that criminals can exploit and misuse for financial gain.

Infected mobile device will be able to act as spyware in the same way as botclient on desktop computers collecting personal information and send it to the attacker. Waledac is a plug-in based

botnet and it is easy to add plug-ins to extend functionality. There will be a possibility to create a plugin that scans the mobile device. The result can be reported back to the C&C in an XML document formed as an MMS.

Most people are trained to enter private data like social security number and credit card number on the mobile device using the mobile network. Is this safe anymore as the traffic in UMTS and newer technologies are routed within the IP-network?

3. MITIGATION STRATEGIES AGAINST MOBILE BOTNETS

Techniques as antivirus scanning, intrusion detection system (IDS) and packet filtering may be used together to stop malware spreading. Like Waledac hiding its traffic through legitimated web traffic, Waledac on mobile devices can hide its traffic through legitimate MMS and SMS traffic making it harder for the researchers to spot botnets on the mobile network. Botclients communicate with each other and would adopt the same resilient behavior as botnets on the Internet.

Waledac has a predefined structure on the messages. If you know what to look for there is possible to search and analyze the network traffic after specific signatures that Waledac or other known botnet creates.

In front of email systems there may be an antivirus scanner that scans every incoming and outgoing email for malware. This requires extra resources, but may be necessary to protect end users against attack.MMS work in similar way as the email system where MMS messages is send from an MMS client to an MMS proxy server where it is converted to standard Internet MIME format to permit various media components to be carried over the Internet environment. The receiving MMS proxy server will convert it back to MMS format before delivering the message to the MMS server which stores MMS message. To protect users from spam and malware attack the mobile network operator should use an antivirus scanner in front of their MMS server scanning incoming and outgoing messages to stop malicious messages to be delivered to the end users.

Security on mobile devices are not safeguarded in the same way as it is on desktop computers making the mobile devices easier to exploit. Compromised items inside a corporate network will constitute a threat since many security systems on the network can easily be bypassed. Mobile devices are often ignored.

Jansen and Scarfone write in their article that most corporate networks do not have centrally managed system to take care of the security policy and security update on mobile devices. It may also be difficult to update the mobile device due to lack of knowledge about how to do it. Many mobile device applications do not have update managers like applications on desktop computers

do. A system is only as secure as its least secure component and mobile devices often fall into this category.

To overcome this problem Janson and Scarfone suggest several techniques. By using PIN codes the access to the SIM card is protected. Additional memory cards will not have the sama protection since they can be taken out and put into other devices. By turning off interfaces like Bluetooth, Infrared (IR), WLAN and other wireless access protocols until they are needed, the attack surface on the mobile device can be reduced. Install antivirus software can provide protection against incoming malware, but the application will drain battery power very fast. Use certificate based solutions to sign applications even if the newest worm Sexy Space shows that this too is exploitable.

Compared to malware and botnet evolution on the Internet, mobile platform are 7-8 year behind. Experience from malware on the Internet can be transferred to malware on mobile device. Norman predicts that malware on mobile device will evolve faster since techniques are already explored. Security mechanism on mobile devices there for have to follow the same evolution as those on desktop computers. Windows Update for Windows Mobile 6 exists already. Adobe Updater, Java Update Manager exists for desktop applications so why not implement these on mobile devices too?

In the past the security of mobile networks has been relatively well controlled by the networks operators. By turning mobile devices into general purpose computation and communication platforms, new security vulnerabilities will emerge. The possibility of downloading Internet application to mobile devices also brings the risk of malware infection. People need to become aware that mobile devices are vulnerable to being infected by malware, and thereby can be turned into a botclient as part of a botnet. In this paper we have shown that there are potential profitable business models for exploiting mobile botnets. It is therefore necessary to start thinking about methods for reducing the threat of botnets on mobile networks.

V. EXPLOITATION OF SOCIAL NETWORKING

Social networking sites, such as Twitter and Facebook, have become mainstays of electronic information sharing. Information sharing often occurs with an unwarranted, inherent trust among users, as they blindly share and accept data from unauthenticated parties. Uniform Resource Locators (URLs) are constantly being exchanged within social networks as users share items of interest. Since a Twitter user is limited to 140 characters when posting an update, sharing a brief statement accompanied by a traditional URL may be impossible. The capability to significantly shorten a URL is provided by several different websites and is often integrated in social

networking applications to happen automatically. Shortened URLs are invaluable in this case because they allow a URL with 137 characters to be shortened to 17 characters. For example:
http://brainstormtech.blogs.fortune.cnn.com/2010/02/12/help-wanted-obamas-twitterer-filibusterers-need-not-apply/?source=cnn_bin&hpt=Sbin
becomes http://u.nu/72q95.

These services provide value, but they also make cyber criminals' goals much easier to achieve. Since the original URL is completely replaced, a user cannot know the destination of the shortened link without clicking on the link. Legitimate URLs are indistinguishable from those that are malicious, providing phishers with an effective cover. This tactic could lure a victim into unwittingly downloading malware or visiting a fraudulent site. It is highly likely that unsuspecting users would not think twice before clicking on the URLs.

Over the course of 2009, Facebook and Twitter experienced a 112% and 347% increase in mobile users, respectively.[48] This growing trend in mobile social networking provides an avenue for the exploitation of mobile devices.

Number of Mobile Subscribers Accessing Facebook, MySpace and Twitter via Mobile Browser
3-month average ending Jan. 2010 vs. Jan. 2009
Total U.S. Age 13+
Source: comScore MobiLens

	Total Audience (000)		
	Jan-09	Jan-10	% Change
Facebook.com	11,874	25,137	112
MySpace.com	12,338	11,439	-7
Twitter.com	1,051	4,700	347

According to a comScore report, both Twitter and Facebook have experienced significant increases in mobile browser access over the past year. "Social networking remains one of the most popular and fastest-growing behaviors on both the PC-based Internet and the mobile Web," said Mark Donovan, comScore senior vice president of mobile, in the company's[49]. "Social media is a natural sweet spot for mobile."

Mobile Browser Access to Social Networking: Smartphone vs. Feature Phone
3-month average ending Jan. 2010 vs. Jan. 2009
Total U.S. Age 13+
Source: comScore MobiLens

	Percent of Subscribers Accessing Social Networking via Mobile Browser		
	Jan-09	Jan-10	Point Change
All Mobile Phones	**6.5%**	**11.1%**	**4.6**
Smartphone	22.5%	30.8%	8.3
Feature Phone	4.5%	6.8%	2.3

Just over 30% of smartphone users access social networking sites using a mobile browser, comScore reports, up from 22% just a year ago. Access to Facebook using a mobile browser grew 112% while Twitter grew a whopping 347%.

What do these numbers mean in terms of actual number of visitors? According to comScore, Facebook saw 25.1 million mobile users in January

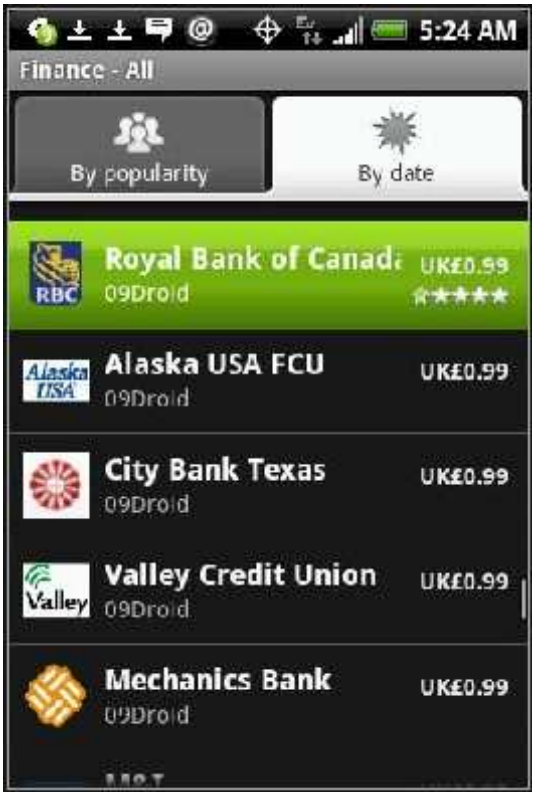
[48] Mike Melanson. Twitter Sees 347% Growth in Mobile Browser Access. 2010. Retrieved March 23, 2010 from http://www.readwriteweb.com/archives/twitter_sees_347_growth_in_mobile_browser_access.php
[49] http://www.comscore.com/Press_Events/Press_Releases/2010/3/Facebook_and_Twitter_Access_via_Mobile_Browser_Grows_by_Triple-Digits

2010, Myspace had 11.4 million and Twitter 4.7 million. As the report points out, "these figures do not include access of the social networking services by the nearly 6 million mobile phone owners who do so exclusively through mobile applications."

As smartphones continue to grow in popularity, social networking services will get more and more traffic from mobile use, and we wouldn't be surprised to see mobile access overtake other methods of access at some point in the future.

VI. EXPLOITATION OF MOBILE APPLICATIONS

Mobile applications, commonly called apps, provide enhanced convenience and functionality. Developers have created myriad mobile applications for various uses and activities, which is contributing to the proliferation of modern mobile devices. Anyone can potentially develop and distribute mobile applications with little oversight, making apps a potential attack vector for cyber criminals.



Several major banking institutions provide legitimate mobile applications that allow customers to conveniently check balances, pay bills, transfer funds, or locate automated teller machines (ATMs) and banking centers. However, banks are not the only ones creating banking-related apps. In early 2010, Google found potentially fraudulent banking applications in their Android Market. An anonymous developer known as "09Droid" sold a collection of banking applications that were not authorized by the banks for which they were seemingly developed.[50] It is unclear if the apps were used to gain access to users' confidential banking information. 09Droid published applications for approximately 40 different banking institutions, all of which Google removed from the Android Market.[51]

A similar incident occurred when Symbian unwittingly distributed the Sexy Space mobile worm as a legitimate, digitally signed application.[52] This malware steals subscriber, device, and

[50] Dan Raywood. Google finds apparently fraudulent banking applications on its Android Marketplace. 2010. Retrieved February 1, 2010 from http://www.scmagazineuk.com/google-finds-apparently-fraudulent-banking-applications-on-its-android-marketplace/article/161047/.

[51] F-Secure. Warning On Possible Android Mobile Trojans. 2010. Retrieved February 13, 2010 from http://www.f-secure.com/weblog/archives/00001852.html.

network information from victims and has the capability to build a botnet. It propagates via spam text messages that are sent from a compromised device to the victim's contacts. The messages, exchanged at the expense of the victims, contain a link to a website hosting malicious applications that will infect the phone if executed. Currently, the Sexy Space mobile worm affects only Symbian mobile devices.

The validation and approval process for mobile applications varies by vendor. The following table provides a brief description of the policies of some of the more popular vendors.

Vendor	*Application Store*	*Application Development Policy*
Apple	App Store	Apple requires developers to enroll in the iPhone Developer Program. Every application submitted to the App Store is evaluated by at least two reviewers for bugs, instabilities, unauthorized content, and other violations.
Google	Android Marketplace	No requirements exist for publishing applications in the Android Marketplace. Once developers register, they have complete control over when and how they make their applications available to users.
Microsoft	Windows Marketplace for Mobile	Developers must register with Windows Marketplace for Mobile. All applications sold on Windows Marketplace for Mobile must meet technical standards, be code signed, and pass policy checking and geographic market validation before they can be certified.
RIM	Blackberry App World	Developers must create a vendor account to submit applications to the Blackberry App World. RIM reviews all submitted applications for content suitability and performs technical testing to ensure applications abide by the Blackberry App World Vendor Guidelines.
Symbian	Horizon	Symbian Horizon is a publishing program and directory of Symbian Signed applications. To publish applications here, developers must obtain a Publisher

[52] John Leyden. Sign mobile malware prompts Symbian security review. 2009. Retrieved February 23, 2010 from http://www.theregister.co.uk/2009/07/23/sms_worm_analysis/.

		ID and run the full Symbian Signed Test Criteria on applications before they can be made publicly available.

Many applications are regularly submitted to vendors for use on these platforms, including some that are malicious. Currently, the Apple App Store contains over 100,000 applications and receives about 10,000 new submissions each week. Apple has received applications that will steal personal data or are otherwise malicious and has rejected them during the review process.[53] As the volume of applications rises, it could be difficult to maintain high confidence in their integrity, regardless of the platform or policy.

VII. EXPLOITATION OF M-COMMERCE

M-commerce, or mobile e-commerce, is another growing trend with mobile devices. Consumers can use mobile devices from any location to research product information, compare prices, make purchases, and communicate with customer support. Retailers can use mobile devices for tasks such as price checks, inventory inquiries, and payment processing. For example, Apple Retail Store employees use modified versions of the iPod Touch that allow them to scan barcode labels and accept credit card payments from customers.

The ability to read credit cards with a mobile device is not limited to retailers alone. A quick search for "credit card" in the Apple App Store reveals a number of different applications for accepting credit card payments. Third-party iPhone attachments for swiping credit cards are also available. "Square" is a small device that plugs into the iPhone's headphone jack and can transfer credit card swipe information to the supporting application. It also allows users to authorize payments in real-time via text message. The Mophie "marketplace" is another credit card reader for the iPhone.

[53] Arik Hesseldahl. Apple's Schiller Defends iPhone App Approval Process. 2009. Retrieved February 13, 2010 from http://www.businessweek.com/technology/content/nov2009/tc20091120_354597.htm

Smartphones' credit card reader functionality has the potential to enable criminal activity such as "skimming" and "carding." Skimming is the theft of credit card information using card readers, or skimmers, to record and store victims' data. This activity is often accomplished in conjunction with otherwise legitimate transactions. Carding is the process of testing the validity of stolen credit card numbers. It can be done on websites that support real-time transaction processing to determine if the credit information can be successfully processed. The capability of a single compact hand-held device to perform each of these tasks will further enable malicious intentions.

VIII. PROTECTING PORTABLE DEVICES

1. DEFENDING CELL PHONES AND PDA AGAINST ATTACK

As cell phones and PDAs become more technologically advanced, attackers are finding new ways to target victims. By using text messaging or email, an attacker could lure you to a malicious site or convince you to install malicious code on your portable device.

What unique risks do cell phones and PDAs present?

Most current cell phones have the ability to send and receive text messages. Some cell phones and PDAs also offer the ability to connect to the internet. Although these are features that you might find useful and convenient, attackers may try to take advantage of them. As a result, an attacker may be able to accomplish the following:

- **abuse your service** - Most cell phone plans limit the number of text messages you can send and receive. If an attacker spams you with text messages, you may be charged additional fees. An attacker may also be able to infect your phone or PDA with malicious code that will allow them to use your service. Because the contract is in your name, you will be responsible for the charges.
- **lure you to a malicious web site** - While PDAs and cell phones that give you access to email are targets for standard phishing attacks, attackers are now sending text messages to cell phones. These messages, supposedly from a legitimate company, may try to convince you to visit a malicious site by claiming that there is a problem with your account or stating that you have been subscribed to a service. Once you visit the site, you may be lured into providing personal information or downloading a malicious file.
- **use your cell phone or PDA in an attack** - Attackers who can gain control of your service may use your cell phone or PDA to attack others. Not only does this hide the real attacker's identity, it allows the attacker to increase the number of targets.
- **gain access to account information** - In some areas, cell phones are becoming capable of performing certain transactions (from paying for parking or groceries to conducting larger

financial transactions). An attacker who can gain access to a phone that is used for these types of transactions may be able to discover your account information and use or sell it.

What can you do to protect yourself?

- **Follow general guidelines for protecting portable devices** - Take precautions to secure your cell phone and PDA the same way you should secure your computer.
- **Be careful about posting your cell phone number and email address** - Attackers often use software that browses web sites for email addresses. These addresses then become targets for attacks and spam. Cell phone numbers can be collected automatically, too. By limiting the number of people who have access to your information, you limit your risk of becoming a victim.
- **Do not follow links sent in email or text messages** - Be suspicious of URLs sent in unsolicited email or text messages. While the links may appear to be legitimate, they may actually direct you to a malicious web site.
- **Be wary of downloadable software** - There are many sites that offer games and other software you can download onto your cell phone or PDA. This software could include malicious code. Avoid downloading files from sites that you do not trust. If you are getting the files from a supposedly secure site, look for a web site certificate If you do download a file from a web site, consider saving it to your computer and manually scanning it for viruses before opening it.
- **Evaluate your security settings** - Make sure that you take advantage of the security features offered on your device. Attackers may take advantage of Bluetooth connections to access or download information on your device. Disable Bluetooth when you are not using it to avoid unauthorized access.

2. DATA SECURITY

Why do you need another layer of protection?

Although there are ways to physically protect your laptop, PDA, or other portable device, there is no guarantee that it won't be stolen. After all, as the name suggests, portable devices are designed to be easily transported. The theft itself is, at the very least, frustrating, inconvenient, and unnerving, but the exposure of information on the device could have serious consequences. Also, remember that any devices that are connected to the internet, especially if it is a wireless connection, are also susceptible to network attacks.

What can you do?

- **Use passwords correctly** - In the process of getting to the information on your portable device, you probably encounter multiple prompts for passwords. Take advantage of this

security. Don't choose options that allow your computer to remember passwords, don't choose passwords that thieves could easily guess, use different passwords for different programs, and take advantage of additional authentication methods.

- **Consider storing important data separately** - There are many forms of storage media, including CDs, DVDs, and removable flash drives (also known as USB drives or thumb drives). By saving your data on removable media and keeping it in a different location (e.g., in your suitcase instead of your laptop bag), you can protect your data even if your laptop is stolen. You should make sure to secure the location where you keep your data to prevent easy access. It may be helpful to carry storage media with other valuables that you keep with you at all times and that you naturally protect, such as a wallet or keys.
- **Encrypt files** - By encrypting files, you ensure that unauthorized people can't view data even if they can physically access it. You may also want to consider options for full disk encryption, which prevents a thief from even starting your laptop without a passphrase. When you use encryption, it is important to remember your passwords and passphrases; if you forget or lose them, you may lose your data.
- **Install and maintain anti-virus software** - Protect laptops and PDAs from viruses the same way you protect your desktop computer. Make sure to keep your virus definitions up to date. If your anti-virus software doesn't include anti-spyware software, consider installing separate software to protect against that threat.
- **Install and maintain a firewall** - While always important for restricting traffic coming into and leaving your computer, firewalls are especially important if you are traveling and using different networks. Firewalls can help prevent outsiders from gaining unwanted access.
- **Back up your data** - Make sure to back up any data you have on your computer onto a CD-ROM, DVD-ROM, or network. Not only will this ensure that you will still have access to the information if your device is stolen, but it could help you identify exactly which information a thief may be able to access. You may be able to take measures to reduce the amount of damage that exposure could cause.

3. PHYSICAL SECURITY

Many computer users, especially those who travel for business, rely on laptops and PDAs because they are small and easily transported. But while these characteristics make them popular and convenient, they also make them an ideal target for thieves. Make sure to secure your portable devices to protect both the machine and the information it contains.

What is at risk?

Only you can determine what is actually at risk. If a thief steals your laptop or PDA, the most obvious loss is the machine itself. However, if the thief is able to access the information on the computer or PDA, all of the information stored on the device is at risk, as well as any additional information that could be accessed as a result of the data stored on the device itself.

Sensitive corporate information or customer account information should not be accessed by unauthorized people. You've probably heard news stories about organizations panicking because laptops with confidential information on them have been lost or stolen. But even if there isn't any sensitive corporate information on your laptop or PDA, think of the other information at risk: information about appointments, passwords, email addresses and other contact information, personal information for online accounts, etc.

How can you protect your laptop or PDA?

- **Password-protect your computer** - Make sure that you have to enter a password to log in to your computer or PDA.
- **Keep your laptop or PDA with you at all times** - When traveling, keep your laptop with you. Meal times are optimum times for thieves to check hotel rooms for unattended laptops. If you are attending a conference or trade show, be especially wary—these venues offer thieves a wider selection of devices that are likely to contain sensitive information, and the conference sessions offer more opportunities for thieves to access guest rooms.
- **Downplay your laptop or PDA** - There is no need to advertise to thieves that you have a laptop or PDA. Avoid using your portable device in public areas, and consider non-traditional bags for carrying your laptop.
- **Be aware of your surroundings** - If you do use your laptop or PDA in a public area, pay attention to people around you. Take precautions to shield yourself from "shoulder surfers"—make sure that no one can see you type your passwords or see any sensitive information on your screen.
- **Consider an alarm or lock** - Many companies sell alarms or locks that you can use to protect or secure your laptop. If you travel often or will be in a heavily populated area, you may want to consider investing in an alarm for your laptop bag or a lock to secure your laptop to a piece of furniture.
- **Back up your files** - If your portable device is stolen, it's bad enough that someone else may be able to access your information. To avoid losing all of the information, make backups of important information and store the backups in a separate location. Not only

will you still be able to access the information, but you'll be able to identify and report exactly what information is at risk.

What can you do if your laptop or PDA is lost or stolen?

Report the loss or theft to the appropriate authorities. These parties may include representatives from law enforcement agencies, as well as hotel or conference staff. If your device contained sensitive corporate or customer account information, immediately report the loss or theft to your organization so that they can act quickly.

CONCLUSION

The user's limited awareness and subsequent unsafe behavior may be the most threatening vulnerabilities for mobile devices. It is critical to understand that a mobile device is no longer just a phone and cannot be treated as such. Unlike the previous generation of mobile phones that were at worst susceptible to local Bluetooth hijacking, modern Internet-tethered mobile devices are susceptible to being probed, identified, and surreptitiously exploited by hackers from anywhere on the Internet. Many mitigation techniques for mobile devices are similar to those for PCs. US-CERT recommends the following best practices to help protect mobile devices:

- ✓ Maintain up-to-date software, including operating systems and applications;
- ✓ Install anti-virus software as it becomes available and maintain up-to-date signatures and engines;
- ✓ Enable the personal identification number (PIN) or password to access the mobile device, if available;
- ✓ Encrypt personal and sensitive data, when possible;
- ✓ Disable features not currently in use such as Bluetooth, infrared, or Wi-Fi;
- ✓ Set Bluetooth-enabled devices to non-discoverable to render them invisible to unauthenticated devices;
- ✓ Use caution when opening email and text message attachments and clicking links;
- ✓ Avoid opening files, clicking links, or calling numbers contained in unsolicited email or text messages;
- ✓ Avoid joining unknown Wi-Fi networks;
- ✓ Delete all information stored in a device prior to discarding it;
- ✓ Maintain situational awareness of threats affecting mobile devices.

Anti-virus software exists for some mobile devices, which is one component of a layered defense. However, it can only assist in protecting against known threats. Users need to understand the threats and proactively take steps to avoid them. A high degree of vigilance is necessary to successfully prevent and mitigate future threats to mobile devices.

REFERENCES

- Aliya Sternstein - http://www.nextgov.com/nextgov/ng_20101228_6846.php
- Arik Hesseldahl. Apple's Schiller Defends iPhone App Approval Process. 2009. Retrieved February 13, 2010 from http://www.businessweek.com/technology/content/nov2009/tc20091120_354597.htm
- Bill Brenner. Proof-of-concepts heighten mobile malware fears. 2006. Retrieved February 17, 2010 from http://searchexchange.techtarget.com/news/article/0,,sid43_gci1171168,00.html.
- Cyrus Peikari. Analyzing the Crossover Virus: The First PC to Windows Handheld Cross-infector. 2006. Retrieved February 17, 2010 from http://www.informit.com/articles/article.aspx?p=458169&seqNum=3.
- Cyrus Peikari. Analyzing the Crossover Virus: The First PC to Windows Handheld Cross-infector. Retrieved February 17, 2010
- Dan Raywood. Google finds apparently fraudulent banking applications on its Android Marketplace. 2010. Retrieved February 1, 2010 from http://www.scmagazineuk.com/google-finds-apparently-fraudulent-banking-applications-on-its-android-marketplace/article/161047/.
- Flexispy Ltd. FlexiSpy Homepage. 2010. Retrieved February 17, 2010 from http://flexispy.com/.
- F-Secure. Warning On Possible Android Mobile Trojans. 2010. Retrieved February 13, 2010 from http://www.f-secure.com/weblog/archives/00001852.html.
- F-Secure. Worm:iPhoneOS/Ikee.B. 2009. Retrieved February 16, 2010 from http://www.f-secure.com/v-descs/worm_iphoneos_ikee_b.shtml.
- John Leyden. Sign mobile malware prompts Symbian security review. 2009. Retrieved February 23, 2010 from http://www.theregister.co.uk/2009/07/23/sms_worm_analysis/
- Ken Dunham, et al. Mobile Malware Attacks and Defense. 2009. Burlington, MA: Syngress Publishing, Inc.
- Kreaty Ferguson - http://www.shaswatpatel.com/mobile-malwares-a-big-threat-for-smartphone-users-in-2011/
- Mike Melanson. Twitter Sees 347% Growth in Mobile Browser Access. 2010. Retrieved March 23, 2010 from http://www.readwriteweb.com/archives/twitter_sees_347_growth_in_mobile_browser_access.php.

- US-CERT Technical Information Paper-TIP-10-105-01 - Cyber Threats to Mobile Devices, from http://www.us-cert.gov/reading_room/TIP10-105-01.pdf
- US-CERT – Virus Basics and Frequently Asked Questions
- US-CERT Cyber Security Tip ST04-020 – Protecting Portable Devices: Data Security
- US-CERT Cyber Security Tip ST05-017 – Cybersecurity for Electronic Devices
- US-CERT Cyber Security Tip ST06-001 – Understanding Hidden Threats: Rootkits and Botnets
- US-CERT Cyber Security Tip ST06-007 – Defending Cell Phones and PDAs Against Attack

CYBER SECURITY

CAPT Ioan Claudiu COMICI

"Cyber threats are asymmetric because attacks may be perpetrated by the few upon the many, with little cost and resources. Cyber attacks are typically anonymous, launched from any of billions of sources worldwide. Impacts may be immediate and obvious, or dormant and subtle, eluding recognition for years. Degrees of damage can range from inconvenient downtime of personal systems to the life-threatening destruction of critical infrastructures." - U.S. Naval Institute, Cyber Threats to National Security, July 2010

INTRODUCTION

THE CYBER DIMENSION

The economy and national security depend greatly and increasingly on the global cyber infrastructure. Cyber infrastructure enables all sectors' functions and services, resulting in a highly interconnected and interdependent global network of CIKR (Critical infrastructure and key resources).

A spectrum of malicious actors routinely conducts attacks against the cyber infrastructure using cyber attack tools. Because of the interconnected nature of the cyber infrastructure, these attacks could spread quickly and have a debilitating effect.

Cybersecurity includes preventing damage to unauthorized use of, or exploitation of electronic information and communications systems and the information contained there in to ensure confidentiality, integrity, and availability. Cybersecurity also includes restoring electronic information and communications systems in the event of a terrorist attack or natural disaster.

The use of innovative technology and interconnected networks in operations improves productivity and efficiency, but also increases the country's vulnerability to cyber threats if cybersecurity is not addressed and integrated appropriately.

The interconnected and interdependent nature of the country's CIKR makes it problematic to address the protection of physical and cyber assets independently.

The NIPP (National Infrastructure Protection Plan) addresses reducing cyber risk and enhancing cybersecurity in two ways:

(1) as a cross-sector cyber element that involves DHS **(**Department of Homeland Security**)**, SSAs (Sector-Specific Plans) and Government Coordinating Councils (GCCs), and private sector owners and operators;

(2) and as a major component of the Information Technology Sector's responsibility in partnership with the Communications Sector.

Cyber infrastructure includes electronic information and communication systems, and the information contained in these systems. Computer systems, control systems such as Supervisory Control and Data Acquisition (SCADA) systems, and networks such as the Internet are all part of cyber infrastructure.

Information and communications systems are composed of hardware and software that process, store, and communicate data of all types. Processing includes the creation, access, modification, and destruction of information. Storage includes paper, magnetic, electronic, and all other media types. Communications include sharing and distribution of information.

Information Technology (IT) critical functions are sets of processes that produce, provide, and maintain products and services. IT critical functions encompass the full set of processes (e.g. manufacturing, distribution, upgrades and maintenance) involved in transforming supply inputs into IT products and services.

I. CYBER SECURITY

I. 1 WHAT IS CYBER SECURITY ?

A computer security is a branch of computer technology known as information security as applied to computers and networks. The objective of computer security includes protection of information and property from theft, corruption, or natural disaster, while allowing the information and property to remain accessible and productive to its intended users. The term computer system security means the collective processes and mechanisms by which sensitive and valuable information and services are protected from publication, tampering or collapse by unauthorized activities or untrustworthy individuals and unplanned events respectively. The strategies and methodologies of computer security often differ from most other computer technologies because of its somewhat elusive objective of preventing unwanted computer behavior instead of enabling wanted computer behavior**.**

Cyber security is an activity or set of activities that involves, protecting, tracking and resolving the cyber threat and also explains how to respond them on web and computers and its networks such as communication through emails and cell phones, entertainment, data transfer and for record keeping and maintenance. Excess use of computer and internet have made the security essential in every field because online attackers and hackers are also the part of computer network which we never know can see our personal information or can steal our confidential data. It's very important to understand the potential risks and basic security measures in order to protect ourselves from the cyber threats.

I. 2 WHAT RISKS ARE INVOLVED IN CYBER SECURITY?

I am going to deal with the long list of risky factors which makes you feel unsafe while dealing with cyber technologies:

a. **Downloading:** One of the most practiced thing or task is the downloading. Internet users allover the world downloads the different kinds of soft wares, music audios/videos, movies, games and many more in bulk quantity every day. Many of them forgets about the risks involves in downloading the file your saving to your computer system straight from the internet. In spite of using the best anti virus you still need to be conscious about the security of your computer and files.

b. **Unauthorized purchases**: You have often heard about the fraudulent companies and online platforms that work online and steal the confidential information of the customers who shop, sell or buy stuff online. Your important information or your data can be theft online. In worst scenarios people loss millions of dollars while selling and purchasing through unauthorized websites.

c. **Infected or malicious programming snippets**: These are the small codes that are developed o transfer viruses or worms of different types to your computer system. They can be generated by the professional computer hackers. These code shave unique features in them viruses are such programs that travel through your network and usually hits your system when some one form you clicks them, but the worms are so powerful that they automatically penetrates into your computer and may erase the important system files may make your system is halted while working. They propagate themselves automatically when you get online.

d. **Trojan horses:** These are such type of dangerous software that damages you behind the screen. It means that at front interface they claim the speeding up or enhancing your RAM cache or downloading speed but at the back end they start stealing and copy your confidential information from your email accounts or computer.

Precaution to stay safe: Although it's always risky to work and deal online but if some rules of cyber security are continuously practiced then cyber threat risk could be minimized to some extent here are some steps to follow:

- The foremost step is that you must understand the importance and risks of cyber security.
- After that what kind of hackers or intruders may involve in your dealings. Motivate yourself against the online hackers and be aware of them all the tie even sending a single mail message from your email account.

- Avoid instant or random clicking in attractive stuff that you see around your mail box screen that may contains win cash prize adds or wok online etc. these fascinating ads may carry infections for your computer.
- Always choose the authorized companies for selling and purchasing stuff online don't forget to see the privacy policy that they are offering

II. CYBER THREAT SOURCE DESCRIPTIONS

Cyber threats to a control system refer to persons who attempt unauthorized access to a control system device and/or network using a data communications pathway. This access can be directed from within an organization by trusted users or from remote locations by unknown persons using the Internet. Threats to control systems can come from numerous sources, including hostile governments, terrorist groups, disgruntled employees, and malicious intruders. To protect against these threats, it is necessary to create a secure cyber-barrier around the Industrial Control System (ICS). Though other threats exist, including natural disasters, environmental, mechanical failure, and inadvertent actions of an authorized user, this discussion will focus on the deliberate threats mentioned bellow:

a. National Governments
b. Terrorists
c. Industrial Spies and Organized Crime Groups
d. Hacktivists
e. Hackers
f. GAO Threat Table

Activities could include espionage, hacking, identity theft, crime, and terrorism.

a. National Governments

National cyber warfare programs are unique in posing a threat along the entire spectrum of objectives that might harm the country's interests. These threats range from propaganda and low-level nuisance web page defacements to espionage and serious disruption with loss of life and extensive infrastructure disruption. Among the array of cyber threats, as seen today, only government-sponsored programs are developing capabilities with the future prospect of causing widespread, long-duration damage to the country's critical infrastructures.

b. Terrorists

Traditional terrorist adversaries, for example of the U.S., despite their intentions to damage U.S. interests, are less developed in their computer network capabilities and propensity to pursue cyber means than are other types of adversaries. They are likely, therefore, to pose only a limited cyber threat. Since bombs still work better than bytes, terrorists are likely to stay focused on

traditional attack methods in the near term. We anticipate more substantial cyber threats are possible in the future as a more technically competent generation enters the ranks.

Their goal is to spread terror throughout the U.S. civilian population. Their sub-goals include: attacks to cause 50,000 or more casualties within the U.S. and attacks to weaken the U.S. economy to detract from the Global War on Terror.

c. Industrial Spies and Organized Crime Groups

International corporate spies and organized crime organizations pose a medium-level threat to the country through their ability to conduct industrial espionage and large-scale monetary theft as well as their ability to hire or develop hacker talent.

Their goals are profit-based. Their sub-goals include attacks on infrastructure for profit to competitors or other groups listed above, theft of trade secrets, and gain access and blackmail affected industry using potential public exposure as a threat.

d. Hacktivists

Hacktivists form a small, foreign population of politically active hackers that includes individuals and groups with anti-country. motives. They pose a medium-level threat of carrying out an isolated but damaging attack. Most international hacktivist groups appear bent on propaganda rather than damage to critical infrastructures. Their goal is to support their political agenda. Their sub-goals are propaganda and causing damage to achieve notoriety for their cause.

e. Hackers

Although the most numerous and publicized cyber intrusions and other incidents are ascribed to lone computer-hacking hobbyists, such hackers pose a negligible threat of widespread, long-duration damage to national-level infrastructures. The large majority of hackers do not have the requisite tradecraft to threaten difficult targets such as critical country's networks and even fewer would have a motive to do so. Nevertheless, the large worldwide population of hackers poses a relatively high threat of an isolated or brief disruption causing serious damage, including extensive property damage or loss of life. As the hacker population grows, so does the likelihood of an exceptionally skilled and malicious hacker attempting and succeeding in such an attack.

In addition, the huge worldwide volume of relatively less skilled hacking activity raises the possibility of inadvertent disruption of a critical infrastructure.

For the purposes of this work, hackers are subdivided as follows:

- Sub-communities of hackers;
- Script kiddies are unskilled attackers who do NOT have the ability to discover new vulnerabilities or write exploit code, and are dependent on the research and tools from others. Their goal is achievement. Their sub-goals are to gain access and deface web pages;

- Worm and virus writers are attackers who write the propagation code used in the worms and viruses but not typically the exploit code used to penetrate the systems infected. Their goal is notoriety. Their sub-goals are to cause disruption of networks and attached computer systems;
- Security researcher and white hat have two sub-categories; bug hunters and exploit coders. Their goal is profit. Their sub-goals are to improve security, earn money, and achieve recognition with an exploit;
- Professional hacker-black hat who gets paid to write exploits or actually penetrate networks; also falls into the two sub-categories-bug hunters and exploit coders. Their goal is profit;

Hackers and researchers interact with each other to discuss common interests, regardless of color of hat. Hackers and researchers specialize in one or two areas of expertise and depend on the exchange of ideas and tools to boost their capabilities in other areas. Information regarding computer security research flows slowly from the inner circle of the best researchers and hackers to the general IT security world, in a ripple-like pattern.

f. Threat Table

The following table is an excerpt from NIST 800-82, "Guide to Supervisory Control and Data Acquisition (SCADA) and Industrial Control System Security (SME draft)", provides a description of various threats to CS networks:

Threat	Description
Bot-network operators	Bot-network operators are hackers; however, instead of breaking into systems for the challenge or bragging rights, they take over multiple systems in order to coordinate attacks and to distribute phishing schemes, spam, and malware attacks. The services of these networks are sometimes made available in underground markets (e.g., purchasing a denial-of-service attack, servers to relay spam, or phishing attacks, etc.).
Criminal groups	Criminal groups seek to attack systems for monetary gain. Specifically, organized crime groups are using spam, phishing, and spyware/malware to commit identity theft and online fraud. International corporate spies and organized crime organizations also pose a threat to the United States through their ability to conduct industrial espionage and large-scale monetary theft and to hire or

	develop hacker talent.
Foreign intelligence services	Foreign intelligence services use cyber tools as part of their information-gathering and espionage activities. In addition, several nations are aggressively working to develop information warfare doctrine, programs, and capabilities. Such capabilities enable a single entity to have a significant and serious impact by disrupting the supply, communications, and economic infrastructures that support military power - impacts that could affect the daily lives of U.S. citizens across the country.
Hackers	Hackers break into networks for the thrill of the challenge or for bragging rights in the hacker community. While remote cracking once required a fair amount of skill or computer knowledge, hackers can now download attack scripts and protocols from the Internet and launch them against victim sites. Thus while attack tools have become more sophisticated, they have also become easier to use. According to the Central Intelligence Agency, the large majority of hackers do not have the requisite expertise to threaten difficult targets such as critical U.S. networks. Nevertheless, the worldwide population of hackers poses a relatively high threat of an isolated or brief disruption causing serious damage.
Insiders	The disgruntled organization insider is a principal source of computer crime. Insiders may not need a great deal of knowledge about computer intrusions because their knowledge of a target system often allows them to gain unrestricted access to cause damage to the system or to steal system data. The insider threat also includes outsourcing vendors as well as employees who accidentally introduce malware into systems.
Phishers	Individuals, or small groups, who execute phishing schemes in an attempt to steal identities or information for monetary gain. Phishers may also use spam and spyware/malware to accomplish their objectives.
Spammers	Individuals or organizations who distribute unsolicited e-mail with hidden or false information in order to sell products, conduct phishing schemes, distribute spyware/malware, or attack organizations (i.e., denial of service).

Spyware/malware authors	Individuals or organizations with malicious intent carry out attacks against users by producing and distributing spyware and malware. Several destructive computer viruses and worms have harmed files and hard drives, including the Melissa Macro Virus, the Explore.Zip worm, the CIH (Chernobyl) Virus, Nimda, Code Red, Slammer, and Blaster.
Terrorists	Terrorists seek to destroy, incapacitate, or exploit critical infrastructures in order to threaten national security, cause mass casualties, weaken the U.S. economy, and damage public morale and confidence. Terrorists may use phishing schemes or spyware/malware in order to generate funds or gather sensitive information.

III. CYBER TRENDS

Charity Scams

Cybercrime trends tend to fluctuate with current events. It's sickening how whenever a natural disaster occurs, pretty soon thereafter you will hear about imposter charities trying to get some of the money that well-meaning people want to donate. You just have to be very careful who you give money or personal information to, especially on the internet.

Before you donate any money to a charity, make sure it is the real deal. With the recent problems following the earthquake in Haiti, a slew of con artists started trying to trick people out of their money. The FBI even put out an alert back in January that warns people if anyone claiming to be an earthquake victim tries to contact them. This just goes to show you how uncaring these criminals really are when they prey on the good will of others.

Tax Scams

During April when income tax season is going full speed in the United States, there are numerous scams to be found when it comes to tax returns and filings. The most common tax scam is a basic phishing attempt built around tax returns. You'll get an email that looks like it is from the IRS (Internal Revenue Service), and in the email they will claim to need more information from you in order to process your return. To help butter you up, they might promise thousands of dollars in tax return money if you give them what they want. With many people filing their taxes electronically, and some perhaps not knowing any better, this can be an easy scam to catch the unsuspecting.

To help combat this annual problem, the Internal Revenue Service maintains a web page full of information regarding all the various tax scams that pop up throughout the year.

Facebook Scams

Social networking sites like Facebook are still extremely popular these days, so they make a prime target for scammers looking to con people into giving up their private information. Where MySpace used to be overrun with spammers from porn sites and crappy bands trying to pad their list of friends, Facebook has the occasional surge of bogus apps that redirect users to sites that try to make them download things like fake security software.

I am a pretty avid Facebook user and have seen firsthand the results of a bogus app that tried to trick people into thinking they could see who was looking at their profile. I don't remember the exact wording, but the idea was that if you allowed this app onto your profile, it would let you see if anyone that wasn't your friend had been searching for you or looking up your information. If such a thing did exist, it would be nice to have, but it does not exist. Instead, people who signed up for the app would get redirected to a website outside of Facebook where it tried to make them install what turned out to be fake security software that took over their system.

Sophisticated campaign tracking and dramatically increased use of social networking technologies, such as Facebook and Twitter, were two of the top trends in cybercrime in 2009, according to a new report. Criminal attacks using social networking sites increased by 500 percent between 2008 and 2009, according to the Blue Coat Web Security Report for 2009, by application delivery network provider Blue Coat Systems. That makes those sites the top focus for cybercriminals' activities.

Fake Security Software

When it comes to cybercrime the most bothersome of them all is the widespread infections of fake security software. These programs are designed to look like real virus scanners or antispyware software, and they bombard you with pop-up messages and disable certain system utilities by saying your computer is infected and the only solution is to buy an upgrade to get rid of the viruses and malware. If you give in and make the purchase, then you are giving your credit card information away to scammers. What is annoying about these type scams is how commonly found they can be, and most of them come in through advertising feeds that aren't monitored by the sites than run the ads.

How to Avoid Being a Victim of Cybercrime

The best way to protect yourself from cybercriminals is to be smarter than them. The key to doing so is to be suspicious of everything that seems even the slightest bit untrustworthy. Don't respond to email solicitations for any reason. Don't even respond to phone calls to your home, because scammers often take the direct approach that way, too. If you keep abreast of all the criminal activity going on out there, then you'll be better able to recognize a potential thread and keep yourself from being the next victim.

IV. BEATING BACK THE BOTNETS

Why Organizations Should be Utilizing Security Information and Event Management (SIEM) Systems to Ferret Out Botnet Infections.

Botnets are insidious. They spread like digital weeds and infect thousands to tens of thousands of machines at a time. Their only purpose is to enrich and empower the botnet owners as they infiltrate endpoints on consumer systems, colleges, and enterprises around the world. These botnets are used to send spam, launch denial of service attacks, and – increasingly – to snoop on corporate systems.

These botnets, vast networks of infected systems under the control and command of criminals, are everyone's problem – and it's time we all did more to eliminate them from the Internet and our networks. One of the challenges in combating botnets, however, is the ease of infecting endpoints. Within the time it takes to view a web page, open an attachment, or load a picture, a user can get infected. Attackers are using vulnerabilities in web browsers and traditional client systems to infect users – who don't even know they've been infected.

When it comes to infection and botnets, what does that mean? It means that the attacker managed to turn the endpoint into a zombie. These zombies are called such because they don't do anything until they are so ordered by their remote commander. Once an attacker has infected enough systems, it then will use those zombies to send spam or kill the availability of unsuspecting web servers. Many botnets also are designed to steal end user authentication data, such as those used to log onto financial services and web sites.

Some of the more infamous botnets are BredoLab, which is estimated to have 30 million infected hosts, Mariposa (12 million), Conficker (10.5 million), and Zeus (3.6 million). Each of these botnets is capable of sending billions of spam messages every day. Many believe these botnets are a consumer security problem. That's a bad assumption.

Increasingly, according to a 2009 DarkReading story, botnet operators have been crafting smaller botnets designed to target specific businesses and people.

That's troubling news for enterprises that don't have the right defenses in place. For instance, botnets easily can be designed to avoid anti-virus, spyware, intrusion detection systems, and many other anti-malware technologies. The good news is that there are ways to spot zombie-infected systems on enterprise networks if you know how to look. For instance, botnets often try to obfuscate the data they're collecting as well as their connections to the controlling hosts. But by using the appropriate monitoring software and integrating that with a security event monitor, it's possible to spot this malicious traffic and identify infected systems.

What's needed is a way to see the patterns of infected systems spewing spam, collecting information, and/or trying to relay information or accept commands from its master. Some of the

data you need may lie within e-mail servers, system server logs, server and endpoint firewalls, and even domain name server requests to countries where your business may not be engaged.

Clues to botnet infection can reside in any of these places. Finding the clues and putting together the picture they make is the challenge. In fact, it can be almost impossible to do manually. There is just too much information to have to sort through. That's where a **Security Information and Event Management (SIEM)** system can help to ferret out botnet infections.

When SIEMs scour logs, they don't get bored or distracted. They help organizations to make sense of the volumes of data that all of their systems generate. SIEMs can perform highly sophisticated analysis so IT teams quickly can recognize new trends and attacks – and this is exactly what is needed to find clandestine botnets communicating on busy networks.

One way they can work is by correlating user identities with the actions of the systems they're using. For example, an endpoint used by an executive assistant shouldn't (typically) be sending e-mail all day that appears to be coming from some Eastern bloc nation. It also would be unusual for someone in accounting to be sending thousands of requests to the same web server all day.

Those are two good examples of how botnets act. And, by having the ability to spot anomalous behavior provided by a SIEM, enterprises will be leveraging one of the most powerful tools in their information security toolbox to stomp out botnet infections.

CONCLUSIONS

The cyber threat is one of the most serious economic and national security challenges a nation can face and a country's economic prosperity in the 21st century will depend on cybersecurity.

Why This is Important

Cyberspace touches nearly every part of our daily lives. It's the broadband networks beneath us and the wireless signals around us, the local networks in our schools and hospitals and businesses, and the massive grids that power the nation. It's the classified military and intelligence networks that keep us safe, and the World Wide Web that has made us more interconnected than at any time in human history. We must secure our cyberspace to ensure that we can continue to grow the country's economy and protect our way of life.

What Must Be Done

A country's cybersecurity strategy is twofold: (1) improve the resilience to cyber incidents and (2) reduce the cyber threat.

Improving the cyber resilience includes: hardening the digital infrastructure to be more resistant to penetration and disruption; improving the ability to defend against sophisticated and agile cyber threats; and recovering quickly from cyber incidents—whether caused by malicious activity, accident, or natural disaster.

Where possible, cyber threats must be reduced. They can be reduced by working with allies on international norms of acceptable behavior in cyberspace, strengthening law enforcement capabilities against cybercrime, and deterring potential adversaries from taking advantage of remaining vulnerabilities.

Underlying all of these efforts is the need to acquire the best possible information about the state of our networks and the capabilities and intentions of cyber adversaries. Critical cybersecurity information must be made available to and usable by everyone who needs it, including network operators and defenders, law enforcement and intelligence agencies, and emergency management officials in the State, governments, private industry, and allied governments.

All these actions are taken to secure the networks, in a manner that preserves and enhances personal privacy and enables the exercise of civil liberties and fundamental freedoms. In the 21st Century, the digital networks are essential to a safer way of life around the world and are an engine for freedom.

Near Term Actions

Some near term actions to support the cybersecurity strategy:

1. Appoint a cybersecurity policy official responsible for coordinating the country's cybersecurity policies and activities.
2. Prepare for the government's approval an updated strategy to secure the information and communications infrastructure.
3. Designate cybersecurity as one of the government's key management priorities and establish performance metrics
4. Conduct interagency-cleared legal analyses of priority cybersecurity-related issues.
5. Initiate a national awareness and education campaign to promote cybersecurity.
6. Develop an international cybersecurity policy framework and strengthen international partnerships.
7. Prepare a cybersecurity incident response plan and initiate a dialog to enhance public-private partnerships.
8. Develop a framework for research and development strategies that focus on game-changing technologies that have the potential to enhance the security, reliability, resilience, and trustworthiness of digital infrastructure.
9. Build a cybersecurity-based identity management vision and strategy, leveraging privacy-enhancing technologies for the country.

REFERENCES

1. Government Accountability Office (GAO), Department of Homeland Security's (DHS's) Role in Critical Infrastructure Protection (CIP) Cybersecurity, GAO-05-434 (Washington, D.C.: May, 2005).
2. National Infrastructure Protection Plan Partnering to enhance protection and resiliency, 2009.

WEBLINKS:

1. http://www.dhs.gov/xlibrary/assets/privacy/privacy_cybersecurity_white_paper.pdf
2. http://en.wikipedia.org/wiki/Cyber_security
3. http://www.wifinotes.com/security/why-cyber-security-a-problem.html
4. http://www.whitehouse.gov/administration/eop/nsc/cybersecurity
5. http://www.securityweek.com/beating-back-botnets#
6. http://www.wifinotes.com/security/why-cyber-security-a-problem.html
7. http://www.esecurityplanet.com/trends/article.php/3874206/Trends-in-Cybercrime-Report.htm
8. http://www.us-cert.gov/control_systems/csthreats.html
9. http://www.dhs.gov/xlibrary/assets/NIPP_Plan.pdf

ELLIPTIC CURVE CRYPTOGRAPHY

1st LT Eng. Mihaita IVASCU

INTRODUCTION

Asymmetric cryptography is a marvellous technology. Its uses are many and varied. For many situations in distributed network environments, asymmetric cryptography is a must during communications. If you're taming key distribution issues with a public key infrastructure(PKI), you are using asymmetric cryptography. If you are designing or employing any kind of network protocol or application requiring secure communications, to come up with a practical solution, you're going to use asymmetric cryptography.

Asymmetric cryptography has, in fact, proved so useful for securing communications that it has become pervasive in modern life. Every time you by something on the Internet, if the vendor is using a secure server, you are using asymmetric cryptography to secure the transaction.

But this type of cryptography is demanding and complex, by its very nature. The hard problems in number theory – the key to the algorithm's functionality – are all intrinsically difficult enough that the processor cycles you must throw at doing it, and/or the chip space you must allocate to the implementation, inevitably far outstrip the resources you must dedicate for doing symmetric cryptography.

That way if you need asymmetric cryptography, you should be considering elliptic curve cryptography(ECC).

- ECC offers considerably greater security for a given key size
- The smaller key size also makes possible much more compact implementations for a given level of security, which means faster cryptographic operations, running on smaller chips or more compact software. This means less heat production and less power consumption – all of which is of particular advantage in constrained devices, but of some advantages anywhere.
- There are extremely efficient, compact hardware implementations available for ECC exponentiation operations, offering potential reductions in implementation footprint even beyond those due to the smaller key length alone.

In short asymmetric cryptography is demanding. But if you're looking for the cryptosystem that will give you the most security per bit, you want ECC. ECC is an approach - a set of algorithms for key generation, encryption and decryption – to doing asymmetric cryptography.

I. ASYMMETRIC CRYPTOGRAPHY

I.1. The need for asymmetric cryptography

In symmetric cryptography, the same key is used for both encryption and decryption. This approach is simpler in dealing with each message, but less secure since the key must be communicated to and known at both sender and receiver locations. Asymmetric cryptography or public-key cryptography is cryptography in which a pair of keys is used to encrypt and decrypt a message so that it arrives securely. Initially, a network user receives a public and private key pair from a certificate authority. Any other user who wants to send an encrypted message can get the intended recipient's public key from a public directory. They use this key to encrypt the message, and they send it to the recipient. When the recipient gets the message, they decrypt it with their private key, which no one else should have access to. The critical feature of asymmetric cryptography, which makes it useful, is the key pair - and more specifically, a particular feature of the key pair: the fact that one key cannot be obtained form the other.

In cases where the same algorithm is used to encrypt and decrypt, such as in RSA, a message can be securely signed by a specific sender: if the sender encrypts the message using their *private* key, then the message can be decrypted only using that sender's *public* key, authenticating the sender.

This also allows for the exchanging of securely signed and one-to-one messages, as follows. The sender encrypts the message using the common algorithm and his own secret key. They then sign the result, encrypt it again (with their signature in cleartext) using the recipient's public key, and send it. The recipient decrypts the received message using their own secret key, identifies the sender from their now-cleartext signature, and then decrypt the result using the sender's public key. This ensures the recipient that whoever composed the message had access to the sender's private key, and that nobody tampered with the message or read it along the way.

I.2. Authentication with asymmetric cryptography

In the case of asymmetric authentication methods – the core technology behind digital signatures and certificates – we normally speak of a private key (in the possession of the entity wishing to prove its identity) and the public key(in the possession of anyone who wishes to verify the identity of the entity possessing the private key).

You may, with the public key, verify that an entity has knowledge of the private key – but you cannot derive the private key from the public. This is the critical feature of the asymmetric cryptography schemes that makes them so useful.

This property is useful for a number of things: it greatly simplifies key exchange, as one example, and it solves one critical problem symmetric cryptography cannot solve – the problem

of guaranteeing unique authentication and non-repudiation. Symmetric hashing/authentication methods – ones for which there is only one key, and both parties in the exchange use it both for authentication and non-repudiation. Symmetric hashing/authentication methods — ones for which there is only one key, and both parties in the exchange use it both for authentication and for signature generation — have the distinct disadvantage that they do not, on their own, offer any way to distinguish which party to the exchange signed a given message. If both or all parties must know the key, based on cryptography alone, you cannot distinguish which signed any given message, because any of them could have. In asymmetric authentication schemes, only one party knows the private key, with which the message is signed. Any number may know the public key. Since the private key cannot be derived from the public, the signature serves as a unique identifier. If the message verifies as having been signed by the person with knowledge of the private key, we can narrow down who sent the message to one. But any number of people may have knowledge of the public key, and all of them can therefore verify the identity of the sender.

Symmetric cryptography Asymmetric cryptography

Figure 1: Symmetric vs. asymmetric keys in a meshed network

I.3. How asymmetric cryptography is used in digital signatures and certificates

Digital signatures and certificates are particularly common applications of authentication with asymmetric cryptography. A digital signature is a transform performed on a message using the private key, whose integrity may be verified with the public key.

Figure 2: A digital signature

Again, the unique properties of asymmetric cryptography make it particularly useful for generating digital signatures. The correct signature may only be generated with the private key. Knowledge of the public key is only useful for verifying the signature. Any number of people

may have knowledge of the public key for verification purposes, without compromising the private key. Since only one entity knows the private key, the private key serves as proof of the sending party's identity, and guarantees the integrity of messages they send.

A digital certificate is a piece of information which is digitally signed by a trusted third party, or certificate authority (CA), and which contains critical identification information, vouching for the identity of an entity. Digital certificates often themselves contain a public key corresponding to the private key the entity itself uses to prove its identity—the well-known web server certificates are examples of this type.

Figure 3: A certificate is made up of a server URL and a server public key

I.4. Encryption using asymmetric cryptography

Asymmetric encryption schemes are used in a variety of applications. Probably the most visible, well-known application is in encrypted email, in peer-to-peer 'keyring' schemes such as Pretty Good Privacy (PGP).

In asymmetric encryption schemes, the public key is used for encrypting messages; these messages, once encrypted, can only be decrypted with the private key. So the recipient publishes or distributes the public key corresponding to a private key of which only they have knowledge. Anyone wishing to communicate securely with the holder of the private key encrypts his or her message using the public key. Only the recipient may decrypt and use the message; other holders of the public key cannot.

This feature of asymmetric cryptosystems greatly simplifies key exchange. In a large network of N communicating entities, if it is fully meshed, maintaining unique symmetric keys for each communicating pair of entities would require the management of (nxn-1) ÷ 2 keys. Using asymmetric cryptography, this quantity can be reduced to N key pairs. In a group of 1,000 users, it's the difference between managing 1,000 key pairs or 499,500 keys.

Functions whose inverse is significantly more difficult

In all asymmetric cryptographic schemes, this property — the property that one key is used for encryption, and another for decryption, and the decryption key cannot be found from the encryption key — is derived from the use of mathematical functions whose inverse is extremely difficult to calculate.

You may understand an asymmetric cryptographic key pair as a pair of numbers which have some relationship associated with a mathematical function which is relatively easy to compute in one direction, but whose inverse is in practical terms intractable. This feature — the function which is tractable in one direction, but intractable in the other, is common to all asymmetric cryptosystems, including ECC.

II. RSA

II.1. Introduction

The RSA cryptosystem, invented by Ron Rivest, Adi Shamir and Len Adleman was first publicized in the August 1977 issue of *Scientific American*. The cryptosystem is most commonly used for providing privacy and ensuring authenticity of digital data. These days RSA is deployed in many commercial systems. It is used by web servers and browsers to secure web traffic, it used to ensure privacy and authenticity of email, it is used to secure remote login sessions, and it is at the heart of electronic credit-card payment systems. In short, RSA is frequently used in applications where security of digital data is a concern.

II.2. The integer factorization problem

The first asymmetric cryptosystem to have seen widespread use is also one of the most accessible illustrations of this principle in action.

RSA gets its security from the difficulty of factoring very large numbers. The difficulty of getting the plaintext message back from the ciphertext and the public key is related to the difficulty of factoring a very large product of two prime numbers.

As an illustration of this: imagine you were to take two very large prime numbers — say, 200 digits long, and were then to multiply them together. Now the result you get has two particular properties:

- it is very large(about 400 digits in length)
- it has two, and exactly two factors, both prime numbers – the two primes you just multiplied together

You can easily — given the two prime numbers from which you start — find the product. But finding the primes given only the product is more difficult. So much more, in fact, that once the numbers get adequately large, it is almost impossible to find them. You simply cannot assemble enough computing power to do so.

So the multiplying of two large prime numbers together is the (relatively) easy forward function in this asymmetric algorithm. Its inverse — the factor finding operation — is considerably more difficult, and in practical terms, it's intractable.

The RSA system employs this fact to generate public and private key pairs. The keys are functions of the product and of the primes. Operations performed using the cryptosystem are arranged so that the operations we wish to be tractable require performing the relatively easy forward function — multiplication.

Conversely, the operations we wish to make difficult — finding the plaintext from the ciphertext using only the public key — require performing the inverse operation — solving the factoring problem.

III. THE DIFFIE HELLMAN/DSA CRYPTOSYSTEMS

III.1. Introduction

Diffie Hellman — along with the Digital Signature Algorithm (DSA) based on it — is another of the asymmetric cryptosystems in general use.

The Diffie-Hellman key agreement protocol (also called exponential key agreement) was developed by Diffie and Hellman in 1976 and published in the ground-breaking paper "New Directions in Cryptography." The protocol allows two users to exchange a secret key over an insecure medium without any prior secrets.

DSA is based on the discrete logarithm problem and is related to signature schemes that were proposed by Schnorr and ElGamal. While the RSA system can be used for both encryption and digital signatures the DSA can only be used to provide digital signatures.

In DSA, signature generation is faster than signature verification, whereas with the RSA algorithm, signature verification is very much faster than signature generation (if the public and private exponents, respectively, are chosen for this property, which is the usual case). It might be claimed that it is advantageous for signing to be the faster operation, but since in many applications a piece of digital information is signed once, but verified often, it may well be more advantageous to have faster verification.

ECC, in a sense, is an evolved form of Diffie Hellman. So to understand how ECC works, it helps to understand how Diffie Hellman works first.

III.2. The discrete logarithm problem

Diffie Hellman uses a problem known as the discrete logarithm problem as its central, asymmetric operation. The discrete log problem concerns finding a logarithm of a number within a finite field arithmetic system.

Prime fields are fields whose sets are prime — that is, they have a prime number of members. These are of particular interest in asymmetric cryptography because, over a prime field,

exponentiation turns out to be a relatively easy operation, while the inverse — computing the logarithm — is very difficult.

To generate a key pair in the discrete logarithm (DL) system, therefore, you calculate:

$y=(gx) \bmod p$

where p is a large prime — the field size. x and g are smaller than p. y is the public key. x is used as the private key. In Diffie Hellman, again, the operations we wish to make 'easy', or tractable, we harness to the operation in the field which is (relatively) easy — exponentiation. So encryption using the public key is an exponentiation operation. Decryption using the private key is as well. Decryption using the public key, however, would require performing the difficult inverse operation — solving the discrete logarithm problem.

The discrete logarithm problem, using the values in the equation above, is simply finding x given only y, g and p.

Expanding that thought slightly: someone has multiplied g by itself x times, and reduced the result into the field (performed the modulo operation) as often as necessary to keep the result smaller than p. Now, knowing y, g and p, you're trying to find out what value of x they used.

It turns out that for large enough values of p, where p is prime, this is extraordinarily difficult to do — much more difficult than just finding y from g, x and p.

If you grasp what's going on in the operations above, you're now in a position to grasp the basic math behind the DSA and discrete logarithm systems. And, by extension, you also understand some of the principles behind ECC. ECC — as we'll discuss in greater detail a little later — also uses a discrete log problem in a finite group. The difference is that ECC defines its group differently. And it is, in fact, the difference in how the group is defined — and particularly how the mathematical operations within the group are defined — that give ECC its greater security for a given key size.

III.3. Asymmetric cryptography as a fine balance

As noted above, in all of the asymmetric cryptosystems, the fact that the system works at all relies upon the comparative difficulty of doing two types of operations — a 'forward' operation which must be tractable, and an 'inverse' operation which must be in practical terms intractable. In fact, in all cases, the degree of difference between the difficulties of these operations actually depends in a quite precise way on the size of the key pairs that are being generated. Both operations get more difficult as the key is made longer. The inverse operation, however, gets much more difficult, much more rapidly.

In all asymmetric cryptosystems, as mentioned above, the key length is the parameter that determines how difficult are both the forward and inverse algorithms.

As described in the preceding sections, the common characteristic of all asymmetric cryptosystems is a function whose inverse is significantly harder. We are now in a position to expand upon this: in all cases, the hardness of the forward and inverse operations is actually defined as two functions on the key length — two functions describing an 'order of growth' of the difficulty of the forward and inverse algorithms.

So, to make more precise our previous description: asymmetric cryptosystems work because the inverse operation rapidly gets more difficult as key length increases than does the forward operation.

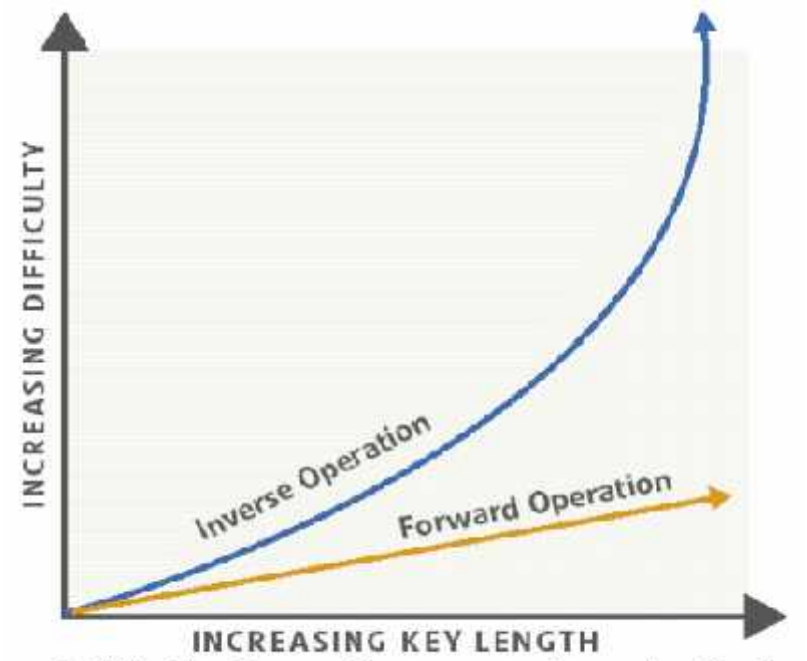


Figure 4: Difficulty of forward, inverse operation against key length

What this means in practical terms is: asymmetric cryptography is always a fine balance. Somewhere, for any given application, and any given cryptosystem, there is a key length x which is long enough that the users of the system can say the inverse operation is as hard as they need it to be to offer the level of security they desire — and yet key length x is not so long that the forward operations become unnecessarily unwieldy.

IV. ECC

IV.1. ECC as the answer for high security and for the future

We have to consider these three facets of the problem now:

- First, the fact that the security and practicality of a given asymmetric cryptosystems relies upon the difference in difficulty between doing a given operation and its inverse.
- Second, the fact that the difference in difficulty between the forward and the inverse operation in a given system is a function of the key length in use, due to the fact that the

difficulty of the forward and the inverse operations increase as very different functions of the key length; the inverse operations get harder faster.

- Third, the fact that as you are forced to use longer key lengths to adjust to the greater processing power now available to attack the cryptosystem, even the 'legitimate' forward operations get harder, and require greater resources (chip space and/or processor time), though by a lesser degree than do the inverse operations.

If you understand these three things, you are now in a position to grasp the advantages ECC offers over other asymmetric cryptosystems. ECC's advantage is this: its inverse operation gets harder, faster, against increasing key length than do the inverse operations in Diffie Hellman and RSA. What this means is: as security requirements become more stringent, and as processing power gets cheaper and more available, ECC becomes the more practical system for use. And as security requirements become more demanding, and processors become more powerful, considerably more modest increases in key length are necessary, if you're using the ECC cryptosystem — to address the threat. This keeps ECC implementations smaller and more efficient than other implementations. ECC can use a considerably shorter key and offer the same level of security as other asymmetric algorithms using much larger ones. Moreover, the gulf between ECC and its competitors in terms of key size required for a given level of security becomes dramatically more pronounced, at higher levels of security.

NIST guidelines for public key sizes for AES

ECC KEY SIZE (Bits)	RSA KEY SIZE (Bits)	KEY SIZE RATIO	AES KEY SIZE (Bits)
163	1024	1 : 6	
256	3072	1 : 12	128
384	7680	1 : 20	192
512	15 360	1 : 30	256

Supplied by NIST to ANSI X9F1

Figure 5: Equivalent key sizes for ECC and RSA

IV.2. What ECC is

Elliptic Curve Cryptography, as described above, is a relative of discrete logarithm cryptography. The DL system, as described above, relies upon the difficulty of the discrete logarithm problem — a logarithm problem calculated within the multiplicative group of a finite field — to frustrate attempts to use the public key to compromise the private one. ECC uses groups and a logarithm problem too.

What ECC also offers, however, is a difference in the method by which the group is defined — how the elements of the group are defined, and how the fundamental operations on them are

defined. It's the difference in the way the group is defined—both the numbers in the set and the definitions of the arithmetic operations used to manipulate them—that give ECC its more rapid increase in security as key length increases. To clarify this point, we'll describe briefly how elliptic curves are defined.

IV.3. Elliptic Curves

The way that the elliptic curve operations are defined is what gives ECC its higher security at smaller key sizes. An elliptic curve is defined in a standard, two dimensional x,y Cartesian coordinate system by an equation of the form:

$y^2 = x^3 + ax + b$ (1)

The graphs turns out to be gently looping lines of various forms.

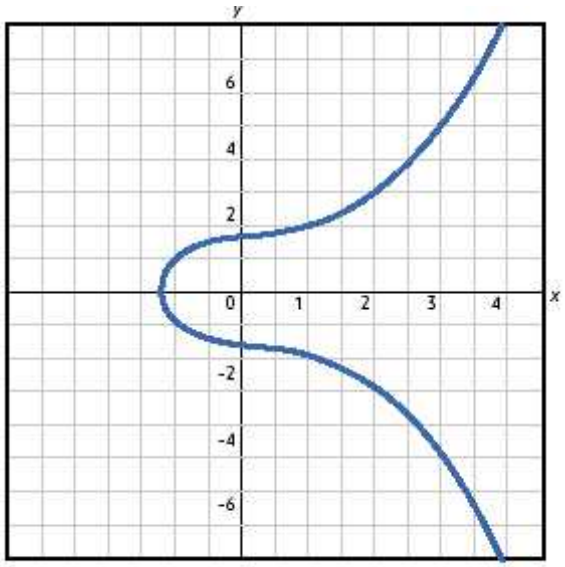


Figure 6: an elliptic curve

In elliptic curve cryptosystems, the elliptic curve is used to define the members of the set over which the group is calculated, as well as the operations between them which define how math works in the group.

Public-key cryptography is based on the intractability of certain mathematical problems. Early public key systems, such as the RSA algorithm, are secure assuming that it is difficult to factor a large integer composed of two or more large prime factors. For elliptic-curve-based protocols, it is assumed that finding the discrete logarithm of a random elliptic curve element with respect to a publicly-known base point is unfeasible. The size of the elliptic curve determines the difficulty of the problem. It is believed that the same level of security afforded by a RSA-based system with a large modulus can be achieved with a much smaller elliptic curve group. Using a small group reduces storage and transmission requirements.

For current cryptographic purposes, an elliptic curve is a plane curve which consists of the points satisfying the equation (1) along with a distinguished point of infinity, denoted ∞.(The

coordinates here are to be chosen from a fixed finite field of characteristics not equal to 2 or 3, or the curve equation will be somewhat more complicated.) This set together with the group operation of the elliptic group theory form an Abelian group, with the point at infinity as identity element. The structure of the group is inherited from the divisor group of the underlying algebraic variety.

The entire security of ECC depends on the ability to compute a point multiplication and the inability to compute the multiplicand given the original and product points. The U.S. National Security Agency has endorsed ECC by including schemes based on it in its Suite B set of recommended algorithms and allows their use for protecting information classified up to top-secret level with 384-bit keys. While the RSA patent expired in 2000, there are patens in force covering certain aspects of the ECC technology, though some argue that the Federal elliptic curve digital signature standard and certain practical ECC-based key exchange schemes(including ECDH) can be implemented without infringing them.

IV.4. Point multiplication

The dominant operation in ECC cryptographic schemes is point multiplication. This is the operation which is the key to the use of elliptic curves for asymmetric cryptography – the critical operation which is quite simple but whose inverse(the elliptic curve discrete logarithm problem) is very difficult. ECC arranges itself so that when you wish to perform an operation the cryptosystem should make easy – encrypting a message with the public key, decrypting it with the private key – the operation you are performing is point multiplication.

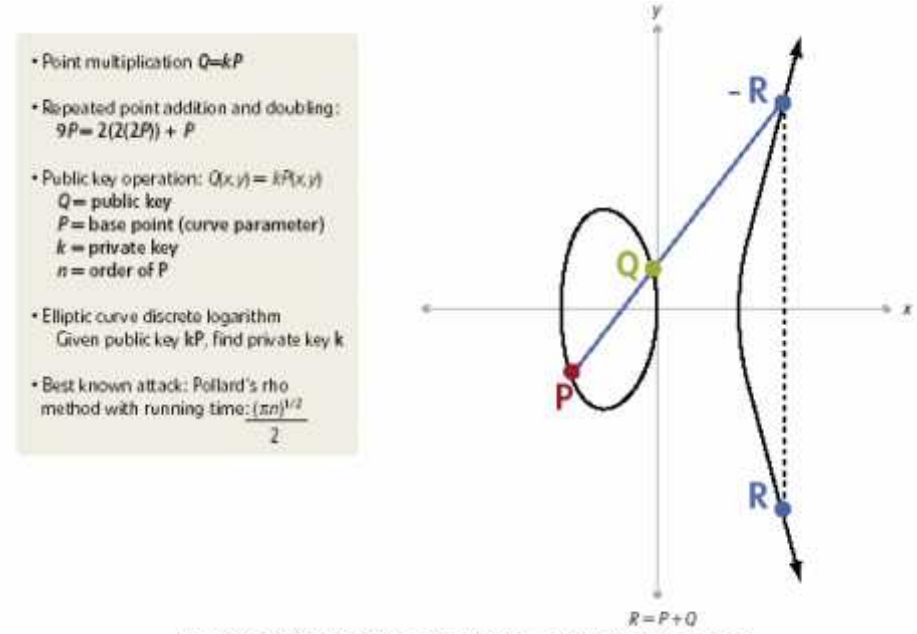


Figure 7: Eliptical curve cryptography

IV.5. The elliptic curve discrete logarithm problem

The inverse operation to point multiplication – finding a log in a group defined on a elliptic curve over a prime field – is defined as follows: **given points Q and P, find the integer k such that Q= kP.**

This is the elliptic curve discrete logarithm problem – and this is the inverse operation in the cryptosystem – the one you effectively have to perform to get the plaintext back from the ciphertext, given only the public key.

Now naively the obvious, certain way of finding k would be to perform repeated addition — operations — stepping through P, 2P, 3P, and so on, until you find kP. You'd start by doubling P, then adding P to 2P finding 3P, then 3P to P finding 4P and so on. This is the brute force method. The trouble with this is if you use a large enough prime field, the number of possible values for k becomes inconveniently large. So inconveniently large that it's quite practical to create a sufficiently large prime field that searching through the possible values of k would take all the processor time currently available on the planet thousands of years.

IV.6. Cryptographic schemes for ECC

Several discrete algorithm-based protocols have been adapted to elliptic curves, replacing a finite field with an elliptic curve:

- the elliptic curve Diffie-Hellman key agreement scheme is based on the Diffie-Hellman scheme.
- The elliptic curve digital signature algorithm is based on the digital signature algorithm
- The ECMQV key agreement scheme is based on the MQV key agreement scheme.

V. THE FUTURE OF ELLIPTIC CURVE CRYPTOGRAPHY

V.1. Present implementations of ECC

At the time of its discovery, the ECC algorithm was described and placed in the public domain. What others found was that while it offered greater potential security it was slow. Certicom focused its efforts on creating better implementations of the algorithm to improve its performance. After many years of research, Certicom introduced the first commercial toolkit to support ECC and make it practical for use in a variety of applications.

Other cryptographers have also become interested in ECC. Today Certicom sponsors the Centre for Advanced Cryptographic Research (CACR) at the University of Waterloo, Ontario along with the Canadian government, Mondex, MasterCard International, and Pitney Bowes. Each year the Centre sponsors an ECC workshop attended by over 100 top cryptographers to discuss advances in the field of elliptic curve cryptography.

At the RSA Conference 2005, the National Security Agency(NSA) announced Suite B which exclusively uses ECC for digital signature generation and key exchange. The suite is intended to protect both classified and unclassified national security systems and information.
Recently, a large number of cryptographic primitives based on bilinear mappings on various elliptic curve groups, such as the Tate and Weil pairings, have been introduced. Schemes based on these primitives provide efficient identity-based encryption as well as paired-based signatures, signcryption, key agreement and proxy re-encryption.

V.2. Future implementation of elliptic curves in cryptography

Future implementation of elliptic curves in cryptography target the following :

- Implementation of an elliptic curve crypto library and security architectures for various platforms ranging from small sensors to high-performance web servers
- Implementation of a common hardware architecture for accelerating ECC as well as RSA
- Enabling broad industry adoption of ECC by
    1. promoting ECC standardization within SSL, the dominant security protocol used on the Internet
    2. contributing ECC technology to OpenSSL and NSS/Mozilla the two most open source cryptographic libraries

ECC for portable devices and applications

Because ECC can achieve better results than RSA with smaller keys makes it a stronger option than the RSA and discrete logarithm systems for the future. And this, in the end, is why ECC is such an excellent choice for doing asymmetric cryptography in portable, necessarily constrained devices right now.
For example, the recommended RSA key size for most applications is 2048 bits. For equivalent security using ECC, you need a key of only 224 bits. The difference becomes more and more pronounced as security levels increase(the hardware gets faster and the security key size increases). A 384-bit ECC key matches a 7680-bit RSA key for security.

CONCLUSION

Elliptic Curve Cryptography(ECC) was discovered in 1985 by Victor Miller(IBM) and Neil Koblitz(University of Washington) as an alternative mechanism for implementing public-key cryptography. Public-key algorithms create a mechanism for sharing keys among large numbers of participants or entities in a complex information system. Unlike other popular algorithms such as RSA, ECC is based on discrete logarithms that is much more difficult to challenge at equivalent key lengths.

The smaller ECC keys mean the cryptographic operations that must be performed by the communicating devices can be squeezed into considerably smaller hardware, that software applications may complete cryptographic operations with fewer processor cycles, and operations can be performed that much faster, while still guaranteeing equivalent security. This means, in turn, less heat, less power consumption, less real estate consumed on the printed circuit board, and software applications that run more rapidly and make lower memory demands. Leading in turn to more portable devices which run longer, and produce less heat.

In short, if you're trying to make your devices smaller—and if you need to do asymmetric cryptography, you need ECC. If you're trying to make them run longer on the same battery, and produce less heat, and you need asymmetric cryptography, you need ECC. And if you want an asymmetric cryptosystem that scales for the future, you want ECC. And if you just want the most elegant, most efficient asymmetric cryptosystem going, you want ECC. If you just want the most elegant, most efficient asymmetric cryptosystem going, you want ECC.

REFERENCES

1. http://www.deviceforge.com/articles/AT4234154468.html
2. http://en.wikipedia.org/wiki/Elliptic_curve_cryptography
3. http://searchsecurity.techtarget.com/sDefinition/0,,sid14_gci836964,00.html
4. http://www.faqs.org/rfcs/rfc3278.html
5. Implementing elliptic curve cryptography – Michael Rosing, Manning Publications 1999.

AN OVERVIEW ON FUTURE INTRUSION DETECTION SYSTEMS

CPT eng. Cosmin IVAN

INTRODUCTION

Traditionally, firewalls and anti-virus programs try to block attacks, while Intrusion Detection Systems, (IDSs) identify attacks as they occur. Such techniques are crucial to network security, but have limitations. A firewall can stop attacks by blocking certain port numbers, but it does little to analyse traffic that uses allowed port numbers. IDSs can monitor and analyse traffic that passes through open ports, but do not prevent attacks.

The next generation systems do not just detect attacks, they try to stop them. Intrusion Prevention Systems (IPSs), are similar to IDSs in that both systems aim to distinguish unauthorized activity from normal activity. An IPS, like an IDS, has a set of signatures or predefined conditions that, when met, trigger a response. Those systems have a similar manner of processing with IDSs. The response itself, however, differs, and is what mostly differentiates an IPS from an IDS. With the proliferation of sophisticated attacks and the discovery of new vulnerabilities, new methods are needed to protect precious data and network resources. Intrusion Prevention Systems use new *proactive* approaches to stop intrusions before any damage is done.

I. INTRUSION DETECTION SYSTEMS

Intrusion Detection Systems are designed to alert a human to potentially unauthorized activity. The underlying concept of an intrusion detection system is that a human must be present in the system to determine when activity is truly unauthorized. IDSs can be divided into two main categories, based on the mechanism that triggers the alarm: **anomaly detection-based** (or statistical anomaly detection-based) IDSs and **misuse detection-based** (or knowledge based, signature based, pattern matching) IDSs. The first ones compare observed activity against expected normal usage profiles (for users, groups of users, applications, etc). Audit event records which fall outside the definition of normal behaviour are considered anomalies. The second ones seek attack signatures in the audit data which announce known misuse. They are based on a set of rules that match typical patterns of exploits used by attackers. Snort is such a system and actually the most widely deployed intrusion detection technology worldwide. Another system is *Bro,* which is a stand-alone system for detecting intruders in real-time by passively monitoring a

network link over which the intruder's traffic transits. We now briefly describe some techniques that are used in each of these two categories of intrusion detection systems.

I.1. Anomaly detection systems

Threshold monitoring sets values for metrics defining acceptable behaviour (e.g. fewer than some number of failed logins per time period). Thresholds provide a clear, understandable definition of unacceptable behaviour, but it is difficult to establish proper threshold values and time intervals over which to check. Approximation can result in a high rate of false positives or high rate of false negatives across a non-uniform user population.

User work profiling maintains individual work profiles to which the user is expected to adhere in the future. If users are assigned to specific work groups that demonstrate a common work pattern and hence a common profile, then we are talking about **group work profiling**. A group profile is calculated based upon the historic activities of the entire group.

Executable Profiling seeks to monitor executables' use of system resources, especially those whose activity cannot always be traced to a particular originating user. Viruses, Trojan horses, worms, trapdoors, logic bombs and other such software attacks are addressed by profiling how system objects such as files and printers are normally used, not only by users, but also by other system subjects on the part of users.

Neural networks can also be trained to recognize abnormal traffic. Radial basis function networks can be used to improve the performance of intrusion detection in anomaly detection with a high detection rate and a low false positive rate.

I.2. Misuse detection systems

Simple systems that scan for byte – sequence signatures. This is the simpler way to detect misuse.

Expert systems may be used to code misuse signatures (even simple byte – sequences signatures) as if-then implication rules. Signature analysis focuses on defining specific descriptions and instances of attack-type behaviour to flag.

Data mining techniques can be used to discover consistent and useful patterns of system features that describe program and user behaviour, and use the set of relevant system features to compute (inductively learned) classifiers that can recognize known intrusions.

Model based reasoning attempts to combine models of misuse with evidential reasoning to support conclusions about the occurrence of a misuse. This technique may be useful for identifying intrusions which are closely related, but whose audit trails patterns are different.

Keystroke monitoring is a very simple technique that monitors keystrokes for attack patterns.

I.3. Advantages and Disadvantages

The advantage of anomaly detection is that it can detect previously unknown attacks and insider attacks, without the need for signatures. On the other hand, the large number of false positives is the most important shortcoming of such systems. Furthermore, besides being complicated and hard to understand, building and updating profiles also require a lot of work.

Misuse detection is considered more accurate, since there is a known database of exploits and so there are few false positives. However, this database has to be updated continuously to keep up with new attacks. Furthermore, misuse detection systems are unable to detect any future (unknown) intrusions that have no matched patterns stored in the system. Insider attacks may also go undetected. Besides the general evaluation done so far, it is expected that every implementation has its own shortcomings. For example, EarlyBird, a content-based IDS that works on all incoming packets and uses content prevalence to determine worm substrings, cannot handle polymorphic attacks. Moreover, since it handles all incoming packets, it needs to use sampling and estimation in computing address dispersion and content prevalence, both of which may lead to misdetecting worms[54].

Other systems like Autograph may not be suitable for UDP based attacks like Slammer. Finally, as said before, IDSs do not prevent attacks. They just sit on a network or host, silently monitor the traffic and only alert when an attack is detected. They cannot stop or even slow down an attack in progress.

II. INTRUSION PREVENTION SYSTEMS

Any device, hardware or software, which has the ability to detect attacks, both known and unknown, and prevent an attack from being successful, is an Intrusion Prevention System. IPSs are *proactive*, *inline* devices that can drop packets or even disconnect connections before reaching the host and block all traffic with the same IP source, if they detect illegal activity. They rapidly end the intrusion and minimize the overall time before the network is back to normal. Through using multiple detection methods and utilizing its position in the line of network traffic, an IPS can detect attacks and intrusions more accurately and reliably. By relying less on signatures and more on intelligent methods of detection, the IPS generates far fewer false alarms.

[54] P. Folga, M. Sharif, R. Perdisci, O. Kolesnikov, W. Lee, "Polymorphic Blending attacks", *pages 241-256, 2006*

II.1. Requirements

Some requirements of an IPS are the following:

- **Accuracy** is one of the most important requirements in an IPS. Having false positives may be extremely annoying in an IDS, but it is absolutely unacceptable in an IPS. False positives are typically generated by systems that rely on a single detection method, and by ones that cannot be configured at different levels to fit into the operational environment. If legitimate traffic is blocked, then problems arise for authorized users. This creates *self-inflicted* DoS attacks, Denial of-Service attacks that originate from the prevention system itself. Sometimes a valid business transaction may act like an attack. In such a case, an "offending" packet may first be dropped and then the entire dataflow. If the source IP is that of a critical business partner, the partner will be prevented from accessing resources.
- **Performance** is also important for IPSs. The problem with inline intrusion prevention is that it tends to become a network bottleneck. All network traffic needs to flow through these devices, and if they don't operate quickly enough, they drop packets, increasing the possibility of false negatives. Thus, they have to work at wire speed.
- **Anticipation of Unknown Attacks and Easy Signature Update for New Attacks:** An IPS must provide flexible methods to update new attack signatures, as well as capabilities to respond to entirely new classes of attacks. In addition, IPS systems should have methods that can respond to new attacks without requiring signature updates. Such methods may include *inverse exclusion*, where all requests, except those that are legal for a given destination, are dropped. Another method is *protocol validation*, where illegal request methods are dropped. Attack-independent blocking is another method where hostile attackers are identified, and all traffic from the attacker is dropped, regardless of whether the attacks are known or not.
- An IPS should be **reliable** and high **available**. Reliability refers to the ability of a system to perform its functions properly without interfering with other systems on the network. Availability is the amount of downtime of the system, due to shutdown, crashes, or maintenance. An IPS gives the network security administrator many options, since it is capable of not only detecting attacks and intrusions, but also directly affecting network traffic through limiting or blocking. It must give the administrator an **easy interface** for setting and changing configurations on the devices.
- IPSs should also **cooperate** with firewalls, antivirus systems, etc.

II.2. Classification

Intrusion Prevention Systems can be divided in two main categories, each of which uses different prevention approaches and methods.

Host-based IPSs (HIPSs)

Host-based intrusion prevention systems are similar to antivirus products, but actively respond to any observed intrusion activity. An IPS usually sits between the kernel and the application utility software that issues requests to the kernel of the O.S. Actions of a HIPS include blocking the request or denying access to the kernel, for activities with high certainty as an intrusion. Some prevention approaches are:

- **Software-based heuristics:** This approach is similar to IDS anomaly detection using neural networks to act against new or unknown types of intrusion.
- **Sandbox approach:** Mobile code like ActiveX or Java applets is quarantined in a sandbox, an area with restricted access to the rest of the system. This system then runs the suspect mobile code in the sandbox and monitors its behaviour. If the code violates a predefined security policy, it is stopped and prevented from executing.
- **Kernel-based protection:** The kernel controls access to system resources like memory, input/output devices and CPU. Programming errors enable exploits like buffer-overflow attacks to overwrite kernel memory space and crash or take over computer systems. To prevent these types of attacks, a software agent is loaded between the user application and the kernel. The software agent intercepts system calls to the kernel, inspects them against an access control list defined by a policy and then either allows or denies access to resources.
- **Address space randomization** is a technique used to fortify systems against buffer overflow attacks. The idea is to introduce artificial diversity by randomizing the memory location of certain system components.

Network-based IPSs (NIPSs)

The other class of IPSs, network-based intrusion prevention systems, generally consists of appliance-based systems that sit *inline* and block suspicious traffic upon detecting an attack. They statefully analyze packet content and block certain packets that match a signature and alert on others. A NIPS protection is based on the content of packets. Some NIPS prevention methods are presented below.

- **Protocol anomaly detection** is used to ensure that packets adhere to the protocol and have no ambiguities. For example, by IP spoofing of FTP PORT commands, the attacker can tell the FTP server to open a connection to a victim's IP address and then transfer a Trojan horse to the victim. Checking for a match between the IP address in the FTP

PORT command and the client's IP address can prevent this anomaly. Protocols should be well defined, thus permitting deviations from the standard use to be detected with good accuracy. Furthermore, fragmented packets could slip through the network and when reassembled at the host, unleash their evil intent. *Normalization*, the process of removing exploitable ambiguities in network traffic before the traffic is evaluated for malicious code, can combat this tactic and ensure that traffic interpreted by the NIPS is the same as that seen by the host end.

- **Traffic anomaly detection** is based on deviations from expected behaviour. Attackers often conduct a port or network scan as a precursor to an attack. Optimizations in the scanning techniques have made it possible that worms can affect the entire vulnerable population in 10s of seconds; so fast that no human-mediated counter response is possible[55]. NIPS implement frequency and threshold triggers that alert to such scanning activity, increasing the likelihood that an attack can be prevented.
- **State-based signature detection** looks at relevant portions of traffic by tracking state, and based on the context specifies by the user, detects attacks. It is not completely automated as the user needs to have some prior knowledge about the attack. For example the Love Letter worm can be detected by a rule that would read as follows: "Look for 'ILOVEYOU' in the subject field only, ignore this string anywhere else in the email". Of course false positives can be generated in this case, since harmless emails with the same title may have been sent.
- **Pattern matching using regular expressions** can detect attack patterns that are slightly different from the fixed ones. A minor change like a space or a tab in the attack could be enough to evade detection. Regular expressions provide wild-card and complex pattern matching, and are able to prevent attacks.
- **Signature detection** is used in conjunction with the above mention techniques to ward off combined attack types, which are very common on today's networks.

III. IMPLEMENTATIONS

We will briefly take a look into some implementations of Intrusion Prevention Systems.

SafeCard is a Gigabit IPS, able to cope with all levels of abstraction in communication (packets, streams, layer-7 data units etc), designed as a compound, pipelined IPS built from independent function elements. The functional architecture of this system is shown below.

Each stage in the pipeline drops traffic that perceivers as malicious. In the first stage, packets are

[55] S. Staniford, V. Paxson, and N. Weaver, "How to own the internet in your spare time", *Symposium,* 2002

filtered based on the header fields (protocols, ports). In stage 2, TCP streams are reconstructed. Then the streams are matched against Snort-like patterns using Ruler, a pattern matching language. Unmatched traffic is further inspected by Prospector, a protocol specific detector that can stop polymorphic buffer overflow attacks. In stage 5, behavioural aspects of traffic are taken into account. Finally, the clean traffic is transmitted.

In SNORT (IDS), every captured packet goes through the following steps: header information decoding at the different layers, application of preprocessor functions like IP fragment or TCP stream reassembly, evaluation of a subset of rules according to the information from step one, and finally if a match is found, the corresponding action is carried out. We see that SNORT does not take into account any behavioural aspects of traffic e.g. unusual amounts of traffic. Moreover, SafeCard uses superior regular expression matching and checks packets for higher protocol specific rules if they exist.

CardGuard is a signature detection system for intrusion detection and prevention that scans the entire payload of packets for suspicious patterns, and is implemented in software. It is non-intrusive in the sense that no cycles of the host CPUs are used for intrusion detection and the system operates at Fast Ethernet link rate. Again, TCP flows are first reconstructed before they are scanned with the Aho-Corasick algorithm.

Some other implementations are available in the market already. Radware provides solutions that block attacks and malicious activity before they get anywhere near applications, with advanced security intelligence based on, signature vulnerability, behaviour-based traffic anomaly and protocol anomaly. Cisco is also in the market with true Intrusion prevention systems (IPS Sensor Software). Systems that detect and slow down the malicious code based on its behaviour, even in the form of an unknown attack, do not fit our definition of IPS. Such a system is the Virus Throttle module from HP, which mitigates harm to other systems, and other systems that focus on the harm already done to an individual machine.

IV. DISCUSSION

Intrusion prevention systems may seems like a great idea in papers, but when it comes to practice, then problems show up. The accuracy of such systems plays a very important role. Thus, techniques used in IDSs are simply not enough and more sophisticated forms of analysis are needed. As mentioned before, a valid transaction can be flagged as malicious. Same results may come from peer to peer applications. The fact that multiple connections are attempted or established to the same host may look suspicious and lead to blocking the traffic. A self inflicted denial of service attack is very possible. EarlyBird has a solution; it computes a hash based on the content as well as the destination port to distinguish worms from valid p2p traffic and avoid

false positives.

The following is also possible: an attacker may *spoof* an IP address and then try to infect a system. After detecting the malicious code and dropping some packets or the entire data flow, some IPSs may also block the IP from further reaching the system. Every other data coming from the same host will not be processed. But what if this IP belongs to a trusted machine? Until when will it be blocked and banned from accessing the resources of the system? It is clear so far that IPSs can be DoSsed. Moreover, they can be detected and bypassed. So IPSs still have a long way to go for improvement.

Other forms of attack may include false training in a way to allow attacks in the future. A black hat attacker may include in his emails pieces of code of an attack just to train the system *not* to recognize it as suspicious or malicious, and increase the chances that later, attacks with the same content will go undetected.

Another serious problem with IPS, especially with NIPSs as they sit inline, is that they automatically become a single point of failure for the network. In case the system fails, unacceptable values of latency and/or self inflicted DoS conditions may be observed. However, it is the case that attacks cannot get through the system even in a failure, in contrast with IDSs, where attacks may go undetected in similar cases.

Many implementations of IPSs are software based. Minos is a hardware project that employs taint analysis (attempts to avoid false positives) to discover illegitimate use of 'foreign' data. By looking at physical addresses only, it may detect certain exploits, such as a register spring attacks. Once misbehaviour is detected, Minos makes no attempt to generate signatures. One of the reasons for this is that by aiming at a hardware solution, Minos has had to sacrifice flexibility for performance, as the amount of information on the hardware level is very limited. Another implementation is Intrusion detection analysis which is distributed to the network node IDS, running in hardware on the end hosts. An NNIDS, when implemented on the network interface hardware, can function independently of the host operating system to provide better protection with less overhead than software implementations.

As traffic volumes and rates continue to race forward, the requirement for inline processing exists and more sophisticated forms of analysis are needed. Given these pressures, the nature of using hardware to support network security analysis should be rethought[56]. Using massively parallel computing elements can provide the necessary performance and avoid the single point of failure feature.

---

[56] V. Paxson, K. Asanovic, S. Dharmapurikar, J. Lockwood, R. Pang, R. Sommer, N. Weaver , "Rethinking Hardware Support for Network Analysis and Intrusion Prevention", *1st Workshop on Hot Topics in Security (HotSec '06), Vancouver, Canada, July 2006*

CONCLUSIONS

Intrusion prevention systems can be considered an evolution of IDS technology. Their proactive capabilities help keep the networks safer from more sophisticated attacks. However, there is still much to be done. The battle against false positives is not easy; neither is it easy to handle all the traffic at wire speed and perform operations without adding latency. Todays attacks will always be a threat with a difficult cure to find.

Bulletproof security does not exist. Attacks can still slip through such systems and no amount of automation can replace trained and vigilant personnel. No matter what, we can't expect some piece of software or hardware to fix everything for us.

REFERENCES

1) Paul Innella, "The Evolution of Intrusion Detection Systems", http://www.securityfocus.com/infocus/1514
2) P. de Boer, M. Pels, "Host-based Intrusion Detection Systems", www.os3.nl, February 2005
3) Charles Iheagwara and Andrew Blyth, "Future Directions in the Development of Intrusion Detection Systems", https://isaca-washdc.sharepointsite.net/webresources/Articles/Future%20Directions%20in%20the%20Development%20of%20Intrusion%20Detection%20Systems.htm
4) Thomas Goeldenitz, "IDS – Today and tomorrow", http://www.sans.org/reading_room/whitepapers/detection/ids-today-tomorrow_351, January 2002
5) Neil Desai, "Intrusion Prevention Systems: the Next Step in the Evolution of IDS", http://www.symantec.com/connect/articles/intrusion-prevention-systems-next-step-evolution-ids, February 2003
6) Charalampos Zois, "Intrusion detection systems", http://www.few.vu.nl/~czois/docs/IPS_final.pdf, fall 2006.

BIOMETRICS AND SECURITY

LT Cornel ANTOCHE

INTRODUCTION

Imagine you're James Bond, and you have to get into a secret laboratory to disarm a deadly biological weapon and save the world. But first, you have to get past the security system. It requires more than just a key or a password -- you need to have the villain's irises, his voice and the shape of his hand to get inside.

You might also encounter this scenario, minus the deadly biological weapon, during an average day on the job. Airports, hospitals, hotels, grocery stores and even Disney theme parks increasingly use **biometrics** – "consists of methods for uniquely recognizing humans based upon one or more intrinsic **physical** or **behavioral** traits. In computer science, in particular, biometrics is used as a form of *identity access management* and *access control*. It is also used to identify individuals in groups that are under surveillance"[57] -- for added security.

In this work, I'll explain about biometric systems that use handwriting, hand geometry, voiceprints, iris structure and vein structure. You'll also see why more businesses and governments use the technology and whether Q's fake contact lenses, recorded voice and silicone hand could really get James Bond into the lab (and let him save the world).

You take basic security precautions every day, you use a key to get into your house and log on to your computer with a username and password. You've probably also experienced the panic that comes with misplaced keys and forgotten passwords. It isn't just that you can't get what you need, if you lose your keys or jot your password on a piece of paper, someone else can find them and use them as though they were you.

I. HOW BIOMETRICS WORKS

Instead of using something you have (like a key) or something you know (like a password), biometrics uses **who you are** to identify you. Biometrics can use **physical characteristics**, like your face, fingerprints, irises or veins, or **behavioral characteristics** like your voice, handwriting or typing rhythm. Unlike keys and passwords, your personal traits are extremely difficult to lose or forget. They can also be very difficult to copy. For this reason, many people consider them to be safer and more secure than keys or passwords.

[57] http://en.wikipedia.org/wiki/Biometric

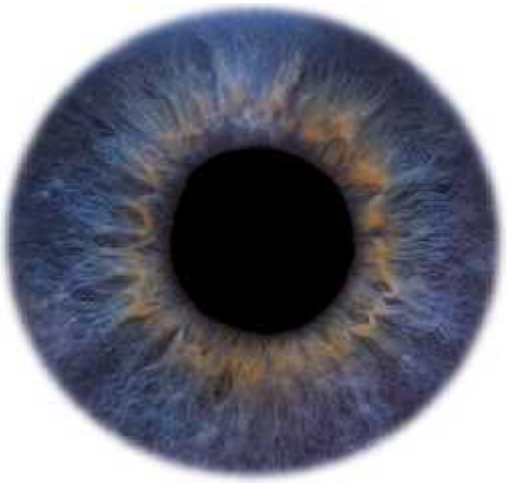


Figure 1 - Biometrics uses unique features, like the iris of your eye, to identify you.

Figure 2 - Fingerprint scanner, bringing biometric security to the home

Biometric systems can seem complicated, but they all use the same three steps:

- **Enrollment**: The first time you use a biometric system, it records basic information about you, like your name or an identification number. It then captures an image or recording of your specific trait.
- **Storage**: Contrary to what you may see in movies, most systems don't store the complete image or recording. They instead analyze your trait and translate it into a code or graph. Some systems also record this data onto a smart card that you carry with you.
- **Comparison**: The next time you use the system, it compares the trait you present to the information on file. Then, it either accepts or rejects that you are who you claim to be.

Systems also use the same three components:

- A **sensor** that detects the characteristic being used for identification
- A **computer** that reads and stores the information
- **Software** that analyzes the characteristic, translates it into a graph or code and performs the actual comparisons

Biometric security systems, like the fingerprint scanner available on the IBM ThinkPad T43 (right), is becoming more common for home use.

Handwriting

At first glance, using handwriting to identify people might not seem like a good idea. After all, many people can learn to copy other people's handwriting with a little time and practice. It seems like it would be easy to get a copy of someone's signature or the required password and learn to forge it.

But biometric systems don't just look at how you shape each letter; they analyze the act of writing. They examine the pressure you use and the speed and rhythm with which you write. They also record the sequence in which you form letters, like whether you add dots and crosses as you go or after you finish the word.

Figure 3 - Tablet PC has a signature verification system

Unlike the simple shapes of the letters, these traits are very difficult to forge. Even if someone else got a copy of your signature and traced it, the system probably wouldn't accept their forgery.

A handwriting recognition system's sensors can include a touch-sensitive writing surface or a pen that contains sensors that detect angle, pressure and direction. The software translates the handwriting into a graph and recognizes the small changes in a person's handwriting from day to day and over time.

Hand and Finger Geometry

People's hands and fingers are unique -- but not as unique as other traits, like fingerprints or irises. That's why businesses and schools, rather than high-security facilities, typically use hand

and finger geometry readers to **authenticate** users, not to **identify** them. Disney theme parks, for example, use finger geometry readers to grant ticket holders admittance to different parts of the park. Some businesses use hand geometry readers in place of timecards.

Systems that measure hand and finger geometry use a digital camera and light. To use one, you simply place your hand on a flat surface, aligning your fingers against several pegs to ensure an accurate reading. Then, a camera takes one or more pictures of your hand and the shadow it casts. It uses this information to determine the length, width, thickness and curvature of your hand or fingers. It translates that information into a numerical template.

Figure 4 - A hand geometry scanner

Hand and finger geometry systems have a few strengths and weaknesses. Since hands and fingers are less distinctive than fingerprints or irises, some people are less likely to feel that the system invades their privacy. However, many people's hands change over time due to injury, changes in weight or arthritis. Some systems update the data to reflect minor changes from day to day.

Voiceprints

For higher-security applications, biometric systems use more unique characteristics, like voices. Your voice is unique because of the shape of your vocal cavities and the way you move your mouth when you speak. To enroll in a voiceprint system, you either say the exact words or phrases that it requires, or you give an extended sample of your speech so that the computer can identify you no matter which words you say.

When people think of voiceprints, they often think of the wave pattern they would see on an oscilloscope. But the data used in a voiceprint is a sound **spectrogram**, not a wave form. A spectrogram is basically a graph that shows a sound's frequency on the vertical axis and time on the horizontal axis. Different speech sounds create different shapes within the graph. Spectrograms also use colors or shades of grey to represent the acoustical qualities of sound.

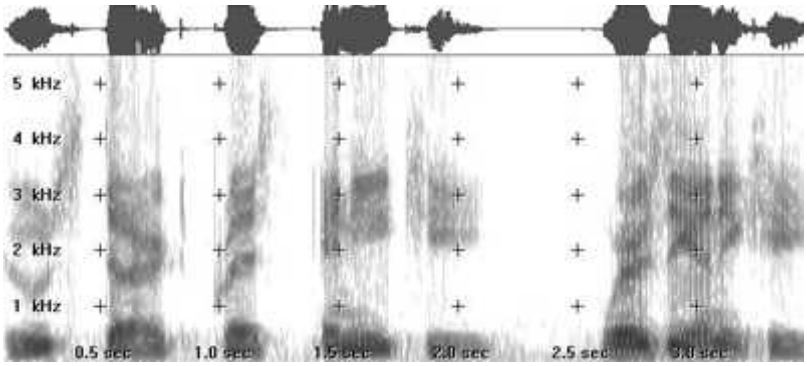


Figure 5 - Speaker recognition systems use spectrograms to represent human voices

Some companies use voiceprint recognition so that people can gain access to information or give authorization without being physically present. Instead of stepping up to an iris scanner or hand geometry reader, someone can give authorization by making a phone call. Unfortunately, people can bypass some systems, particularly those that work by phone, with a simple recording of an authorized person's password. That's why some systems use several randomly-chosen voice passwords or use general voiceprints instead of prints for specific words. Others use technology that detects the artifacts created in recording and playback.

Iris Scanning

Iris scanning can seem futuristic, but at the heart of the system is a simple CCD digital camera. It uses both visible and near-infrared light to take a clear, high-contrast picture of a person's iris. With near-infrared light, a person's pupil is very black, making it easy for the computer to isolate the pupil and iris.

When you look into an iris scanner, either the camera focuses automatically or you use a mirror or audible feedback from the system to make sure that you are positioned correctly. Usually, your eye is 10 centimeters to one meter from the camera. When the camera takes a picture, the computer locates:

- The center of the pupil
- The edge of the pupil
- The edge of the iris
- The eyelids and eyelashes

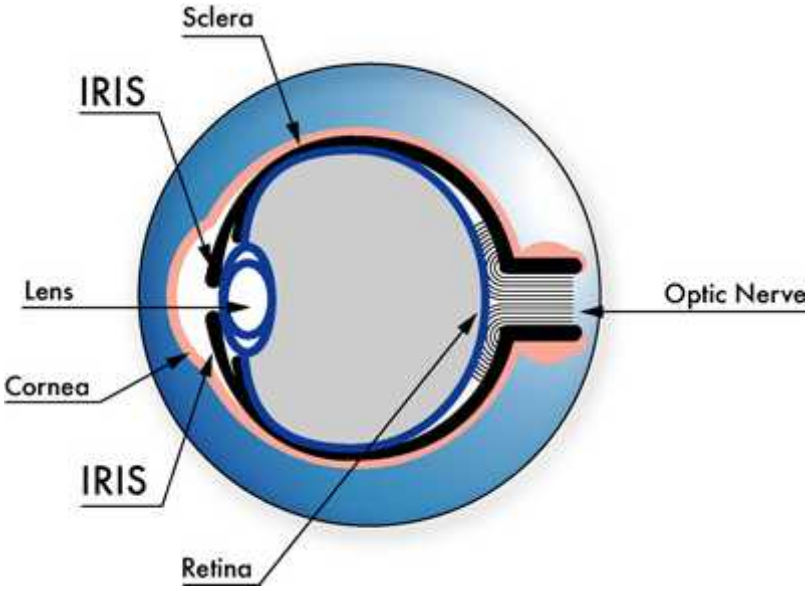


Figure 6 - Eye anatomy

It then analyzes the patterns in the iris and translates them into a code.

Iris scanners are becoming more common in high-security applications because people's eyes are so unique (the chance of mistaking one iris code for another is 1 in 10 to the 78th power[58]. They also allow more than 200 points of reference for comparison, as opposed to 60 or 70 points in fingerprints.

The iris is a visible but protected structure, and it does not usually change over time, making it ideal for biometric identification. Most of the time, people's eyes also remain unchanged after eye surgery, and blind people can use iris scanners as long as their eyes have irises. Eyeglasses and contact lenses typically do not interfere or cause inaccurate readings.

Vein Geometry

As with irises and fingerprints, a person's veins are completely unique. Twins don't have identical veins, and a person's veins differ between their left and right sides. Many veins are not visible through the skin, making them extremely difficult to counterfeit or tamper with. Their shape also changes very little as a person ages.

To use a vein recognition system, you simply place your finger, wrist, palm or the back of your hand on or near the scanner. A camera takes a digital picture using near-infrared light. The hemoglobin in your blood absorbs the light, so veins appear black in the picture. As with all the other biometric types, the software creates a reference template based on the shape and location of the vein structure.

[58] http://archives.cnn.com/2000/TECH/computing/07/24/iris.explainer/index.html

Figure 7 - Vein scanners use near-infrared light to reveal the patterns in a person's veins.

Scanners that analyze vein geometry are completely different from vein scanning tests that happen in hospitals. Vein scans for medical purposes usually use radioactive particles. Biometric security scans, however, just use light that is similar to the light that comes from a remote control. NASA has lots more information on taking pictures with infrared light.

Privacy and Other Concerns

Some people object to biometrics for cultural or religious reasons. Others imagine a world in which cameras identify and track them as they walk down the street, following their activities and buying patterns without their consent. They wonder whether companies will sell biometric data the way they sell e-mail addresses and phone numbers. People may also wonder whether a huge database will exist somewhere that contains vital information about everyone in the world, and whether that information would be safe there.

At this point, however, biometric systems don't have the capability to store and catalog information about everyone in the world. Most store a minimal amount of information about a relatively small number of users. They don't generally store a recording or real-life representation of a person's traits - they convert the data into a code. Most systems also work in only in the one specific place where they're located, like an office building or hospital. The

information in one system isn't necessarily compatible with others, although several organizations are trying to standardize biometric data.

In addition to the potential for invasions of privacy, critics raise several concerns about biometrics, such as:

1) **Over reliance**: The perception that biometric systems are foolproof might lead people to forget about daily, common-sense security practices and to protect the system's data.
2) **Accessibility**: Some systems can't be adapted for certain populations, like elderly people or people with disabilities.
3) **Interoperability**: In emergency situations, agencies using different systems may need to share data, and delays can result if the systems can't communicate with each other.

II. CASE STUDY - USING BIOMETRICS TO ACHIEVE IDENTITY DOMINANCE IN THE GLOBAL WAR ON TERRORISM

A fingerprint match identified the 20th hijacker. In December 2001, U.S. military forces detained Mohamed Al Kahtani as an enemy combatant on the field of battle in Southwest Asia. During repeated interrogations Kahtani denied being a combatant and offered an innocent explanation for his presence in the region. While Kahtani was in military custody, an FBI team fingerprinted him in much the same way law-enforcement officials routinely fingerprint criminal suspects in the United States. They took Kahtani's 10 "rolled" fingerprints; that is, one fingerprint of each digit recorded from nail to nail. This collection of biometric data eventually led U.S. investigators to believe Kahtani was the missing 20th hijacker in the terrorist attacks of 11 September 2001. The 9/11 Commission concluded that Kahtani was "[t]he operative likely intended to round out the team" for Flight 93, which crashed in Somerset County, Pennsylvania.[59]

Kahtani was identified because U.S. authorities matched the fingerprints taken from him in December 2001 to his fingerprints of 4 August 2001, when he arrived at Orlando International Airport on a Virgin Atlantic flight from London. During the immigration inspection at the airport, Kahtani, despite holding a valid U.S. visa, raised the suspicions of an alert immigration official. According to the 9/11 Commission, "Kahtani was denied entry by immigration officials because he had a one-way ticket and little money, could not speak English, and could not

[59] *The 9/11 Commission Report, The Final Report of the National Commission on Terrorist Attacks upon the United States*

adequately explain what he intended to do in the United States."[60] He received a "voluntary departure," which, in practical terms, meant officials placed him on a flight and returned him to Dubai. As part of the voluntary departure process, officials took prints from his two index fingers.

Once U.S. authorities biometrically linked Kahtani, the detainee in December 2001, to Kahtani, the foreigner who tried to enter the United States in August 2001, they had a valuable lead to pursue for counterterrorism and homeland security purposes. The Kahtani match raised an intriguing possibility: Investigators knew Mohamed Atta had been in Florida in August 2001. Could Atta be linked to Kahtani? Based on their review of surveillance camera footage taken at the airport on 4 August 2001, investigators matched a license plate to a car rented by Atta. Other corroboration established that Atta was at the airport terminal at the time Kahtani's flight arrived. Of course, Kahtani never volunteered this information during his many military interrogations. He stuck to his cover story. The fingerprint match provided the necessary actionable intelligence. As a result, a person the military encountered on a foreign field of battle was linked to a terrorist activity - the 9/11 attacks. This case study illustrates the importance of "identity dominance," which the U.S. military must embrace.

What is Identity Dominance?

Just as the U.S. military has established its superiority in other arts of war, now, working with other U.S. Government organizations, it must strive for identity dominance over terrorist and national security threats who pose harm to American lives and interests. In the context of the Global War on Terrorism (GWOT), identity dominance means U.S. authorities could link an enemy combatant or similar national-security threat to his previously used identities and past activities, particularly as they relate to terrorism and other crimes.

The U.S. military needs to know whether a person encountered by a warfighter is a friend or foe. To do so, we need to answer the following questions:

- Has the person previously ?
- Been arrested in the United States or other countries?
- Used aliases or fraudulent "official" documents?
- Been detained by U.S. or coalition forces?
- Been refused entry into the United States?
- Been linked to a terrorist activity?
- Had his fingerprints found on the remnants of an improvised explosive device (IED)?
- Been seen within a crowd committing terrorist acts?

[60] *Ibid*

To the extent the U.S. military is forced to rely solely on a purported name or on "official" documents provided by someone, answers to these questions remain elusive. We cannot reliably find the answers if we use only the name the person provides and his "official" documents. Foes, particularly terrorists, will provide aliases and will often have the necessary fraudulent documents to back them up. A terrorist will also have a cover story that explains his actions in seemingly harmless terms. Fortunately, biometric technologies, based on a person's physiological or behavioral traits, can indelibly link a person to an identity or event. Names can be changed and documents forged, but a biometric is much less susceptible to alteration and forgery. Moreover, although many people have the same or similar names and many documents look alike, a person's biometrics tend to be robust and distinctive.

Biometric Technology Support

To achieve identity dominance, the U.S. military must make maximum use of biometric information and the technologies that collect, process, store, and search data. The military must work in cooperation with other U.S. Government partners, most notably the FBI, the Department of Homeland Security, the Department of State, and the intelligence community. Cooperation must also extend to state and local law-enforcement officials, who serve on the front lines of homeland security, and to our international allies as well.

Identifying individuals

Biometric technologies take automated measurements of certain physiological or behavioral traits for purposes of human recognition. Human recognition consists of verification: *Is this person who he claims to be?* and identification: *Who is this person?* These technologies can search a biometric data-base to verify a person's identity by doing a one-on-one match: *Does this needle match that needle?* And they can identify a person by doing a one-to-many search: *Is this needle in any haystacks?* This identification capability is critical for identity dominance because finding terrorists is like finding a needle in the midst of many haystacks.

Thanks to advances in computer technologies, pattern recognition, and algorithm development, some biometrics can search through large databases reliably and quickly. For example, the FBI's Integrated Automated Fingerprint Identification System (IAFIS), established in 1999, contains in an electronic database the 10 rolled fingerprint records of approximately 48 million individuals who have been arrested in the United States on felony or serious misdemeanor charges. When police make an arrest, they routinely submit the arrestee's fingerprints to IAFIS to determine if the person has a prior criminal record. The FBI processes an average of 25,000 such criminal identification submissions daily. Over 95 percent of the time, the search result is returned to the police in less than 2 hours.

Just as fingerprints can be found at crime scenes, fingerprints can be found at terrorist sites. Forensic examiners can harvest these latent prints and search them against the Integrated Automated Fingerprint Identification System database and its counterparts. Because a latent fingerprint contains much less data than a set of 10 rolled fingerprints, the system returns a candidate list of possible matches as opposed to a firm, highly reliable, match/no match result. A latent fingerprint examiner must then review the list for a final determination.

The IAFIS experience is instructive for the Department of Defense (DOD). Just as domestic law enforcement takes 10 rolled fingerprints (and other biometrics) from arrestees, U.S. military units must take 10 rolled fingerprints (and other biometrics) from Red Force members (enemy combatants and national security threats). Just as Integrated Automated Fingerprint Identification System stores arrestees' fingerprints in an interoperable format, DOD must store Red Force biometric data. Just as law-enforcement officials routinely search arrestees' fingerprints against Integrated Automated Fingerprint Identification System, so too must DOD routinely search Red Force members' fingerprints (and other biometric information) against all relevant databases to find the terrorist "needle."

The military needs reliable answers to several questions to enable it to identify people who are or might be national security threats. To get such reliable answers regarding previously used names and past activities, the U.S. military, working with other U.S. Government organizations and allied governments, must fully leverage the power of biometrics to ensure identity dominance. In doing so, some important and related functions would be served:

• Force protection - keeping U.S. and coalition personnel safer.

• Actionable intelligence - gaining information to use to detect, detain, disrupt, and deter terrorists.

• Law enforcement - recording legally admissible evidence to use to prosecute terrorists through the judicial system, if that path is pursued.

• Homeland security - safeguarding Americans and the nation.

Emerging foes

The U.S. military has always faced the challenge of identifying friend or foe. In the Global War on Terrorism, this challenge is all the more difficult because we face a highly mobile, elusive enemy who deliberately engages in tactics to conceal his true affiliation and allegiance. Terrorists use aliases to hide who they really are, and they have fraudulent official documents to support their claimed identities. Assistant Secretary of Defense for Homeland Security Paul McHale explains: "Our enemy today is no longer in uniform, our enemy today is no longer in

combat formation. Our enemy is probably wearing civilian clothes and is virtually indistinguishable from innocent counterparts throughout civilian society."[61]

The mobility of terrorists poses a serious challenge for the United States and its allies. Terrorists have demonstrated they can enter Western countries, blend into society, and remain elusive. They take advantage of our free and open societies to plot and carry out operations intended to destroy our countries. The 9/11 plotters planned and supported their attacks from the United States, Germany, Spain, Malaysia, Saudi Arabia, and other free countries.

Ensuring Identity Dominance

How can we better identify and target this elusive enemy? The Defense Science Board Task Force on Identification Technologies recently advised Secretary of Defense Donald Rumsfeld that "the Global War on Terrorism cannot be won without a 'Manhattan Project'-like tagging, tracking, and locating" program for national security threats.[62] A critical component for identifying national security threats is for the U.S. military to process biometric data taken from Red Force members using the Automated Biometric Identification System (ABIS), an interoperable enterprise approach modeled after and interoperable with the FBI's highly successful Integrated Automated Fingerprint Identification System. This approach is multitheater, multiservice, multifunctional, and multibiometric.

Multitheater

The Automated Biometric Identification System capability must reach across all theaters of operation for the U.S. military and international allies. Biometric data must be taken to standards that ensure interoperability so biometric data collected in any theater of operation can be searched against all relevant databases for possible matches.

Multiservice

DOD cannot afford to permit each military service to do its own thing with respect to biometric data. For example, U.S. Army troops in Najaf should take biometric data from Red Force members and forward it to the central Automated Biometric Identification System database; Navy units performing maritime interception operations in the Persian Gulf or U.S. Marines patrolling in Fallujah could later access and search the same biometric data.

Multifunctional

---

[61] *Paul McHale, Assistant Secretary of Defense for Homeland Defense, "Homeland Security Defense: An Update," 4th Global Homeland Security Conference and Expo: Protecting the Nation's Critical Infrastructure and Key Assets, E.J. Krause and Associates and Deloitte Consulting Conference, Bethesda, Maryland, 23 November 2004.*

[62] *http://cryptome.quintessenz.org/mirror/dsb101504.txt*

The Automated Biometric Identification System approach serves multiple functions, which means U.S. military forces can gather biometric data for use by a Department of Homeland Security inspector at a port of entry for foreigners visiting the United States, by a Department of State diplomat issuing visas, or by law-enforcement personnel carrying out arrests. Because it contains biometric data taken from Red Force members, the ABIS is a true national resource for homeland security purposes.

Multibiometric

The Automated Biometric Identification System approach must include multiple biometric records or modalities, such as fingerprints; mug shots (face); DNA; and iris, voice, and palm prints. DOD's immediate focus must be on fingerprints as the essential modality for an identity dominance capability.

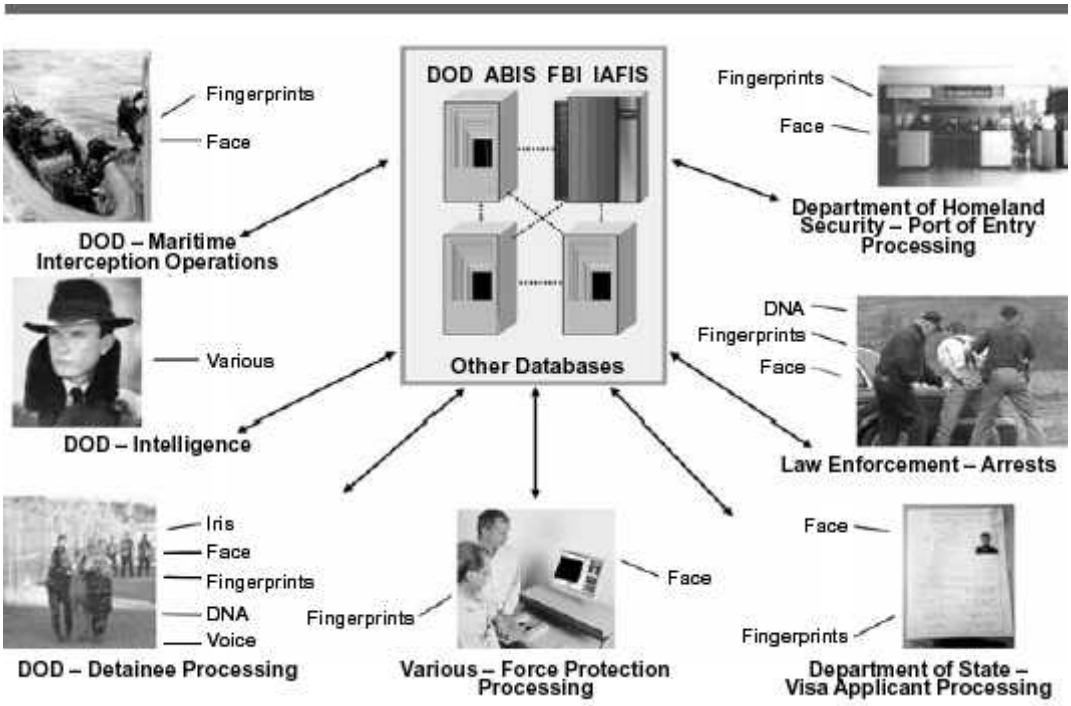


Figure 8 – Identity dominance

Several factors account for this focus on fingerprints:

• **Established biometric.** Since the late 19th century, fingerprints have been recognized as distinctive, ubiquitous, and robust. Nearly everyone has fingerprints, fingerprints do not change over time, and the legal system has long accepted fingerprints as evidence of identity.

• **Established technology**. Since 1999, searching and matching fingerprint data has become a highly accurate, automated process based on a standard that ensures interoperability. The keystone to this process is the FBI's Integrated Automated Fingerprint Identification System.

• **Established databases**. There are already many fingerprint databases. Integrated Automated Fingerprint Identification System, with its computerized records on approximately 48 million

people, is the leading example. Many states have their own fingerprint databases. Moreover, many foreign countries have national fingerprint databases.

• **Established benefits**. Fingerprints might be left behind at criminal or terrorist sites. Forensic investigators routinely harvest latent fingerprints from such sites, which are subsequently searched against databases for possible matches.[63]

While face-recognition technology does not perform as well as fingerprint technology, it is improving and can be used as a valuable screening mechanism. With state of the art surveillance cameras, we can capture an image of a person's face clandestinely and from a distance. As with fingerprints, there are many legacy databases of mug shots, which are routinely taken during the police booking process and used for many other forms of vetting, such as visa applications.

Other biometric modalities, such as iris images, palm prints, and voiceprints, should also be incorporated into the ABIS approach. Doing so would improve and expand our identity-dominance capability by allowing our allies and us to search multiple biometric modalities on suspected national security threats.

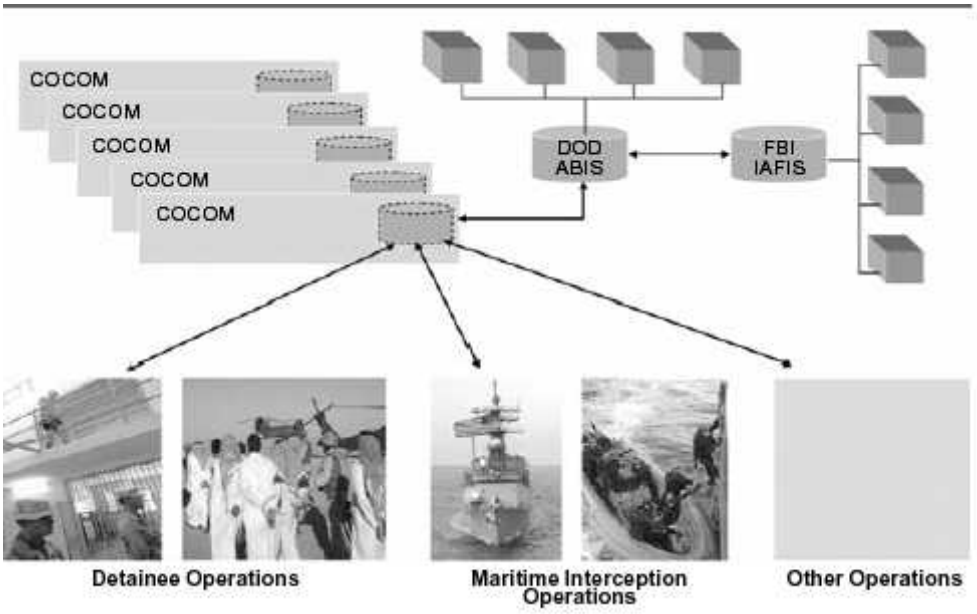


Figure – 9 Conceptual DOD ABIS architecture

A multimodal approach maximizes the use of biometric data, but identity dominance requires a single, virtual database in the form of a network of interoperable databases. For example, the Integrated Automated Fingerprint Identification System and ABIS databases must be interoperable. This seamless approach would make any standard query from another entity

[63] *Peter T. Higgins, "Fingerprints and Hand Geometry," in Biometrics: Identity Assurance in the Information Age (2003): chap. 3; Colin Beavan, Fingerprints: The Origins of Crime Detection and the Murder Case that Launched Forensic Science (New York: Hyperion, 2001).*

transparent. That is, it would be forwarded to the portal of the national security database and then searched among all relevant databases. The response would be returned to the user in a similarly transparent fashion.

Enhancing Identity Dominance

To enhance its identity dominance capability, DOD must take immediate steps in four critical areas: standards, policy, operations, and architecture.

Standards

First and foremost, military units processing Red Force members must collect fingerprints in the correct internationally accepted format - the 10 rolled fingerprints. Fingerprints taken in this way are interoperable with other fingerprint databases, such as Integrated Automated Fingerprint Identification System and IAFIS. In February 2004, the DOD chief information officer mandated that DOD organizations conform to the Electronic Fingerprint Transmission Specification (EFTS) derived from American National Standards Institute/National Institute of Standards and Technology, ITL 1-2000.[64]

In response, Lieutenant General Steven Boutelle, the executive agent for biometrics, issued new standing operating procedures (SOPs) for biometric collection from detainees that requires collecting EFTS-compliant fingerprints, mug shots based on NIST best practices, and DNA samples from detainees. The SOP also encourages collecting iris patterns and voice recordings from Red Force members. My hope is that we can expand this biometric collection in the future. The military should also collect additional modalities such as palm prints and voice recordings from Red Force members.

Policy

Thanks to McHale's leadership, DOD has a policy in place to permit routine sharing of Red Force biometric data with the FBI. This policy needs to be broadly applied to permit other organizations to submit searches to ABIS. For example, federal, state, and local law-enforcement officials submit approximately 25,000 criminal search requests per day to Integrated Automated Fingerprint Identification System. These front-line responders should be able to search the fingerprints of criminal arrestees against ABIS.

DOD policy must also encourage military units to collect biometric data from foreigners who access U.S. installations in places like Iraq or who interact with U.S. forces. In this way these foreigners, known as Grey Force, can be better vetted as security risks. Similarly, DOD policy

---

[64] *U.S. Department of Defense Chief Information Officer Memorandum, "Department of Defense Compliance with the Internationally Accepted Standard for Electronic Transmission and Storage of Fingerprint Data from 'Red Force' Personnel,"*

must enable military services, like the Navy, to collect biometric data from foreign seafarers stopped in international waters as part of maritime interception operations. This data could then be rapidly searched against ABIS, IAFIS, and related databases for matches. Ideally, the Navy's biometric capability also would be integrated into a U.S. Coast Guard biometric capability.

As an urgent priority, DOD also needs a policy to ensure effective use of biometric data it collects from Red Force members. Specifically, the military should not release a detainee from custody until the detainee's fingerprints have been searched with negative results against ABIS (to identify recidivists or match fingerprints left at a terrorist scene) and IAFIS (to identify someone who has a U.S. arrest record). In this way, the military could show that it recorded the detainee's fingerprints to FBI standards and received the results of a search (negative ABIS; negative IAFIS). Thus, DOD would ensure it has a good set of fingerprints before releasing a detainee from custody. This approach will also quickly identify detention centers in places like Iraq and Afghanistan that have not been upgraded with proper equipment and/or training. If a police department in the United States did not take fingerprints of arrestees, it would be committing a dereliction of duty. There is a lesson in this for DOD.

Operations

The military must exploit biometric data left behind on IEDs and in terrorist safe houses and other terrorist sites. The military should use both U.S. and foreign forensic investigators to harvest latent fingerprints found at terrorist scenes and outinely search latent prints against ABIS and IAFIS for possible matches, indicating, for example, that the same person was involved in multiple IED bombings. Such pattern analysis would provide useful intelligence.

Architecture

In 2004, the DOD Biometrics Fusion Center, with the support of the U.S. Northern Command, the Army Chief Information Officer/G6, DOD leaders, and other organizations, established the DOD ABIS, which is interoperable with IAFIS. DOD has a state-of-the-art system in place to process biometric data. DOD now needs to improve ABIS to push its capabilities closer to the warfighter, which would mean DOD must encourage development of rugged, lightweight, portable biometric-collection devices that can capture and transmit biometric data for rapid searching. The next generation of devices must also be fairly easy to use. As recent experience in Iraq demonstrates, it is extremely difficult for the military to provide extensive training during hostilities. Therefore, the devices must be intuitive and reliable.

The Future

In the Global War on Terrorism, the relevance of biometric technology has grown exponentially. The military must achieve identity dominance, where military forces have the distinct ability to

separate friend from foe by linking people to their previous identities and past terrorist or criminal activities. We can use biometric technology to achieve identity dominance and must deploy it to meet the requirements of force protection, actionable intelligence, and law enforcement. Establishing identity dominance through a comprehensive ABIS will enable the U.S. military to identify friend or foe to keep world safer.

CONCLUSION

The current system of passwords and pin numbers needed to access financial services has drawn a lot of criticism of late due to the increasing incidents of hacking. The system is at the mercy of hackers, who use the hacked data to draw funds from the victims account. This is where Biometrics with its foolproof system comes in.

Many South East Asian countries like Japan and South Korea have gone ahead with Biometric security in a big way, installing Biometric Access at ATM's and other public facilities, in order to safeguard financial data.

Biometric access control can also be used to improve attendance in governmental organizations, plagued by rampant instances of absenteeism. In a recent move to check absenteeism, the Municipal Corporation of Delhi tied the salary of its employees to the attendance marked by the Biometric attendance system, installed earlier this year. The system required the employees to record their entry at 9.00 A.M and then at the time of exit at 5.00 PM.

While biometric security systems can offer a high degree of security, they are far from perfect solutions. Sound principles of system engineering are still required to ensure a high level of security rather than the assurance of security coming simply from the inclusion of biometrics in some form.

The risks of compromise of distributed database of biometrics used in security applications are high, particularly where the privacy of individuals and, hence, non-repudiation and irrevocability are concerned. It is possible to remove the need for such distributed databases through the careful application of biometric infrastructure without compromising security.

The influence of biometric technology on society and the potential risks to privacy and threats to identity will require mediation through legislation. For much of the short history of biometrics, the technological developments have been in advance of the ethical or legal ones. Careful consideration of the importance of biometric data and how they should be legally protected is now required on a wider scale.

REFERENCES

1) http://en.wikipedia.org/wiki/Biometric

2) http://archives.cnn.com/2000/TECH/computing/07/24/iris.explainer/index.html

3) The 9/11 Commission Report, The Final Report of the National Commission on Terrorist Attacks upon the United States

4) Paul McHale, Assistant Secretary of Defense for Homeland Defense, "Homeland Security Defense: An Update," 4th Global Homeland Security Conference and Expo: Protecting the Nation's Critical Infrastructure and Key Assets, E.J. Krause and Associates and Deloitte Consulting Conference, Bethesda, Maryland, 23 November 2004.

5) http://cryptome.quintessenz.org/mirror/dsb101504.txt

6) Peter T. Higgins, "Fingerprints and Hand Geometry," in Biometrics: Identity Assurance in the Information Age (2003): chap. 3; Colin Beavan, Fingerprints: The Origins of Crime Detection and the Murder Case that Launched Forensic Science (New York: Hyperion, 2001).

7) U.S. Department of Defense Chief Information Officer Memorandum, "Department of Defense Compliance with the Internationally Accepted Standard for Electronic Transmission and Storage of Fingerprint Data from 'Red Force' Personnel,"

ALPHABETICAL INDEX OF AUTHORS