Voice over IP Forensics.

Associate Professor Jill Slay
University of South Australia
Defence and Systems Institute
Mawson Lakes
+61883023840
Jill.slay@unisa.edu.au

Matthew Simon
University of South Australia
Defence and Systems Institute
Mawson Lakes
+61883023840
Matt.simon@defence.gov.au

ABSTRACT
With the tremendous growth in popularity and bandwidth of the Internet, VoIP technology has emerged that allows phone calls to be routed over Internet infrastructure rather than the traditional Public Switched Telephone Network (PSTN) infrastructure. The issues faced by law enforcement authorities concerning VoIP are very different from that of traditional telephony. Wiretapping is not applicable to VoIP calls and packet capturing is negated by encryption. This paper presents and discusses experimental work carried out to explore methods by which electronic evidence may be collected from systems where VoIP conversations play an important role in suspected criminal activity or communications

Categories and Subject Descriptors
C.2.2 [**Network Protocols**]: Applications

General Terms
Security, Legal Aspects.

Keywords
Voice over Internet Protocol , Forensic Computing

1. INTRODUCTION
This paper reports on a pilot study designed to provide significant input into the current concern regarding the security and privacy implications of widespread adoption of Voice over Internet Protocol (VoIP) for personal and business telecommunications. The aims of our research were to:

1. Examine the potential threat to the privacy of telecommunications users' by the capture and reassembly of VoIP packets from a computer or network after a VoIP conversation has taken place and
2. Evaluate the potential use of such reassembled packets in forensic computing investigations.

e-Forensics 2008, January 21-23, 2008, Adelaide, Australia.

2. BACKGROUND
In our previous work [1] we have illustrated how VoIP technology, while still not prominent, is set to radically change the way voice data is communicated, and thus to revolutionise the Australian and International Telecommunications industry. With the growth in popularity and speed of the Internet, this technology is emerging rapidly, allowing phone calls to be sent via Internet infrastructure rather than the traditional Public Switched Telephone Network (PSTN). There are many advantages to using VoIP technology instead of the current PSTN system. The primary benefit is cheaper call costs for local, long distance and international calls. VoIP is also an advantage in terms of regional and remote users since it avoids large-scale roll-out of cable and cut costs in large organisations with extensive internal phone systems.

Like any new and emerging technology, many potential problems have been raised with regard to security, and thus to privacy. Recently, the Voice over IP Security Alliance (VoIPSA) [2] released a detailed review of threats faced by VoIP technology. The most serious of the threats are denial of service, host and protocol vulnerability exploits, surveillance of calls, hijacking of calls, identity theft of users, eavesdropping and the insertion, deletion and modification of audio streams.

The purpose of our ongoing research is to inform policymakers on the legislative issues surrounding privacy, telecommunications interceptions and electronic evidence preservation and also to advise technically, as to whether this technology should be used in restricted environments, and to drive future technological security control developments. The corollary to this issue is that insecure implementations of VoIP may easily provide valuable electronic evidence and this issue needs to be made known to law enforcement.

Our pilot study [1] began to examine the potential threat to privacy by the capture and reassembly of VoIP packets by a hacker (or other criminal or terrorist) from a computer or network after a VoIP conversation has taken place. We have applied memory forensics to address some of the concerns with the use of VoIP. The results of the limited number of experiments conducted with one particular implementation of VoIP (SIP) on one specific operating system show that it is possible to recover packets from memory after the completion of a VoIP call. Although very few packets remained in memory in our pilot study, there may be enough evidence with these few packets to prove that a call has actually been placed and between whom.

We have found little other published research in this area of IT Security / Forensic Computing. Neumann, Tillwick, & Olivier [3] explore the information exchanged in VoIP call control messages and the implications this has on personal privacy. Chen Wang & Jajodia [4] examine the privacy and security aspects of peer-to-peer (P2P) VoIP calls and show how the use of VoIP has substantially shifted the previous balance between privacy and security that exists in traditional PSTN calls. These researchers and others though focus on networks and protocols to explore VoiP phenomena. Our pilot breaks new ground in examining the implications of captured packets (whether captured by hackers or Law Enforcement) after the call has taken place.

3. FORENSIC COMPUTING

Forensic Computing is an extremely important cross-disciplinary research domain based on computer science and drawing on telecommunications and network engineering, law, justice studies, and social science. Research within the discipline must be performed vigilantly in order to keep the field current. Neither the pace of technological development nor the presence of criminal activities within our society will cease or taper. There is a continual need to extend forensic computing theory, tools and knowledge by conducting high quality research in relevant aspects of the field. Voice over Internet Protocol (VoIP) is a relatively new digital technology that has many implications for forensic computing. Criminals and criminal organisations can exploit the strengths of this technology because it allows a degree of anonymity and security superior to that of traditional telecommunications. The long-standing technique of wiretapping a physical telephone line to eavesdrop on a conversation is a concept that is not applicable to VoIP communications. Alternative methods for obtaining evidence are required to help fill the void that has been created by the switch from traditional telephony to VoIP.

Forensic computing encompasses multiple sources of electronic data that could possibly contain evidence. The volatile memory of a computer, called Random Access Memory (RAM), is a source of data not often exploited by forensic investigators in the search for evidence. Issues with obtaining the data in memory are one of several reasons for this. The practice of examining data within the memory of a target machine is called memory forensics.

3.1 Developments

Forensic computing is an emerging field that is currently in a transitional phase that will see the field mature into a solid and respected science [5],[6], [7]. New technologies are constantly emerging into the marketplace and the forensic computing field must be vigilant in updating practices and knowledge to account for such technologies. In recent years, there has been a massive increase in the number of devices with embedded logic and storage. There have also been remarkable increases in data storage capabilities in household computers, laptops and dedicated servers. The field of forensic computing must constantly research new technologies as they emerge, as well as review methods and tools that already exist in order to keep methods, practices and knowledge up to date. Many common devices have embedded processing and storage capabilities. These devices could potentially be used during criminal activity and as such, may contain data that could be used as evidence. Devices such as personal digital assistants (PDA), mobile phones, household computers, laptop computers, iPod's, USB storage devices, removable hard drives and other proprietary technologies (e.g. a Blackberry) have the potential to hold information of evidentiary value that could be used in the course of a criminal, civil or internal investigation.

3.2 Forensic Computing Research

Research within the field of forensic computing is multifaceted. New technologies must be thoroughly examined to determine how they work. This will provide a base for procedures to be developed, allowing a process where relevant data can be extracted in a forensically sound manner. Developing forensically sound tools with which to perform this procedure is an important element of the process.

The goal of research in forensic computing is to determine the '*what'* and '*how*' of the data on the device or computer. The *'what'* is to determine what data is relevant and can be used as evidence in an investigation. The *'how'* is to determine the means of collecting and storing the data. If techniques for extracting and storing data cannot be proven forensically sound, it will likely loose integrity and becoming inadmissible in a court of law.

Research within forensic computing needs to be versatile as multiple techniques may be required for any particular situation during an investigation. Forensic computing investigators often face the problem of encountering live systems, that is, a computer or device that is switched on. Forensic investigations are rarely conducted using the target system or on original evidence sources [8], [9]. The target system usually needs to be powered down so that the data on the disk drives cannot be altered, and can subsequently be removed (if practical), secured, and then entered as evidence. The investigator must consider all factors when deciding whether to power down the relevant devices. The computer or device may be powered down using a number of different methods from shutting down normally to removing the battery or plug. If the computer or device is left on, it may be possible for the investigator to extract information from the main memory, although this is more likely with a computer than a proprietary device. Tools and methods have been developed to gain information from memory that is lost once the machine is powered down (e.g. current running processes, memory resident worms etc) [8].

Situations often arise in forensic computing investigations that require special procedures or tools because a particular technology or situation encountered. Forensic computing research requires constant investigation into new devices and technologies, as well as new methods and procedures for existing technologies. This will equip law enforcement bodies and other forensic investigators with the knowledge and abilities to allow thorough and safe investigation of digital devices. This dissertation will research the viability of using newly created methods of obtaining evidence from digital devices for forensic purposes.

3.3 Volatile and Memory Forensics

Memory forensics is a relatively unexplored area of computer forensics. There are numerous reasons for the lack of use of this technique but is mainly due to the fact that the process of imaging the memory is not verifiable. Little investigation into tools and processes for sound memory dumping techniques has been performed in forensic computing research [10]

The volatile memory of a system potentially contains tremendous amounts of information about the system including (but not limited to) open files, active processes, terminated processes and device drivers. This source of data is lost when the target system is turned off in the course of securing the non-volatile data sources. Situations can occur in forensic investigations where a target system cannot be turned off. Servers are a prime example as downtime generally leads to loss of revenue or essential services. Imaging the memory of a target machine allows evidence to be collected without any downtime. Imaging memory in general investigations allows an extra source of data from which evidence can be inferred.

Dumping of the memory from a system requires some method of interfacing with the data. Linux Systems contain a memory device similar to devices representing hard disks. Memory can be imaged by copying the data from this memory device [11], [12].. Windows XP does not have an equivalent memory device and must use a section object to gain access to the memory. A section object allows a mapping to memory pages which can subsequently be accessed by multiple processes [10], [14].

Imaging memory for forensic purposes is gaining recognition as an area in need of further research. Burdach [12], [13] has published a number of papers on techniques of finding evidence in both Windows and Linux memory images. The scope of memory forensics is large because and has potential to add a diverse source of potential evidence.

4. TELEPHONY AND VOIP PROTOCOLS

The Internet and the common telephone are two instrumental communication tools in today's society. These technologies have vastly different methods of transmitting data between locations, both use a variety of common media and have a digital architecture but little other similarities.

With the tremendous growth in popularity and bandwidth of the Internet, technology has emerged that allows phone calls to be routed over Internet infrastructure rather than the traditional Public Switched Telephone Network (PSTN) infrastructure. The technology, called Voice over Internet Protocol (VoIP), uses the Internet Protocol (IP) to route packets containing small portions of voice conversations between the callers. Some of the advantages of using VoIP are: cheaper call costs, better ability to integrate new services and a significant increase in scalability. In contrast, there are also disadvantages and potential problems with using VoIP such as lower quality of service, increased security risks and determining physical location of callers.

Voice over Internet Protocol is telephony using the same technology as the foundation of the Internet; this is a packet switched network. Packet switched networks take small sections of data, called packets, which are sent from the end-point through a series of routers each of which forward the packet closer to its destination. Many different data streams can be sent over the same connections because the bandwidth is not reserved for one particular stream of data. Essentially, VoIP sends data packets between smart devices (e.g. computers) over a relatively unintelligent network whereas traditional circuit-switched telephony sends voice data between dumb devices (telephones) through a super smart central network [6].

VoIP telephony is becoming very attractive to users and telecommunication providers alike for a number of reasons. The primary reason for the growth of VoIP is decreased costs. Switching data across a packet switched network is much cheaper than establishing a circuit over a circuit switched network because less bandwidth is required. Using the Internet infrastructure is also cheaper as it is not owned solely by a few private companies; many companies, governments and organisations own and maintain parts the infrastructure. VoIP services are often used in conjunction with PSTN by using media gateways. In the case of a PSTN to SIP network, the media gateways convert the PSTN in band signalling to SIP messages and voice data to RTP packets (or the other way around if travelling from IP to PSTN [7]. Long distance and international calls benefit from VoIP as the majority of the distance can be travelled via IP networks and rejoin the PSTN at a local call distance from the target destination. One major telecommunication company in the United Kingdom is said to be converting from a PSTN backbone to IP by 2007. Customers using the company will likely not know the calls are travelling via the Internet [8]. Telecommunication companies also see VoIP as a method of offering existing customers additional multimedia services such as video-calls [7].

Another reason for the increased use of VoIP is the extensibility and scalability. VoIP is capable of integrating almost any feature imaginable; the limitations are in the end devices and protocols, not in the network (as is the case with PSTN). With an IP network, in-band signalling is not required, separate data streams are used for data and signalling [7]. In the PSTN network, to integrate new features, the complex central switches have to be reprogrammed which is a very involved task [6].

VoIP telephony can use various protocols and methods to establish calls and transmit data. Skype software, a common implementation, uses proprietary protocols which are also encrypted. Many VoIP implementations however use Session Initiation Protocol (SIP) and Real-time Transport Protocol (RTP). SIP is the protocol used for setting up and tearing down calls and sending other call related data during the conversation; this is analogous to DTMF, the PSTN in band signalling mechanism. It is a text based, application level protocol and relies heavily on other protocols for transport (such as IP and UDP) [7]. VoIP implementations that use SIP generally rely on a SIP proxy server into which the users must login. This proxy is also used to route call and signalling data. Clients can find each other and forward SIP messages via this proxy [7]. SIP messages are not only used for initiating and tearing down calls, they are also used for changing call parameters or other features such as integrating more callers into a conference session. SIP registrars are additional servers used for to locate other users; generally the SIP proxy also acts as the registrar.

4.1 VoIP and Crime

The popularity of VoIP is increasing as the cost savings and ease of use is realised by a wide range of people and corporations. The technology is attractive to criminals, especially the non-carrier VoIP, as it often does not require verification of any details to commence using the service. The security of placing such calls may also be appealing to criminals, as many implementations use strong encryption to secure both the voice payload as well as control messages. Skype uses 256 bit AES encryption (*Skype Privacy FAQ* 2006) while Google Talk does not encrypt its

payload (but will support encryption in the future) (*Google Talk: Frequently Asked Questions* 2006).

The following hypothetical situation illustrates the ease of which this technology can be used in criminal activity:

An organised crime ring operates within Australia, distributing illicit drugs throughout the country. The operators of the crime ring decide that by using Skype software they can anonymously communicate when necessary. From a criminal perspective, there are several disadvantages of using the traditional PSTN telephone system. There is the possibility that a law enforcement body could wiretap the connection should they become suspicious, all calls made and received are logged by the service provider, and using a PSTN phone fixes someone to a given location (i.e. the physical location of the phone). As an alternative solution, criminals can use laptops running Skype, create profiles in the same way as a regular user and communicate when necessary. If law enforcement should investigate, there is no line to wire tap, no call logs and no ability to tie a person to a specific geographic location. The criminals using Skype can also be contacted or make contact from different locations, providing flexibility.

Although this particular situation is only hypothetical, it is possible. The smallest through to the largest criminal organisations, including international terrorists, could potentially communicate using VoIP, as it incorporates the flexibility of email, the richness of voice and the safety of a decentralised system using strong encryption algorithms.

It is essential that computer forensic research evaluate the use of VoIP technology and devise methods to allow law enforcement agencies to overcome some of the aspects that are advantageous to criminals. Wire-tapping is not applicable to VoIP communications and therefore other methods of recovering evidence and information are required.

5. RESEARCH METHODOLOGY

In our research, several techniques were used in the process of recovering VoIP evidence. Our scenario is 'after the event forensics' where we imagine that a suspects computer has been used to make VoIP calls and has not been powered off after the conversation.

Calls of varying lengths were made using different operating systems and under a range of conditions.

The first step in the process is to acquire the memory image from the target system. The subsequent steps involve finding packets within the image, identifying the type of packets, extracting the payload from RTP packets and recreating audio files by reconstituting consecutive packet payloads into one file. Each step in this process requires a practical solution and consequently both existing and new solutions have been either used or devised.

The technique of imaging memory uses an existing tool created for acquiring memory images under *Microsoft Windows XP*. The algorithm for finding packets in a binary image is a newly devised method created to address this specific problem. Details of the the algorithm, and the process used to recover evidence from the memory image of a target machine are currently not open for disclosure.

However, details of the protocol format (e.g. fields that are present in the protocol structure) are important in understanding how the identification and reconstruction of packets is possible

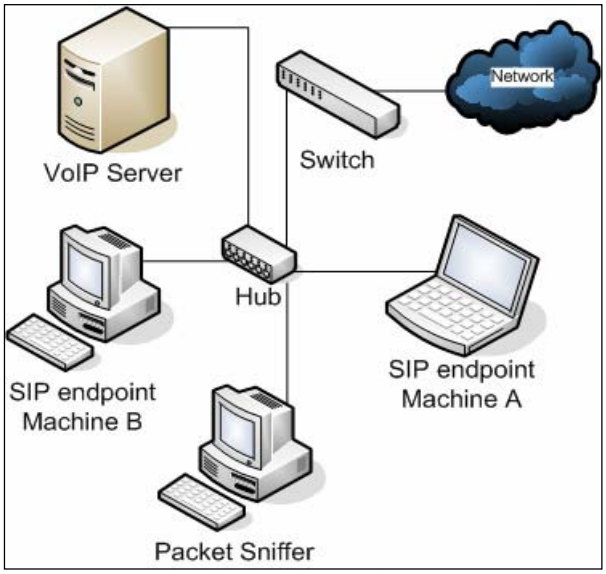

Figure 1: Experimental Setup

and some discussion of this is presented here.

5.1 Session Initiation Protocol

SIP is a signalling protocol that is used for signalling between SIP nodes. VoIP implementations often use SIP as a means of setting up and tearing down calls. SIP is not responsible for carrying any audio payload as this is performed by another protocol such as RTP. A number of SIP message types are used for communication between SIP nodes (note that communication is always routed via the server). The format of a SIP address is similar to an email addresses in that it is in the format of *username@host*. The string 'sip:' is appended to the beginning to indicate it as a SIP address rather than an email address. The *username* is the user's extension that is in the same format as an email address username (i.e. a combination of number, letter and the hyphen character). The host can be either a URL or IP address and indicates the location of the SIP proxy.

There are two types of SIP messages, requests and responses. Request messages are generated by a SIP client and responses are generated by a SIP proxy in response to a request message. SIP is a flexible protocol and as a result the full implementation is quite complex. Two important fields in the header are the 'to' and 'from' fields. These fields contain the SIP address of the endpoint that is sending the message and the SIP address of intended recipient. The 'to' and 'from' fields can also contain other information such as the user's display name. Additional information in the header fields is dependant on the VoIP implementation.

5.2 Real-Time Transport Protocol

RTP is a highly flexible protocol that is used to carry delay sensitive payloads over the Internet. RTP is often used in VoIP implementations to carry audio payload and is an Internet standardised protocol as described in RFC 3550. RTP packets are transported within UDP packets, which is typical of protocols that

carry delay-sensitive data (if the packet is delayed for too long then retransmission is of no use). The RTP protocol contains several fields that allow the payload of each packet to be reconstituted back into its original form. The Synchronous Source Identifier (SSRC) is a 32-bit number that is distinct for each conversation stream. Note that one conversation between two SIP endpoints consists of four different conversation streams and four unique SSRC (each endpoint has a distinct incoming and an outgoing stream). The Sequence Number field is 16 bits in length and increments by one for every distinct RTP packet sent. The starting value of the sequence number is chosen randomly and is different for each conversation stream (each unique SSRC). The timestamp field increases in each packet relative to the amount of time that the payload in the packet covers. It is similar to the sequence number and will usually increment at a constant rate relative to the sequence number, especially with voice calls. The starting value for the timestamp is chosen randomly.

5.3 Memory Imaging

The physical memory of a system can be imaged in two different ways under *Microsoft Windows XP*; one of these methods is hardware based and the other software based [13]. The hardware method is the most reliable and verifiable but is not a practical solution for forensic investigation needs. It requires the use of a special purpose PCI card that must be installed into the machine prior to it being switched on. This requirement renders this method ineffective for forensic investigations (Carrier & Grand 2004). Another hardware method exists and requires that the target system have a Firewire/IEEE port. This method has been shown to be highly effective but does not work on *Microsoft Windows* machines [14]. The software-based solution, while not ideal, is the only practical solution for acquiring a memory image from a target computer at the scene of a crime.

Microsoft Windows XP does not have a memory device object so the physical memory of the machine needs to be accessed through a section object. The section object maps to pages in memory allowing access by processes that do not own the memory page.

6. RESULTS AND DISCUSSION

The result of work was that we were able to search for packets in a memory image taken from a computer running Windows XP or Mac OS 10 and sort, order and output the packets and eventually, under specific circumstances to reconstitute the packets to an audio file (audio could be heard).

This means that potentially a user's privacy might be breached by a hacker using software similar to that which we have developed as part of this research. The hacker would be able to copy remnants of the data stream from the computer's volatile memory and recreate small portions of the conversation.

It is possible to extend the software developed in this project as a tool for users to mitigate against the threat of a hacker who was able to breach their physical and logical security and extract remnant VoIP packets from memory space (in specific situations). This also means the same kind of software could be used as a forensic tool.

The experimentation conducted in this research is a small section of a large domain. Additional research is required to improve the confidence of the conclusions drawn from the results of the experimentation.

Other situations need to be tested, such as varying the length of time between termination of the call and imaging of the memory. This should also be conducted a number of times to increase the accuracy of the results. A foreseeable problem with this course of experimentation is that it is difficult to control the actions of the operating system, which introduces another element of randomness.

Using different SIP clients should also be explored. It is unclear exactly how much influence the VoIP application has on the amount of remnant packets, and how much is reliant on the operating system.

The exploration into other signalling and transport protocols is an important future direction for this research. Many VoIP implementations do not use the SIP and RTP protocols such as *Google Talk* and *Skype*. Adapting to other protocols in some cases should be trivial and almost impossible in others. *Google Talk* uses the open protocol called Jabber and currently does not implement encryption. Alternatively, *Skype* uses a proprietary protocol that incorporates heavy encryption.

The use of encryption ties closely to the protocol used. Both SIP and RTP can easily be encrypted (RTP needs to be substituted with SRTP). The limitation in this situation is that packet data is unlikely to hold the unencrypted version of an encrypted payload, as the payload is encrypted before being inserted into a packet, or decrypted after being read from the packet. The unencrypted payload may exist in the memory space but will not be found using the method presented in this research. This research would play a key role in this situation if the encryption key to decrypt the encrypted packets was found.

A major problem with this research is the inability to verify the resulting memory images obtained during the memory imaging process. It is difficult to ascertain how much of a change the process of acquiring the image makes to the image itself. Investigation into the amount of change that is caused by this factor is difficult. Using special-purpose memory image hardware such is the only foreseeable method of conducting such an investigation. It may be possible to conduct a 'before and after' experiment that compares the hardware-acquired memory state before and after the process of software-based memory imaging specified in this research.

7. ACKNOWLEDGMENTS

Our thanks to the Department of Communications, IT and the Arts who funded part of this work with a Telecommunications Research Grant.

8. REFERENCES

[1] Simon, M & Slay, J 2006, 'Voice over IP: Forensic Computing Implications', 4th Australian Digital Forensics Conference, Edith Cowan University, School of Computer and Information Science, December 4, 2006.

[2] Kuhn, DR, Walsh, TJ & Fries, S 2005, *Security Considerations for Voice over IP Systems*, National Institute of Standards and Technology, Gaithersburg.

[3] Neumann, T., Tillwick, H & Olivier, MS 2006 "Information Leakage in Ubiquitous Voice-over-IP," in S Fischer-Hübner,

S Furnell and C Lambrinoudakis (eds), *Trust, Privacy and Security in Digital Business,* LNCS 4083, 233-242, Springer.

[4] Ahuja, SR & Ensor, R 2004, 'VoIP: What is it Good for?' *Queue,* vol. 2, no. 6, September, 2004, pp. 48 - 55.

[5] Beckett, J.J & Slay, J 2007, 'Digital Forensics: Validation and Verification in a Dynamic Work Environment', HICSS-40, January 3rd 2007 Hawaii.

[6] Patel, A & Ó Ciardhuáin, S 2000, The impact of forensic *computing on telecommunications*, pp. 64-67.

[7] Broucek, V & Turner, P 2001, 'Forensic Computing: Developing a Conceptual Approach for an Emerging Academic Discipline.' paper presented at the 5th Australian Security Research Symposium, Perth, Australia.

[8] Kruse II, WG & Heiser, JG 2002, Computer Forensics: Incident Response Essentials, Addison-Wesly, New York.

[9] Marcella, AJ & Greenfield, RS (eds) 2002, Cyber Forensics: A Field Manual for Collecting, Examining, and Preserving Evidence of Computer Crimes, Auerbach, New York.

[10] Jones, KJ, Bejtlich, R & Rose, CW 2006, *Real Digital Forensics*, Addison Wesley, Upper Saddle River, NJ.

[11] Burdach, M 2005, *Digital Forensics of the Physical Memory*, Warsaw, <http://www.forensicfocus.com/digital-forensics-of-physical-memory>.

[12] Burdach, M 2006, *Physical Memory Forensics*, viewed 28 September 2006, <http://strony.aster.pl/forensics/pdf/mburdach_physical_memory_forensics_bh06.pdf#search=%22memory%20forensics%20XP%22>.

[13] Bellovin, SM, Blaze, M & Landau, S 2005, 'The real national-security needs for VoIP', *Commun. ACM,* vol. 48, no. 11, p. 120.

[14] Vidstrom, A 2006, *Forensic memory dumping intricacies - PhysicalMemory, DD, and caching issues*, viewed August 2006, <http://ntsecurity.nu/onmymind/2006/2006-06-01.html>