INFORMATION AND PRIVACY COMMISSIONER OF ALBERTA

Report of an investigation of a malicious software infection affecting health information

December 9, 2009

Alberta Health Services

Investigation Report H2009-IR-007

(Investigation H2793)

Introduction

[1] On June 2, 2009 Alberta Health Services (AHS) contacted the Office of the Information and Privacy Commissioner to report that its Edmonton area computer network had been infected with malicious software. AHS reported that health information may have been exposed to outside parties while the malicious software was active.

[2] On July 8, 2009 AHS issued a news release stating that 11,582 individuals' health information had been affected by the malicious software and that it would notify these individuals directly by mail.

[3] The Commissioner authorized me to conduct an investigation under section 84(a) of the *Health Information Act* (HIA, or "the *Act*"). Section 84(a) allows the Commissioner to conduct investigations to ensure compliance with any provision of the HIA.

[4] This report lays out the findings and recommendations resulting from my investigation.

Background

[5] AHS identified the malicious software as a variant of "Coreflood," a Trojan horse program.

[6] According to the Information Systems Audit and Control Association's (ISACA) *Glossary of Terms*, malicious software, or "malware" is "software designed to infiltrate, damage or obtain information from a computer system without the owner's consent.[1]" Malware commonly includes categories of programs known as

[1] *ISACA Glossary of Terms*, Information Systems Audit and Control Association, June 1, 2008, page 43, http://www.isaca.org/glossary.pdf.

viruses, worms, Trojan horses, and spyware. The ISACA Glossary defines a Trojan horse program as "purposefully hidden malicious or damaging code within an authorized computer program.[2]" Computer users often activate Trojan horse programs unknowingly when they click on a link or program in an email or webpage that they think will be benign or useful, but that actually contains hidden malware that installs itself on their computer. Trojan horse programs often create a "back door" that allows an external party to take control of and/or steal data from the affected computer.

[7] Coreflood has been known to computer security experts since 2002. The authors of Coreflood have created many variations over the years that have served different purposes, from attacking internet chat users, to providing anonymization services, to stealing sensitive information.[3] The variant discovered at AHS was designed to steal data from an infected computer and send it to a server controlled by an unauthorized party (also known as a command and control or C&C server).

[8] In this case, Coreflood captured text viewed by users during secured web-browser sessions, encrypted[4] it, stored it on the infected computers and periodically sent it to a computer outside of AHS' network. Once the encrypted data was transmitted, Coreflood deleted it from the infected computer. This cycle of data collection, transmission and deletion would continue until Coreflood was removed from the infected computer. Coreflood, in some cases, may have also captured usernames and passwords needed to log into websites. This means that whenever an AHS computer user logged into a website from an infected computer, Coreflood could have captured the person's username and password and text data from whatever screens they looked at while they were logged on and passed this information on to the C&C server.

[9] The only clinical application affected by Coreflood was Alberta Netcare Portal, which is an internet application run by Alberta Health and Wellness.

[10] According to Alberta Health and Wellness, the Alberta Netcare Electronic Health Record (EHR) is "a secure lifetime record of an Albertan's key health information available for consultation by authorized health service providers.[5]" Users access Alberta Netcare EHR through an internet portal, or website. This application is commonly known as "Netcare Portal." Authorized users of Netcare Portal, depending on their privileges, may access the following health information about individual patients:

i. patient demographic information
ii. prescribed and dispensed drugs
iii. known allergies and intolerances

[2] *Ibid*, page 72.

[3] *Coreflood/AFcore Trojan Analysis*, Joe Stewart, SecureWorks, June 30, 2008, http://www.secureworks.com/research/threats/Coreflood/.

[4] Encrypted data is scrambled so that it cannot be read if intercepted. An encryption key is needed to de-crypt, or de-scramble the data. Only the author of the Trojan horse program would have the decryption key. However, security experts were working to break Coreflood's encryption. A software program to break Coreflood's encryption was publically released on May 22, 2009.

[5] *What is the Alberta Netcare EHR?*, Alberta Health and Wellness, http://www.albertanetcare.ca/9.htm. Refer to the Information and Privacy Commissioner's Investigation Report H2008-001 for a detailed description of Netcare Portal at http://www.oipc.ab.ca/downloads/documentloader.ashx?id=2256.

iv. immunizations
v. laboratory test results
vi. diagnostic imaging reports
vii. other medical reports, such as discharge summaries and consultations
(not all of the above health information is available in Netcare for all patients)

[11] Most organizations, including AHS, employ anti-virus systems to prevent, detect and remove all forms of malware, including viruses, Trojan horses, spyware and others. Anti-virus systems work by scanning computers on the network and attempting to identify known malware. Malware programs, like any software, contain unique characteristics, patterns and structures within their underlying programming. Those that study and catalogue malware have identified these characteristics, known as signatures. Antivirus systems detect whether a given piece of software contains malicious programming by scanning for these signatures.

[12] Malware evolves over time. Its authors frequently repackage or disguise malware programs so their signatures cannot be recognized by anti-virus systems. An organization can only prevent and detect malware if it receives frequent updates of the latest malware signatures via download from its anti-virus system vendor. These updates are known as "virus definitions." Further, malware must first be discovered and analyzed by anti-virus experts before a definition and a way to remove the malware can be developed. This means that anti-virus vendors are always one step behind malware authors. Further, some organizations get hit by malware before its related definition is released.

Application of HIA

[13] The HIA applies to "health information" in the custody or control of a "custodian."

[14] Alberta Health Services is a "custodian" under section 1(1)(f)(iv) of the HIA.

[15] I examined a sample of the Netcare data files that were captured by Coreflood. These files contained registration information about identifiable patients, their diagnostic, treatment and care information (as outlined at paragraph 10), and health services provider information, identifying those who treated them. This information all falls within the definition of "health information" set out in section 1(1)(k) of the HIA.

Issue

[16] Did the Custodian fail to safeguard health information, in contravention of section 60 of the *Health Information Act*?

[17] Before analyzing this issue, it is useful to review the actions taken by AHS in response to the Coreflood outbreak.

Timeline and summary of actions taken by the custodian

[18] On March 26, 2009 the AHS information services help desk began receiving calls that a financial software application was crashing. Upon investigation AHS detected evidence of Coreflood activity on its systems. However, its anti-virus system had not detected this outbreak. AHS contacted its anti-virus vendor and provided technical details on what they suspected was a new variant of Coreflood.

[19] The next day, AHS' anti-virus vendor confirmed that the malware affecting its systems was a new variant of Coreflood. This variant of Coreflood was so new that the vendor had not yet created a certified definition file that would have detected it. The vendor provided an early release of a virus definition file that would allow AHS to detect the new variant of Coreflood. An early release definition file is used only in urgent situations because it is not fully tested and certified by the anti-virus vendor.

[20] Over the following days AHS was able to detect and remove Coreflood from 3,498 computers. Unfortunately, this was not the end of the story.

[21] On May 22 AHS discovered a second new variant of Coreflood after detecting an unusual increase in network activity including a large amount of outbound traffic. Once again AHS reported this discovery to its antivirus vendor. Once again, the vendor had not yet provided a definition file that could have detected this variant. On May 25, AHS received a definition file that allowed it to detect the second variant of Coreflood, but the fix provided by the antivirus vendor was not able to completely remove the malware from all computers. Despite previous efforts, it appeared that Coreflood was still active and transmitting data outside of AHS. As mentioned previously (see paragraph 8) Coreflood encrypts the data it captures, so it was not possible at this point to determine exactly what data was being captured and transmitted.

[22] AHS began a forensic analysis on May 26 to determine the full effect of the two Coreflood outbreaks. A software program to decrypt Coreflood data was publically released on an internet security site on May 22, and identified by AHS on May 27. AHS was now able to decrypt the data captured by Coreflood. By May 29, after decrypting Coreflood data in a secure test environment, AHS knew that Coreflood had been sending Netcare data to an external party. AHS' forensic review also identified the address of the server that was receiving data sent by Coreflood. AHS blocked all data from being sent to that server on May 29.

[23] Further analysis at AHS showed that Coreflood had been transmitting data to an external computer server between May 14 and May 29. Both outbreaks of Coreflood (in March and in May) were controlled from the same external computer. The March outbreak of Coreflood didn't transmit any health information outside of AHS' network, but it likely transmitted the second new variant of Coreflood to other computers, which was then was activated in May.

[24] AHS was able to narrow down those at risk to two groups: Patients whose health information was accessed in Alberta Netcare from an infected computer and employees who accessed personal banking or email accounts from work using an infected computer.

[25] AHS conducted a privacy risk assessment which considered the risk of harm to employees and patients resulting from the Coreflood outbreak. AHS noted its belief that the health information captured by Coreflood would not, by itself, be useful to conduct identity theft or fraud. However, it concluded that the compromised data was still sensitive health information and made a decision to notify those patients whose records were captured by Coreflood by letter. AHS also notified its employees whose computers were infected and advised them on June 2 by email to change passwords of any banking or personal email systems they may have accessed from work.

[26] To determine which patients to notify, AHS considered a number of factors. Based on its analysis of outgoing network traffic, AHS knew that data was being transmitted to the C&C server between May 14 and May 29. AHS was also able to identify each infected computer. By analyzing access logs[6] in Netcare, AHS was able to identify those patients whose records were accessed from infected computers during the period Coreflood was transmitting data. AHS identified 11,582 individual patient records in Netcare that met these criteria.

[27] It is worth noting that not all of these 11,582 records would have been transmitted to the C&C server. As mentioned earlier, Coreflood captures information, stores it on the infected computer in an encrypted data file and transmits it periodically. Once the information is transmitted, Coreflood erases the data file. Therefore, some of these records were transmitted, while others were stored in encrypted data files awaiting transmission, but did not get transmitted before AHS cut off communication to the C&C server on May 29. AHS decided to take the more cautious approach and notify everyone who was potentially affected.

[28] AHS issued a news release describing the incident on July 8 and began notifying the 11, 582 individuals by mail the following week.

Analysis and findings

Notification

[29] In general, I agree with AHS' analysis of privacy risk (see paragraph 25). The health information that was transmitted by Coreflood would not be useful by itself for fraud. However, the information contained names, addresses, phone numbers, birth dates, and provincial health numbers, along with the diagnostic, treatment and care information outlined earlier (see paragraph 10). This information could be used as a starting point to put together an individual profile for fraudulent purposes. Further, health information is inherently sensitive. People deserve to know that their health information may have been exposed.

[6] Netcare keeps logs of which users access individual patient records. By reviewing access logs for a given time period, it is possible to identify exactly which patient records were accessed by any particular user and computer terminal.

[30] The HIA does not require custodians to notify individuals whose health information has been disclosed inappropriately. I believe AHS took a prudent and responsible course of action by notifying the patients whose Netcare records may have been exposed. In my opinion, any AHS staff that accessed their personal banking from one of the infected computers is at greater risk. These employees should take the warnings they received from AHS seriously and I hope they have already changed their passwords and are monitoring their accounts for suspicious activity. The Information and Privacy Commissioner supports AHS' decisions to notify staff and affected patients about this breach.

Duty to protect health information

[31] Custodians have a duty to protect health information from technological threats to confidentiality, such as malware. The relevant parts of 60 of the HIA read as follows:

> **Duty to protect health information**
>
> **60(1)** A custodian must take reasonable steps in accordance with the regulations to maintain administrative, technical and physical safeguards that will
>
> (a) protect the confidentiality of health information that is in its custody or under its control and the privacy of the individuals who are the subjects of that information,
>
> …
>
> (c) protect against any reasonably anticipated
>
> (i) threat or hazard to the security or integrity of the health information or of loss of the health information, or
>
> (ii) unauthorized use, disclosure or modification of the health information or unauthorized access to the health information,
>
> …
>
> **(2)** The safeguards to be maintained under subsection (1) must include appropriate measures
>
> (a) for the security and confidentiality of records, which measures must address the risks associated with electronic health records, …

[32] The above subsections mean that custodians need to identify threats to patient confidentiality and implement reasonable measures to mitigate the risk presented by these threats. Further, the HIA places particular emphasis on mitigating the risks associated with electronic health records, such as Netcare. Subsection 60(1)(c) refers to "reasonably anticipated threats." This means that custodians must take steps to protect health information that a reasonable person, faced with similar circumstances would also take. If a threat to confidentiality is generally well known, it is reasonable to expect a custodian would anticipate it and devise a way to mitigate the threat.

[33] The threat to confidentiality posed by malware programs is well known and has been highly publicized in the media. The first computer viruses were written in the early 1970s and malware has been a known threat to computer systems ever since. The motivations of malware authors have changed over the years. The

first virus programs were experiments. Next, programmers authored malware to cause mischief or to advance their reputation. Many observers have noted a trend in recent years where malware is propagated by those with criminal motives, intent on stealing data to commit fraud.

[34] In my opinion, it is reasonable to anticipate the threat to confidentiality of health information posed by malware. Therefore, custodians must take reasonable steps to protect against this threat. Ever since proclamation of the HIA in 2001 our Office has consistently advised custodians to implement malware protection or antivirus systems.[7]

[35] I asked AHS to describe its policies and practices that protect its computer systems against malware.

[36] AHS has a policy entitled "Antivirus Malicious Code," which was implemented in 2002 and was last reviewed in July 2008. The policy says that AHS maintains an antivirus solution on all workstations and servers and includes direction on quarantine procedures and virus removal. I also confirmed that AHS updates its virus definitions daily and also has an ability to update its antivirus system at any time, in response to urgent situations.

[37] It is common practice for large organizations to run malware protection from more than one vendor. The reasoning behind this is that if one brand of malware protection fails to recognize a given virus, another brand might. Most of AHS' infrastructure is protected by a single brand of malware protection; however, AHS does have different brands of malware protection running on some servers. AHS already has a large and complex computer network. AHS says that adding yet another brand of malware protection would add to complexity and affect system performance, making their entire network more difficult to manage. In AHS' view, these factors outweigh the benefit of any incremental protection to be gained.

[38] Referring back to the timeline in paragraphs 18 to 28, I note that Coreflood began transmitting data on May 14, but unusual activity was not detected until May 22. Two other security systems could have detected evidence of a Trojan horse program between May 14 and May 22: a firewall and an intrusion detection system. I will discuss each of these in turn and analyze the role these systems played in AHS' response to this incident.

[39] A firewall is a system that enforces a boundary between networks, typically forming a barrier between a secure environment (like AHS' internal network) and an open environment such as the Internet.[8] In short, a firewall prevents unauthorized transmissions from entering or exiting a computer or network. Generally, network administrators do not have time to inspect every flow of data into and out of their domains and will only review anomalous events that reach a certain threshold.

[7] See, for example, *Health Information - A Personal Matter: A Practical Guide to the Health Information Act*, 2001, Office of the Information and Privacy Commissioner, Alberta, page 26, http://www.oipc.ab.ca/Content_Files/Files/Publications/HIA_Guide.pdf.
[8] *ISACA Glossary of Terms*, Information Systems Audit and Control Association, June 1, 2008, page 32, http://www.isaca.org/glossary.pdf.

[40] AHS has several firewalls throughout its network. AHS' review of firewall logs determined that Coreflood was transmitting data between May 14 and May 29. However, this review was conducted after the fact. I asked AHS why Coreflood transmissions weren't detected at the firewall until May 22. AHS says that the Trojan horse spread slowly at first and was only transmitting a small amount of data in the early days of the infection. As Coreflood spread, it began sending more and more data until the flow of outbound data increased to the point where it was noticed by network administrators on May 22. One could argue that AHS should have noticed the increase in network traffic sooner and responded more quickly to the Coreflood outbreak. At the same time, it was AHS' monitoring of its firewalls that detected Coreflood activity after its anti-malware system failed to do so. Therefore, I cannot conclude that AHS was not monitoring its firewall. I can only say this is an area where increased vigilance may have detected Coreflood activity sooner.

[41] An Intrusion Detection System (IDS) inspects network and computer activity to identify suspicious patterns that may indicate a network or system attack.[9] AHS has an IDS that could have detected Coreflood activity prior to May 22. In fact, AHS' post-incident review of its IDS logs showed Coreflood activity between May 14 and May 22. Closer monitoring of its IDS logs would have allowed AHS to respond more quickly to the Coreflood outbreak.

[42] AHS was maintaining an up-to-date anti-malware system and still became infected by a new variant of Coreflood. AHS could have discovered the Coreflood outbreak sooner than it did, had it been more attentive in monitoring its firewall and IDS. This would have allowed AHS to respond sooner and potentially fewer Albertans would have been impacted. However, it is doubtful that these measures would have prevented the outbreak entirely.

[43] AHS had an anti-malware system, firewalls and an intrusion detection system in place. In my opinion, these are reasonable controls to protect health information against malware. I noted some areas for improvement above, but it is important to understand the HIA holds custodians to a standard of reasonableness, not perfection. Just because a system can be improved does not mean it was unreasonable in the first place. It is possible for a custodian to take reasonable steps to protect against a given threat and still experience a breach of privacy. In fact, I believe it would be unreasonable to expect custodians to prevent viruses for which no definition is available.

[44] Therefore, I find the custodian did not fail to safeguard health information in contravention of section 60 of the *Health Information Act*.

[45] This is not to say there is no room for improvement in AHS' stance with regard to malware. AHS' malware detection and response measures can be improved. AHS acknowledges this and has presented several recommendations to improve its defenses, which I will review in the next section.

[9] *Ibid*, page 40.

AHS Internal Review and Recommendations

[46] AHS conducted an internal forensic review of this incident and presented a report and recommendations to the Information and Privacy Commissioner (as he had requested in a news release from our office on July 8).

[47] The HIA requires that custodians periodically review their controls that protect the confidentiality of health information.[10] An incident such as the Coreflood outbreak can be used as an opportunity to review weaknesses in technical controls and identify areas for improvement. I am pleased to report that AHS has done so in this case. AHS identified a number of areas, outlined below, where it can improve its malware defenses, and has committed to taking remedial action.

Intrusion Detection System monitoring and management

[48] As noted earlier, AHS could have detected the Coreflood outbreak sooner had it been monitoring its IDS more attentively. AHS will conduct a review of its IDS monitoring and management procedures. This review will clarify responsibility for identifying and reporting threats discovered through the IDS.

Elevated permissions review

[49] Some AHS staff members require elevated, or administrative, permissions to make changes to computers and networks within the larger AHS network infrastructure. Coreflood, and many other viruses, need to reach a computer that is running with elevated permissions so they can spread to all of the other computers the administrative user controls. AHS will review administrative accounts to identify users who have been granted elevated privileges that are not needed and remove those elevated privileges. Those users with a legitimate business need for elevated privileges will be given a normal computer account for day to day use and will only log on to their administrative account when necessary. At any given time some administrator accounts will necessarily be active so this measure will not completely prevent the spread of all future virus outbreaks; however it will serve to contain the damage they cause.

Major threats task force

[50] AHS responded to the Coreflood outbreak by assembling a team with the technical expertise to remove the threat. This team was assembled under AHS' normal incident-response process. AHS' investigation identified a need to establish a task force with specialist knowledge in computer forensic practices to respond to major threats to its network infrastructure. AHS will identify and train staff to be members of this task force.

Forensic tools

[51] In response to this incident, AHS created an ad hoc secure computer environment to test the behavior of Coreflood and decrypt data. Valuable time was lost while AHS assembled the hardware and software to perform this

[10] See *Health Information Regulation*, section 8(3).

analysis. AHS will proactively set up a secure test environment and obtain forensic tools to remain on standby until needed in similar situations in the future.

Anti-malware rapid deployment

[52] As noted earlier, anti malware system vendors release virus fixes that have not been fully tested (see paragraph 19). These early releases can be implemented in an emergency situation to prevent the spread of malware. At the same time, organizations are naturally hesitant to implement these fixes as they may interfere with other computer applications. This risk is especially relevant in complex network environments such as AHS'. AHS will develop an emergency risk assessment process to evaluate early releases of virus fixes to determine whether they can safely be deployed on its network during malware outbreaks.

Online resources

[53] Information on malware is freely available on the internet. However, more detailed analysis of the latest malware is available through paid subscription services. AHS relied on free information in responding to this breach. AHS will pay to subscribe to the more comprehensive online resources to better respond to emerging threats.

[54] I agree that the above measures will improve AHS' malware prevention, detection and response. AHS has committed to implementing all of the above measures by March 31, 2010.

Recommendations

[55] In light of the commitments AHS has made to improve its processes to protect against malicious software, I have no further recommendations.

[56] AHS has agreed to meet with the Information and Privacy Commissioner in six months to review its progress in meeting the commitments outlined in this report.

Conclusion

[57] This investigation established that Alberta Health Services had reasonable measures in place to protect against malicious software. Despite this, AHS became infected with a new variant of the Coreflood Trojan horse program and Albertans' health information was disclosed to an unknown party. AHS responded responsibly by performing a thorough forensic investigation, informing those affected by this breach and committing to improving its practices. I would like to thank AHS for its full cooperation with my investigation.

Brian Hamilton
Portfolio Officer, Health Information Act
Office of the Information and Privacy Commissioner of Alberta