My topic will address the security issues and flaws as well as a different implementation approach of an Active Directory authentication method. As I currently use Windows AD for work, I would like to answer what kind of threats exist for this form of authentication on various levels, such as the user level like the server level or even the software level. The papers I chose delve into a new implementation of enhanced security, demonstrate some attacks on the software and what these can perpetrate, or provide an approach to monitor the security logs in a fashion that aids the system administrator. When finished, I plan to take the information and practically implementing in my workplace so that I may strengthen my office's security as well as the software we rely on to aid end users.

Paper 1:

Active Directory and Related Aspects of Security

Afnan Binduf, Hanan Othman Alamoudi, Hanan Balahmar, Shatha Alshamrani, Haifa Al-Omar, Naya Nagy
IMAM ABDULRAHMAN BIN FAISAL UNIVERSITY
Dammam, Saudi Arabia
fbinduf@hotmail.com, hananamoudi2@hotmail.com, hanan50@outlook.com, shtha.7.homoud@gmail.com, Haifaomar.ho@gmail.com, nmnagy@iau.edu.sa

Abstract

This paper discusses active directory that is used across many organizations to centralize control of users' logins to organization resources and network. A Saudi company that does not implement active directory which enables centralized, secure management of an entire network is analyzed in this paper. Active directory servers have some vulnerabilities that affect the security of active directory. There are many guidelines recommended to address security issues. Active directory has built-in support of security controls, and there are many possible ways to enhance the security of active directory and the network. Not implementing active directory in large organizations lead to the loss of control over user's resources and information which might result in serious security threats.

Keywords: Active directory, Group Policy Preferences (GPP), Kerberos protocol, Domain controller, IPSec protocol, change auditor, failover cluster instances, Domain admin, AlwaysOn availability group, Azure active directory

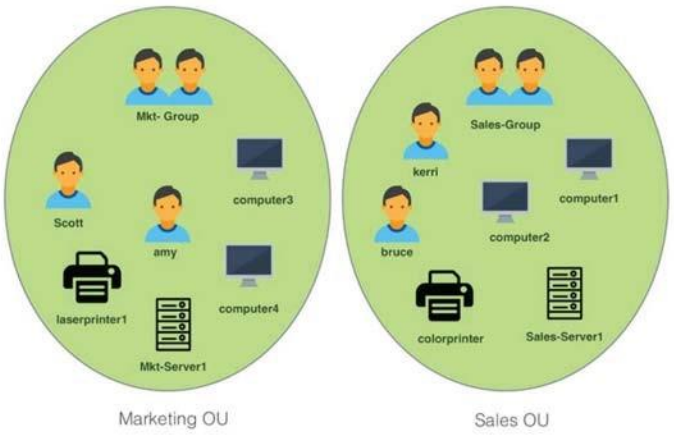
I. INTRODUCTION

Nowadays most of the companies use active directory, and it is hard for any company to work without the active directory. Active directory is a central repository for information of all company's resources that exist in the network, like employees, groups, devices, printers, programs, and documents. So, the administrators of the Active directory can efficiently manage the company's information from a central repository. [1] Although most of the companies use active directory, however, only a few of the companies know how to use it securely and know how to avoid vulnerabilities in active directory. Information system auditing has been done on a Saudi company. The communication was with a manager in the company to perform this audit. The audit that has been done focused on important risks that affect the security of the company. This company uses the different system similar to active directory to manage financial department only. This is considered one of the risks that threaten the security of the organization. Active directory comes with windows server and it can be used to manage entire company. It designed to work with window operating system. It provides scalability, security, and central management. The company requested to remain anonymous for security reasons. This company is producing indoor and outdoor lighting products. The result of this audit recommended that active directory should be implemented in all department of the company, so the main goal of this research based on IS auditing experience with this company is to discuss the important role of active directory in managing users and resources of the organization and maintaining acceptable levels of security in the system. In this paper, vulnerabilities of active directory servers are discussed in section 2. The third section discussed the security of active directory which is the alert feature, maintain integrity and confidentiality in active directory network, user authentication in active directory and active directory availability

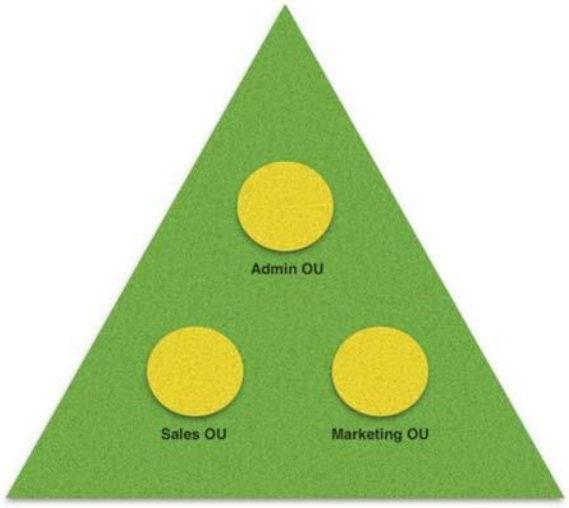
group. At the end of the paper, future work and conclusions are described.

II. BACKGROUND

The active directory provides central services because it contains all contents of organization database such as resources, service, user accounts, shared folders, etc. To manage network resources. Active directory domain service is a service provided by Windows, started in Windows Server 2000 and evolved over the years through the versions of Windows and has reached Window Server control authentication. [2] There are two parts of the active directory structure: the first part is a physical structure which consists of location and servers configured as domain controllers and the second part is a logical structure which consists of four organizing components that are considered as a container, based on the geographical area shown on figure 1. An organizational unit (OU) represents a city, the domain represents the state, a tree represents the country, and a forest represents the continent. Since the company's scope ensures several elements that need to be grouped together, it means all the information present in the central location that facilitates the work process in the organization. [3]. If the active directory is not implemented by the organization, it will lose the following advantages:



I Active Directory organisation unit



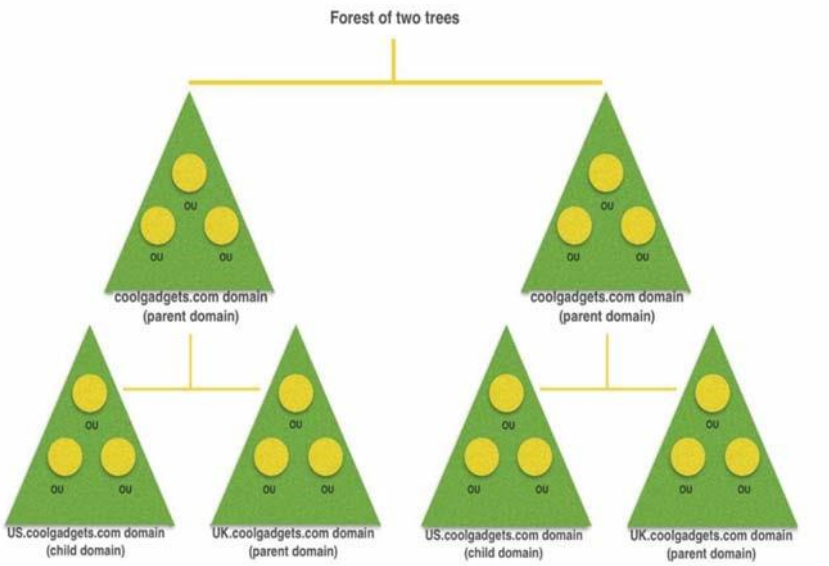
II An Active Directory domain and OUs

- Simplifies network resource management and security policy management in a hierarchical organization of active directory. [3]
- The ability to meet the increased and growing needs of the organization.

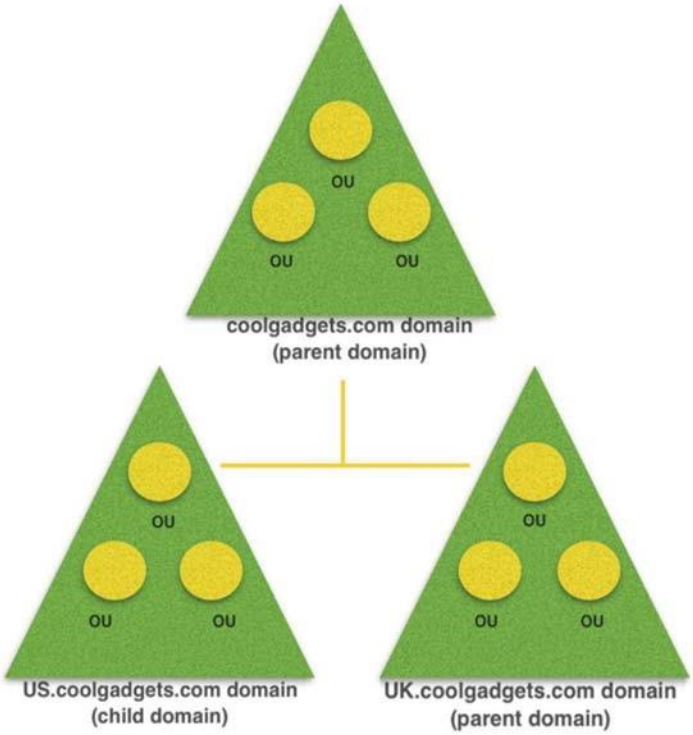
  Therefore, after the active directory installed it allows modifying the properties and adding objects. [3]
- Allow managing the organization from one point. [3]
- It enhanced security by providing a secure login more than another directory service. It's used IPSec protocol in Windows Server 2000/2003 [4] and Kerberos protocol used in Windows Server 2012. [1]

The following papers [5], [6] discuss another aspect related to active directory which are different from the topics discussed in this paper. The first article discusses continuous auditing tool of Active Directory.

The second paper design method that are used to improve the confidentiality of active directory service.



IV An Active Directory forest



III An Active Directory tree

III. Vulnerabilities of Active Directory Servers

The Saudi organization under scrutiny does not use an active directory to control the employee's access resources, so the organization should choose an appropriate type of active directory server because many of these servers have a vulnerability that Microsoft organization still make security updates to resolve the vulnerabilities. Windows Server 2000 and 2003 has a vulnerability like remote code execution and prone to denial of service attack(DoS). Remote code execution it happens when the attacker executes some arbitrary code on the system to allow the attacker to control the system by creating the account, changing, deleting data. Denial of service attack happens when the attacker sends malicious queries to a system. The attacker can exploit the vulnerability and make the Active Directory service become unresponsive [7]. Denial of service does not stop here it also attacks Windows Server 2008 R2, Windows Server 2012, and Windows Server 2012 R2 [8].

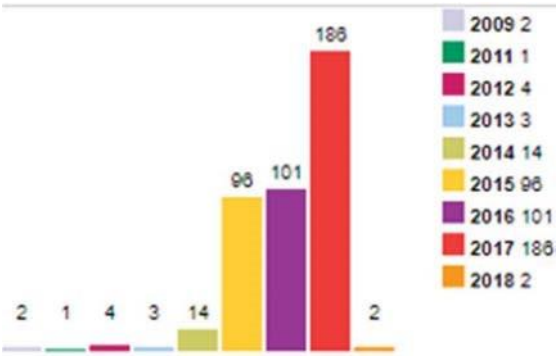


figure 2: analysis of vulnerabilities by year that mentioned in https://www.cvedetails.com/version/121761/Microsoft-Windows-Server-2008-.html.

Figure 2 shows that exploit vulnerabilities increase over time, as applied to the year 2017. From figure 3 and 4, the number of vulnerabilities is 186 and the most common type of vulnerability in windows server 2008 is gain information and gain privileges in windows server 2012.

figure 1: Types of active directory logical structure

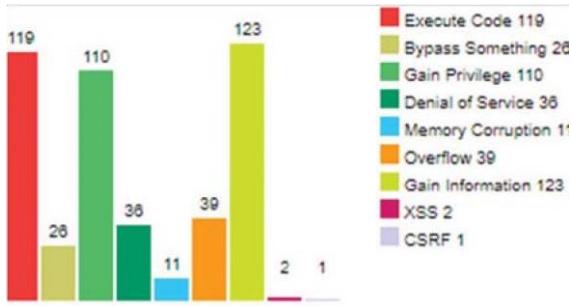


figure 3: analysis of vulnerabilities by type that mentioned in https://www.cvedetails.com/version/121761/Microsoft-Windows-Server-2008-.html

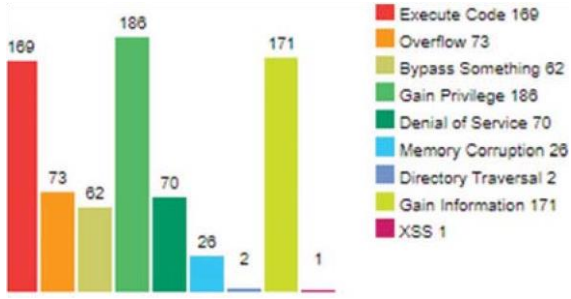


figure 4: analysis of vulnerabilities by type that mentioned in https://www.cvedetails.com/product/23546/Microsoft-WindowsServer-2012.html?vendor_id=26.

From the figure 3, it shows the other vulnerabilities that affect active directory like Cross-site scripting (XSS) vulnerability in Active Directory Certificate Services. Cross-site scripting exists on Microsoft Windows Server 2003 SP2 and Server 2008 Gold, SP2, R2, and R2 SP1.Cross-site scripting happens when the attacker injects a web script code. The attacker can exploit this vulnerability by sending to the client/user a link and make them visit the vulnerable website by clicking on the link [9]. Second vulnerability that affects Microsoft Active Directory is a buffer overflow. Exploits this vulnerability can allow attackers to execute arbitrary code with network service privileges. When the attacker fails to exploit the Cross-site scripting vulnerability it will cause a denial of service and it can happen in Microsoft Windows XP SP2 and SP3, Windows Server 2003 SP2, Windows Vista SP2, Windows Server 2008 SP2, R2, and R2 SP1, and Windows 7 Gold and SP1 [10].

IV. Security Issues of Active Directory

There are a lot of security issues that give the attackers opportunities to gain access to the Active Directory. Here are the most common active directory security issues.

Service Accounts has over-permissioned.

A service account is an exceptional type of account that often provide a lot of privileges and permit services to contact with the underlying operating system. Further privileges added to this Account can be used maliciously to heighten of access rights which are a serious security danger. It is very important to make sure that each Service Account is deputize required right. The service that running underneath service account has a credential in LSASS (protected memory) which can be extracted by an attacker. If the stolen credential has admin rights, the domain could be compromised [11].

A lot of Domain Admins.

The entire administrative rights of " all workstations, servers, Domain Controllers, Active Directory, Group Policy" are done by the members of Domain Admins. By default, this is too excessive power for any account in a company. Typically, Domain Admins involve Service Accounts and further groups that are not immediately linked to Active Directory administration. Preferably the Domain Admin group should be idle to make sure that each role has only the required right to execute tasks associated with that role. Only Active Directory administrators require privileges Domain. Anyone is not handling Active Directory in an active way, must not be in Domain Admins anymore [11].

The same passwords for a local Administrator account on all systems.

Most system administrators using the same username and password for local administrator account, this allow administrator to log on to other accounts using the same username and password, which is not good idea from a security point of view, because this makes it easy for the attacker to gain access to all systems once the attacker get the credentials of one local administrator account, so it prefer to have a unique credentials for each local administrator account [11].

Domain Controllers are running an old version of OS.

Further advance of security improvements appears with every consecutive version of Windows Server, and prior security defect is patched. If the newer release of the operating system is not installed, some security dangers will appear. For instances, Domain Controllers faced security risks when it is running in the older version of Windows Server. [11]

Using Group Policy Preferences (GPP) to handling credentials. More functionality is providing to the system administrators by Group Policy Preferences (GPP). Group Policy Preferences (GPP)can change the password of the local administrator account, create local account and services, etc. the password that stored in the XML file which placed in SYSVOL share is created significant issues because any domain can gain access to the files in SYSVOL. If the credential is previously configured in GPP, eliminate it directly and Remove the files [11].

The password length of Service Accounts is less than 20 characters.

It is simple to request data that has been encrypted with the password of Service Account. It is feasible to decode the data by using brute forced offline and reveal the password for the account if the password is supported by the Kerberos network authentication protocol. This issue can be mitigated when the password of Service Accounts is more than twenty characters [11]

V. The Security of Active Directory *A-Alert Feature*

Most companies want to know exactly who did a specific action in their workplace? The main issue is that they need to be informed by some alert method when someone changes such a file, and this can happen using Active Directory Auditing Agents. Knowing every single event in the system increases the level of accountability. Auditing, protection, and alert are advantages provided by Active Directory Change Auditor to enhance the accountability. When an event occurs, people who have the right will be alerted automatically. One more

example that admins set the configuration for tracking the use of the account. If you install a new service in the organization, usually the service will be available for pre-selected devices. With change auditor service alert, can be configured to maintain the accounts used with other nodes and admins can be notified [12].

B-Maintain Integrity & Confidentiality in Active Directory Network

As active directory database contains the most sensitive data focusing on the secure active directory and its network traffic is important. IPsec which stand for IP security is network layer protocol that secures the traffic on the network and protects from different network attacks. IPsec used to add a security layer to the network and therefore participate in protecting the active directory. IPsec can secure the transmission between two ends by encrypting packets or by securing transmission path between two IP addresses. The encryption provided by IPsec protect against eavesdropping attack that might affect the confidentiality through the network and cause further attacks. IPSec provides a checksum to ensure that packet did not alter or modified through its transmission and to detect attacks such as session hijacking and man in the middle. Domain controllers host active directory database and because of that protecting domain controllers is necessary. Domain controller must be placed in a secure physical place to protect it from unauthorized access and modification of the files or configurations. Setting a strong password to administrative account is also an

important step because this account has access to the whole network. Renaming or disabling administrator account may provide more security since it is targeted by hackers. Installing security updates is necessary to patch vulnerabilities that might be exploited against the system. Antiviruses play important role to protect domain controllers from malicious malware and should be up to date to detect new viruses and malware [4].

C-User Authentication in Active Directory

Users access to network resources are controlled through logon process where the user must provide his or her credential to gain access to services and applications [13]. User authentication mechanism in active directory carried by Kerberos protocol. Kerberos protocol is security protocol that provides flexible authentication. Instead of sending user credentials over the network, key is created for the user session and used for short limited time. Authentication of the user while using Kerberos is required only one time and once the user authenticated he can access services and applications without the need to log in again. Sign in for each application independently lead to some problem [14]. Authentication process starts with authentication service request (AS_REQ packet) which contains the client username, service name and the current time of authentication request that used by domain controller to make sure that the logon is current and avoid a replay attack. figure 5 shows AS_REQ packet

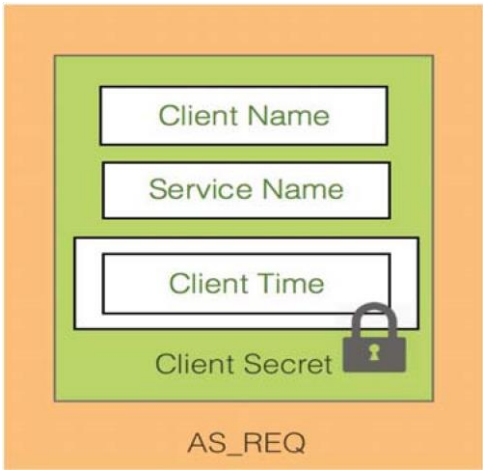


Figure 5: The authentication service request packet

The domain controller is responsible for validating user authentication request [15] and issuing a ticketgranting ticket (TGT) which cached and used by windows, so the user does not have to logon to services again. Authentication service response (AS_REP packet) have TGT encapsulated with it. Figure 6 shows AS_REP packet.

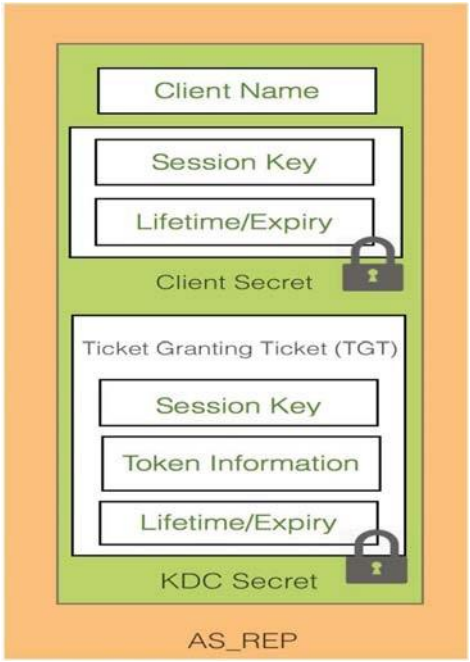


Figure 6: The authentication service response packet

Session key used to communicate with the domain controller. Lifetime /Expiry is a limited period defined by TGT when it expires TGT must be renewed or authentication request must be done again. Session key and lifetime/Expiry are encrypted using user password hash. TGT contain token information which is about user information like his access right, groups he belongs to. TGT encrypted using a hash of the KDC's secret which is the hash of krbtgt account credential for domain controllers [1].

D-Active Directory Availability Group

When we are looking for availability, the newest Azure Active Directory Availability group provide several services that enhance and secure the active directory model. Regrettably, considering availability often be the last step after the project is finished. Azure usually puts the high availability in the top feature need. In azure active directory data are kept in triple redundancy. However, services are not redundant and losing one node can be less expensive than losing one application. This can be done when companies are using failover cluster instances (FCIs) with the provided AlwaysOn Availability Groups feature. Availability groups in Azure serve the following features. Services like Availability sets, which allow working on the data in multiple locations shown in figure 6 Other services like Management services, Notification hubs, AutoScale and Virtual machines [12].

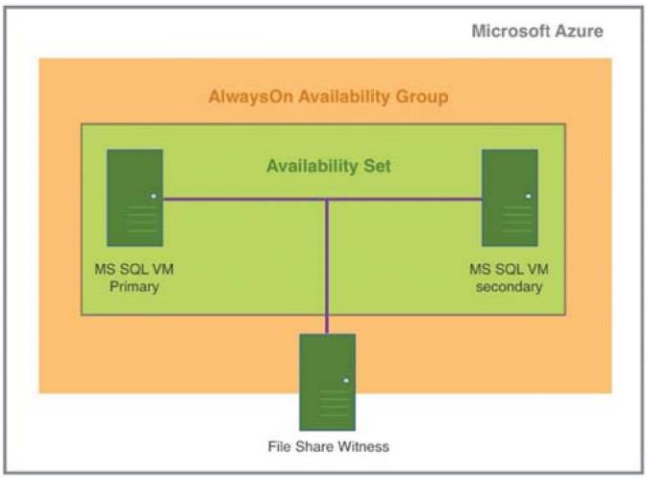


figure 7: Availability groups in Azure, using an availability set.

VI. CONCLUSION

Active directory is a very common solution for any company that wants to control and manage information and resources. It is integrated into windows servers which have some vulnerabilities that could affect active directory in the organization. maintaining the security of windows server enhance the security of active directory. There are different features of the active directory that might raise security concerns if not used properly. Active directory has pre_built in security measures and many capabilities that help in managing security in the organization. securing the environment where the active directory is placed is the responsibility of the company. While this paper discussed active directory as the important system for organization security, there are too fighting chances to make this scope exceed the level that has been reached. Developing a business case for the active directory can be presented in future papers to refine our findings. Comparing and a contrast between active directory and other similar available systems should be encountered. In addition, a plan to implement active directory based on benchmarking can be developed.

REFERENCES

[1] B. Desmond, J. Richards, R. Allen and A. G. Lowe-Norris, *Active Directory*, 5th ed. Sebastopol: O'Reilly Media, Inc, 2013, p. 1-281.

[2] P. Pengsart, A. R. X. Belo, J. X. Vaz, J. B. S. Marques and E. Junior, "ADFS Authentication for Healthcare System," in *International Conference on Information Technology*, Nakhon Pathom, 2017.

[3] G. Tomsho*, MCTS Guide to Configuring Microsoft Windows Server 2008 Active Directory*. 2009, pp. 76-77.

[4] L. Hunter, *Active Directory Field Guide*. Apress, 2005, pp. 149-176.

[5] M. Derek, "REVOLUTIONIZING CONTINUOUS AUDITING OF DIRECTORY", vol. 29, no. 4, pp. 42-44, 2018.

[6] P. C. R. V. Parmi, "An Advanced approach of Active Directory Techniques," *International Journal of Information and Technology (IJIT),* p. 7, 2015.

[7] "Microsoft Windows Active Directory Denial of Service Vulnerability", *Tools.cisco.com*, 2018. [Online]. Available: https://tools.cisco.com/security/center/viewAlert.x?alertId=53262. [Accessed: 05- Feb- 2018].

[8] (MS07-039) VULNERABILITY IN WINDOWS ACTIVE DIRECTORY COULD ALLOW REMOTE CODE EXECUTION (926122) - Threat Encyclopedia - Trend Micro US", *Trendmicro.com*, 2018. [Online]. Available: https://www.trendmicro.com/vinfo/us/threatencyclopedia/archive/security-advisories/(ms07-039)%20vulnerability%20in%20windows%20active%20directory %20could%20allow%20remote%20code%20execution%20(926122). [Accessed: 02- Feb- 2018].

[9] "Active Directory Certificate Services Vulnerability - oval:org. mitre. oval:def:12749", *Itsecdb.com*, 2018. [Online]. Available: http://www.itsecdb.com/oval/definition/oval/org.mitre.oval/def/12 749/Active-Directory-Certificate-Services-Vulnerability.html. [Accessed: 02- Feb- 2018].

[10] "Active Directory Buffer Overflow Vulnerability - oval:org. mitre. oval:def:14037", *Itsecdb.com*, 2018. [Online]. Available: http://www.itsecdb.com/oval/definition/oval/org.mitre.oval/def/14 037/Active-Directory-Buffer-Overflow-Vulnerability.html. [Accessed: 02- Feb- 2018].

[11] S. Metcalf, "The Most Common Active Directory Security Issues and What You Can Do to Fix Them – Active Directory Security", *Adsecurity.org*, 2018. [Online]. Available: https://adsecurity.org/?p=1684. [Accessed: 03- Feb- 2018].

[12] M. Copeland, J. Soh, A. Puca, M. Manning, and D. Gollob, *Microsoft Azure: planning, deploying, and managing your data center in the cloud:* New York: Apress, 2015.

[13] D. J. R. K. Jaroslav Kadlec, "Implementation of an Advanced Authentication Method Within Microsoft Active Directory Network Services," in *2010 Sixth International Conference on Wireless and Mobile Communication*, Brno, 2010.

[14] H. Wang and C. Gong, "Design and Implementation of Unified Identity Authentication Service Based on AD," *2016 8th International Conference on Computational Intelligence and Communication Networks (CICN)*, 2016.

[15] C.-M. L. C.-H. M. a. T.-C. K.-C. L. Chih-Hung Hsieh, "AD2: Anomaly Detection on Active Directory Log Data for Insider Threat Monitoring," in *2015 International Carnahan Conference on Security Technology (ICCST)*, Taipei, 2015.

Paper 2:

AD^2: Anomaly Detection on Active Directory Log Data for Insider Threat Monitoring

Chih-Hung Hsieh
Institute of Informaiton Industry
Taipei, Taiwan
Email: chhsieh@iii.org.tw

Chia-Min Lai, Ching-Hao Mao, and Tien-Cheu Kao
Institute of Informaiton Industry
Taipei, Taiwan
Email: {senalai, chmao, tckao}@iii.org.tw

Kuo-Chen Lee*
Institute of Informaiton Industry
Taipei, Taiwan
Email: kclee@iii.org.tw

Abstract—What you see is not definitely believable is not a rare case in the cyber security monitoring. However, due to various tricks of camouflages, such as packing or virutal private network (VPN), detecting "advanced persistent threat"(APT) by only signature based malware detection system becomes more and more intractable. On the other hand, by carefully modeling users' subsequent behaviors of daily routines, probability for one account to generate certain operations can be estimated and used in anomaly detection. To the best of our knowledge so far, a novel behavioral analytic framework, which is dedicated to analyze Active Directory domain service logs and to monitor potential inside threat, is now first proposed in this project. Experiments on real dataset not only show that the proposed idea indeed explores a new feasible direction for cyber security monitoring, but also gives a guideline on how to deploy this framework to various environments.

Keywords—Active Directory Log Analysis, Anomaly Detection, Behavioral Modeling, Machine Learning, Advanced Persistent Threat.

I. INTRODUCTION

The so called "advanced persistent threat"(APT) is a set of stealthy and continuous computer hacking processes. Typically, APT processes take a high degree of covertness over a long period of time, use malware of sophisticated techniques to exploit vulnerabilities in systems, and continuously monitor or extract confidential data from specific targets, once attackers get the control of victim systems. According to cyber security technical reports from various organizations or companies, it takes averagely at least 346 days for more than 81% victims to aware that they have been hacked [1]. Yet, due to various tricks of camouflages, such as packing, obfuscated shellcode, or virtual private network (VPN), signature-based malware detection technique, such as intrusion detection system (IDS), is getting less useful, especially on identifying advanced persistent threat(APT) at post-compromised stage. Facing the above challenges, considering account's subsequent behaviors may gives cyber security engineers additional context-based evidences and smart chance to detect threats from insiders [2].

In this report, instead of expert rules with only a few pre-defined signatures, a threat detection method regarding account's behavior sequences recorded as Active Directory (AD) logs was proposed. The AD domain controller of an organization monitors all related information when any intranet accounts try to allocate or acquire various resources and services. To the best of our knowledge so far, none of previous works are dedicated to threat identification by using sequential *: Corresponding Author

AD data modeling. The proposed framework 1) takes the AD logs as time-series input data providing chronological evidences, 2) emphasizes on sequential context mining from collected AD log, and 3) for each account, build the probability Markov model where best depicts the corresponding likelihoods of different behaviors occurring.

For real world feasibility concerning, a real dataset of AD logs, from real organization in Taiwan containing large amount of employees, was collected. After proper pre-processing and behavior learning, the performance of the proposed framework is measured with cross-validation manner for the sake of objectively evaluating the effectiveness and robustness. A fair N-fold cross validation experiment shows that the learnt behavioral Markov model gives successful results in terms of outstanding recall rates as well as fair precisions. The merits and contributions of this paper are fourfold. 1) the inside threat detection problem is first formularized as a sequential behavior modeling problem regarding with time-series AD log data mining. 2) The learnt model has good ability at monitoring the post-compromised anomaly behavior in terms of 66.6% recall and 99.0% precision rate. 3) the useful domain knowledge provided by well-known cyber security company, TrendMicro, was encoded as an annotation profile and indeed helps building accurate model; 4) a practical guideline for future users doing parameter selection is also proposed based on series discussions of how the framework parameters effect the detection results.

In the remaining parts of this report, section II briefly introduces the target Active Directory domain service. Section III describes the kernel methods of the proposed framework, including Markov model, the encoded

representation for domain knowledge, and the designed anomaly detection mechanism. The performance of the proposed framework is evaluated in Section IV. At last, Section V concludes this project.

II. BACKGROUNDS & MATERIALS

A. Active Directory Domain Service

The Active Directory domain service is a directory service that Microsoft developed for Windows domain networks [3][4]. An AD domain controller authenticates and authorizes all users and computers in a Windows domain type networkassigning and enforcing security policies for all computers and installing or updating software. For example, when a user logs into a computer that is part of a Windows domain, Active ~~Directory~~ checks the submitted password and determines ~~whether the~~ user is a system administrator or normal user [5]. The AD domain controller of an organization monitors all related information when any intra-net accounts try to allocate or acquire various resources and services. To the best of our knowledge so far, none of previous works are dedicated to threat identification by using sequential AD data modeling. Figure 1 is an illustration example of how a "Kerberos" process, one variant of AD domain, works and interacts with intra-net accounts. The process goes as the following steps, detail about this "Kerberos" process example can be referenced from [6]:

Stage 1 The Authentication Service Exchange

Step 1 Kerberos authentication service (KRB AS REQ) request

Step 2 Kerberos authentication service response (KRB AS REP)

Stage 2 The Ticket-Granting Service Exchange

Step 3 Kerberos ticket-granting service (KRB TGS REQ) request

Step 4 Kerberos ticket-granting service response (KRB TGS REP)

Stage 3 The Client/Server Exchange

Step 5 Kerberos application server request (KRB AP REQ)

Step 6 Kerberos application server response (optional) (KRB AP REP)

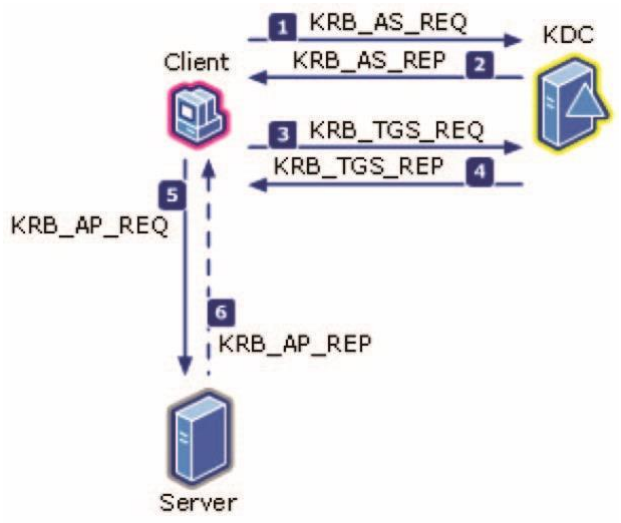


Fig. 1. An illustration of AD domain controller works and interact with intra-net accounts.

B. The Real Dataset

The performance of the proposed framework is evaluated in this section. In this experiment, a certain government organization of 95 employees in Taiwan was selected as the deployment environment of the proposed method. The Active Directory domain service using version of Windows Servers 2008 R2 is mounted on this organization's domain network and keeps monitoring all service requests and resource allocations raised ~~from int~~ra-net account. In this circumstance, Total 12,310,519 logs w~~it~~h size of 22.5 giga-bytes (GBs) was collected during two months, from 2014/11/26 to 2015/01/02. Among the 22.5 GBs data, logs with respect to accounts of real employees was left to be analyzed and forms an original dataset D in the following experiment. At last, D consists of 2,887,504 logs recorded in a 5.2 GBs file from 95 employees. It shoud be mentioned that because the number of event code "4624" and "4634" extremely dominate than those of others, and may lead to a biased Markov model. For the dataset D in the following experiment, those two event codes are safely ignored as considering model state.

III. METHOD

The ultimate goals of this project are as followings: 1) to collect Active Directory log data which is generated once any account try to access the Active Directory domain service; 2) to build behaivoral model capable of describing personal tendency for each user; and 3) to estimate the resulted likelihood given one's model to generate certain subsequent event codes. Based on the requirements mentioned above, the structure of the proposed framework can be made up of three major modules. The first is responsible for pre-processing raw data composed of whole access logs caused by all intranet accounts and for forming the input dataset of following analytic usages. In the second module, The Markov model, the famous machine learning algorithm and well-known as a consequent state changing modeling tool, is then adopted to be the kernel approach to summarizing the user's behaviors. The last one is designated as an anomaly detection process to determine the likelihood that a certain account's model produce the given input sequences of event codes. Figure 2 shows the whole framework comprising of the three modules. The remain parts of this section give the brief introduction of adopted Markov model, the usage of prior knowledge, and the details about those three major modules.

A. Markov Model

In probability theory, a Markov model is a stochastic approach to model randomly changing systems where it is assumed that future states depend only on the present state and not on the sequence of events that preceded it (that is, it assumes the Markov property)[7]. Generally, the reason of taking this assumption is because it enables subsequent reasoning and computation regarding the model that would otherwise be intractable. The simplest Markov model is the Markov chain. It models the state of a system with a random variable that changes through time. In this context, the tendency of every transition form one state to another is described by a probability. The Markov property suggests that the distribution of this probability depends only on the distribution of the previous state. Due to the advantage of being good at describing consequent state changing, there are lots of applications, such as speech recognition [8], hand-written text recognition [9], gesture recognition [10], and cyber security intrusion detection[2], based on Markov model or its popular variant, hidden Markov model. Figure 3 is an illustration example of a Markov model with 3 states [11].

B. Generic Markov-model-state Annotation profile

Because the proposed method try to build the behavioral model for each account and to detect anomaly once account

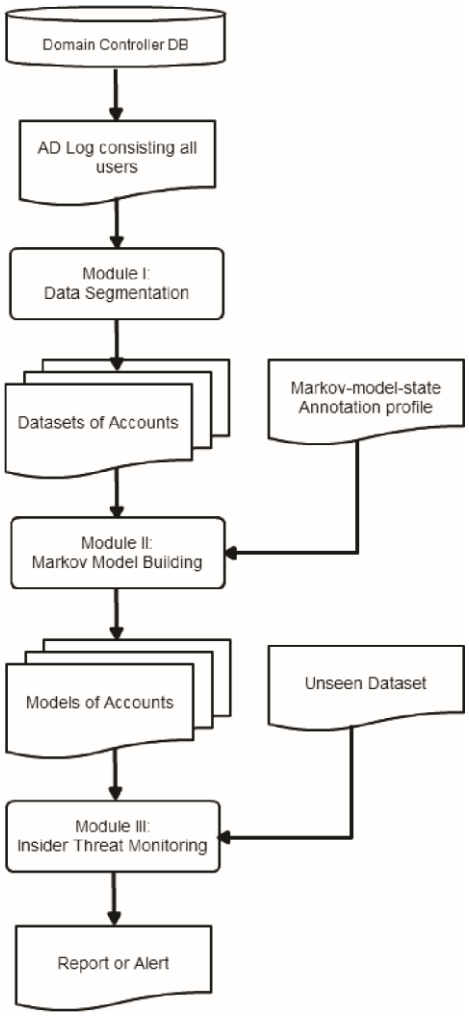


Fig. 2. The proposed framework of inside threats monitoring

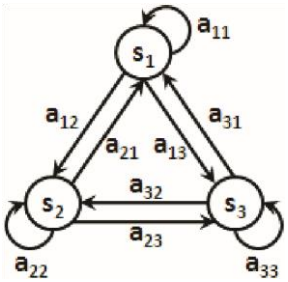


Fig. 3. An example of 3-state Markov model.

do not act like themselves compared to their historical daily routine. For this purpose, instead of considering only Active Directory code, making use of additional information as more as possible may significantly improve the model's representative degree. In this subsection, we define a Markovmodel-state annotation (MMA) profile to encode proper prior knowledge by describing that which event code should be co-considered with certain specific attributes as a complete Markov state. For example, according to the Active Directory domain service specification provided by Microsoft, event codes "4771" represents that an account was pre-

authentication failed for some reason. Therefore, it will provide much useful information about details of failure when event code “4771” be co-considered with the attribute “result code”. Figure 4 is an illustration of a Markov-model-state annotation profile. It includes three annotations for three different event codes, respectively. For example, event code of no.4623 should be co-considered with both two fields, “xxx” and “yyy”, to be formed as one state of Markov model. Note that for any event code which is unlisted in profile, it means that this event code is adopted default setting where implies using only event code itself as Markov state.

Annotation 1:

$$Event\_Code = 4623\ with$$
$$Field\_Name_1 = xxx,$$
$$Field\_Name_2 = yyy.$$

Annotation 2:

$$Event\_Code = 4624\ with$$
$$Field\_Name_1 = xxx,$$
$$Field\_Name_2 = yyy,$$
$$Field\_Name_3 = zzz.$$

Annotation 3:

$$Event\_Code = 4723\ with$$
$$Field\_Name_1 = aaa,$$
$$Field\_Name_2 = bbb,$$
$$Field\_Name_3 = ccc.$$

Fig. 4. An example of Markov-model-state annotation profile.

C. Module I: Generate Dataset of Event Code Chains.

Because the collected raw data is mixed with AD log data coming from multiple different accounts and the proposed method try to build separate behavior model for each account, one function of Module I is to partition original AD log data into different files of dataset, one for each account. Besides, the adopted Markov approach models one account’s multiple instances of different operations as a probabilistic model. Different operation instances also need to be divided from a long consequent log sequence of event code into multiple shorter event code segments. The hypothesis used here to separate out those event code segments is that considering idle time in naive but real circumstance, the time intervals inter two independent segments of operations are usually longer than those intra one operation. For this reason, during the idle time, the time intervals between two subsequent AD logs are usually longer than those in working periods. In module I, a real-valued parameter θ is adopted as a cutting threshold. θ is longest allowed time interval and is used to divide a event sequence into two segments. If a time interval of two subsequent AD logs are longer than θ, it will lead two ~~different~~ event segments generated by cutting out the idle ~~t~~ime.

D. Module II: Build Markov Model given an Event Chains Dataset

This section defines what components constitute adopted Markov model and how to build Markov model given the dataset consisting of event chains of an employee.

Given the dataset D_i containing event chains of the i^{th} employee, $i = 1,\ldots,E$, the resulted Markov model based on the dataset should include following components:

1) A finite state set $S = \{s_1,s_2,\ldots,s_{ns}\}$ that contains all possible states of Markov model defined by the MMA mentioned in previous section and derived from D_i. Note that ns is the total number of derived states in Markov model;
2) An $1 \times ns$ initial probability (IP) vector, IP = $[ip_1,\ldots,ip_{ns}]$ where ip_i represents the probability that i^{th} state is the initial state of an event code segment, and $\sum_{i=1}^{ns} ip_i = 1$.

   $\forall\ i = 1,\ldots,ns,$

   $$ip_i = \frac{\#\text{event chains starting with } s_i \text{ in } D_i}{\#\text{event chains in } D_i}$$

3) An $ns \times ns$ transition probability (TP) matrix, as following:

$$TP = \begin{bmatrix} tp_{1,1} & \ldots & tp_{1,j} & \ldots & tp_{1,ns} \\ \ldots & \ddots & \ldots & \ldots & \ddots & \ldots \\ tp_{i,1} & \ldots & tp_{i,j} & \ldots & tp_{i,ns} \\ \ldots & \ddots & \ldots & \ldots & \ddots & \ldots \\ tp_{ns,1} & \ldots & tp_{ns,j} & \ldots & tp_{ns,ns} \end{bmatrix}$$

   where for each i and j, $tp_{i,j}$ represents the transition probability from i^{th} state to j^{th} state, with constraints that

$$\sum_{j=1}^{ns} tp_{i,j} = 1, i = 1, \ldots, ns$$

$$\forall\ i, j = 1, \ldots, ns,$$

$$tp_{i,j} = \frac{\#\text{transitions starting from } s_i \text{ to } s_j}{\#\text{transitions from } s_i \text{ in } D_i}$$

Given an observed event code segment, $c = [o_1,o_2,\ldots,o_T]$ of length T, o_t means the t^{th} observed Markov state in c. And the model probability (MP) that Markov model of (S,IP,TP) generates event code segment c can be calculated by:

$$MP(c) = ip_{o_1} \prod_{t=1}^{T-1} tp_{o_t,o_{t+1}}.$$

E. Module III: Probability Estimating given Markov Model

The anomaly detection mechanism for inside threat monitoring is implemented in Module III. After generating training dataset by using Module I, the output of Module II is personal behavioral model accompanied with a referenced probability, P_{ref}. This referenced probability is calculated by estimating the likelihood that this Markov model generates the training dataset of itself which is used to build the corresponding model. P_{ref} provides referenced usages that how well this ~~resulted~~ model fits the used training dataset and what is the likelih~~o~~od that this employee produce corresponding logs when he did his daily routine jobs. Based on the P_{ref} and a userdefined threshold parameter, δ, the following equation (1) is designed to detect the anomaly.

$$Cond._i : \frac{NMP_i(Tr_i) - NMP_i(Dataset_{unseen})}{NMP_i(Tr_i)} \geq \delta$$

$$NMP_i(Dataset) = \sqrt[s(Dataset)]{\prod_{c \in Dataset} \sqrt[l(c)]{MP_i(c)}} \quad (1)$$

$s(Dataset)$ = size of $Dataset$ $l(c)$ = length of c

Tr_i represents training dataset for model building of employee $i = 1,\ldots,E$ and E is the maximum index of employees. Assume that $Dataset$ is a set of event segments given as input of employee i's learnt Markov model and c is any event segment belonging to $Dataset$. $NMP_i(\cdot)$ and $MP_i(\cdot)$ return the normalized and original model probabilities where i's learnt Markov model can generate the whole $Dataset$, respectively. The reason of using normalized probability is that Markov model has two characteristics: 1) the longer the length of an event segment has, the smaller resulted probability is; and 2) the more segments a dataset consists of, the smaller resulted probability of this dataset is. To cope with this situation and provide a fair evaluation between datasets or event segments with different sizes, we normalize the resulted probability, $MP_i(\cdot)$, of a $dataset$ to $NMP_i(\cdot)$ not only according to length of each event segment but also according to size of each of dataset with equation (1). Note that $NMP_i(Tr_i)$ is exactly the referenced probability P_{ref} for i^{th} employee, mentioned in the beginning of this subsection.

The idea of using equation (1) as anomaly detection mechanism is relative intuitive. Once the condition $cond._i$ is true, it means the probability of i^{th} employee's Markov model generating unseen dataset ($NMP_i(Dataset_{unseen})$) is relative lower than the i^{th} employee's referenced probability (P_{ref} or $NMP_i(\cdot)$) by a given ratio threshold δ. In this circumstance, it is quite unlikely that this $Dataset_{unseen}$ came from the i^{th} employee. Due to this hypothesis, when the $cond._i$ is true, it should trigger an alert of anomaly.

IV. PARAMETER SELECTION & PERFORMANCE EVALUATION

A. Experiment Settings

Considering both effectiveness and robustness of proposed method, the performance is fairly measured with a N-fold cross validation manner. Assume that D is the dataset containing event chains of all employees, then $D_1 \cup D_2 \cup \ldots \cup D_E = D$, where D_i is the dataset of event chains for i^{th} employee, $i = 1,\ldots,E$, and E is now 95 in current experiment. In N-fold cross validation, each D_i will first be partitioned into N folds. In each fold, one of the N parts is used for validation, named as $InnerV\ a_{ij}$, while the other $N-1$ parts are combined and formed as the so called $InnerTr_{ij}$ for model building. Note that $InnerTr_{ij} \cup InnerV\ a_{ij} = D_i$ for $j = 1,\ldots,N$.

According to our N-fold cross validation setting, in each fold j, for each employee i, the corresponding $InnerV\ a_{ij}$ will be used as the $Dataset_{unseen}$ in equation (1) to evaluate the model trained by $InnderTr_{ij}$. The model trained

$InnderTr_{ij}$ will then try to differentiate whetehr the given input $Dataset_{unseen}$ belongs to corresponding account or not. In this experiment, i = 1,…,E, j = 1,…,N, while E and N are set to be 95 and 5. It will results in total $5 \times 95 \times 95 = 45{,}125$ testing cases.

In our proposed framework, there are three kinds of parameters needed to be determined. Following are the brief reviews of them, and the ranges of parameter value to be tuned in the following parameter selection.

1) θ: This is the longest allowed time interval between two subsequent event codes in one event code segments, and is used to divide a event sequence into two segments. In this experiment, θ is set to be 3 minutes, 6 minutes, and 9 minutes.
2) δ: The usage of this parameter is a anomaly detection threshold included in equation (1). In this experiment, δ is set to be 1%, 5%, and 10%.
3) MMA: a Markov-model-state annotation profile specifies which event code should be co-considered with certain attributes as a state in the Markov model. Note that for event code not to be listed in MMA means that use the default setting. The default setting now is using event code itself only. Note that, in this experiment, the possibilites of candidated MMA can be classified into two categories. The first category is without using any domain knowledge, such that MMA can be: a) for every event code using none of attribute (None Used MMA, i.e. default setting), b) for every event code using all kinds of attribute (All Used MMA), and c) for every event code using randomly selected attributes (Random MMA). The second one is based on the domain knowledge given by our cooperated domain experts working in TrendMicro company (TrendMicro MMA). Figure 5 shows the used domain knowledge-based TrendMicro MMA which annotates event codes 4768, 4769, and 4771 with additional attributes, respectively.

B. Results & Discussions

As mentioned above, there are total 45,125 cases testing if the given input $Dataset_{unseen}$ belongs to Markov model being evaluated. The predicted result could be positive or negative. The positive case means the predicted label is "$Dataset_{unseen}$ does not belong to this account" and is also the anomaly case where the Management Information System (MIS) engineers are interested. On the other hand, the negative one represents that the input data belongs to this account. The following measurements are used to evaluate the proposed method and corresponding parameter setting. Table I, Table II, and Table III shows the different performance under different settings of maximum interval time θ = 3, 6, and 9 minutes. Based on the results of 5-fold cross validation , it can be observed that:

1) The behavioral modeling seems to take advantage of additional annotations when co-considering event code with annotated attributes as a Markov state. The effect is shown by that the TrendMicro MMA and All Used MMA obviously

Annotation 1:

$$Event\_Code = 4768\ with$$
$$Field\_Name_1 = \text{“}return\ code\text{”}$$
$$Field\_Name_2 = \text{“}failure\ code\text{”},$$
$$Field\_Name_3 = \text{“}service\ name\text{”}.,$$

Annotation 2:

$$Event\_Code = 4769\ with$$
$$Field\_Name_1 = \text{“}return\ code\text{”}$$
$$Field\_Name_3 = \text{“}service\ name\text{”}.,$$

Annotation 3:

$$Event\_Code = 4771\ with$$
$$Field\_Name_1 = \text{“}return\ code\text{”}$$
$$Field\_Name_2 = \text{“}failure\ code\text{”},$$
$$Field\_Name_3 = \text{“}service\ name\text{”},$$

Fig. 5. The Markov-model-state annotation profile from TrendMicro company.

outperform than the other two, None Used MMA and Random MMA, in terms of recall and accuracy. Although the TrendMicro and All Used MMA perform almost exactly the same, the former is still more feasible than the latter in the realistic deployed environment. Because the space complexity of Markov model state is about $O(n^m)$ while n is the number of attributes, and there are m different values in each attribute. The MMA co-considering with all possible attributes may result in the number of Markov states exponentially increasing. Therefore, the TrendMicro MMA, incorporating significant prior knowledge, provides the best and rational results.

2) In this experiment, the longer the interval time θ is, the better performance of Markov model built can deliver. The idea of the proposed framework is to model user's complete operations as multiple instances of possible behaviors. The time interval of 3 minutes may be more likely too short to contain a complete operation than using 9 minutes as maximum allowed idle time. However, this setting should be customized according to the scenario of different deployed environments based on appropriate experiments for parameter selection.
3) It is no wonder that smaller values of δ will cause the whole anomaly detection mechanism more sensitive by making the $cond._i$ in equation (1) more easily to be trigger (i.e. increasing recall rate). On the other hand, despite larger values of *delta* increase the threshold of trigger an anomaly alert, it still indeed enhances the certainty grades once any anomaly alerts are triggered (i.e. increasing precision rate). When setting the values of δ, it should concern the inevitable trade-off situation between recall and precision. Generally speaking, because the cyber security attacking causes huge amount of damage in the most cases, to make sure an acceptable recall rate is our first thumb rule.
4) Although, combining the prior domain knowledge form TrendMicro, AD^2 can not only produce highest performance in terms of recall and accuracy, but also may significantly reduce the number of possible states ~~in Markov~~ model state set (S) compared to all-used MMA~~. However,~~ AD^2 with TrendMicro MMA still can only produce about 66% recall rate or accuracy. It shows us that anomaly detection only based on AD log may be limited. Due to this reason, the future work of this study is inspired with that combining AD log with other various logs or contexts has opportunity to improve the performance of detecting anomaly.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

TP : True Positive, TN : True Negative, (5) FP : False Positive, FN : False Negative.

TABLE I. PERFORMANCE EVALUATION WITH θ = 3 MINUTES.

MMA	TrendMicro			All Used		
δ	1%	5%	10%	1%	5%	10%
Recall	64.81%	61.99%	57.98%	64.82%	61.99%	57.98%
Precision	99.09%	99.19%	99.29%	99.09%	99.19%	99.29%
Accuracy	64.60%	61.89%	58.01%	64.60%	61.89%	58.02%

MMA	None Used			Random		
δ	1%	5%	10%	1%	5%	10%
Recall	54.00%	51.77%	48.90%	56.86%	51.70%	45.14%
Precision	99.01%	99.10%	99.21%	99.02%	99.13%	99.21%
Accuracy	53.95%	51.82%	49.06%	56.76%	51.77%	45.36%

TABLE II. PE FORMANC EVALUA ION WITH θ = 6 MI UTES.

MMA	TrendMicro			All Used		
δ	1%	5%	10%	1%	5%	10%
Recall	64.88%	62,15%	58.30%	64.88%	62,15%	58.30%
Precision	99.10%	99.15%	99.28%	99.10%	99.15%	99.28%
Accuracy	64.67%	62.02%	58.32%	64.67%	62.02%	58.32%
MMA	None Used			Random		
δ	1%	5%	10%	1%	5%	10%
Recall	54.27%	52.17%	49.44%	56.20%	50.15%	42.11%
Precision	98.99%	99.04%	99.18%	99.13%	99.18%	99.26%
Accuracy	54.20%	52.18%	49.56%	56.17%	50.26%	42.40%

TABLE III. PERFORMANCE EVALUATION WITH θ = 9 MINUTES.

MMA	TrendMicro			All Used		
δ	1%	5%	10%	1%	5%	10%
Recall	66.60%	64.38%	61.37%	66.60%	64.38%	61.37%
Precision	99.07%	99.16%	99.25%	99.07%	99.16%	99.25%
Accuracy	66.34%	64.21%	61.32%	66.34%	64.21%	61.32%
MMA	None Used			Random		
δ	1%	5%	10%	1%	5%	10%
Recall	54.55%	52.92%	50.59%	53.22%	48.11%	41.64%
Precision	98.98%	99.07%	99.15%	99.10%	99.12%	99.20%
Accuracy	54.47%	52.92%	50.68%	53.24%	48.24%	41.93%

V. CONCLUSION

Not only because APT attacking takes a high degree of covertness over a long period of time, but also it usually cause lots of human efforts or financial damage. Efficient and effective inside threat monitoring becomes a hot issue during recent decade. In this project, unlike just using of expert rules with only a few pre-defined signatures, the idea of most likely ~~state ch~~anging estimation is leveraged as a behavioral modeling techni~~q~~ue. The logs of Active Directory domain service from every intra-net accounts was collected. And an anomaly detection framework based on famous Markov model algorithm was proposed to analyze AD log and to build the personal model for each account. Further a novel Markov-model state annotation (MMA) profile was also be incorporated during the model training and testing phases. Experiments on a dataset from a real environment of 95 employees shows that the proposed Markov-model based approach combined with TrendMicro prior knowledge will give the best performance of about 66.6% recall and 99.0% precision rates compared to model without using domain knowledge. The major advantages of of using TrendMicro MMA than using all-used MMA is that TrendMicro MMA consists of only a few Markov-model-state annotations such that it can significantly reduced the number of possible Markov states being concerned, compared to all-used MMA. However, even combining the prior domain knowledge, AD^2 only can produce about 66% recall rate or accuracy. That may gives us another conjecture that anomaly detection based on analyzing AD log may be limited by information which AD log can tell. This observation inspires our team that combining AD log with other various logs or contexts may be helpful to detect anomaly. A brief guideline of how to set up the parameters included in this framework is also provided according to the experimental result. The future works

include: 1) keep improving the recall rate without sacrificing accompanied precision; 2) make a clustering analysis on the intra-net accounts to see whether different people behave like a group or not.

REFERENCES

[1] "Trend micro white paper on advanced persistent threat(apt)," Trend Micro Inc., Tech. Rep., 2013.

[2] H.-K. Pao, C.-H. Mao, H.-M. Lee, C.-D. Chen, and C. Faloutsos, "An intrinsic graphical signature based on alert correlation analysis for intrusion detection," in *Technologies and Applications of Artificial Intelligence (TAAI), 2010 International Conference on*. IEEE, 2010, pp. 102–109.

[3] "Directory system agent," Microsoft, MSDN Library, Tech. Rep., 2014, [Online; accessed: 6-May-2014]. [Online]. Available: https://msdn.microsoft.com/en-us/library/ms675902(v=vs.85).aspx

[4] M. E. Russinovich and D. A. Solomon, *Microsoft Windows Internals: Microsoft Windows Server (TM) 2003, Windows XP, and Windows 2000 (Pro-Developer)*. Microsoft Press, 2004.

[5] "Active directory collection: Active directory on a windows server 2003 network," Microsoft, TechNet Library, Tech. Rep., 2015, [Online; accessed: 6-May-2015]. [Online]. Available: https://technet.microsoft.com/en-us/library/cc780036(WS.10).aspx

[6] "How the kerberos version 5 authentication protocol works," Microsoft, TechNet Library, Tech. Rep., 2015, [Online; accessed: 6-May-2015]. [Online]. Available: https://technet.microsoft.com/enus/library/cc772815(v=ws.10).aspx

[7] J. R. Norris, *Markov chains*. Cambridge university press, 1998, no. 2.

[8] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[9] M.-Y. Chen, A. Kundu, and J. Zhou, "Off-line handwritten word recognition using a hidden markov model type stochastic network," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 5, pp. 481–496, 1994.

[10] A. D. Wilson and A. F. Bobick, "Parametric hidden markov models for gesture recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, no. 9, pp. 884–900, 1999.

[11] "Markov model and hidden markov model," 2015, [Online; accessed: 1-May-2015]. [Online]. Available: **http://www.csie.ntnu.edu.tw/ u91029/HiddenMarkovModel.html**

2010Sixth International Conference on Wireless and Mobile Communications

Paper 3:

Implementation of an Advanced Authentication Method Within Microsoft Active Directory Network Services

Jaroslav Kadlec, David Jaros, Radek Kuchta
Dept. of Microelectronics, FEEC
Brno University of Technology
Brno, Czech Republic
kadlecja | jarosd | kuchtar @feec.vutbr.cz

***Abstract*— This paper describes a new approach for developing and implementing an advanced authentication method within active directory network services. For advance authentication process a new type of user multi-factor authentication based on the classical three-factor authentication extended by the position information and time is described in this paper. The main objectives of our applied research are extended security features for more robust and more secure user's authentication process. Application scenario of advanced multi-factor authentication method within corporate networks based on the Microsoft Active Directory network services is presented. Five different factors for user's authentication provide more secured access control layer for current corporate networks with Microsoft Active Directory with only small implementation costs.**

Keywords- multi-factor authentication; position; credential provide.

978-0-7695-4182-2/10 $26.00 © 2010IEEE
DOI 10.1109/ICWMC.2010.48

INTRODUCTION

User identity is the most valuable information in this digital age and person's digital identity has to be trusted all times. Authentication is the process by which end users identify themselves to a network and customized access capabilities are given based on the role they serve in the organization. Policy Manager uses an Active Directory domain server [1, 2], which includes an authentication authority to dynamically assign a policy (or role) to a user or a device, based on the end user's login or MAC (Media Access Control) address. User can be verified by several factors [3]. Conventional authentication systems are based on the one factor authentication typically by the shared secret, e.g. password or PIN (Personal Identification Number). The newest authentication systems add next factor. Smart cards or tokens are quite trustworthy ways for user authentication but can be stolen and abused [4]. On the other hand biometrics can unique identify person with minimal risk of identity replacement. Problem can be with storing biometrics information in digital representation and securing this very sensitive user data to prevent it from possible misuse and also with higher implementation costs [5].

Classic multifactor authentication combines following factors:

- Something you know – password or PIN
- Something you have – token or smart card (two-factor authentication)
- Something you are – biometrics, such as a fingerprint (three-factor authentication)

Five-factor authentication adds another two indicators, which can help to identify users and secure user's sensitive data. The next two factors are:

- Where you are - position information (four-factor authentication)
- When you are – time information (five factor authentication)

Time can limit locking of the user's account. If the user has fixed working time than accessibility of user's account in another time except regular working time is unwanted. The same situation is with position. System administrator can restrict login area only to several locations for example user's office or company buildings. Combination of these two additional factors provides one more advantage. If a user logoffs in his office and next request to login is from another city or country five minutes later it is probably attempt to attack user's account. Position also can serve for authorization to restrict accessibility of confidential data only to fix location, e.g. user's office. Knowledge of user's working times and locations, gives us another possibilities how to secure his account and prevent possible attacks to his private credentials [6].

In the paper, a basic application scenario is described in second section. Next section describes implementation to the Microsoft Windows Vista Credential Security Service Provider and, at the end of the paper, future work and conclusions are described.

APPLICATION SCENARIO

Our application of five factor authentication is divided into the three basic levels. The first level is the most robust implementation on a thick client. The thick client has connected MAD (Multifactor Authentication Device), which provides biometrical user identification and information about current position. A user can logon through the thick client to the network by all five factors and connection to AD (Active Directory) controller is the most trusted. AD controller verifies user's credentials and according to user's current location and time on the AD controller sets user's policies and access rights.

Figure 1. Application scenario of multifactor implementation

WINDOWS LOGON IMPLEMENTATION

Authentication protocols are implemented in Windows by security service providers. Windows Vista introduces a new authentication package called the Credential Security Service Provider, or CredSSP, that provides a single sign-on (SSO) user experience when starting a new Terminal Services session. CredSSP enables applications to delegate users' credentials from the client computer (by using the client-side security service provider) to the target server (through the server-side security service provider) based on client policies [7].

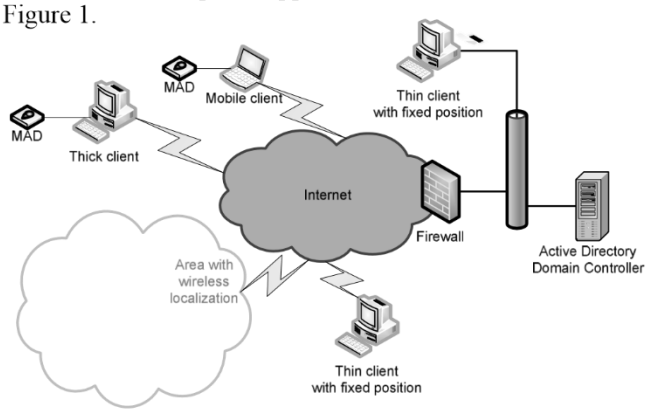
Credential providers [7] are in-process COM (Component Object Model) objects that are used to collect credentials in Windows Vista and run in local system context. In summary, the logon UI (User Interface) provides interactive UI rendering, Winlogon provides interactive logon infrastructure, and credential providers help gather and process credentials.

After all providers have enumerated their tiles, the logon UI displays them to a user. The user interacts with a tile to supply his or her credentials. The logon UI submits these credentials for authentication. Combined with supporting hardware, credential providers can extend the Microsoft Windows operating system to enable users to logon through biometric (fingerprint, retinal, or voice recognition), password, PIN, smart card certificate, or any custom authentication package a third-party developer wants to create.

Credential providers are not enforcement mechanisms.

The second level is not so secure level without MAD connected to a thin client. The thin client has in his AD profile fixed location, which is used for all logon requirements from it. A user is verified only by login and password. In this case only three factors are used (shared knowledge, location and time). Therefore implementation of the thin client can be done only to trustworthy computers.

Third level is implementation in area with wireless localization. Complete application scenario is shown in Figure 1.



They are used to gather and serialize credentials. The LSA and authentication packages enforce security [10, 11, 12, 13].

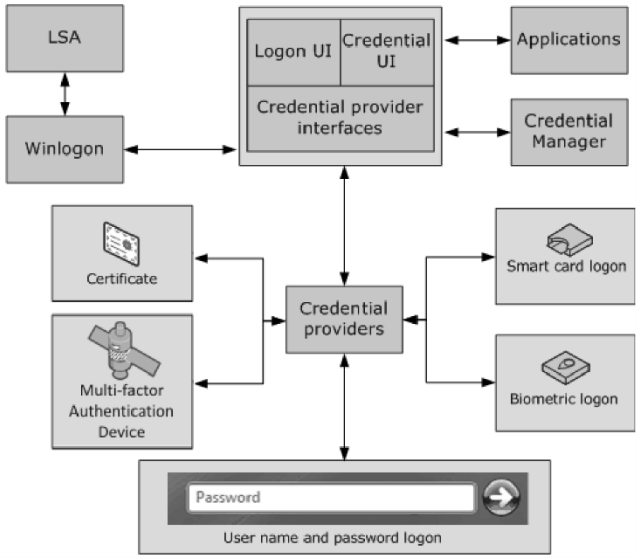


Figure 2. Windows Vista hybrid credential provider architecture with integrated MAD CredSPP [8]

- Packaging credentials for interactive and network logon.

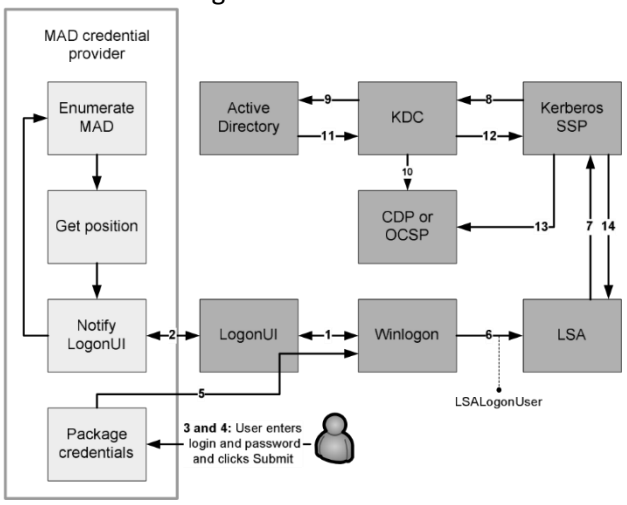


Figure 3. Modified logon flow [8] for MAD authentication process

The hybrid credential provider architecture is shown in

Figure 2. The hybrid credential provider API (Application

Credential providers are registered on a Windows Vista computer and are responsible for:

- Describing the credential information required for authentication.
- Handling communication and logic with external authentication authorities.

Programming Interface) does not design UI but describes which controls need to be rendered to windows logon screen. The hybrid credential provider interfaces with the Windows Smart Card API or Biometric API both directly and indirectly. The direct interface is via public

routines, which allow the detection of connected biometrics or smartcard devices or even detection of inserted card. The indirect interface is via the custom APIs specific for each connected devices, which allow the credential provider to read a user credential directly from the device. The MAD credential provider uses own MAD API for low-level communication. Obtained user's credential from MAD through MAD API are combined with password from logon UI and sent to credential provider interface.

AUTHENTICATION PROCESS

Authentication process with connected MAD is a combination of standard Windows logon process with custom scripts executed by the Active Directory (AD) user's policies [9]. Flow sequence of logon process within Multifactor Authentication Device (see Figure 3):

1. WinLogon requests the logon UI credential information. Asynchronously, our multifactor authentication resource manager starts. The multifactor authentication credential provider:
   a. Gets a list of multifactor authentication devices (uses our MAD API).
   b. Get position information from connected multifactor authentication devices, the MAD credential provider copies it into a temporary secure cache on the terminal.
   c. Notifies the logon UI that new credentials exist.
2. The logon UI requests the new credentials from the MAD credential provider. As a response, the MAD credential provider provides to the logon UI actual position information. The user selects a multifactor authentication device logon title, and Windows displays a logon dialog box.
3. The user enters his login and password and clicks Go.
4. The credential provider that resides in the LogonUI process (system) collects login, password and position. As part of packaging credentials in the MAD credential provider, the data is packaged in a KERB_INTERACTIVE_LOGON structure. The main contents of the KERB_INTERACTIVE_LOGON structure are User Name, Domain Name and Password.
5. The credential provider now wraps the data (such as encrypted PIN, container name, reader name, and position information) and sent them back to LogonUI.
6. Data from Logon UI are now presented by Winlogon for LSALogonUser.
7. LSA calls Kerberos Authentication Package (Kerberos SSP) to create a Kerberos Authentication Service Request (KRB_AS_REQ) containing a pre-authenticator [10].
8. The Kerberos SSP sends an authentication request [10] to the Key Distribution Center (KDC) service that runs on a domain controller, to request a Ticket Granting Ticket (TGT).
9. The KDC finds the user's account object in the active directory and uses the user's credentials to verify the user identity.
10. The KDC validates the user's key to ensure that the credential information come from a trusted source.
11. The KDC service retrieves user account information from Active Directory. The KDC constructs a TGT based on the user account information that it retrieves from Active Directory. The TGT includes the user's security identifier (SID), the SIDs for universal and global domain groups to, which the user belongs, and (in a multi-domain environment) the SIDs for any universal groups of, which the user is a member. The TGT's authorization data fields include the list of SIDs.
12. The domain controller returns the TGT to the client as part of the KRB_AS_REP response.
13. The response is as per RFC 4556.
14. The client validates the reply from the KDC (time, path and revocation status).
15. Now that a TGT has been obtained, the client obtains a Service Ticket to the local computer in order to log on to the computer.
16. On success, LSA stores the tickets and returns success to the LSALogonUser. On this success message, user profile, last logon time and position information are obtained.
17. Custom login script for multifactor authentication device is called from AD login policies. The MAD custom script serves as an intelligent decision algorithm, which compares current position with last logon position and last logon time with current time on AD authentication server from Kerberos authentication packet. Using authentication server time prevents changing time cheating.

Based on these comparisons user access is allowed or denied.

a. In case of successful authorization logon process continues normally according to user's policies. Last login time and position in AD is actualized to current values.
b. If user access is denied WinLogon returns to original state and waits for another user logon attempts.

Preconditions for successful login into AD are customized user's properties in AD extended by login position and time information. These values are validated against position and time of MAD used for user authorization.

Logon UI for the thin client with implemented multifactor authentication is shown in Figure 4. The thin client does not obtain user's credentials from MAD, but allows only weakest authorization by three factors. User is challenged for his username and password. These credentials are expanded by the fixed position information of the thin client and AD authorization authority runs modified authorization process, which was described before. Difference of thin and thick client Logon UI implementation

4554

is the thick client offers only password input box for entering password. All other necessary information is read from connected MAD (position, username obtained by the biometric validation).

Figure 4. Microsoft Windows Vista logon screens with integrated support of Multifactor Authentication Device (MAD connected-obtained position information, dialog used for user login)

FUTURE WORK AND CONCLUSIONS

In the paper, a new idea of multi-factor authentication process extended by position information and time were described. Our main research effort was focused on the application scenarios of user authentication process extended by two new factors and implementation this new approach to the currently used corporate networks. For implementation of this new multi-factor authentication method we chose Microsoft Active Directory as one of the most used corporate network technologies.

Currently we are working on design of new authentication devices that will provide additional user's authentication data. We are also preparing new authentication modules for Microsoft servers that allow to process and set authentication policies for new designed multifactor authentication techniques.

The paper was mainly focused to the description of basic use-cases of five-factor authentication process and description of possible way of implementation into the newest network authentication process' structure. Developed solution of authentication with the help of position information described in the paper is mainly focused to the field of corporate networks but it could be also used in many different applications.

ACKNOWLEDGMENT

This research has been supported by the Czech Ministry of Education, Youth and Sports in the frame of MSM

0021630503 *MIKROSYN New Trends in Microelectronic Systems and Nanotechnologies* Research Project, partly supported by the Czech Ministry of Industry and Trade in the project FR-TI1/057 *Automatic stocktaking system*, partly supported in the project GA 102/09/1897 *Car Transport Safety – BAD* and in the 2C08002 Research Project *KAAPS Research of Universal and Complex Authentication and Authorization for Fixed and Mobile Computer Networks* in the frame of the National Program of Research II.

REFERENCES

[1] Shin, J., W., Park, S., T., and Hwang, C., S.: Domain-based Key Management Scheme for Active Network, Proceedings of World Academy of Science, Engineering and Technology, no. 14, pp. 33-36, Aug. 2006, ISSN: 1307-6884.

[2] Koshutanski, H., Lazouski, A., Martinelli, F., and Mori, P.: Enhancing grid security by fine-grained behavioral control and negotiation-based authorization, International Journal of Information Security, vol. 4, no. 8, pp. 291-314, Aug. 2009, ISSN: 1615-5262.

[3] Wang, L., W., He, L., Y., Liao, X., K., and Wang, H., M.: Research on control flags-based weighted authentication trustworthiness model, 11th Pacific Rim International Symposium on Dependable Computing, Proceedings, no. 1, pp. 369-373, Dec. 2005, ISBN: 07695-2492-3.

[4] Falk, R., Goudalo, W., Chen, E., Y., Savola, R., and Popescu, M.: Multi-level Authentication Scheme Utilizing Smart Cards and Biometrics, 3rd International Conference on Emerging Security Information, Systems and Technologies, no. 1, pp. 93-98, Jun. 2009, ISBN: 978-1-4244-4308-6.

[5] Sutcu, Y., Li, Q., and Memon, N.: Protecting biometric templates with sketch: Theory and practice, IEEE Transactions on Information Forensic and Security, vol. 3, no. 2, pp. 503-512, Sep. 2007, ISSN: 1556-6013.

[6] Zhang, Y., C., Liu, W., Lou, W., J., and Fang, Y.-G.: Location-based compromise-tolerant security mechanisms for wireless sensor networks, IEEE Journal on selected areas in communications, vol. 2, no. 24, pp. 247-260, Feb. 2006, ISSN: 0733-8716.

[7] Kiaer, M.: Multifactor authentication in Windows - Part 2: Preparing Devices on XP and Windows 2003. *WindowSecurity.com.* [Online] 12. 2. 2008. [Cited: 17. 6 2009.] http://www.windowsecurity.com/articles/Multifactor-authenticationWindows-Part1.html.

[8] Mysore, S. H.: Windows Vista Smart Card Infrastructure. *Microsoft Download Center.* [Online] 16. 8 2007. [Cited: 17. 6 2009.] http://www.microsoft.com/downloads/details.aspx?familyid=AC2014 38-3317-44D3-9638-07625FE397B9&displaylang=en.

[9] Griffin, D.: Create Custom Login Experiences With Credential Providers For Windows Vista. *MSDN Magazine.* [Online] 7. 1. 2007. [Cited: 5. 6 2009.] http://msdn.microsoft.com/enus/magazine/cc163489.aspx.

[10] Zhu, L. and Tung, B.: Public Key Cryptography for Initial Authentication in Kerberos (PKINIT). *RFC4556.* http://www.ietf.org/rfc/rfc4556.txt: Microsoft, June 2006.

[11] Microsoft. How the Kerberos Version 5 Authentication Protocol Works. *Microsoft TechNet.* [Online] 6. 5. 2008. [Cited: 17. 6 2009.] http://technet.microsoft.com/en-us/library/cc772815.aspx.

[12] Harrison, E. R.: Lightweight Directory Access Protocol (LDAP): Authentication Methods and Security Mechanisms. *RFC:4513.* http://www.rfc-editor.org/rfc/rfc4513.txt: Novell, Inc., 2006.

[13] Melnikov, A. and Zeilenga, K.: Simple Authentication and Security Layer (SASL). *RFC4422.* http://www.ietf.org/rfc/rfc4422.txt: OpenLDAP Foundation, 2006.

4565