

Methods of data transfer protection in the Internet network

Mariusz Sojak, Szymon Głowacki, Paweł Policewicz

Faculty of Production Engineering

Warsaw University of Life Sciences

St. Nowoursynowska 166, 02-787 Warszawa, Poland

mariusz_sojak@sggw.pl

Abstract: The aim of this paper was to show how data transferred by means of a data communications network may get into the wrong hands. The paper shows both the selected methods of attacking transmission of data sent via the Internet and the selected security methods, which may be used to protect information.

Keywords: security methods, authorization, authentication, encryption of information, digital signature, digital certificate.

1 Introduction

A computer network is a set of many interconnected computers. The transmission medium that connects the computers in a network may be cables: STP - shielded twisted pair, FTP – foiled twisted pair, SFTP - shielded foiled twisted pair), UTP – unshielded twisted pair, coaxial 50 Ω , telephone lines, optic fibre, (single-mode fibre and multiple-mode fibre). Transmission may as well be wireless (e.g. Wi-Fi, WiMAX, satellite connections) (Sojak 2009). Internet was formed out of many LAN, MAN and WAN networks and it is “a network of all networks”. During the research on the ARPAnet network (Advanced Research Projects Agency network) a set of TCP/IP protocols (Transmission Control Protocol/Internet Protocol) was created, which are a language that allows the computers to communicate in a network. This set is called a “TCP/IP protocol”. It has been popularized since 1982 and it still is a basic set of protocols used for the information exchange in the Internet.

2 A general principle of how the Internet works

The main task of a computer network is data exchange between computers, it is done by sending small portions of information called packets. Packets, apart from

a piece of information, contain a description in the form of a header, which contains information concerning i.a. sender's address, recipient's address, packet size, checksums, and the location of a given piece of information in the document that is being transmitted. The two latter elements are significant as regards the reassembly of the document when it reaches the recipient. The recipient, in order to check whether there are errors in the data, checks the correctness of the received packet on the basis of the checksums. The receiving computer reassembles the incoming pieces of information into a whole on the basis of the information describing the location of each piece of information in a document. The recipient confirms each correctly received packet. If the confirmation of packet receipt does not reach the sending computer within a given time limit, the sending computer retransmits the lost packet, which guarantees correct transfer of information. Packets may reach the recipient by different routes. Routers are intermediaries in the information transfer. Their tasks include sending data and finding the best route that connects the sender and the recipient. (Wojciechowski 2006).

Routing protocols are used by routers to determine the route of data transfer in a WAN network; it allows to create routing tables dynamically. Routing algorithms determine which path is better. The following parameters are taken into account: bandwidth, delay, hop count, route cost, load, MTU (Maximum Transmission Unit), reliability and communication cost. If a router uses load balancing mechanisms, there may appear a few best routes in the routing table. The router will use them all at the same time, distributing the load evenly between the routes. There are a lot of divisions of routing protocols, e.g. interior (IGP - Interior Gateway Protocol, IGRP/EIGP - Interior Gateway Routing Protocol / Enhanced IGRP, OSPF - Open Shortest Path First, RIP - Routing Information Protocol, IS-IS - Intermediate System to Intermediate System), exterior (EGP - Exterior Gateway Protocol, BGP - Border Gateway Protocol), multicast transmission (DVMRP - Distance Vector Multicast Routing Protocol, IGMP - Internet Group Management Protocol). The division of protocols may be based on the method of operation, taking into account the distance vector (RIP, IGRP), link condition (OSPF, IS-IS/Integrated IS-IS, ES-IS), hybrid – they take into account both the distance vector and link condition (EIGRP), path-vector – they describe the routes by means of attributes (EGP, BGP, IDR - Inter-Domain Routing Protocol) (pl.wikipedia.org). Figure 2.1 shows the division of networks used by the OSPF protocol.

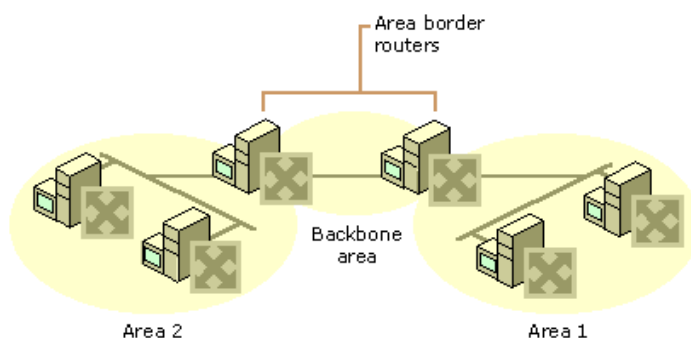


Figure 2.1 Division of internetworks introduced by OSPF (Source: www.microsoft.com)

As we scale (enlarge) a network, the size of its database storing link conditions becomes larger, more memory is required, and it takes longer to calculate the route. In order to solve this problem, OSPF protocol divides the internetwork into areas (collections of neighbouring networks), which are interconnected by means of the backbone area. Each router stores a database of links condition for the areas that are connected directly to this router. Area Border Routers ABR connect the backbone area with other areas (technet.microsoft.com).

3 The working principle of a system based on DNS or DDNS servers

DNS (Domain Name System) is used to change IP addresses from the numeric form into the domain form, and it was introduced in 1984 in order to facilitate using the Internet and simplify computer naming system. DNS system is supervised by two organizations IANA (Internet Assigned Numbers Authority, www.iana.org) and ICANN (Internet Corporation for Assigned Names and Numbers, www.icann.org). They coordinate the rules for assigning IP numbers and top level domains (e.g. .pl, .gov, .com, .eu) between countries or selected organizations, and they transfer the rights concerning the management of these domains to these entities. In Poland, Scientific and Academic Computer Network (NASK - Naukowa i Akademicka Sieć Komputerowa, www.nask.pl) operates the .pl domain. It is possible to distinguish two types of top level domains: TLD - Top Level Domains – they are functional domains, (gTLD – generic TLD) and country domains (ccTLD – country code TLD). Due to DNS servers it is not necessary to memorize addresses in the form of numbers - 212.77.100.101 but it is possible to type the address e.g. www.wp.pl. If a computer where the requested address was typed, obtains the response from the DNS server with its translated numeric IP address, then it establishes the connection. Thanks to that, a client may communicate with this server in two ways, via the domain address, and directly, using the numeric address. In the latter case, so called in-addr.arpa entries in the reverse DNS server tables play a significant role.

DNS servers operate on the basis of fixed addressing of computers using IP numbers – a server is required to have its fixed IP address.

Unfortunately, some Internet users do not have fixed IP addresses. Quite a few Internet Service Providers - ISPs e.g. Neostroda do not guarantee their customers fixed IP addresses, as they assign numbers on the basis of DHCP servers. In this case, customers cannot provide server services e.g. www on the basis of DNS servers.

DDNS servers service comes to rescue to such users.

DDNS (Dynamic Domain Name System) – is a service which has a database with domain-name entries that updates itself dynamically. Due to that a www server can be reached under one fixed name regardless of the IP address it has. Due to the fact, communication with the server takes place using the domain address translated by the DDNS server. Unfortunately, determining the IP address change by the DDNS takes a few minutes. The update information is sent to the DDNS server, and

then the server updates the data in a DNS server. This process takes another few minutes. The diagram of establishing connection using DDNS was presented in figure 3.1.

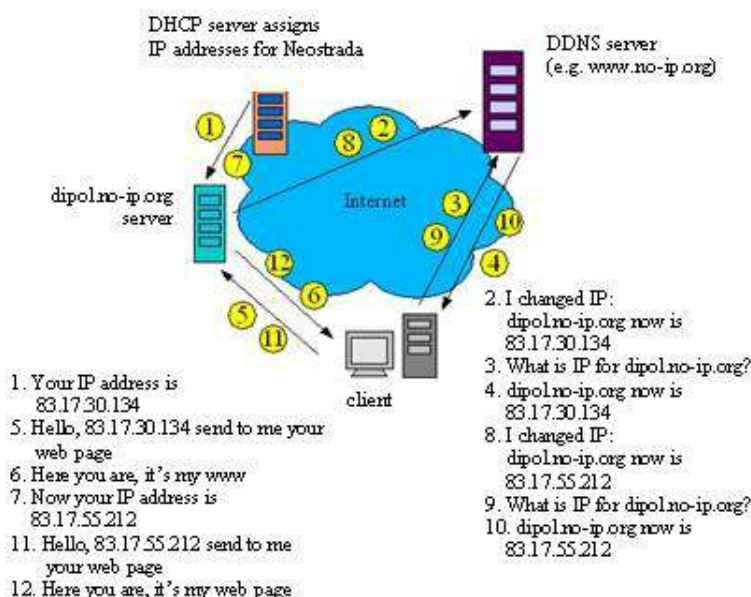


Figure 3.1 Diagram showing the operation of DDNS (Source: www.dipol.com.pl)

Companies providing DDNS services include Nettica.com, DyNS, dyn.pl, gratisdns.de, ChangeIP.com, dDNS.pl, dDNS.pl, home.pl, NASK.

4 Selected methods of attacks on data transferred in a network

Dangers, which threaten computers connected to the Internet network include i.a. sniffing, spoofing, SYN-flooding and port scanning. Each of these techniques is based, to an extent, on the method of establishing connections. The intruder communicates with the victim's system in order to obtain the desired results. There are two ways of communication. Unidirectional for sniffing and bidirectional. In a bidirectional communication, the system that is attacked responds to the forged SYN packet, which initiated the connection. The reply content, depending on the method used, is a source of information and may be employed to use up the resources of a remote server, so called DoS, DDoS, DRDoS attacks. Selected methods of attacks and security methods were discussed further.

Sniffing - is generally a method used by sniffers to intercept and analyze information transferred via the TCP/IP network. This technique involves setting the network card in the promiscuous mode; thanks to this, an unauthorized person may see all the packets in a network, including those that are not destined for him/her. When the sniffer is installed on the router, changing the mode of the network card is not necessary. After packet interception and data analysis, an intruder may determine the name and the version of the operating system and services on the victim's computer. Sniffing is a form of attack that is very difficult to detect as it leaves no traces, which could prove it took place.

Sniffing usually takes place in LAN networks. In a LAN network computers are connected with the router via a hub or a switch. The method of connection determines who will be able to read the messages we send and receive. In case of a network that uses a hub, all computers in a LAN network may use all the packet sent by our computer. All that a sniffer who wants to intercept packets must do is to run a sniffing program (Szmit 2005).

In networks using switches, communication of computers comes down to sending a packet between a client and the server, thus eliminating other hosts. It makes sniffing more difficult. Figure 4.1 presents the difference in sending packets in a network with a hub and a switch. A good method protecting our computers against sniffing is encryption of data transmission based on SSL or SSH protocols. It is also possible to use special programs called anti-sniffers to defend ourselves from sniffers.

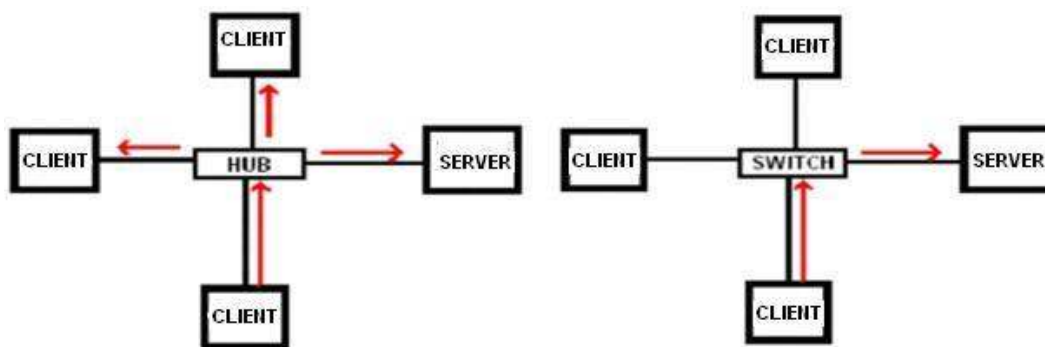


Figure 4.1 Differences in transmitting packets in a network with a hub and a switch
(Source: www.webhosting.pl)

Spoofing – involves forging the source IP address in a packet transmitted via the Internet. It leads to hiding the identity of the sender. This method is used during Distributed Denial of Service attacks (DDoS) as a part of SYN-flooding. Spoofers use this method to pass themselves off as a trusted IP address (e.g. from an intranet) in order to obtain the rights that a given network address possesses. Passing oneself off as someone else is rarely used as it is difficult to predict initial TCP/IP sequence numbers. The only efficient method to prevent such attacks is configuring the network in such a way that it rejects external packets that come from the outside of the network, and whose headers contain the address that indicates a computer in a local area network. This task is performed by a router, which analyzes incoming packets, and, depending on its configuration and the content of the packets, does not allow access into a network (www.haxite.org). Older firewalls do not protect against spoofing. However, the new ones, equipped with mechanisms analyzing packets from the data-link layer to the application layer of the OSI model, make it possible to conduct complex security policy.

SYN-flooding – is a popular network attack. It aims at blocking services provided by a given server. It involves sending a huge number of packets to the server. In each packet the SYN (synchronization) flag is set, which informs the server about the attempt to communicate with it. As a reply to SYN packets, the server sends back packets with set synchronization and acknowledgement flags (SYN, ACK) to the

computer initiating the connection. The next stage after the server sends the packet is waiting for the reply of the computer which initiated the connection. But the computer does not reply. It makes the server record the information about the attempts to establish connection in its states table. Tables are stored in the RAM memory. Therefore, their size is limited. Sending huge numbers of packet with the SYN flag set leads to using up the server's resources. In consequence, it does not respond to other senders. In some cases the attack makes the victim's system hang or crash. The working principle of SYN-flooding was presented in figure 4.2 (www.webhosting.pl).

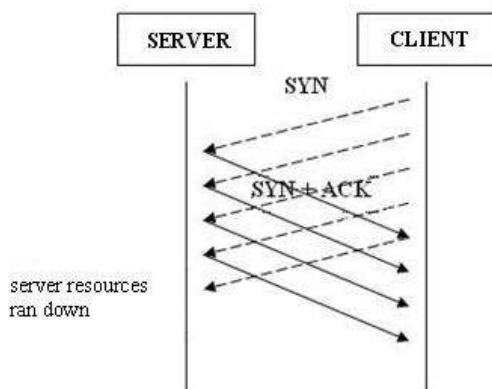


Figure 4.2 SYN-flooding attack

In case of SYN-flooding, a method of efficient defence may be installing a firewall. Windows system users should install a personal firewall e.g. ZoneAlarm, Sygate Personal Firewall, Eset Smart Security (the packet contains a firewall) and XP Windows users should use the built-in firewall. Security methods were described in a separate point (www.webhosting.pl). Another types of attacks involving so called buffer overflow are: *ICMP flooding (PING flooding)* – this method involves flooding ports with undesirable ICMP packets, which are in turn generated by e.g. a ping program. *MAC flooding* – this method uses a limited size of the Content Addressable Memory (CAM) table, which, at the point of the overflow makes the switch act as a hub, and, in this way, it makes sniffing in a given network easier for the intruder.

Port scanning – prior to breaking into a computer a hacker looks for gaps in security. In order to do this he scans ports. Scanning ports involves sending TCP packets to one or more computers in order to obtain information about open ports and available services. A port scanner is used for this purpose. It sends an empty packet with a SYN flag set to a given server, then it waits for the confirmation. As a result of the confirmation it obtains a packet with the set SYN/ACK flags. This packet contains the information which allows for identification of the system and the services that are made available on it. Port scanning is similar to sniffing, it can indicate a remote attack on the system. An nmap program may be used to scan ports.

5 Security methods

When we use a computer at home we do not think about its physical security and guaranteed system workflow. Security factor plays an important role as regards

commercial servers. These systems should allow for limited access of the third parties who can connect directly with them. The following methods, described below, are used to achieve this.

Authorization and authentication – granting access to computer resources in the Internet to some users after confirming their identity. Authorization is a process which confirms that a given user is a person s/he pretends to be and has the rights to use the resources s/he seeks. Authorization process is performed using a mechanism granting the users, recognized by the system, access to specified resources. Users, who want to access a given www webpage, must first give their name and correct password. Authorization method is widely known as it is simple to use and inexpensive to carry out. A certain kind of authorization that uses password are PINs (Personal Identification Numbers). They are usually used in banking systems (e.g. authorization of cash withdrawal from a cash dispenser). The aim of authorization is to verify whether the cardholder is its rightful owner, i.e. whether s/he knows the PIN number. Other methods of authorization are: recognition of the user's fingerprints or the pattern of veins in the palm of the hand, identity recognition on the basis of the retina of the eye and voice recognition. Authorization methods' feature is lack of possibility to grant access rights to resources to another person (Wojciechowski 2006).

Information encryption – encryption is a means of protecting communication and data on the hard disk in the form of files. In order to both decrypt and encrypt data, so called keys for the encryption algorithm are necessary. This key is usually a number in a form of a phrase or a sequence of characters. Secret information should be encrypted during the transmission in the computer network. Files saved to disk should be treated in the same way. Only persons who possess the decryption key will be able to read the secret data. The recipient of the information may read the message only if s/he knows this key. Modern encryption algorithms have two keys called private and public key. Private key is held back by the user. It is created by the owner and used to encrypt the information s/he sends. Public key is used for decryption. It should be known to all people who will read encrypted information sent by us. It is also assumed that, apart from the public key, the recipient also knows the algorithm used for encryption. Thus, both the sender and the recipient use cryptographic tools handling the same algorithm. A characteristic feature of modern cryptographic algorithms is asymmetry of private and public keys. It is only possible to decipher information encrypted with the private key when we know the public key. A person who does not have a private key cannot encrypt or digitally sign any message. Encryption also takes place when a web browser communicates with a WWW server via HTTPS protocol (HyperText Transfer Protocol Secure). An encrypted connection guarantees safe and secure communication between the server and our computer. The browser informs us about it by displaying an icon of a locked padlock at the bottom of the window. It makes it impossible to read our data by unauthorized persons (Wojciechowski 2006).

SSL links and HTTPS connected with them - there are a few methods used to protect information transmitted using HTTP connections. The most popular include: SSL (Secure Socket Layer) designed by the Netscape company and TLS (Transport Layer Security). SSL works with application layer protocols such as Telnet, HTTP, FTP, as well as the TCP/IP protocol. SSL and TLS are based on data transmission encryption and authentication of both the client and the server. They operate as an additional layer between the TCP/IP transport layer and the application layer. All transmissions between the client and the server are encrypted and decrypted by the SSL or TLS layer. SSL and TLS encryption is based on SSL digital certificates.

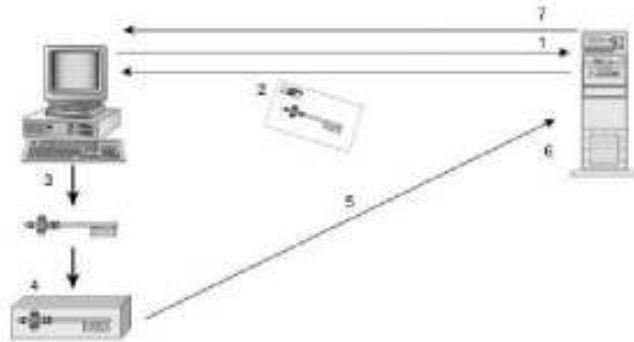


Figure 5.1 The process of SSL acknowledgements exchange. (Source: Komar 2002)

Schema description:

1. The client requests the server to issue safeguards, it also sends i.a. a random number.
2. A server sends back the certificate to the client, together with its public key, cipher option and its own random number. At this stage, after obtaining the server's certificate an error may appear in the client's system for two reasons. The first is the discrepancy between the name of the certificate with the full, qualified name of the WWW server. It is a result of moving the webpage to another server that has a changed name. The second reason for the lack of communication may be the fact that the www server's certificate is not present on the client's list of trusted certificate issuers. The certificate is accepted if there is a superior certifying institution, which the client and server may consider trusted. One of such Certifying Offices is Thawte Digital Certificate Services, which owns the SSL algorithm and a VeriSign organization. After connecting with a secure server it is possible to see its certificate confirming the server's identity.
3. Generating the session's initial key by the client.
4. The session's initial key is encrypted with by the client using the public key of the server obtained in point 2. The client's session's initial key protected in this way may only be deciphered using the private key that is stored on the server.
5. The client sends the encrypted session's initial key to the server.
6. The client's session's initial key is deciphered by the server using its own private key. ON the basis of the two random numbers generated by the server and the client (sent in points 1 and 2) and on the basis of the session's initial key

(defined by the client) both parties generate the session key used for the actual data exchange.

7. Authentication of the server on the client's side takes place by sending an encrypted message using the session key. In this way, both parties of the connection are authenticated.

Many encryption mechanisms use the system of private and public key. A packet encrypted using a private key may only be decrypted by using a matching public key and the other way round; a packet encrypted using a public key may only be decrypted using a matching private key. Encryption and decryption of all further data is performed using the main key. Using a different session key for each transmission makes it impossible for the hacker to use data obtained during one transmission to attack another transmission. A 128 bit cipher is used for the encryption, which results in so high number of operations needed to decrypt the data that it puts off potential intruders. In the USA the decryption mechanism uses 521 bits, which guarantees a higher security level.

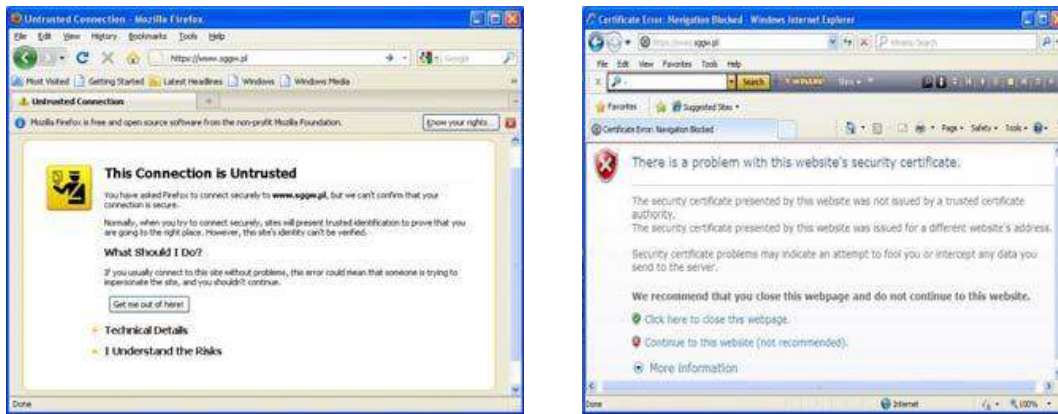


Figure 5.2 Error caused by the absence of server's certificate on the client's list of trusted certificate issuers in the Mozilla Firefox and Internet Explorer browsers

All webpages equipped with the SSL mechanism have an URL address that begins with HTTPS://. Webpages that do not begin with such address are not encrypted. Therefore, when they are used, valuable information such as credit cards numbers should not be provided. Web browsers also have additional indicators of secure SSL connection. In such browsers as Internet Explorer there is a padlock at the bottom of the window, and in Netscape Navigator there is a key at the bottom of the window. Since then data exchanged between the Internet user and the server are encrypted and protected. (Komar 2002, Bowen 2009).

Digital signature – digital signature is also called electronic signature. It allows to attach to the signed document information, created on the basis of the private key that belongs to the person signing the document. The validity of the signature may be established by the owner of the signatory's public key. The signature has three important functions: it guarantees integrity control – makes it possible to find any unapproved changes in the message content and its attachments; it guarantees unde-

niability – the author of the message cannot deny having sent and signed it; it proves the authenticity of the sender – only the owner of the unique private key may send a message (authorization of the sender, authentication) (Wojciechowski 2006).

PGP (Pretty Good Privacy) is a software encryption system. It is used to protect e-mail messages and attachments. Mail programs i.e. Microsoft Outlook Express, Eudora and Microsoft Outlook may use modules extending PGP. These modules are installed in a mail program and they can be found in the menu options or they are buttons in the window of this program. PGP allows to sign letters digitally, which guarantees authenticity of e-mails that are sent. Each document and user have a unique digital signature. In a situation when one user wants to send two different documents, they will be assigned two different digital signatures because their summaries (on the basis of which checksums will be created) will be different (Scrimger et. al. 2002, Karbowski 2007).

Digital certificate – is a document that contains information together with the public key. Certificates are issued by the so called Certifying Institutions, which act as a third party that is trustworthy. The certificate is saved to a specified data carrier e.g. the microprocessor card. Thanks to the digital certificate it is possible to identify a person using a given digital signature. A computer program, generating a pair of asymmetric keys, uses the information saved by the user itself exclusively, and cannot verify it. Keys control involves certification of the authenticity of the third party's signature by a trusted person. The user, after generating his or her pair of asymmetric keys, sends the public key to the person issuing the certificate. After verification of the user's information's authenticity, the entity issues the proper digital certificate. The certificate is inseparably connected with the public key. Verification of the certificate involves verifying the digital signature using its public key. The key should also be certified in order to prevent a hacker pretend to be a certifying centre. The certificate contains the following information: specification of the user's, association's or server's name, for whom the certificate was issued (depending on the kind of certificate), the certificate holder's public key, the name of the issuing body that issued the certificate, the issuer's identifier, the certificate's serial number, the expiry date of the certificate, the digital signature of the certification centre (krystian.jedrzejczak.webpark.pl). Among certifying institutions we can find Thawte Digital Certificate Services, VeriSign, Centrum Technologii Internetowych (Internet Technologies Centre), Cybertrust, GlobalSign, Entrust, Aladdin.

Firewall – connecting LAN with the Internet is connected with the dangers from the outside. Firewall is a system that protects computer networks against break-in. Firewalls protect against access of unauthorized persons to the local area network. Therefore, they are installed on the border between networks. Some firewalls prevent all external packets from entering a network. However, they grant access to packets used by e-mail programs, and allow for safe communication between the user and the external network. Another advantage of firewalls is introducing and looking through all the incoming packets. Points all the data cross on their way to the local area network and outside it are posts called bastion hosts. They are the

front line of defence. Firewalls protect against sniffing and passing off as someone else, and they protect against viruses and Trojan horses.

Firewalls can be divided into software and hardware firewalls. Software firewalls are special programs. They help us control the whole incoming and outgoing packets traffic. They include the following programs: ZoneAlarm, Outpost Firewall, Sygate Personal Firewall, Ashampoo FireWall, Comodo Personal, GeSWall, Jetico Personal, Kerio Personal Firewall. Hardware firewalls are special external devices. They filter and analyze packets using special integrated circuits and internal software. We distinguish three types of firewalls: a firewall working at the network level, a firewall working at the application level and a firewall working at the transmission level.

Main advantages of firewalls include: system protection, allowing the whole network to use one common IP address (Proxy server), allowing systems that use other protocols than TCP/IP to connect to the Internet, monitoring WAN connections and network traffic, Proxy Cache Server optimizes load on the WAN link during intensive work with www. Disadvantages of firewalls: they need to be updated frequently, and make remote network control more difficult, low efficiency of proxy servers decreases network efficiency (Pieńkowski 2003).

6 Conclusions

1. Functioning of the Internet network depends on many mechanisms such as internet protocols, DNS servers and DDNS routers.
2. Internet user is exposed to numerous dangers such as: sniffing, spoofing, SYN-flooding or port scanning. However, there are methods, which allow to counteract and prevent these dangers such as authorization and authentication, encryption of information, digital signature, digital certificate and firewall.
3. Despite using different kinds of security measures (the ones mentioned above as well as anti-spam software and anti-spyware), a computer will never be 100% safe. Therefore, common sense should be used when sending different kinds of information using computer networks.

References

1. Bowen R., Coar K. 2009. *Apache. Receptury*. Wyd. 2. Helion. ISBN: 978-83-246-1549-0.
2. Karbowski M. 2007. *Podstawy kryptografii*. Wyd. 2. Helion. ISBN: 978-83-246-1215-4.
3. Komar B. 2002. *TCP/IP dla każdego*. Helion. ISBN: 83-7197-782-4.
4. Praca zbiorowa. 2003. *Administrator PC Przewodnik Telepracy*, skład: Jerzy Pieńkowski, Fundacja Pomocy Matematykom i Informatykom Niepełnosprawnym Ruchowo, Wydanie I, Warszawa.
5. Scrimger R., LaSalle P., Leitzke C., Parihar M., Gupta M. 2002. *TCP/IP Biblia*, Tłumaczenie: Adam Jarczyk. Wydawnictwo Helion. ISBN: 83-7197-668-2.

6. Sojak M., Głowacki Sz. 2009. *Wi-Fi networks in small and medium sized businesses (SMB)*. *Studia Informatica, Systems and information technology*. 1(12)2009, s. 55-66, ISBN: 1731-2264.
7. Szmit M., Gusta M., Tomaszewski M. 2005. *101 zabezpieczeń przed atakami w sieci komputerowej*. Helion. ISBN: 83-7361-517-2.
8. Wojciechowski A. 2006. *Usługi w sieciach informatycznych*. Wydawnictwo Naukowe PWN S.A., Warszawa. ISBN-10: 83-01-14751-2, ISBN-13:978-83-01-14751-8.
9. pl.wikipedia.org, www.microsoft.com, technet.microsoft.com, www.dipol.com, www.webhosting.pl, www.nettica.com, www.dyns.cx, dyn.pl, www.gratisdns.de, www.changeip.com, www.ddns.pl, www.haxite.org, home.pl, www.nask.pl, krystian.jedrzejczak.webpark.pl, www.thawte.com, www.verisign.com, www.securitynet.pl, www.cybertrust.com, www.globalsign.com, www.globalsign.com, www.entrust.com, www.aladdin.com