An Outlook to Security and Trust in Internet Communications

Robert Logozar
Polytechnic of Varazdin
J. Krizanica 33, HR-42000 Varazdin, Croatia
robert.logozar@velv.hr

Petra Koruga
Faculty of Organization and Informatics, University of Zagreb, Pavlinska 2, HR-42000 Varazdin, Croatia
petra.koruga@foi.hr

Abstract. *We address the key questions on security and trust in Internet communications by using a synthetic approach that connects technological and human aspects of the subject. As first, a brief outline of the Internet security technology is given. This serves as a ground for the exposition of the general concepts and principles, like the security pillars and the security threats. The achieved solutions provide the basis for building of the user trust. The trust is then proportional to the user's perception of the achieved security level. The omnipresence of Internet in all human activities today, including financial transactions, e-commerce, trade, auctioning, and other, proves by itself that the trust of majority of online users is won. The mechanisms of user protection provided by online service companies are improving, as well as the overall legal support. The general efforts to educate the Internet community and to provide the relevant statistics are on a rise.*

Key words: *Internet security infrastructure, end-to-end security, technology-related and human-related aspects, security perception, trust, security statistics.*

1 Introduction

The modern era is characterized by a widespread use of many different communication systems, among which *computer networks* seem to be the most versatile and the most complex. The computer networks have grown globally and locally, penetrating the inner organizational structures and occupying the world-large scales. They serve extremely large number of users, connecting them to local and universal communities. They are known of their diversification and of huge quantities of information transmitted over them. All computer networks are unified today under the notion of *Internet* — a global "network of networks" that is based on its ubiquitous TCP/IP communication protocols.

Data transfer and various applications rely more and more on the networks, which become an unavoidable part of the computing infrastructure. In this context, Internet may be viewed as a vast number of abstract *information channels* with many different *security threats*. Along with the benefits of the intensive networking, the need for thorough security solutions emerges as more and more crucial. Via Internet we can acquire a new antivirus program or a patch for our operation system, but in the same time we can expose our computer to severe security threats.

The technical complexity of Internet often contributes to mystification, misunderstanding, exaggeration or underestimation of the security issues. As we approach the third decade of widespread use of Internet in all spheres of life, we must say that the trust of majority of users is improving. It is based on both, the widespread trends and the personal experience. Still, when we ask simple questions, such as:

- Can Internet be securely used for human communication?
- Can privacy and identity of Internet users be protected?

broad users may be perplexed. On one hand the answers should be affirmative — judging by the mere fact on enormous number of delicate online transactions happening as we speak. On the other hand, the reports on Internet frauds can cause disbelief and mistrust, and raise the questions about the reliability and credibility of implemented security.

The aim of this paper is to describe the existing infrastructure of the Internet security, and to help the readers answer the above questions. We start with a low-level description of the security technology infrastructure aimed at wider audience. Upon that we build an interdisciplinary approach which shows that in a complex system aimed for human communication and interaction, all participants are important. Every single computer counts! It may be a firm brick in the local and global security wall, a hole in it, or a source of unintentional or intentional threats and danger. With respect to that, the technological solutions provide a condition sine qua non, but they cannot be considered as a full security foundation. The human influence is unavoidable and it calls for a *synthetic* and *interdisciplinary* approach, also known as *multilateral* or *multidimensional security of communications*. The field outgrows the mere technological domain. It also deals with psychological, ethical, economical, legal and political issues [1, 2]. Through such a synthesis, the security aspects common to all communication systems will emerge, making evident that Internet is not at all that specific. Most of the security problems are universal. They have just emerged as more blatant on Internet due to its vast communication potentials. And like any other media, Internet has its advantages and its limitations.

Such a comprehensive thinking will lead us to a more general treatment of the security issues from which other communication channels can largely benefit, too. Finally, our intention is to provide a better insight into the technological security aspects to the readers with social and humanistic background, and to outline the importance of human and social aspects to the technically educated audience.

2 A crash review of Internet security infrastructure

In this section we briefly interpret some technological features of Internet and relate them to the communication security aspects. Although it may seem as futile to even try to present this large subject in a single article, this attempt goes along with our thesis that only a "knowledgeable" Internet user can fully appreciate the achieved security solutions, and can be fully aware of the remaining risks. Only a "well-educated" user will be able to answer the questions from the beginning of this article.

2.1 The Internet and OSI model

From the technical standpoint, Internet is described by the *TCP/IP model* (also called *Internet model*), or *Internet Protocol Suite*. The first name is after the Internet two most important protocols. The model can be divided into four *abstraction layers*, which are outlined from the top down in Table 1 [3,4].

Table 1. TCP/IP model, or Internet Protocol Suite.

L a y e r (protocols)	
4	Application Layer (FTP, HTTP, SMTP, SSH, SSL, TLS, …)
3	Transport or Host-to-Host Layer (TCP, UDP,…)
2	Internet or Inter(Network) Layer (IP, IPv6, IPsec, …)
1	Link, or Host-to-Network (ARP, PPP, DSL, ISDN, FDDI, …)

Communication protocol, or shortly *protocol*, is a procedure that precisely describes how the communication is to be done. Typical protocols corresponding to each of the layers above are listed within the parenthesis. The Transport and Internet layers, and the corresponding TCP (Transmission Control Protocol) and IP (Internet Protocol), present the core of the Internet as we know it. There are a few more protocols in these two layers, as there are several more protocols in the top-most, application layer, and in the lowest, link layer.

The concept of network layers was not encouraged in the original Internet specifications. It was introduced later by the OSI (Open Systems Interconnection) model which was initiated by ISO (International Organization for Standardization). The OSI model remained mostly within theoretical realms, but its good solutions largely influenced the way of analyzing and developing of the computer networks. Although strict comparisons are not fully justified, most authors try to provide some mapping between the two models. The seven layers of the OSI model and their "rough" relation to the abstract layers of the TCP/IP model are shown in Table 2.

It is mainly due to the OSI model that the link layer of the Internet model is now divided into two sublayers: the *Data Link* (or *Network Interface*) *Layer* and *Physical* (or *Hardware*) *Layer*, bringing the number of Internet layers to 5.

Starting from the bottom, the Internet *Link* (*Host-to-Network*) *Layer* corresponds to OSI layers number 1 and 2 – the *Physical Layer*, and *Data Link Layer*. The *Internet Layer* corresponds to OSI layer 3 – the *Network Layer*. The *Transport* (*Host-to-Host*) *Layer* is mapped to the OSI layer 4 with the same name but with a different strict definition, and also partly to the OSI layer 5. The Internet *Application layer* roughly corresponds to the OSI layers 5 – 7: *Session*, *Presentation* and *Application* layers. In the OSI presentation layer the encryption was predicted, allowing the syntax of the application layer to be independent from the selected security solutions, and also from the functions of other lower layers (confer 2.4).

Now we can follow the "layer stack" of the TCP/IP model and describe its components. The *stack* structure suggests that the function of every higher layer is facilitated by the services of the layers below it.

Table 2. The abstract layers of the TCP/IP and OSI models and their rough relation.

TCP/IP model		*OSI model*
5. Application Layer		7. Application Layer
		6. Presentation Layer
		5. Session Layer
4. Transport Layer		4. Transport Layer
3. Internet Layer		3. Network Layer
Link Layer	2. Data Link L.	2. Data Link Layer
	1. Physical L.	1. Physical Layer

The physical layer specifies electrical properties of the network devices and their interfaces to the transmission media (copper lines, optical fibers, radio frequency electromagnetic waves), through which it sends the bits of data. It defines the connectors' pinouts, voltages, clock-rates and other technical details of network hubs, repeaters, network interface cards, routers, and other devices.

The data link layer functionality relies on the services of the lower, physical layer, and provides the transmission of digital data organized in *frames* between the hosts on the same network (LAN, WAN, confer 2.6). The frames travel from one end of the transmission media to the other end. The data link layer provides a service interface to the network layer above it by checking and correcting the transmission errors. It also regulates the data flow on the basis of

physical addressing, taking into account the capacities and speeds of the sending and receiving devices.

The Internet (or **network**) **layer** provides the transfer of *data packets* from a source host to a destination host, within the same network, or on different networks (also known as *internetworking*). Each packet has a standard header with the necessary data, among which are the source and destination IP addresses. The packets are transferred independently from each other, possibly through different nodes and via different paths. This is realized by *packet switching* — the underlying technical concept which enables better overall usage of the available channel *bandwidth* or *throughput*. The net result is *routing* of the packets on a path to the final destination, from one network node to the next, till the final node is reached. Here, the network and Internet topology must be known so that the packets can be transported via routes that avoid congested communication lines and routers. The Internet layer is the lowest layer that provides the end-to-end connectivity. Its functionality is today mostly provided by the Internet Protocol.

The transport layer uses services of the network layer to ensure the end-to-end transfer of the messages from a process on a source computer to a process on a destination computer. This layer assures flow control, congestion control, and application addressing through the software constructs known as port numbers. The layer provides a necessary abstraction level for the work of application software in the layer above, by assuring its independence from the lower layers. The main protocols of the layer are TCP and UDP. The TCP provides the so called *connection-oriented* data transmission, and UDP provides the *connectionless* transmission of *datagrams*.

The application layer is used by applications for specific communication tasks. The layer presents the higher-level protocols: FTP, SMTP, HTTP, etc. Also of our interest are security-providing protocols, like SSH, SSL, TLS, which will be specifically mentioned in 2.4. Generally, the application data is formatted and coded according to these protocols, and is then *encapsulated* into the protocols of the lower transport layer. They in turn use the services of the protocols which are lower in the layer stack.

2.2 The Lack of security in the basic Internet layers

Internet misses a true and convincing security concept in its fundamental Internet and Transport layers, and the corresponding IP and TCP protocols. As was already stated above, IP deals with data packets — the self-contained, independent chunks of information that bear the IP addresses — and the underlying mechanism of packet switching. This basic concept provides much of the Internet functionality, like the optimal use of resources, great flexibility and low cost, but it also introduces additional security risks. As opposed to the *circuit switching* found in telephone connections, the traveling path of information on Internet is more arbitrary and not at all certain. Since the information is in digital, "electronic", form, it is furthermore prone to low-cost and easy-to-be-done subversions and attacks. Namely, with today's digital technology, the electronic digital data are not only the most easily stored, transferred, received, and protected from noise — comparing to all other forms of data presentation and physical realization, like those written on paper, or analog signals modulated in radio waves — but are also the most easily copied, altered, multiplied or forged. Because of that, the proper protection of data and implementation of security mechanisms is of utmost importance (confer also section 3).

The examples of the low-cost threats are:

- *Password sniffing* — searching for non-encrypted passwords by programs installed on servers, possibly on those with intensive traffic;
- *IP spoofing* — finding the IP address information within the packet IDs and using them maliciously;
- *Password stealing* — e.g. by Trojan horses thrown into the system.

All these attacks can be performed in every node of the network that is traversed by the data packets of a message.

Internet is known to be open both horizontally, for free spreading of the network, and also vertically, meaning that new protocols and layers can be added (2.4). But the vertical openness could require changes in the infrastructure, which is hard and expensive to implement. Also it could present a source of incompatibility and restrictions for its horizontal openness.

The basic TCP/IP architecture of Internet is non-cryptic in its nature. This immediately allows for the possible loss of secrecy and loss of integrity, because of the potential attacks performed in any of the Internet layers. The usual, unsecured Internet services, such as electronic mail and file transfer, are unprotected from such attacks. Yet another common problem, unsolvable by the original Internet infrastructure, is the lack of *authentication*.

The authentication is the act of verifying the genuineness of an entity, i.e. the security process of establishing that the entity is what it claims to be, and that it can act as a known subject (person, organization, process, computer). Only after the authentication, the *authorization* should be done.

The authorization is the process of verification that a known subject is allowed to perform certain actions and access certain resources (confer also 3.1).

It is obvious that without authentication and authorization any higher forms of business communication cannot be realized. Both of these basic security actions can be realized by the adequate use of cryptographic mechanisms.

In the mid 1990s there was an attempt to remedy the Internet security defects by introduction of the end-to-end security protocols IPsec and IPv6. They ensure security mechanisms in the Internet layer by authenticating and encrypting each IP packet. The idea was to alleviate the burden of the security implementation from the application software. However, the need for implementing and maintaining the

dedicated software for this protocol on every remote computer, resulted that the security solutions in the application layer prevailed (2.4).

2.3 Cryptographic solutions

To make further discussion clearer, we shall briefly outline the fundamental cryptographic concepts (for more details see e.g. [5,6]). There are two basic cryptographic systems in use: *symmetric cryptosystem* with secret keys, and *asymmetric cryptosystem* with private and public keys.

The symmetric cryptosystem was used in DES (Data Encryption Standard), a former American standard from 1977. In the late 1990s DES was replaced by Triple DES. In early 2000s AES (Advanced Encryption Standard) superseded DES and Triple DES with its longer 128-bit code blocks, and longer keys (128, 192, and 256 bits). A disadvantage of the system is that a safe channel must be used for the distribution of the *secret keys*. Though the need for an extra safe channel can be regarded as a technical burden, by establishing such a connection between a known and certified sender and recipient, the proper authentication of the communicators can be simultaneously solved.

The asymmetric cryptosystem eliminates the need for another safe channel by introducing a pair of keys consisting of: a *private key*, which must be kept secret by the sendee, and a *public key*, which is to be disseminated to possible senders. The sendee (recipient) who wants to receive an encrypted message distributes her or his public key to the other side(s). The other side—the sender—uses it for encryption of the message to be sent back to the sendee. There's no fear that the message will be understood by any third party. The method ensures that decryption cannot be done without having the sendee's private key. Also, the private key cannot be retrieved from the public key because of the properties of the mathematical functions involved in the cryptographic process. Thus, only the sendee who owns the private key will be able to decrypt the message. The RSA system, named after its inventors Rivest, Shamir and Adleman, is such an asymmetric system. Because there is no need for additional safe channel, this is an ideal solution for Internet. A disadvantage is that it requires much higher computing resources than the symmetric cryptography.

Another problem of the asymmetric cryptography that arises from the nonexistence of a safe side channel (in which, by definition, a proper authentication of the communicators would be accomplished), is that an intruder can take someone else's, or generally false, identity, and abuse it. So, in this case the need for a proper authentication emerges as crucial.

The problem is solved by implementation of the *Public Key Infrastructure* (PKI). Within the PKI an institution called *Certificate Authority* (CA) is authorized to issue the PKC (*Public Key Certificate*), or *Identity Certificate*, which binds a public key to the identity of a communicator (the name of a person or organization, their address, etc.). It is done by the mechanism of digital signature (see below). The PKCs are distributed by institutions like Trusted Third Party (TTP), Public-Key Manager (PKM), Public-Key Distribution Center (PKDC), and similar. An alternative to this can be a scheme like *Web of Trust*, where users sign the PKC by themselves or by endorsement of other peer communicators. If both sides trust the PKCs, a cryptographic simulation of the safe channel is provided. This is then appropriate for the dissemination of the symmetric keys.

As was just suggested, the secure protocols and mechanisms use both cryptographic systems in order to get the optimal results. Since the asymmetric cryptography is about two orders of magnitude (or even more) slower than the symmetric one, it is used only for the crucial parts of communication — for the authentication and for the encryption of the secret symmetric keys. After the symmetric keys are exchanged, the rest of the communication is protected by much faster symmetric encryption.

In the *digital envelope* data itself are encrypted symmetrically, while the asymmetric cryptography (simulating the safe channel) is used for transmission of the symmetric key only. Thus much greater speed of secure communication is achieved. Digital envelope ensures data secrecy, but not data integrity. Namely, although information remains secret to an intruder, it can be illicitly damaged or altered.

Digital signature solves the problem of the message integrity by calculating the *hash function* or *message digest* out of it, and then applying the asymmetric encryption to the digest. Both, the encrypted digest and the original message are sent. If the message is changed, the recipient will know that by comparing the original digest (after decrypting it), and the newly calculated digest from the received message. Only if the two digests match, the message is genuine. The mechanisms of digital signature and digital envelope can be combined together to provide the joint secrecy and integrity. If public keys were distributed properly, as pointed out before, the digital signature ensures the authenticity, secrecy and integrity. Usually by the name of digital signature all these security mechanisms are assumed.

2.4 End-to-End security — cryptography in the application layer

The simplest way of introducing the security to Internet, while leaving the lower layers of the TCP/IP model untouched, is to implement it in the highest, application layer, by means of cryptography. This is known as ***End-to-End* (*EtE*) *Security in the Application Layer***. Although the attacks in the lower layers are not prevented, they are made futile with respect to many security aspects.

This can be interpreted as the introduction of a new, *Security Layer*, as is shown in Table 3 (after [4], confer also OSI presentation layer in 2.1). The new layer is justified by the appearance of new security protocols, like the highly successful SSL. The SSL was a secure protocol within the Netscape suite of network applications in mid 1990s.

Table 3. The introduction of Security Layer [4].

	Layer
6	Application (HTTP)
5	**Security (SSL, TLS)**
4	Transport (TCP)
3	Network (IP)
2	Data link (PPP)
1	Physical (DSL, ADSL, cable TV)

The concept of the EtE security and a separate security layer is implemented through the several protocols:

- SSL (Secure Socket Layer) [7], already mentioned above, and nowadays replaced with the newer:
- TLS (Transport Layer Security) [8];
- HTTPS (Hyper Text Transfer Protocol Secure) that is the HTTP over SSL or TLS;
- Secure Shell [9];
- PEM (Privacy Enhanced Mail), now replaced with:
- S/MIME (Secure/Multipurpose Internet Mail Extensions);
- PGP (Pretty Good Privacy) data encryption and decryption software relying on the Web of Trust;
- GnuPG (GNU Privacy Guard) free cryptographic software;
- EBICS (Electronic Banking Internet Communication Standard), a transmission protocol aimed for the users of the Internet banking services.

All the above protocols are examples of a successful protection of user data from the first two threats listed in 3.1 below: the loss of secrecy and the loss of integrity. E.g. S/MIME ensures the secrecy of e-mail communication by the use of secret keys and symmetric cryptosystem, under the assumption that the local Internet servers are secure for the key handling. PGP uses the asymmetric RSA encryption with public keys for the critical data. The public keys are kept on the client's computer and are disseminated by the concept of web of trust (2.3).

2.5 The achieved security level

The way of measuring the achieved technical security level is by finding the computation work needed by an adversary to breach the applied security and undermine the system. It can be expressed as *Intrusion Work* W_I:

$$W_I = P_{Cmp} \times t \,. \tag{1}$$

Here P_{Cmp} is the computing power or speed of the intruder's computer expressed in some suitable measure, and t is the time spent on breaking the security by *brute force*, i.e. by systematically trying all the possible keys. It is implied that the intruder is using the most efficient algorithms known.

In the early 1990s, when the computers had the processor power of the order of about 10 – 100 MIPS (1 MIPS = 1Mega IPS, IPS = Instruction Per Second), the W_I used to be roughly 10 – 100 MIPS×Years. Because of the huge growth of the computing power in the last two decades, the intrusion work of 10^2 MIPS×Years is not impressive for quite some time, even for the commonly accessible computers. In the meantime, the ways of measuring the computing speed and the corresponding benchmark tests have also changed, in order to more accurately reflect the performance of computer systems as a whole. So instead of the previous IPS concept, now the more actual SPECint and SPECfp measures are used. Roughly, we can say that the computing power grew by the factor of $10^3 – 10^4$. Thus, in order to keep the intrusion time at the same value as before, the required intrusion work should be enlarged by the same factor. This is to be achieved by raising the security level of the cryptographic protection, which is directly related to the key length.

The intrusion work W_I needed to "break" a key is highly dependent on the key length. For symmetric cryptography the rise is close to exponential, and the chances to break it by brute force are extremely low. There are no reports of successful cracks by now.

The asymmetric cryptography requires longer keys for the same level of security than the symmetric one. The factor is roughly ten, with the tendency of getting bigger as the key lengths increase (e.g. confer [6]). In the combined cryptography where the RSA system is used for the asymmetric encryption of the symmetric keys, the RSA is considered as the weakest link, the most vulnerable to attacks. There are only a few credible reports on cracking down the RSA systems with shorter keys, by the use of abundant computing resources. One such testifies on breaking the PGP — as one of the most popular hybrid systems using the RSA — by enormous computer power, organized specifically for the testing purposes [10, 11].

Anyhow, longer and longer keys and improved algorithms are in use to protect against brute force and other methods of attacks. Back in 1996 the symmetric keys of 75 bits were advised, with suggestions to enlarge them to 90 bits to compensate for the rise of computer power. After the DES was changed with Triple DES, and nowadays with AES, the key lengths are at least 128-bit for standard applications, and 256-bit for the critical ones.

As for the asymmetric encrypting, 1024bit and longer keys are not rare any more. Only a decade ago such cryptography was treated as a high-tech product strictly forbidden for export from the USA. The high-level security needed in banks requires asymmetric keys of 2048 bits and even longer [12]. The safer solutions will require a slightly higher investment for the security user, while resulting in a much higher intrusion work and bigger time expense for the security breacher.

As a conclusion, by taking the cryptography security level to follow the demands of an application, the intrusion work can be designed high enough to make the attacks not worth the effort [13]. If the cryptographic EtE security solution requires several months or even years of computing to be broken by brute force, then by ensuring a simple policy of changing the keys on a regular basis, together with the other

mechanisms of recognizing and stopping such attacks, we can make them fruitless.

We shall summarize this subsection with the following important note: *In the context of building the communication trust, the use of cryptography must be done completely consistent and without exceptions.* Furthermore, the cryptography should be standardized and regulated more thoroughly, which is generally not the case. Poor cryptography was often put in large software packages, perhaps under the pressure of restrictive export regulations [14]. As a final result, non-secure products could appear on the market, justifyingly adding to the loss of user trust.

2.6 Intranets, firewalls and local security

As opposed to the global uncertainty of Internet, LANs and WANs (Local Area Networks, Wide Area Networks) and *intranets*, as a sort of inner, localized Internet, present the proprietary networks in which security polices can be established and enforced rigorously. Here the security on the **technical level** can be made highly predictable. The general defects of Internet can be, if not completely mended, at least kept under control. The intranet is interesting because it can use the standard Internet infrastructure (protocols) and applications, while enabling the full supervision of all the servers and clients within the localized network. Besides that, intranets can use other specialized protocols (like X.25) and networking solutions that can highly improve security.

For intranet and other private networks, the basic security principle of connecting them to the "wilderness of Internet", is by doing that **only** via strictly controlled protecting systems called *firewalls*. The firewall is a hardware or software component, or combination of both, used to control the communication between different segments of network and specifically between the intranets and Internet. It is done on the basis of the established set of rules and policies. Mostly, the firewalls are set to control the traffic from some insecure and hard-to-control parts of the network, like Internet, to the local secure networks or home computers. They should protect the "inner side" from the unauthorized accesses and threats from the "outer side", while allowing the desired and approved data transfer. Also, they should restrict the transfer of the secret data from inside to the outside world.

In order to be able to update and fine-tune the rules and policies, the firewall must be complemented with the intrusion detection system (IDS).

The firewall can be organized as one or more of the following:

i. *Packet filter*, which filters out the packets with respect to their departing and arriving IP addresses, and requested TCP ports (application services). The filtering is done according to the list that specifies the addresses and services which are forbidden, those which are allowed, and the rules of actions for the rest of the packets.
ii. *Application Layer Firewall*, which acts through the application software by controlling the IP packets coming to particular applications, like Web browsers, FTP clients, and others.
iii. *Firewall on Proxy Servers* acts similarly to the application layer firewall, but since they are organized as servers, either on separate computers or as software, they offer their clients additional level of security.
iv. *Firewall with Network Address Translation* (NAT) mechanism protects the computers behind itself by hiding their true IP addresses. This is usually combined with the standard role of the NAT (the enlargement of the number of IP addresses within local networks).

The firewall mechanisms violate the standard protocol layering and highly influence the network traffic. However, this seems to be a necessary sacrifice to remedy the Internet security deficiency (confer 2.2). A well-configured firewall proved to be a good protection from the outside intrusions. Of course, the practice shows that the term "well-configured" is often not given its full dimension — at least until the first hostile attacks happen, or until the users complain on the denial of service.

Two or more localized networks can be connected by means of a safe channel. As was shown in 2.4, such a channel can be established via unsafe Internet by the use of safe protocols. In a general situation of a *distributed information system* that requires a complete and integral security, the systems like Kerberos should be implemented [15]. Besides the authentication, the appropriate authorization of all the participants is performed on a strict schedule basis (compare 2.2). The firewalls which are backed-up with such secure authentication and authorization mechanisms allow much greater flexibility and connectivity of proprietary networks to Internet, while still maintaining a high security level.

An example of a specialized network is SWIFTNet. This is a system for exchanging financial messages between banks on the worldwide scale which is operated\by SWIFT (Society for Worldwide Interbank Financial Telecommunication) [16]. In the early 2000s the company moved from a network based on X.25 protocol to a new infrastructure based on a suite of their proprietary SWIFTNet protocols. The client banks had to use Bilateral Key Exchange (BKE) system. By 2008 the yet newer SWIFT Phase 2 protocol was introduced, and the BKE was replaced by more secure and easier to update Relationship Management Application (RMA) system.

The whole philosophy of SWIFT is to provide a system for financial messaging of maximal possible security. Via SWIFTNet, only messages in the form of *payment orders* are transferred. The true fund transactions, as described and initiated by the payment orders, are to be done separately between the banks themselves, over their mutual accounts on a certain timely basis. The messaging system is centralized and operates in the form of *store-and-forward* principle. The payment order from a sending bank is sent securely by using cryptography into the SWIFT intranet area. A copy of the order is stored in the

SWIFT premises and an action of an authorizing side is required. After the authentication and authorization is done, the order is sent to the receiving bank. The security is guaranteed by redundancy on all levels: from the redundant networking and computing facilities, exhaustive checking procedures, rigorously written and checked software, up to the most important—the highly trained and abundant personnel. This is an excellent example of the use of proprietary protocols, LANs, intranets, safe side channels, key exchanging systems, as well as the human aspects (see section 3), for the implementation of a secure network.

2.7 Summary of the technology related security mechanisms

As a conclusion of this section we give a short list of the technology based security mechanisms [17]:

- Firewalls for the end-connection to the network protection and intrusion detection systems;
- Proxy servers for the access management [4];
- Content managers for control of the data brought into and sent out of the information system;
- Virus protection tools for incoming and outgoing emails and files;
- Service monitors for checking of the service usage, and early detection of the hostile procedures;
- Fail-over systems, for prevention from the loss of availability;
- Encryption implemented in the online applications (EtE Security), and applied to sensitive files;
- Authentication systems: passwords and IDs, physical tokens, cryptographic certificates;
- Digital signatures for verifying the sources of Internet contents.

3 Security and trust

After studying the basic technological security aspects of Internet, we should be closer to the rank of the "knowledgeable" and "well-educated" user from the introduction. Now we can turn to the general security aspects which are independent, not only of the communication channel type, but also of the human activity taking place over the channel. Then we can connect those aspects to the security mechanisms provided on Internet, and think of their counterparts in other communication systems.

It is good to clarify that in our discussion we concentrate on the security aspects that are caused by *human behavior* and human actions. We are not concerned by the properties of the networking and computing devices that are governed by the physical laws, nor do we regard the security aspects that are influenced by the forces of nature. A fire in some crucial network node can cause the loss of service availability and possibly some serious problems to the Internet users, but as long as the source of this calamity is not a human who was specifically attacking the networking or computing infrastructure, it is out of the scope of our topic. The only thing we could do here is to humbly admit that our firewalls won't help us in this matter, and that we should delegate this kind of problem to a true fire brigade. In fact, if the cause of above disaster was a human act against the networking facilities, we could add it to our statistics of bizarre attacks. So, what we are dealing with here are human-initiated actions against computing and networking software and hardware, which can be connected to the behavior described as: *unwanted*, *non-ethical*, *immoral*, *illegal*, *criminal*, and similar.

As we have stressed earlier, these human-related security issues of Internet are essentially not much different from those in other communication channels. They are just more complex and more potentially harmful, primarily due to the following reasons.

- The use of digitalized data that can easily be:
  - *modified*, *altered*, *copied*, *replicated*, *distributed*, etc., as a result of the corresponding malicious activities:
  - *data alteration*, *counterfeiting*, *plagiarism*, *spamming* (confer also the discussion in 2.2).
- The use of a global, diversified, network:
  - which has a multi-layered structure that multiplies the points of intrusion, and thus makes the analysis and control of the weakest links much harder (see 3.2);
  - which (still) lacks the global security standardization and implementation (confer 2.5);
  - which (still) lacks the adequate legal and ethical support from other communication channels and social institutions on the global level (see 3.4).

3.1 The pillars of security

To concretize our discussion, we start by outlining the well known pillars of security. These are [1]:

1. *Authenticity* — the ability to prove the identity of communicators (see the full definition in 2.2);
2. *Secrecy* — the ability to keep the information secret from all unwanted parties;
3. *Integrity* — the ability to preserve the information to be identical to the original, i.e. to keep the information whole and nothing but the whole;
4. *Privacy* — the guarantee, or a set of rules and policies, that the gathered information will be used confidentially, only by the agreed persons and only for the agreed purposes;
5. *Non-repudiation* — the legal obligatoriness for the performed transactions, and ability to provide the undeniable, legally accepted proves of the topics 1 to 4 above.

The above requirements are endangered by the following ***security threats***:

1. *Loss of secrecy* or *loss of confidentiality* — unauthorized acquisition of information;
2. *Loss of integrity* — unauthorized modification of information;
3. *Loss of availability* — unauthorized decreasing of functionality;
4. *Loss of responsibility* or *loss of accountability* — unauthorized loss of control and supervision, a sit-

uation when everything becomes available, with no limitations and no restrictions, and when no one is responsible for the condition of the system.

The threats are to be answered by appropriate implementation of the technological solutions outlined before, in combination with the legal and social support. All these together should establish the five pillars of security from above, and assure the following protection mechanisms against the security threats:

1. **Secrecy protection**. The message contents must stay secret to everybody but to the trusted partner(s) to whom the message was intended (*content secrecy*). Also, the transmitting and the receiving side must be able to stay anonymous (*participant secrecy*). *For solutions see*: 2.3, 2.4.
2. **Integrity protection**. Every manipulation of the contents of the message with an intention to alter and modify it in any way, must be discovered and treated accordingly, in order to reverse the message to its original state, or at least to indicate that it was being corrupted. *For solutions see*: 2.3, 2.4.
3. **Availability protection**. The communication must be available to all the users who demand it, under the condition that their access rights are granted and approved (in other words, they must have proper communication rights according to the system security policy). *For solutions see*: 2.6.
4. **Responsibility protection**. This can be further described in the following three points:
   - 4.1 The receiver of a message must have a possibility to prove some third party (e.g. legal authorities) that the defined entity did send her or him the message. *E.g implemented by SWIFT* (2.6).
   - 4.2 The transmitter of a message must be able to prove the transmission of the message and the authenticity of its contents, and, if necessary, to further prove that the receiver has received the message. *Examples: SWIFT, EBICS, PayPal, eBay* [18, 19].
   - 4.3 The Users (customers) cannot deny their obligation to pay for the services once the provider has sufficient evidences for administering the services in an agreed way. *Examples: SWIFT, EBICS, PayPal, eBay.*

3.2 Security principles

There are two very general and fundamental security principles that are already mentioned in the text:

- **The weakest link principle**. A system is as strong as its most insecure part. A well designed and maintained security system must be nearly equally strong in all of its components, since attacker needs a single (weak) point to break into the system.
- **Every computer and every user of a system is a possible intentional or unintentional attacker**. A single non-secure point seriously weakens the entire system. This can be from both, the inside of an intranet or LAN, like a personal computer of a negligent user, or from the outside world, performed by an attacker, or a person who intentionally or unintentionally helps the attacker [17].

The second principle could be derived from the first, but is nevertheless stated explicitly to emphasize the human aspects of the threats. The problem is to assure validity of the principles in every single component of a complex communication system like an online application.

The originators of the threats can be recognized among the following groups of people [1]:

- Outsiders – who are not allowed to use the system;
- Insiders – the authorized users of the system;
- Operators and managers of the system;
- Maintenance and servicing personnel;
- Manufacturers and vendors.

As an example, we shall quote the findings that although about 3/4 of attacks come from the outside of the firewall, the most damaging and hardest-to-recover-from are the attacks from the inside [4, 20]. The attacker may be an outsider with criminal motivations who managed to get help from an insider. It can also be an insider, or a former insider, like an employee holding a grudge against his or her former company.

3.3 Trust

While the security can and must be related to the technological and other infrastructure (legal, social), *trust* is a notion of highly individual and psychological nature. We shall define it according to [21]: *trust is a certainty of some preferred outcome in the future*. This is close to the common notion of the word[*]. There are also other, different interpretations of the term, which may be suitable to other contexts[†].

Since the main purpose of security measures is to establish the "preferred certainty", the above definition leads us to the concept of ***security-based trust***. It relies on:

- Continuity of regular, desirable, behavior of the surrounding;
- Help of the confidential people and institutions;
- Individual knowledge and ability to control the situation.

These three components of trust are overlapping. The continuity of regular behavior depends largely on the functionality of the surroundings. In the technical environment such as Internet, the regular behavior is maintained by technical and organizational procedures. The latter two points are of typical human character. They are highly dependent on the user's general education, as well as on her or his knowledge of the information technologies and particularly of the topic that we are discussing right here.

[*] According to Random House Unabridged Dictionary, trust is: "1. reliance on the integrity, strength, ability, surety, etc., of a person or thing; confidence. 2. confident expectation of something; hope. …

[†] E.g. a peculiar definition is given in [22]: the trust is believing into a positive outcome of a transaction only in the case of lacking certainty. When certainty is big there is no need for trust, since one can count on assuredness. The bigger is (potential) risk, the bigger must be trust (put in something). This emphasizes the benevolent dimension of the term, and could be called *optimism-based trust.*

Contrary to trust, *distrust* is caused by:

- Discontinuity of acceptable behavior;
- Continuity of unacceptable behavior;
- Helplessness.

3.4 Security perception and trust

Trust as we define it should rely on the achieved security level. Since it is a highly individual notion formed by rational and irrational human factors, it is the ***perceived security*** that must be considered. If we denote the achieved security level by S, than the perceived security S_P should be some function of S:

$$S_P = P(S) . \quad (2)$$

The situation is illustrated in Fig. 1 with three different, arbitrarily chosen, perception functions. P_1 is an ideally realistic perception, for which $S_P = P_1(S) = S$, i.e. P_1 is the identity function. This corresponds to the perception of our knowledgeable and well-informed user. The perception P_2 is consistently pessimistic, and thus still linear, while P_3 is optimistic perception with nonlinear response.

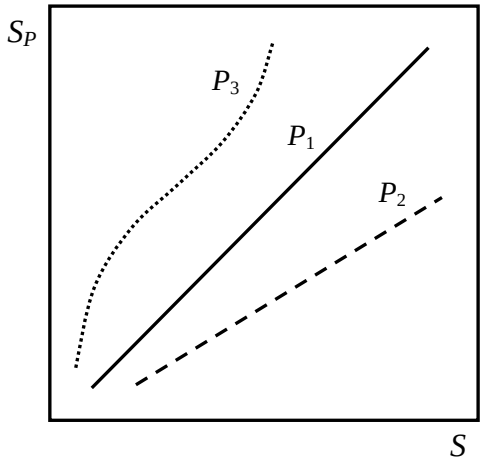

Figure 1. Security level S and its perception S_P. The perception P_1 is realistic, P_2 consistently pessimistic, and P_3 optimistic and nonlinear.

It is clear that many other variables and parameters, besides the security itself, can influence the perceived security. The net contribution of all of them still results in some function similar to those presented here. The goal should be to exclude all irrelevant factors and to achieve a realistic security perception.

After the above deliberation, trust T should be some rising function of the perceived security S_P:

$$T = f(S_P) . \quad (3)$$

One such possible function is illustrated in Fig. 2 to serve our short discourse. Though quite abstract and without any quantitative ambitions, the graph offers a visualization of the security-trust relationship. The simplest relation between the two would be some linear proportion, like:

$$T \sim S_P \ , \ T \approx S_P . \quad (4)$$

If the function drawn on Fig. 2 is assumed, for the same security perception value S_P the point A corresponds to a security-unfounded trust (reflecting the peculiar interpretation of the term in the foot note above), and C is a point of unnecessary caution. The point B presents some "founded trust", of course, presuming that the function $T = f(S)$ is correct.

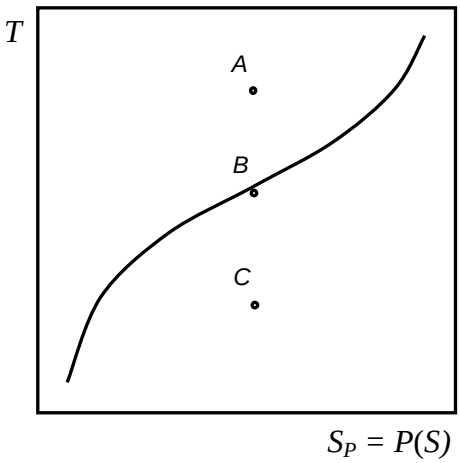

Figure 2. The relation between trust *(T)* and security perception *(S_P)*. The latter should be a realistic estimate (function) of the true security level.

The quantity of trust can be conceptualized through its connection to a certain use or application and the corresponding risks. For such risks we plan the adequate security. As a general rule, the security costs should be some (considerable) fraction of the risk estimates. Let's say that 1/10 is a good starting point. Specialized IT security companies will suggest more precise investment figures for desired applications (confer also 2.5). The costs of all possible damage should be accounted for, including the possible loss of revenues because of the lost client trust.

Regarding the trust in general, we must say that the world outside the communication channel must also be safe. In section 2 we have shown that the technological security solutions do exist. But here we emphasize that the scientists, engineers and technicians cannot do the complete job even if, by a miracle, they would be able to create, technically-wise, a perfectly secure channel. The action and help of all the participants, other professions and the community as a whole is needed. In other words the scientific community is obliged to say that: *the **technological security infrastructure,** given solely by the technical solutions—is not enough*!

The technological security infrastructure should be complemented by the **human security aspects**:

- Adequate legal support and judicial system, effective crime punishment and prevention, a righteous socio-political system;
- Community consensus on the security and technology standards, security-related statistics, and user education.

Nothing of the above can be ignored, as it often happens in practice. Give the best available, and the most user-friendly technology to uneducated people, and the problems will arise. Give powerful technology to underdeveloped, or even worse, ill-developed

society, and you can be sure that all kinds of the system abuses will occur because of various possible reasons. This is in accordance to the often quoted fact about the Internet security chain: *humans and human-related aspects are the weakest point* [1].

Regarding the high level of technological development in all spheres of human society today, the above statement is true also in other communication channels, and also in all other human activities. This dichotomy or, better to say, the dialectics, between the technological and human aspects of communication channels, deserves a more detailed exposure in a separate paper.

4 A glimpse to the present state

Ten, fifteen, years ago, the skeptics would insist that Internet security infrastructure still requires improvements in the consistent implementation of technological solutions, and much, much better support in the human-related spheres. The use of Internet for delicate communication and expensive transactions was considered too risky and was not recommended. In the meantime, the trends and practices showed them to be wrong — if not in predicting many of the risks and possible problems — than in the tempo at which almost all human activities, from all realms of life, transferred to the ubiquitous use of Internet.

The public trends were positive and enterprising. Today even the critical security applications are not exemption from this. ***To not use all the advantages of the Internet seems like a waste of a great opportunity!*** Such a proactive public attitude did boost the ICT security sector. The practice called for immediate implementation of the theoretical solutions, for improvements of the global technology standards, and for the cooperation of the local authorities and institutions in providing better and safer business environment.

4.1 Migration to Internet

The fast development of various kinds of online business communications is for sure witnessed by many of us during the last decade and a half. We got used to the comfort and efficiency of a myriad of Internet services, like: E-banking, E-trading; direct payments, Internet auctions, B2B communications, although being aware of the potential risks. These online transactions have a broad financial range, and the corresponding broad range of security risks. The transactions go from a few dozen of EUR or US$, up to the hundreds, thousands and much higher quotes, and the threats range from the low-level threats of the well-known fraud scenarios, up to the extreme level threats of hacking experts. But we, as knowledgeable users, expect that the security solutions are tailored and maintained according to the application requirements, and that the side assurance mechanisms, starting from the legislation and good practices of the online service companies, will protect us from losses. Every now and then we should also check the validity of all our security assumptions and expectations.

Today it is more than obvious that all of the above listed Internet business activities are here to stay. They will not decrease in volume. As statistics shows, it is just the opposite, and the reasons are obvious:

i. The omnipresence of Internet today is dictated by the advantages that it offers, resulting in the new needs and habits of the consumers.
ii. For most of the users the losses from frauds are within tolerable limits. Switching to other ways of communication and transaction would cost even more in terms of time and money spent, and again would not guarantee a fully risk-free operation.
iii. The user experiences, practices and reports show that the frauds are not fundamentally Internet-generated, nor solely Internet related. It is true that some of the Internet aspects and features are prone to easy-to-be-done immoral and illegal acts (confer 2.2 and 3.1), but these happen almost proportionally in all other communication channels.

Stated shortly, if ten years ago the question was for which communication and business activities to use the Internet and for which not, nowadays the only question left is how to achieve a sufficient security level for just about every kind of online activity that we can, and will do, on Internet.

4.2 The need for relevant statistics

Aside from the fact that online business communication is rolling and cannot be stopped, a serious approach requires in detail statistical analysis as a ground for further discussions and more precise conclusions. Such statistics is still missing. Most of the companies, especially those with large transaction volumes, consider the security data as highly confidential. They fear that the consumer trust could be ruined if the users find out that the security was too low (confer 3.4). So, even if attacked, the big companies would try to solve the problems by themselves as far as they could. Such behavior originates from the early days of Internet business, when online transactions still had to prove its reliability. As the trust of majority of online customers is already won and their habits generally established, one would expect that more relevant data about the frauds and losses are to be available for the broader public.

Some of the companies involved in providing security solutions realize the importance of raising the public awareness by informing them objectively. RSA division and CyberSource are good examples [23, 24]. The latter is one of the rare companies that provide truly relevant statistics of the lost revenues due to online frauds. Based on this, we have made a preliminary estimation of the overall security risk at 0.1% of the total transaction volume, which is close to the risk ratio in offline activities. A more detailed insight and support to this conclusion deserves a separate topic.

According to our investigation, besides the mentioned CyberSource report, not many others, if any at all, are open to public. On the other hand many governmental and nongovernmental organizations, like Fraud Watch and Internet Crime Complaint Center

(IC3) [25, 26], greatly contribute to the public awareness by educational activities and by publishing the fraud statistics based on the Internet users' reports.

5 Conclusion

The Internet has grown from an (idealistically) open, free, and insecure place in its beginnings, to a (realistically) less open and free, but potentially much more secure communication channel. We have outlined its existing security technology infrastructure based on the EtE cryptography concept in the application layer, and discussed a whole palette of security solutions.

These solutions must enable the realization of the security pillars: authenticity, secrecy, integrity, privacy, and non-repudiation, and prevent, or at least make futile, the corresponding security threats. A well designed security infrastructure must include both, the adequate technological solutions, and all important human aspects, like the instruments of financial and legal protection. Upon such, multilateral security foundation, the user trust is built. The trust is proportional to the perceived security level. For the latter to be realistic, a proper education of the Internet users is needed.

Of the above requirements, security technology is available for quite some time. In practice, however, the problems of the consistent implementation, constant maintenance and technology improvement remain. The human security aspects are also improved, both, in the legislations on the national level and through the security policies of international e-commerce corporations.

To complete this analysis of the Internet security, the relevant statistics should be involved. It must give a better insight of the risks of the particular online activities, and the overall risk estimates. A preliminary investigation shows that these risks are already similar to the risks in other communication channels.

Internet and its security aspects can serve us to get a better insight into the problems of a general communication channel. Its human versus technological aspects is a topic that deserves further investigation.

6 Acknowledgments

The first author wishes to thank B. Aurer and Z. Hutinski whose ideas and support initiated this paper. We would also like to thank to V. Sac for his valuable comments on the technical aspects of computer networks.

References

1. Müller, G., Rannenberg, K. editors; Multilateral Security in Communications, Vol 3, Addison-Wesley, München, 1999.
2. Müller, G., Rannenberg, K. editors; Multitlateral Security – Empowering Users, Enabling Applications, The Ladenburger Kolleg "Security in Communication Technology" Annex in [1], 563-570.
3. RFC-1122, Requirements for Internet Hosts -- Communication Layers, R. Braden (ed.), October 1989.
4. Tanenbaum, A.S., Wetherall, D.J., Computer Networks, 5th ed., Pearson, Boston USA, 2011.
5. Simmons, G.J.; Contemporary Cryptology, The Science of Information Integrity, IEEE press, New York, 1992.
6. Netscape, DevEdge Online Documentation, Introduction to Public-Key Cryptography, 1998, http://developer.netscape.com/docs/manuals/security/pkin/index.htm.
7. The SSL Protocol, Version 3.0, http://www.freesoft.org/CIE/Topics/ssl-draft/INDEX.HTM
8. TLS, RFC 5246 http://tools.ietf.org/html/rfc5246
9. The SSH Protocol, http://www.snailbook.com/protocols.html.
10. Unruh, W.; PGP attacks, 1998, http://axion.physics.ubc.ca/pgp-attack.html.
11. Engelfriet, A.; The comp.security.pgp FAQ, 1998, http://www.uk.pgp.net/pgpnet/pgp-faq/
12. Esslinger, B., Fox, D.; Public Key Infrastructures in Banks – Enterprise-wide PKIs, in [1], 283-300.
13. McDermott, J., Attack-Potential-Based Survivability Modeling for High-Consequence Systems, Proc. Third IEEE Int. Inf. Assurance Workshop, 2005, Washington DC, J. Cole, S. Wolthusen eds.
14. Riordan, J.; Patterns of Network Intrusion, in [1], 173-186.
15. Kerberos V5 Installation Guide, MIT, 1990 – 1996 http://www.lns.cornell.edu/public/COMP/krb5/install/install_toc.html.
16. SWIFT, http://www.swift.com/support .
17. Armstrong, I.; Web Commerce – Trading Securely, Security Magazine, October 2000, http://www.scmazine.com/scmagazine/2000_10/feature.html.
18. PayPal, http://www.paypal.com.
19. eBay, http://www.ebay.com, ebay.co.uk, ebay.de.
20. Verizone Business, 2009 Data Breach Investigations Report, Verizon, 2009.
21. Müller, G., Reichenbach M.; Sicherheitskonzepte für das Internet, 5. Berliner Kolloquium der Gottlieb Daimler – und Karl Benz-Stiftung, Springer-Verlag, Berlin, 2001, section 4.3.
22. Braczyk, H.H. et al.; Trust and Socio-Technical Systems, in [1], 425-438.
23. RSA, The Security Div. of EMC^2, www.rsa.com.
24. CyberSource Corp., www.cybersource.com.
25. Fraud Watch, www.fraudwatchinternational.com.
26. IC3, Internet Crime Complaint Center, http://www.ic3.gov.