2011 3rd International Conference on Cyber Conflict
C. Czosseck, E. Tyugu, T. Wingfield (Eds.)
Tallinn, Estonia, 2011 © CCD COE Publications

Towards Next-Generation Intrusion Detection

Robert Koch
Institut für Technische Informatik (ITI)
Universität der Bundeswehr
Munich, Germany
Robert.Koch@UniBw.de

***Abstract*- Today, Intrusion Detection Systems (IDS) are integral components of larger networks. Even so, security incidents are on a day-to-day basis: Numerous data leakage scandals arouse public interest in the recent past and also other attacks like Stuxnet are discussed in the general public. On the one side, the commercial success of the Internet and the possibilities to carry out attacks from a relatively safe distance attracts criminals and made e-Crime to a multi-billion dollar market over the past years. On the other side, more and more services and systems migrate to the Internet, for example Voice over IP (VoIP) or Video on Demand (VoD). This enables new and potential attack vectors.**

With the steadily increasing use of encryption technology, State-of-the-Art Intrusion- as well as Extrusion Detection technologies can hardly safeguard current networks to the full extend. Furthermore, they are not able to cope with the arising challenges of the fast growing network environments.

The paper gives an overview of up-to-date security systems and investigates their shortcomings. Latest security-related threats and upcoming challenges are analyzed.

In the end, requirements for a Next-Generation IDS are identified and current research as well as open issues are presented.

Keywords: Next-Generation Intrusion Detection, Security Threats, Intrusion Detection, Intrusion Prevention, Data Leakage Prevention, Early Warning

I. INTRODUCTION

With the interconnection of computer systems, numerous security threats emerged. One of the first publications towards IDSs was a technical report in 1980 [1]. A first model of a real-time IDS and a prototype had been built, the Intrusion Detection Expert System IDES [2]. Nowadays, plenty of specialized systems exist, but the basic functionality can be differentiated with regard to the detection technique, misuse- (signature) and anomaly detection (behavior). While the former ones search for well-known patterns, the latter ones build a model of the normal network behavior and attacks can be detected by measuring significant deviation of the current status against the behavior expected from the model. Therefore, anomaly-based systems are able to detect new and yet unknown threats at the cost of higher false alarm rates. The placement of the system, host- or network-based, is another aspect. Host-based systems are able to access a wide range of system information, logs, etc., while network-based systems are only able to evaluate the network traffic. However, because of their installation at central points in the network, they are able to detect attacks against the whole network or distributed attacks, which cannot be detected by a host-based analysis.

Other attributes can be used for a more precise classification, like time-based constraints or the degree of interoperability (e.g., [3, 4]).

Today, well-known attacks or new threats like a worm propagation can be detected and obstructed. Anyway, all systems suffer from important real-world problems. Even more, the current technology trends tighten this situation: Yet available systems will not be able to cope with challenges like encryption or increasing bandwidth.

The further paper is organized as follows: In Section 2, a brief overview of the evolution of security threats is given. Section 3 presents State-of-the-Art security systems and research and points out their most important shortcomings. Based on the identified shortcomings, requirements for Next-Generation Intrusion Detection are derived in Section 4. An architecture of a Next-Generation IDS is proposed in Section 5. Concepts under development, which try to address some of the most important current shortcomings, are presented as well. Finally, Section 6 concludes the paper by highlighting the most important open research issues.

II. THREATS AND TENDENCIES

The scope of attackers and malicious programs has changed significantly over the years. The focus of the first computer virus was on the destruction of data, e.g. formatting the hard disk drive or deleting executable files (e.g. [5]). With the development of worms, automated infection over networks was enabled and used to build botnets, consisting of numerous user PCs without the knowledge of the owners. These networks can consist of hundreds of thousands of infected systems and are used to send Spam or to block services by Distributed Denial-of-Service (DDoS) attacks.

The destructive behavior in the beginning changed to commercial-driven reasons. Today, spammers can participate in Affiliate Programs: They sign up in a program and are provided with an unique identifier. If a sale is backtracked to an identifier, the corresponding spammer is rewarded with a commission [6].
The commercial success of the Internet and the possibilities to carry out attacks from a relatively safe distance attracts criminals and makes e-Crime to a multi-billion dollar market (e.g., see [6, 7]).
Therefore, the profitable and relatively risk-free underground market stimulates the proliferation of malicious code by the creation and selling of attack kits. No technical in-deep knowledge is needed any longer to create new, dangerous malicious software [8]. The first attack kit (Virus Creation Lab, 1992) only provided basic functionality, but state-of-the-art kits like Nukesploit are highly professional and sold for several thousand Dollars. Also different service levels are available, for example unlimited support or regular updates [9]. A major difficulty arising with the professional construction kits are the high numbers of new signatures. A new signature appears with every new created code, building malicious code families.
More and more services migrate to the Internet, for example VoD or VoIP. With more and new services, also more new potential attack possibilities arise. For example, some malicious programs encrypt the data on the infected system and the user has to pay for the key. This type of malicious software (ransomware) appeared for the first time in 1989 [10]. Today, Trojan Horses exists which are able to encrypt data based on public key cryptography [11].
Another aspect is the handling of malicious programs: Latest trends show, that the percentage of targeted attacks continuously increases. E.g., the Hydraq Trojan (Aurora): Several large companies had been compromised by attackers using this Trojan [7]. The attacks started by evaluating data about employees, available on the company's website or in social networks: Social Engineering is on the rise again. Social networks like Facebook or Twitter are in the focus of attackers because of their prosperity of information. Many people are easygoing when dealing with sensitive data in social networks. This information is used by social engineers to create attacks, e.g. Emails with malicious attachments, obviously sent by a friend and with a topic related to the latest movements in the social net. So, the probability that the target opens the attachment and infects the system is very high. Targeted attacks are often constructed for a single or few destinations, so no patterns will be available.
The dissemination routes of malicious software are not restricted to networks: E.g., promotional gifts like USB-sticks given away on trade shows are popular instruments [12]. A Trojan is already installed on the stick. By connecting the stick to a computer, the Trojan installs itself on the system. Therefore, the threat is injected directly onto the target system or network, bypassing the security systems. With the help of offline-propagation, also formerly secure systems and networks like Supervisory Control And Data Acquisition Systems (SCADA) can be compromised. Therefore, a protection against attacks from the outside is not enough.

Data leakage has become an important issue for the last years. In contrast to the insider threat, data leakage includes accidental or unintentional data loss in addition to malicious theft [13]. Numerous scandals about data loss arose public interest, for example see [14, 15]. The insider threat is one of the most challenging endangerments today. While governments and the military had been in the spotlight of attacks during the cold war, today the industry is the most important target for espionage. A recent study specified the economic loss for each individual business company in Germany on an average of about 5,57 million Euro in 2009. 61 % of all large-scale enterprises had been hit by business crime in the past two years [16].

The particular endangerment by the insider is based on the authorized access and the knowledge about the security mechanisms. Also, by the widely spread use of data storage mediums like memory sticks, it can be easy for a legitimate employee to extrude confidential data if no protection mechanisms are in place. The released numbers of the percentage of the insider threat compared to all incidents of data loss differ keenly and go up to 80 % and more. The Verizon Data Breach Investigation Report attracted interest in 2008, because their evaluation of the insider threat presented a value of only 18 % [17]. Anyway, in the Report of 2010, Verizon published a proportion of about 48 % incidents caused by insiders after evaluating a wider range of cases [18]. In addition, the estimated numbers of unreported cases based on insider jobs are much higher, because numerous companies do not press charges because of a possible loss of reputation. The detection of data leakage is difficult by nature, but the situation is even worse, because high damage only can be avoided by immediate reactions. Beside this challenges, the technical evolution of the Internet opens up additional problems. More and more services are offering protected access. For example, the well-known Firesheep [19] addon for the Firefox-browser attracted numerous people. It enables easily operated HTTP session hijacking attacks. While these security hole existed for several years without concerning the public interest (because of the complex way to utilize it), the addon is easy to use [20]. Therefore, anybody is able to take over a foreign session. The tool comes with filters for e.g. Facebook, Twitter and GMail. So, numerous services like Facebook announced to switch their services to TLS. The trend towards the use of encryption will also be enforced with the broader application of IPv6 as IPSec is a mandatory component of IPv6 [21]. In February 2011, the last address blocks of IPv4 had been assigned. This should speed-up the utilization of IPv6 in the near future; at the moment, less than 1 percent of all traffic is IPv6 (e.g., [22]). Encryption can train the application of IDSs, therefore being a crucial factor.

Important is the shifting from attacks directed onto the operating systems or network protocols to attacks of vulnerabilities in the application layer. The nonstop evolution of the applications results in complex programs and flawed program code. Today, over 70 % of all attacks are targeting the application layer [7]. Most utilized vulnerabilities are provided by browsers and programs like the Acrobat Reader (e.g., [23]). Based on that, the number of Zero Day vulnerabilities increased in recent years.

Vendors are sometimes delaying patches unnecessarily by using a fixed patch-day policy: Program updates are only published on a regular basis (e.g., [24]). Also the safety awareness of the users is inadequate, many users are overstrained by complex and often changing security mechanisms and program configurations: The most successful exploits are taking advantage of vulnerabilities first reported more than a year ago [7].
Current available system are hardly able to cope with these trends. Summarized, the following threats and tendencies are identifiable and emerging:

1. New and yet unknown attacks (new services, devices, etc.)
2. Increasing number of Zero Days
3. Social Engineering and targeted attacks
4. Exploitation of vulnerabilities in the application layer
5. Increasing insider threat
6. Risk of data leakage
7. Ascending use of encryption technology
8. Users are negligent with security-related tasks

Following, current IDSs and techniques will be considered with respect to these properties.

III. CURRENT SYSTEMS AND SHORTCOMINGS

SNORT [25] is a signature-based Network-IDS (NIDS) and Intrusion Prevention System (IPS) capable of performing real-time traffic analysis and packet logging. To gain acceptable results regarding the false alarm rates, signature-based systems like SNORT have to be configured strongly depending on the hosts and services presented in the network. If the system generates many false alarms, no administrator will pay attention to the IDS after a few days. However, a complete in-depth configuration of all systems and services is time-consuming and difficult. Also, the configuration has to be administered all the time: Small changes like an update can have a significant impact. Therefore, the application of signature-based techniques in big network environments often is not successful.
Signature-based systems are reactive by nature [26] and restricted to already known attacks. Therefore, the efficiency of an IDS relies on the update-rates and response-times of the responsible company.
For example, on January, 20th 2011, the latest IDS signatures of Juniper and Sourcefire were released on January, 18th while the latest signatures of Proventia and IntruShield were published on January, 11th (as seen on [27]). Therefore, the up-to-dateness of the signature databases is a crucial point. Lippmann analyzed the effect of identifying vulnerabilities and patching software with regard to IDSs and the up-to-dateness of their signatures [28]. The signatures for IDSs are often not faster available than the publication of software-patches. However, vendors like Microsoft or Adobe often use patch-days for publishing numerous patches at once, therefore delaying patches unnecessarily. For example, latest threats opened up by the vulnerabilities in the Internet Explorer [29] have not been fixed in the

consecutive patch-day even exploits had been published meanwhile and an easy deployable code had been included in the Metasploit-Framework [30].
Anomaly-based systems do not need a signature database, instead they use a model for the evaluation. The accurate modeling of network behavior is an active field of research [31]. The difficulty of behavior-based models is the possibility of misinterpretation of permitted but unknown legal user actions, resulting in high false alert rates. Often, a learning phase is needed to train the corresponding detection model. Online and offline learning must be differentiated. The former one is an incremental respectively sequential training: Learning is performed piece-by-piece in a serial fashion on one individually (randomly) selected training sample set. The latter one takes the whole problem data in one learning iteration [32]. All data has to be labeled in advance due to its benign or malicious character which can be a difficult task (e.g., see [33, 34]). Because of the time and effort needed, Almgren and Jonsson use active learning methods to reduce the needed amount of labeled training data [35].
Even more, numerous anomaly-based systems use online learning in the productive environment (e.g., [36], [37]). Attacks can take place or malicious code can already be in the network during the learning phase, resulting in an erroneous model [33]. Therefore, the system will recognize the previously learned attacks as normal behavior and no alarm will be raised [38, 39]. A possible countermeasure is the use of malicious rather than benign data for the training. Winter et al. proposed a one-class support vector machine (SVM) for the analysis. The system is trained with malicious network traffic [40]. Anyway, also the usage of negative behavior is difficult in matters of completeness. To countervail the endangerment of a learning phase, using unsupervised learning methods can be a solution. Numerous machine-learning approaches have been developed over the last decades. Examples are statistic-based systems, data mining, expert systems, supervised learning-based approaches like neural networks and unsupervised learning-based approaches like k-means- or self-organizing feature maps (SOM) ([38, 41, 42]). Sometimes, supervised and unsupervised concepts are combined (e.g., see [43]). Casas et al. proposed a robust clustering technique to detect anomalous traffic flows based on a sequentially captured temporal sliding-window basis [44]. The approach is completely unsupervised and able to detect attacks without relying on signatures, labeled traffic or training. The system can be directly plugged-in and starts working from scratch without previous knowledge.
Signature-based systems are using string matching techniques to find known patterns of malicious code. This is a computational complex task and can generate up to 80 percent of the total processing time of the IDS [45]. Payload filtering delay has become the main cause of the reduction of network performance [46]. Because of that, software-based NIDSs are hardly able to keep up with traffic over 200 Mbps [47] when executing a full payload analysis (Deep Packet Inspection (DPI)). Therefore, numerous designs and algorithms for hardware-based acceleration have been proposed in recent years. Today, two categories of hardware approaches can be classified, logic and memory architectures. Logic architectures mostly use on-chip logic resources of Field-Programmable Gate Arrays (FPGA) to convert regular expression patterns into parallel state machines

or combinatorial circuits. Memory architectures compile string patterns to finite-state machines and store the corresponding state transition tables in memory [48]. Therefore, memory architectures are more flexible because they allow on-the-fly pattern update without resynthesis and relayout which is needed by logic architectures. By the use of FPGA, the string matching process can be accelerated strongly (e.g., see [47, 49, 50]). On the other side, a tremendous amount of chip resources is used by the growing rule sizes. A lot of work is done to reduce the required number of logic cells per search character (e.g., see [45, 51]). Kong et al. argue that all on-chip solutions are not scalable in long term [46]. Memory and on-board architectures are more likely able to keep up with the increasing set of rules and bandwidth. For example, pre-filtering [47] or prefix and suffix sharing for the rules [46] enables higher speeds (6.4 Gbps respectively 4 Gbps). Even FPGAs can reach high throughput speeds, Gao argues that already light-weight systems like Snort cost too many hardware resources [52]. Also, with the growth of the signature sets and the design scale, the interconnecting latency increases. Therefore, the operation clock frequency and throughput speed will drop. Gao also mentions, that a larger design requires more time for updating and reconfiguration procedures. During that time, the system is vulnerable, e.g. for a new worm spreading. Therefore, Gao uses Ternary Content-Addressable Memories (TCAM) for signature matching, reaching 2 Gbps (and theoretical much higher values) for the Snort signature set.

Other approaches use the network interface card (NIC) for implementing efficient and fast detection systems. While Otey et al. only use header information for the evaluation [53], Bruijn et al. analyze different levels of abstraction, e.g. packets, streams and aggregates in their system [54].

If the mounds of data are to high for a payload or at least a header inspection, Flow-based evaluation can be fulfilled. IPFIX (RFC 5472 [55]) defines a Flow as a group of packets that share a common set of properties. The Flow is completely specified by that set of values, together with a termination criterion (like inactivity timeout). Two important standards are NetFlow ([56] et al.) and sFlow [57]. Plenty of tools are available for the evaluation, e.g. Scrutinizer [58] or NetFlow Analyzer Professional [59]. Flow-based evaluation enables the possibility to analyze higher bandwidth. However, using Flows introduces a delay, therefore the system is not able to initiate fast, near real-time countermeasures [60]. Even more, attacks on the application layer (which are already the most important attacks and still rise) often cannot be detected by the evaluation of the Flow parameters. Figure 1 gives an overview of the implications and correlations of important threats and technology trends for anomaly- and signature-based systems.

In addition to IDSs, the area of Data Leakage Protection (DLP) and Extrusion Detection Systems (EDS) is an emerging sector in recent years. While IDSs focus on the attacker and malware trying to enter and infect the systems from outside, EDSs are monitoring the outgoing traffic searching for keywords or verifying the compliance of the communication with the policies of the company.

When coping with data leakage, DLP systems can cover up to three areas depending on the functionality, namely data at rest (databases, files, etc.), data in

motion (network traffic) and data in use (data traveling to peripherals like DVD burner or printer) [61]. Therefore, packet inspection, session monitoring, encryption and other techniques are used by a DLP.

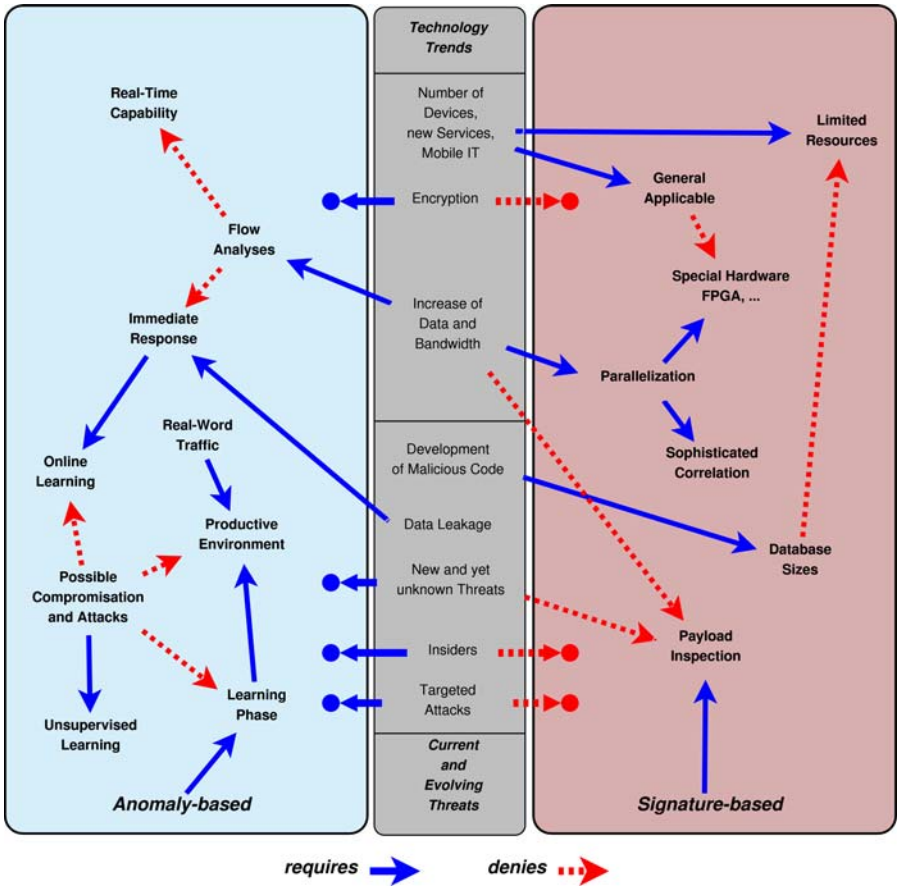

Figure 1. Evolving threats and technology trends and consequences for signature- respectively anomaly-based IDSs.

In the area of DLP, the first implementation was Security Enhanced Linux (SELinux) developed by the National Information Assurance Research Laboratory of the US National Security Agency (NSA), Red Hat and some other companies [62]. SELinux was released as Open Source in 2000 and is included in several Linux distributions today. Plenty other DLP systems and services are available recently, for example from Trend Micro [63] or IBM [64].

Additional to the IDSs, Early Warning Systems (EWS) are in operation. Compared to IDSs, EWSs are larger scaled, monitoring data sources distributed over the whole Internet. The information of attacks in one subnet can be used to alert and

safeguard other sub-networks, which are not yet under attack. For example, the spread of a new worm can be easily detected in the early phase of its run.
In 2001, the Internet Storm Center (ISC) of the SANS Institute was established [65]. It is an analysis and warning service for the Internet and consists of sensors covering 500.000 IP addresses around the globe. Firewall and intrusion detection log entries are collected and sent to the DShield database of the ISC. Abnormal trends and behavior is identified through human volunteers and automated evaluation. Based on that, the handler on duty sets the Infocon level which should reflect changes in malicious traffic and the possibility of disrupted Internet connectivity. Another example is the Arbor Networks Active Threat Level Analysis System (ATLAS) [66].
The NEWS (Network Early Warning System) plugin [67] gathers the collectively provided view of peer computers to detect network anomalies. NEWS uses corroboration from multiple users running in the same area. If enough people see the same problem in the same area, an alarm is raised. The design principles apply to large-scale systems that generate a significant amount of network traffic. Numerous other systems are under development like FIDeS [68] or WOMBAT [69].
The shortcomings and challenges of the current systems are summarized again:

1. Complex configuration
2. Detectability of Zero Days
3. Delays for signature updates
4. Rising bandwidth and data volume
5. Sizes of pattern databases
6. Application-Layer attacks
7. Encrypted network connections

Table I assigns the characteristics to the different systems.

Table I: Shortcomings of current Intrusion Detection / Prevention & Data Leakage Prevention Systems.

	Intrusion Detection						
	Signature-Based		Behavior-Based		Data Leakage		EWS
	Host	Network	Host	Network	Host	Network	*Network*
Configuration	×	×	√	√	×	×	(√)
Zero Days	×	×	√	√	¬	¬	√
Signature Delays	×	×	√	√	¬	¬	√
Bandwidth	×	×	√	(√)	√	(√)	(√)
Database Sizes	×	×	√	√	√	√	(√)
Application Layer	(√)	(√)	(√)	×	√	(√)	√
Encrypted Communication	√	×	√	(√)	√	×	×
Targeted Attacks	×	×	(√)	(√)	¬	¬	×
Distributed Attacks	×	(√)	×	√	¬	¬	√

√ means uncritical while × shows shortcomings of the particular systems, () means restricted applicable, ¬ stands for not applicable. Note that EWS are inherently network-based, therefore there is no column for host-based systems.

IV. REQUIREMENTS FOR A NEXT-GENERATION IDS

Based on the shortcomings of current IDSs defined in Section 3 and the security-related threats and tendencies in the Internet shown in Section 2, the requirements for a Next-Generation IDS are educed. In detail, the IDS has to fulfill the following requirements:

1. Behavior-based analysis: Because of the increasing number of Zero Days, the growth of targeted attacks and the increasing percentage of encrypted communication (benign as well as malicious, e.g. botnet communications), signatures are often not available in time or not possible at all. Even if the application of behavior-based methods is a challenge, sophisticated statistical methods can be used to detect attacks (e.g., see [70], [71]) even in encrypted environments. Other reasons require behavior-based techniques, too: More and more mobile devices like smartphones are participating in the networks. Because of limited computing resources and with regard to the endurance of the batteries, the application of signature-based techniques is not reasonable or even possible. Also in server systems, the necessary near-realtime evaluation of patterns is limited not only by the amount of data to investigate but also by the sizes of the databases and millions of patterns.
2. Abdication of a learning phase: The use of behavior-based techniques often (but not necessarily), requires a learning phase of the system in the productive, real-world environment. Because of the endangerment of the learning phase and the difficult task of creating clean labeled data, this phase must be omitted as far as possible. Unsupervised learning techniques can be used (see Section 3) or the learning phase must be replaced by other techniques. For example, the anomaly-based system developed by Casas et al. [44] does not need signatures, labeled data or training. In the area of neural networks, Moraga examined how to design a neural network only based on knowledge [72]. It is important to understand that the abdication of the learning phase does not transform a behavior-based into a signature-based system: The detection is still fulfilled by the comparison of the measured state of the environment to the prediction of the model.
3. No payload evaluation: For a general applicability the system must be designed without the need of a payload evaluation as far as possible. Even more, the increasing use of encryption denies the use of payload data. Therefore, a Next-Generation IDS cannot rely on the availability of the packet payload.
4. Network-based evaluation and use of agents: Even if a host-based installation has several advantages with regard to the available information (e.g., running processes, decrypted data, log files, etc.), the IDS requires a network-centric design. On the one hand, distributed and sophisticated attacks against the whole network only can be recognized by a network-based installation, on the other side the management of

numerous host-based system is error-prone, complex to manage and often poorly scalable in large environments. Only if it is indispensable, host-based agents should supplement the network-based core system.

5. Cross-evaluation and distribution: The upcoming threats and challenges require an exhaustive use of behavior-based techniques. Therefore, the related false alert rates have to be reduced. By examine ingress traffic and the correlation of anomaly detection alerts of administratively disjoint domains, the false alert rate can be reduced significantly and abnormal data and Zero Days can be detected [73].
6. Active and automated prevention: The system must be able to carry out a completely automated operation. On the one hand, the amounts of data, connections and speed of actions are already too high to be able to permit a reasonable manual interaction. On the other side, especially in the area of DLP, a beginning leakage of data must be stopped as early as possible. The loss of reputation after losing data will often be more expensive (e.g., see [16]) than the costs caused by an misleadingly activated interruption of a single connection. Of course, the probability of a wrongly dropped connection must be very low.

V. ARCHITECTURE OF A NEXT-GENERATION IDS

To fulfill the requirements presented in Section 4, an architecture for a Next-Generation IDS is presented. An abstract view of the architecture is shown in Figure 2.

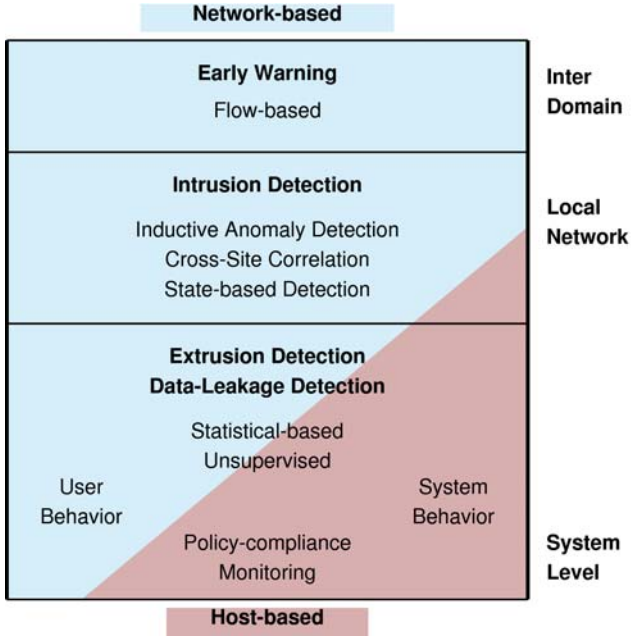

Figure 2. Layers of a Next-Generation IDS.

The system consists of three main parts, Early Warning, Intrusion- and Extrusion Detection. The different parts can be implemented distributed and autonomous. An EWS has to be integrated comprehensive over the Internet. Event correlation, anomaly detection and inter domain cross correlation can be used to detect new threats. This knowledge can then be used to secure other, yet not affected sub-networks in the Internet. The main purpose of the EWS is the detection and prevention of automated and undirected attacks.
The Intrusion Detection is carried out as NIDS. Multiple detection techniques have to be combined: A behavior-based analysis of the network traffic is done to detect known as well as new, yet unknown threats. The needed model has to be built in an unsupervised fashion in such a way, that no endangered learning phase is needed. If the learning phase cannot be eliminated completely, in contrast to most existing systems, malicious instead of benign data can be used (inductive anomaly detection). Cross-site correlation between systems and networks can be used to reduce the false alert rates of the anomaly detection efficiently. Statistical evaluation has to be done to cope with encrypted traffic. Additional, specialized host-based autonomous agents can be used to assist the evaluation. E.g., agents with state-based detection techniques can be used to identify critical states in a SCADA network: The critical states are well-known in industrial systems, therefore an Intrusion Detection can be realized based on a critical state analysis [74].
Extrusion Detection is the last component. It is also integrated into the NIDS, because due to the risk of insiders, manipulation and the administrative outlay with numerous hosts, host-based detection is not enough. Therefore, the user- and system-behavior is monitored by network-based sensors as well as host-based agents.
With respect to the current research and developments, several open issues arise, especially in the area of In- and Extrusion Detection in encrypted environments. Especially the claim of not using payload-related data to be able to cope with encrypted communication, targeted attacks and unknown threats is rarely address in current research (e.g., see [75]).
There are three basic approaches to carry out Intrusion Detection in encrypted communication, namely:

- Protocol-based: Detection of misuse of the encryption protocol
- Intrusive: Modifications of the network infrastructure or the encryption protocol
- Non-Intrusive: Statistical analysis of encrypted traffic

E.g., ProtoMon is a system developed by Joglekar et al. [76] which instruments shared libraries for cryptographic and application level protocols for conducting intrusion detection. Monitoring is integrated into the protocol handling. By that, attacks on the encryption protocol can be detected. Nevertheless, malicious activities hidden inside the encrypted channel could not be detected.

Intrusive techniques are used by Goh et al. They proposed an IDS for encrypted networks which is able to analyze the payload and simultaneously maintaining the confidentiality of the encrypted traffic [77]. The network traffic is replicated and sent to the receiver and also to the Central IDS (CIDS). The protocol is set onto an underlying VPN and adds an additional layer. The system is able to do payload analysis and to keep the confidentiality, but it strongly depends on modifications of the protocols and infrastructure. Also, additional communications protected by e.g. SSH or TLS cannot be analyzed.
Foroushani et al. proposed a system based on the evaluation of the transferred packet sizes and the time intervals between messages [70]. Attacks are detected without decryption by the use of intrusion signatures which are generated from the frequency of accesses and specifications of the TCP traffic. Anyway, because of a high false alarm rate (about 20 % in the best case), the system is not usable for a production environment. The system requires behavior profiles for the target servers and the exchanged information, which are often not available.
Other work addressing IDSs in encrypted environments can be found, but to the best of our knowledge, all of it can be assigned to one of the three categories named before (e.g., see [78] or [79]). Thus, all of these systems are not appropriate for the defined requirements due to the shortcomings already shown.
An important point of all behavior-based systems are the false alert rates. For a comprehensive development of behavior-based techniques in productive environments, false alerts have to be minimized. The idea of a correlation of ingress traffic from different domains is relatively new and shows promising first results. Boggs et al. were able to demonstrate a Proof-of-Concept with pretty small false alert rates [73]. Further investigations are necessary to improve the shown principles and make them usable for the defined requirements.
In the recent area of DLP, most of the proposed systems are host-based and not able to operate only on a network-based installation (e.g., see [80-82]). Extrusion and data leakage detection is a crucial part of a Next-Generation IDS. Therefore, these techniques have to be analyzed regarding the capability to be adapted to network-based systems.

VI. CONCLUSION AND FURTHER WORK

In the paper, an overview of todays most important security threats was given and observable tendencies were shown. State-of-the-Art IDSs, DLPs and EWSs were presented and their shortcomings analyzed. After that, the requirements for a Next-Generation IDS were derived. The papers shows the open issues and wherever available, recent research addressing these topics. The most important and yet unsolved requirements in the area of encryption and behavior-based analysis were lifted out.

ACKNOWLEDGMENT

The authors wish to thank the members of the Chair for Communication Systems and Internet Services at the Universität der Bundeswehr, headed by Prof. Dr. Gabi Dreo Rodosek, for helpful discussions and valuable comments on previous

versions of this paper. The Chair is part of the Munich Network Management Team.

REFERENCES

[1] Anderson, J., *Computer Security Threat Monitoring and Surveillance*, Fort Washington, April 1980

[2] History of Intrusion-Detection Research at SRI's Computer Science Laboratory, `http://www.csl.sri.com/programs/intrusion/history.html`, last seen on January 2011

[3] Sperotto, A., Schaffrath, G., Sadre, R., Morariu, C., Pras, A., Stiller, B., *An Overview of IP Flow-based Intrusion Detection*, IEEE Survey and Tutorials, Third Issue, 2010

[4] Sabahi, F. and Movaghar, A., *Intrusion Detection: A Survey*, 3rd International Conference on Systems and Networks Communications, ICSNC '08, IEEE Computer Society, 2008

[5] *The Jerusalem Virus*, `http://antivirus.about.com/cs/virusencyclopedia/p/jerusalem.htm`, last seen on January 2011

[6] M86 Security, *Security Labs Report*, Jul 2009-Dec 2009 Recap

[7] *Symantec Internet Security Threat Report*, Trends for 2009, Volume XV, April 2010

[8] McHugh, J. and Christie, A. and Allen, J., *Defending Yourself: The Role of Intrusion Detection Systems*, Software, IEEE, Volume 17, Number 5, 2000

[9] *Symantec Report on Attack Kits and Malicious Websites*, 2010

[10] *Ransomware: Extortion via the Internet*, `http://blogs.techrepublic.com/security/?p=2976`, last seen on January 2011

[11] Young, A., Yung, M., *Cryptovirology: Extortion-Based Security Threats and Countermeasures*, Proceedings of the IEEE Symposium on Security and Privacy, 1996

[12] Beckert, Kathrin, *Sicherheitstipp: Wirtschaftspionage per USB-Stick*, FH Gelsenkirchen, `https://www.it-sicherheit.de/ratgeber/it\_sicherheitstipps/tipp/sicherheitstipp-wirtschaftspionage-per-usb-stick/`, last seen on March 26th, 2011

[13] McCormick, M., *Data Theft: A Prototypical Insider Threat*, Advances in Information Security, Volume 39, April 2008

[14] *Ministry of Defence in new data loss scandal*, October 10th, 2008, `http://www.cio.co.uk/news/3225/ministry-of-defence-in-new-data-loss-scandal/`, last seen on January 2011

[15] *Data loss incident affects NASA*, December 10th, 2010, `http://www.backup-technology.com/5451/data-loss-incident-affects-nasa/`, last seen on January 2011

[16] PricewaterhouseCoopers, Martin Luther University Halle-Wittenberg, *Wirtschaftskriminalität 2009*, `http://www.pwc.de/de/risiko-management/assets/Studie-Wirtschaftskriminal-09.pdf`, 2009

[17] Baker, W.H., Hylender, C.D., Valentine, J.A., *2008 Data Breach Investigation Report*, Verizon Business RISK Team, Verizon Business, 2008

[18] Baker, W.H. et al., *2010 Data Breach Investigation Report*, Verizon Business RISK Team, Verizon Business, 2010

[19] *Firesheep Addon for HTTP session hijacking attacks*, `http://codebutler.github.com/firesheep/`, last seen on January 2011

[20] *Firefox extension steals Facebook, Twitter, etc. sessions*, `Firefox extension steals Facebook, Twitter, etc. sessions`, last seen on January 2011

[21] Kaushik, Das, *IPSec & IPv6 - Securing the NextGen Internet*, `http://ipv6.com/articles/security/IPsec.htm`, last seen on March 22th, 2011

[22] Amsterdam Internet Exchange, *sFlow Stats*, http://www.ams-ix.net/sflow-stats/, last seen on March 22th, 2011

[23] The H Security, *Adobe patches 23 holes in Reader and Acrobat*, http://www.h-online.com/security/news/item/Adobe-patches-23-holes-in-Reader-and-Acrobat-1102416.html, last seen on January 2011

[24] The H Security, *SAP introduces a patch day*, http://www.h-online.com/security/news/item/SAP-introduces-a-patch-day-1079976.html, last seen on January 2011

[25] *SourceFire SNORT*, http://www.snort.org/, last seen on January 2011

[26] Ghosh, Anup and Michael, Christoph and Schatz, Michael, *A Real-Time Intrusion Detection System Based on Learning Program Behavior*, Recent Advances in Intrusion Detection, Lecture Notes in Computer Science, LNCS 1907, Springer Berlin / Heidelberg, 2000

[27] *Computer Network Defence Operational Picture (Talisker Radar)* http://www.securitywizardry.com/radar.htm, last seen on January 2011

[28] Lippmann, R., Webster, S., Stetson, D., *The effect of identifying vulnerabilities and patching software on the utility of network intrusion detection*, Proceedings of the 5th international conference on Recent advances in intrusion detection, RAID 2002

[29] The H Security, *Microsoft issues warning about critical IE hole*, http://www.h-online.com/security/news/item/Microsoft-issues-warning-about-critical-IE-hole-1158684.html, last seen on January 2011

[30] Metasploit Framework *ms11_xxx_createsizeddibsection.rb*, https://www.metasploit.com/redmine/projects/framework/repository/revisions/11466/entry/modules/exploits/windows/fileformat/ms11_xxx_createsizeddibsection.rb, last seen on January 2011

[31] Thottan, Marina and Ji, Chuanyi, *Anomaly Detection in IP Networks*, IEEE Transactions on Signal Processing, Volume 51, No. 8, 2003

[32] Bitter, C. and Elizondo, D.A. and Watson, T., *Application of artificial neural networks and related techniques to intrusion detection*, The 2010 International Joint Conference on Neural Networks (IJCNN), IEEE, 2010

[33] Tandon, Gaurav and Chan, Philip and Mitra, Debasis, *MORPHEUS: motif oriented representations to purge hostile events from unlabeled sequences*, Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security, VizSEC/DMSEC '04, ACM, 2004

[34] Li, Yang and Fang, Binxing and Guo, Li and Chen, You, *Network anomaly detection based on TCM-KNN algorithm*, Proceedings of the 2nd ACM symposium on Information, Computer and Communications Security, ASIACCS '07, ACM, 2007

[35] Almgren, Magnus and Jonsson, Erland, *Using Active Learning in Intrusion Detection*, IEEE Computer Security Foundations Workshop, CSFW 04, Volume 17, IEEE, 2004

[36] *Lancope Network Behavior Analysis*, http://www.lancope.com/solutions/network-behavior-analysis.aspx, last seen on January 2011

[37] *FlowMatrix Network Behavior Analysis System*, http://www.akmalabs.com/flowmatrix.php, last seen on January 2011

[38] Debar, Herve and Dacier, Marc and Wespi, Andreas, *A Revised Taxonomy for Intrusion-Detection Systems*, IBM Research, Zurich Research Laboratory, 8803 Rüschlikon, Switzerland, 1999

[39] Bolzoni, D. and Etalle, S., *Approaches in anomaly-based network intrusion detection systems*, Intrusion Detection Systems, Springer, 2008

[40] Winter, Philipp and Hermann, Eckehard and Zeilinger, Markus, *Inductive Intrusion Detection in Flow-Based Network Data Using One-Class Support Vector Machines*, 2011 4th IFIP International Conference on New Technologies, Mobility and Security (NTMS), IEEE, 2011

[41] Liu Hui and Cao Yonghui, *Research Intrusion Detection Techniques from the Perspective of Machine Learning*, 2010 Second International Conference on Multimedia and Information Technology (MMIT), Volume 1, IEEE, 2010

[42] Hu, W. and Hu, W. and Maybank, S., *Adaboost-based algorithm for network intrusion detection*, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Volume 38, Issue 2, IEEE, 2008

[43] Carrascal, Alberto and Couchet, Jorge and Ferreira, Enrique and Manrique, Daniel, *Anomaly Detection using prior knowledge: application to TCP/IP traffic*, Artificial Intelligence in Theory and Practice, IFIP International Federation for Information Processing, Volume 217, Springer, 2006

[44] Casas, Pedro and Mazel, Johan and Owezarski, Philippe, *Steps Towards Autonomous Network Security: Unsupervised Detection of Network Attacks*, 2011 4th IFIP International Conference on New Technologies, Mobility and Security (NTMS), IEEE, 2011

[45] Sourdis, I. and Pnevmatikatos, D.N. and Vassiliadis, S., *Scalable multigigabit pattern matching for packet inspection*, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Volume 16, pages 156 – 166, IEEE, 2008

[46] Kong, Chao and Yang, Bo and Jia, Zhiping and Chen, Zhenxiang, *A Common On-board Hardware Architecture for Intrusion Detection System*, 2009 International Conference on Multimedia Information Networking and Security, IEEE Computer Society, 2009

[47] Korenek, Jan and Kobiersky, Petr, *Intrusion Detection System Intended for Multigigabit Networks*, Design and Diagnostics of Electronic Circuits and Systems, IEEE Computer Society, 2007

[48] Lin, Cheng-Hung and Chang, Shih-Chieh, *Efficient Pattern Matching Algorithm for Memory Architecture*, IEEE Transaction on Very Large Scale Integration (VLSI) Systems, Volume 19, January 2011

[49] Mitra, A., Najjar, W., Bhuyan, L., *Compiling PCRE to FPGA for Accelerating SNORT IDS*, ACM/IEEE Symposium on Architectures for Networking and Communications Systems, 2007

[50] Baker, Zachary K. and Prasanna, Viktor K., *High-throughput Linked-Pattern Matching for Intrusion Detection Systems*, Symposium on Architectures for Networking and Communications Systems, ANCS 05, ACM, 2005

[51] Sourdis, Ioannis and Pnevmatikatos, Dionisios, *Fast, Large-Scale String Match for a 10Gbps FPGA-Based Network Intrusion Detection System*, Field-Programmable Logic and Applications, Springer, 2003

[52] Ming Gao and Kenong Zhang and Jiahua Lu, *Efficient packet matching for gigabit network intrusion detection using TCAMs*, 20th International Conference on Advanced Information Networking and Applications, AINA '06, 2006

[53] Otey, M. and Parthasarathy, S. and Ghoting, A. and Li, G. and Narravula, S. and Panda, D., *Towards nic-based intrusion detection*, Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2003

[54] De Bruijn, W. and Slowinska, A. and Van Reeuwijk, K. and Hruby, T. and Xu, L. and Bos, H., *SafeCard: A Gigabit IPS on the Network Card*, Recent Advances in Intrusion Detection, Lecture Notes in Computer Science, LNCS 4219, Springer, 2006

[55] *IP Flow Information Export (IPFIX) Applicability*, `http://tools.ietf.org/html/rfc5472#section-3.6.2`, last seen on January 2011

[56] *Cisco Systems NetFlow Services Export Version 9*, `http://www.ietf.org/rfc/rfc3954.txt`, last seen on January 2011

[57] *InMon Corporation's sFlow: A Method for Monitoring Traffic in Switched and Routed Networks*, `http://www.ietf.org/rfc/rfc3176.txt`, last seen on January 2011

[58] *Scrutinizer NetFlow & sFlow Analyzer*, `http://www.plixer.com/products/free-netflow.php`, last seen on January 2011

[59] *NetFlow Analyzer*, http://www.manageengine.com/products/netflow/download.html, last seen on January 2011

[60] Lim, Shu Yun and Jones, Andy, *Network Anomaly Detection System: The State of Art of Network Behaviour Analysis*, International Conference on Convergence and Hybrid Information Technology, 2008. ICHIT '08, IEEE Computer Society, 2008

[61] Lawton, G., *New Technology Prevents Data Leakage*, Computer Journal, Volume 41 Issue 9, September 2008, 10.1109/MC.2008.394

[62] *Security-Enhanced Linux*, http://www.nsa.gov/research/selinux/, last seen on January 2011

[63] *Data Loss Prevention Services*, http://us.trendmicro.com/us/products/enterprise/data-loss-prevention/services/, last seen on January 2011

[64] *Fidelis Security Systems appliances and support*, http://www-935.ibm.com/services/us/index.wss/offering/iss/a1031185, last seen on January 2011

[65] *Internet Storm Center* of the SANS (SysAdmin, Audit, Network, Security) Institute, isc.sans.org, last seen on January 2011

[66] *Active Threat Level Analysis System*, Arbor Networks, atlas.arbor.net, last seen on January 2011

[67] Bustamante, F., Choffnes, D., *NEWS plugin for the Vuze BitTorrent Client*, http://www.aqualab.cs.northwestern.edu/projects/NEWS.html, last seen on January 2011

[68] *Early Warning and Intrusion Detection based on Combined AI Methods*, www.fides-security.org, , last seen on January 2011

[69] *Worldwide Observatory of Malicious Behaviors and Attack Threats*, wombat-project.eu, last seen on January 2011

[70] Foroushani, V.A., Adibina, F., Hojati, E., *Intrusion Detection in Encrypted Accesses with SSH Protocol to Network Public Servers*, Proceedings of the International Conference on Computer and Communication Engineering 2008, May 13-15, Kuala Lumpur, Malaysia

[71] Melnikov, N., Schönwälder, J., *Cybermetrics: User Identification Through Network Flow Pattern Analysis*, EMANICS Workshop on NetFlow/IPFIX Usage, Jacobs University Bremen, October 2009

[72] Moraga, C., *Design of Neural Networks*, Knowledge-Based Intelligent Information and Engineering Systems, Lecture Notes in Computer Science, Volume 4692/2008

[73] Boggs, N., Hiremagalore, S., Stavrou, A., Stolfo, S., *Experimental Results of Cross-Site Exchange of Web Content Anomaly Detector Alerts*, IEEE Conference on Technologies for Homeland Security, Boston, 2010

[74] Carcano, A. and Coletta, A. and Guglielmi, M. and Masera, M. and Nai Fovino, I. and Trombetta, A., *A Multidimensional Critical State Analysis for Detecting Intrusions in SCADA Systems*, Industrial Informatics, IEEE Transactions on, 2011

[75] Schaffrath, G., *Network Intrusion Detection Systems & Encryption: Friends or Foes?* , Communication Systems Group, University of Zürich, August 2008

[76] Joglekar, S., Tate, S., *ProtoMon: Embedded Monitors for Cryptographic Protocol Intrusion Detection and Prevention*, Journal of Universal Computer Science, Volume 11, 10.3217/jucs-011-01-0083

[77] Goh, V.T., Zimmermann, J., Looi, M. (2010), *Experimenting with an Intrusion Detection System for Encrypted Networks*, Int. J. Business Intelligence and Data Mining, Vol. 5, No. 2, pp. 172-191

[78] Yasinsac, A., Goregaoker, S., *An Intrusion Detection System for Security Protocol Traffic*, Department of Computer Science, Florida State University

[79] Yamada, A., Miyake, Y., Takemori, K., Studer, A., Perrig, A., *Intrusion Detection for Encrypted Web Access*, AINAW 2007, ISBN 0-7695-2847-3

[80] Papadimitriou, P., Garcia-Molina, H., *Data Leakage Detection*, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Number 1, January 2011

[81] Cui, W., Katz, H., Tan, W. *Design and Implementation of an Extrusion-based Break-In Detector for Personal Computers*, Proceedings of the 21st Annual Computer Security Applications Conference (ACSAC 2005), 1063-9527/05

[82] Martignoni, L., Stinson, E., Fredrikson, M., Jha, S., Mitchell, J., *A Layered Architecture for Detecting Malicious Behaviors*, RAID 2008, LNCS 5230