How to Exchange Security Events? Overview and Evaluation of Formats and Protocols

Jessica Steinberger*‡, Anna Sperotto‡, Mario Golling† and Harald Baier*

*da/sec - Biometrics and Internet Security Research Group University of Applied Sciences Darmstadt, Darmstadt, Germany
Email:{Jessica.Steinberger, Harald.Baier}@h-da.de

†Research Center CODE Faculty of Computer Science, Universität der Bundeswehr München, Neubiberg, Germany
Email:{Mario.Golling}@unibw.de

‡Design and Analysis of Communication Systems (DACS) University of Twente, Enschede, The Netherlands
Email:{J.Steinberger, A.Sperotto, A.Pras}@utwente.nl

***Abstract*—Network-based attacks pose a strong threat to the Internet landscape. Recent approaches to mitigate and resolve these threats focus on cooperation of Internet service providers and their exchange of security event information. A major benefit of a cooperation is that it might counteract a network-based attack at its root and provides the possibility to inform other cooperative partners about the occurrence of anomalous events as a proactive service. In this paper we provide a structured overview of existing exchange formats and protocols. We evaluate and compare the exchange formats and protocols in context of high-speed networks. In particular, we focus on flow data. In addition, we investigate the exchange of potentially sensitive data. For our overview, we review different exchange formats and protocols with respect to their use-case scenario, their interoperability with network flow-based data, their scalability in a high-speed network context and develop a classification.**

I. Introduction

Network attacks pose a significant problem to the Internet landscape, which causes substantial financial losses [1], [2]. To counter these attacks extensive research in the field of detecting anomalous events within computer networks has been conducted over the last decade [3]. Sustainable detection and mitigation of cyber-criminal activities requires efficient and effective network-based solutions that can identify, track, and mitigate malicious traffic. Silva et al. [3], van Eeten et al. [4] and Arbor Networks [5] identified Internet service provider (ISP) networks as key points for this purpose, since they constitute a privileged observation point for large amount of traffic. In order to leverage this key position of ISPs in detection and mitigation of cyber-criminal activities, [6] identifies cooperation between ISPs as a viable solution for attack mitigation. An advantage of detection and mitigation of network-based attacks at ISP level and their cooperation is the possibility to exchange security events and thus improve timeliness and dissemination of network security information. As a consequence, the collaborating ISPs benefit from an increased security level based on shared knowledge.

For ISPs to collaborate, a common data representation to describe security-related data is important. Further, an exchange protocol to transmit security events over network borders is also required. Several exchange formats (e.g., IODEF, ARF, x-arf) and protocols (e.g., RID, IDXP) have been published in the last years. Although, a previous study that we have conducted [7] shows that most of the exchange formats and protocols are unknown to network operators, recent network-based attacks (e.g., Spamhaus) have clearly indicated that there is a need for collaboration.

To the best of our knowledge, only two other publications are related to exchange formats and protocols. First, Koch et al. [8] reviewed exchange formats and protocols used by intrusion detection and response systems, but the authors do not differentiate between a high-level description of functional requirements, an exchange format or an exchange protocol. Further, the authors do not provide information about flow-based interoperability and used security mechanisms of the exchange formats and protocols. Second, Kampanakis [9] focuses on information sharing models and reviewed efforts driven by the US Department of Homeland Security, Mitre and the US NIST. These models cover a wide range of exchange formats shown in Figure 1 and thus do not focus on Intrusion Detection and Incident Management.

In this paper we review the exchange formats and protocols used in context of intrusion detection and incident management. We classify the exchange formats related to their use-case scenario and assign them to the different stages of the *Operations Security Management Process of The Mitre Cooperation* [10] (Figure 1). We analyze both the data representation and the use-case scenario of the exchange formats. Further, we review existing exchange protocols and explain their intended use. As we identified the key position of ISPs in detection and mitigation of cyber-criminal activities, we develop various criteria to assess the exchange formats and protocols specifically in context of high-speed networks. Moreover, we assess the exchange formats for the use in conjunction with flow-based data, because a previous study from Steinberger et al. [7] stated that ISPs focus on detection of anomalous events based on aggregated network data (e.g. NetFlow, IPFIX). The goal of this paper is to provide network operators a hands-on selecting an exchange format and protocol suitable to use in their network. Therefore the main contributions of this paper are: (i) a comprehensive literature survey of 10 exchange formats and 7 exchange protocols that can be used to share security event related information in context of intrusion detection and incident handling, (ii) a structured overview that can be used by network operators when they have to decide what format and protocol should be used, (iii) an assessment of the exchange formats for the interoperability with flow-based data, (iv) a qualitative evaluation and comparison of the formats and protocols in context of high-speed networks and

finally an investigation of how to exchange potentially sensitive information.

The remainder of this paper is structured as follows. In Section II, we review the exchange protocols and formats, explain their structure and their use-case scenarios. In Section III we evaluate existing exchange protocols and formats regarding their strength and weaknesses especially in terms of their interoperability with flow-based data respectivly high-speed networks. Finally, the results are discussed and concluded in Section V.

II. Event Exchange Formats and Protocols

In this section, we introduce various formats and protocols to exchange security events as listed in Table I. First, we introduce the terminology by defining an exchange format and protocol. Additionally we define security event and security incident. Second, we review existing exchange formats and protocols, explain their structure and their use-case scenarios.

A. Terminology

A *format* describes a structure for the processing, storage, or display of data. An *exchange format* defines a representation of information for sharing data. An *exchange protocol* "is a set of rules defining how to interconnect network devices and establish a channel to transmit network datagrams, representing exchange formats, across a computer network [11], [12]".

The analysis of the state of the art has revealed that there are various terminology used to describe the content of an exchange message. In this paper, we adhere to the following: "A security event is defined as an identified occurrence of a system, service or network state indicating a possible breach of information security policy or failure of safeguards, or a previously unknown situation that may be security relevant [13], [14]." In accordance to this definition we summarize security alert, security alarm and security warning to security event. In contrast to a security event, "a security incident is defined as single or a series of unwanted information security events that have a significant probability of compromising business operations and threatening information security [13], [14]".

B. Exchange Formats

In this section we introduce the security event exchange formats that focus on intrusion detection and incident handling. Therefore, we classify the exchange formats related to their use-case scenario and assign them to the different stages of the Operations Security Management Process (OSMP) of The Mitre Cooperation. This classification and assignment are shown in Figure 1. Next, we present the development and data representation of each exchange format. In addtion, we describe what kind of communication the exchange format is intended to facilitate. Last, we describe the involved communication partners of an exchange format message.

[1]Assessment Results Format
[2]Abuse Reporting Format

TABLE I. Overview of exchange protocols and formats

Protocol	OSI layer	Format	Security
CIDF	Transport	CISL messages	symmetric cryptography
RID	Application	IODEF	TLS
XEP-0268	Application	IODEF	TLS
IDXP	Application	IDMEF	TLS
CLT	Transport	CEE	provided by syslog (RFC 5425)
SMTP	Application	CAIF ARF x-arf	None S/MIME Multipart/Signed Multipart/Encrypted
syslog (RFC 3164)	Transport	syslog (RFC 3164)	none
syslog (RFC 5425)	Transport	syslog (RFC 5424)	TLS

1) Common Intrusion Detection Framework: In 1997 the Common Intrusion Detection Framework (CIDF) project was launched by the US governments Defense Advanced Research Projects Agengy (DARPA) to construct an infrastructure that allows Intrusion Detection, Analysis, and Response Systems (IDAR) from different manufacturers to share information [15]. CIDF provides a LISP-like structure to express information about events, attacks, and responses called Common Intrusion Specification Language (CISL) [16]. The syntax of CISL consists of S-expressions, a list-based data structure, which are known for their use in the LISP family of programming languages. Besides S-expressions, CISL includes nouns and verbs to encode security events. Further, CISL is intended to facilitate machine to machine (M2M) communication. CISL is quite powerful, because a message described in CISL has no limitations related to the type of information that is communicated [17]. Despite the fact that CISL is quite powerful, it is presently dormant [18], [19].

2) Incident Object Description Exchange Format: In 2001, initial efforts towards the Incident Object Description Exchange Format (IODEF) started. The Terena IODEF Working Group was dissolved in 2002 and further IODEF development has been transferred to IETF Incident Handling (INCH) Working Group (WG). In 2003 the IETF Extended INCH WG developed the Format for Incident Report Exchange (FINE). FINE defines a format to facilitate the exchange of security incident information between Computer Security Incident Response Teams (CSIRTs). In particular, FINE is a high-level description of functional requirements for a format and thus not an implementation. Referring to the requirements defined in FINE, the data model of IODEF was specified in RFC 5070 [20]. In contrast to CISL, IODEF (or called INCH) focuses on a human-to-human (H2H) communication. IODEF is used to share information commonly exchanged by CSIRTs about computer security incidents. IODEF is written in XML and transferred over network using Real-time Internetworking Defense protocol (RID). Besides RID, IODEF could be transferred using the specification XEP-0268 of the Extensible Messaging and Presence Protocol (XMPP) [21], [22]. IODEF is compatible to the Intrusion Detection Message Exchange Format (IDMEF), but IODEF incident handling (reporting, investigations), storage, statistics and trend analysis result in a longer lifetime of IODEF messages compared to one time use of IDMEF messages.

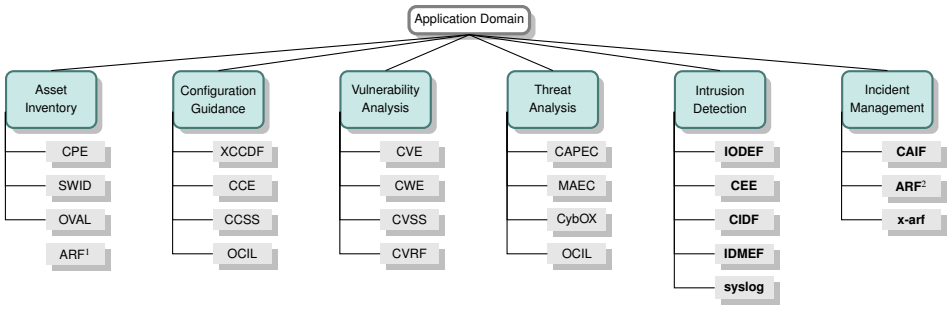

Fig. 1. Application domains of exchange formats based on [10]

3) Common Announcement Interchange Format: In 2002 the development of Common Announcement Interchange Format (CAIF) started. CAIF is an XML-based format to store and exchange security announcements developed by the Stuttgart University's Computer Emergency Response Team (RUS-CERT). CAIF is designed to describe a problem, vulnerability, or exposure that threatens the security of the IT infrastructure. Further, CAIF provides information about problems or solutions (a vendor patch or a workaround) or contains background information about one or more IT security issues. CAIF is able to address more than one security event within a single document and allows to group information for more than one target group of readers as well as multi-lingual textual descriptions within one document. Different renderings of an announcement for the intended target groups addressing one, a sub-set, or all problems multi- or mono-lingual in the languages provided are supported [23]. CAIF messages are generated by vendors or security teams and sent to users or administrators. Therefore CAIF is intended to facilitate a H2H communication. Currently, CAIF is generally considered to be inactive [23].

4) Intrusion Detection Message Exchange Format: The Intrusion Detection Message Exchange Format (IDMEF) development is under control by the Intrusion Detection Working Group (IDWG) under the IETF. There have been numerous drafts of the IDMEF RFC. In 2003 the draft RFC had been submitted to the Internet Engineering Steering Group (IESG). IDMEF has been approved by the IESG as an Informational RFC 4765 [24]. IDMEF is intended to be a standard data format to share information of interest (alerts, alive messages etc.) with the focus on intrusion detection events for IDS/IPS systems and management systems that interact with them. Thus IDMEF is intended to facilitate M2M communication. The purpose of an alert message is to automatically inform a manager about a single or multiple detected events. Alerts occur asynchronously in response to outside events [24], whereas heartbeat (alive) messages are sent in a regular period to indicate the analyzer's current status. Thus heartbeat messages are usually generated by network devices. IDMEF messages are written in XML and transferred over network using the Intrusion Detection Exchange Protocol (IDXP).

5) Common Event Expression: The Common Event Expression (CEE) [25] format was developed by a community representing vendors, researchers and end users. The development of CEE was coordinated by MITRE in 2009. CEE proposes an extensible event syntax, event vocabulary, log transport information and log recommendation information to achieve consensus in event representation, communication and interpretation. CEE uses an XML and a JSON representation. The architecture of CEE consists of three parts: Requirements, which are addressed in the CEE Profile; Events, which are represented as records using the CEE Log Syntax (CLS); and Records, which are shared via a CEE Log Transport (CLT). At the very least, CEE event consists of 4 elements (time, level, id, and message) to formulate a valid CEE message. CEE events are generated through various devices within the network (such as firewalls, IDSs) and transported to a centralized repository that is reviewed by an administrator. Thus, CEE is intended to facilitate machine to human (M2H) communication. The main difference between CEE and IODEF is that IODEF is an incident-related effort and IODEF incident reports often include event logs. These event logs may be provided in the CEE format, and a CEE-defined event may be incorporated into IODEF-defined incidents.

As both, architecture and syntax have not been finalized yet, CEE is still released as a beta version. In 2013 the U.S. Government organization, which sponsored MITRE's work on CEE has decided to stop funding development of CEE. As a result, MITRE has stopped all work on CEE, but offered to transfer all CEE-related specifications, documents and source materials to efforts willing to continue logging standards. The Project DMTF Cloud Audit and Project Lumberjack are continuing the work of CEE [25].

6) Messaging Abuse Reporting Format: Yakov Shafranovich wrote the initial documentation for the Abuse Reporting Format (ARF) to report spam via e-mail. Later, a small group of members of the Messaging Anti-Abuse Working Group (MAAWG), including Yakov Shafranovich, wrote the first draft of ARF in 2005 [26]. In the first draft of the MAAWG, an ARF message was based on email coupled with Multipart Internet Mail Extension. This Multipart MIME message consists of three parts. The first MIME part provides a human-readable description of the report. The second MIME part of the message is a machine-readable format with the content type of "message/feedback-report". The third MIME part contains either the original message in its entirety or a copy of the entire header block from the original message. All MIME parts are required and thus must be included to form a valid ARF message. Following the first draft of ARF from the MAAWG, IETF chartered a working group called Messaging Abuse Reporting Format (MARF). The MARF working group updated the draft of ARF, removed unused report types [27] and published a specification called ARF [28] in 2007. ARF has been published as RFC 5965 and added to the library

of official IETF standard documents [26]. By 2007 ARF was already a de facto standard to report spam via e-mail [26] and is intended to facilitate both M2M and M2H communication.

7) x-arf: In 2012 a competence group called "E-Mail" of eco - Association of the German Internet Industry has presented the reporting format x-arf [29]. The main intention of x-arf is to extend ARF. x-arf is a format to exchange computer security incidents and attack data via e-mail and is intended to facilitate both M2M and M2H communication. x-arf uses header fields of e-mails and two or three MIME-Parts. The content of an x-arf message is defined by a JSON schema. An x-arf message typically consists of three containers, where the third one is optional. The first container conveys textual content, is human-readable and UTF-8 encoded. The second container contains a JSON object (a list of key/value pairs) represented in YAML notation. The fields within the second container of an x-arf message depend on the type of abuse of the x-arf message. The third container is intended to contain log data. x-arf currently provides 5 different JSON schema: abuse, fraud, auth, info and private. In [30] Kohlrausch and Übelacker reported the usage of x-arf for exchanging data between national research networks. Within this use-case scenario they found that the x-arf standard format is not able of aggregating multiple security events. Furthermore, they conclude that x-arf messages are limited to one single destination address. Therefore they published a draft of x-arf specification version 0.2, which enhances x-arf in order to secure end-to-end communication for x-arf messages and to store bulk data within the RFC 2822 container. The RFC 2822 container describes the format of an Internet message, which consists of lines of plain text. The current specification of x-arf is version 0.2 and describes the three types of x-arf messages: plain, unencrypted, unsigned and single abuse message (PLAIN), a signed and/or encrypted via S/MIME and PGP/MIME abuse message and a bulk e-mail abuse message. One of the main changes from specification version 0.1 to 0.2 is the change of the identifier of an x-arf message. The identifier changed from X-ARF: YES to X-XARF: PLAIN, X-XARF: SECURE and X-XARF: BULK.

8) Syslog Message Format: In 2001 Eric Allmann wrote the first draft of the syslog message format that was first standardized as IETF RFC 3164 [31]. The syslog message format is used to log messages generated by various sources e.g., an operating system, a process or an application. A syslog message consists of a priority value (PRI), a header and a free-form message part that provides information about the event (source, severity, host name, time stamp and message) and is limited to 1024 bytes. In 2009 Rainer Gerhards updated the syslog message format to remedy shortcomings of the RFC 3164 such as missing time zones or year within the time stamp. This update was published as RFC 5424 [32] and is not backwards compatible to RFC 3164 [33]. Besides the free-form message part in RFC 3164 and the structured data in RFC 5424, RFC 3165 [34] defines how to encapsulate syslog messages within XML and sends them by means of TCP.

C. Exchange Protocols

In this section we introduce the security event exchange protocols. First, we describe the development and what content the exchange protocol is intended to transmit. Second, we describe to which OSI layer the exchange protocol belongs to and what kind of security mechanisms the protocol ensures.

1) Common Intrusion Detection Framework: The Common Intrusion Detection Framework (CIDF) was developed by DARPA to share information of IDAR systems. CIDF messages are sent over one of four transport mechanisms, which can be negotiated depending on the needs of the components in question. The following types of transport/messaging are supported: (i) CIDF message layer without acknowledgement and re-transmission directly over UDP, (ii) CIDF message layer with reliable delivery (acknowledgement, re-transmission, and duplicate removal) over UDP, (iii) CIDF message layer directly over TCP and (iv) CIDF operations over CORBA [35]. The default transport layer protocol for CIDF messages is reliable CIDF messaging over UDP on port 3295. The CIDF ensures privacy with the use of symmetric encryption algorithms. Besides the transport mechanism and the use of symmetric encryption algorithms, CIDF requires, that each device must possess its own certificate or certificate chain in a X.509v3 format. In 1998 the DARPA Information Assurance community decided to join forces with the IETF [17] and stopped further development of CIDF.

2) Real-time Inter-network Defense: In 2012 the Managed Incident Lightweight Exchange (MILE) WG standardized the initial work on Real-time Inter-network Defense (RID). RID and its transport mechanism (RID-T) outlines a proactive inter-network communication method to facilitate sharing incident-handling data while integrating existing detection, tracing, source identification, and mitigation mechanisms for a complete incident handling solution [36]. RID adds exchange semantics to IODEF messages intended for the cooperative handling of security incidents within consortia of network operators and enterprises [37] and provides a peer-to-peer exchange model. RID is an application-layer protocol that passes incident-handling data over HTTP/TLS. In addition, RID uses XML security features of encryption and digital signatures such as XMLencrypt and XMLsig.

D. Extensible Messaging and Presence Protocol

The Extensible Messaging and Presence Protocol (XMPP) is a real-time communication technology used for instant messaging collaboration and generalized routing of XML data. The specification XEP-0268: Incident Handling defines methods for incident reporting among XMPP server deployments using the IODEF format [22]. XEP-0268 is an application-layer protocol that uses TLS to transmit incident messages over network. XEP-0268 was developed by the IETF's INCH WG and has been deferred [21], [22].

1) Intrusion Detection Exchange Protocol: In 1998 the IETF IDWG started to develop a high-level requirements document, an intrusion detection message format and message transport protocols [17] to define a common way to exchange Intrusion Detection Messages between intrusion detection analyzers and managers across a TCP/IP network. The Intrusion Detection Message Exchange Requirements (called IDP) were published in RFC 4766 [38]. Out of these requirements, the Intrusion Alert Protocol (IAP) was developed [39]. IAP was the first attempt to meet IDP [40] and uses a request-response model of communication similar to HTTP 1.1 [39].

In December 2000 the IDWG decided to look into using Block Extensible Exchange Protocol (BEEP) as basis of the message transport protocol to overcome security related limitation of IAP rather than completing the work on IAP.

BEEP, previously known as BXXP [41] is an Internet Standard from the IETF and a generalized application-level protocol framework. The advantages of BEEP is that it takes care of the connections, the authentication, and the packaging up at the TCP/IP level of the messages. BEEP competes on the same level as HTTP [42].

In a second attempt the IDWG developed the Intrusion Detection Exchange Protocol (IDXP) that is specified as a Block Extensible Exchange Protocol (BEEP) profile. In 2007 the application-layer protocol IDXP was published as RFC 4767 [43]. IDXP transmits data of the following MIME types: text/xml, text/plain and application/octet-stream.

2) CEE Log Transport Protocol: The CEE Log Transport (CLT) [44] was developed by The MITRE Corporation in collaboration with the NATO Consultation, Command and Control Agency (NC3A), information technology (IT) vendors and IT administrators. The CLT defines requirements that a CLT protocol must meet to transmit CEE messages. Therefore CLT also defines transport mappings. Since Syslog is the de facto standard in log transport protocols and it is supported by numerous products, the only mapping published to transmit CEE messages is a specification for sending JSON-encoded events over the RFC 5425 syslog protocol.

3) Simple Mail Transfer Protocol: In 1982 Jonathan Postel developed the Simple Mail Transfer Protocol (SMTP) that was published in RFC 821. Over the years, the RFC of SMTP was updated several times. The new Draft Standard of SMTP is published in RFC 5321 [45]. SMTP is an application layer protocol to transfer mail messages over TCP. The content of a mail message includes a message header and a possibly structured message body. If the body of a mail message is structured, it is defined according to the Multipurpose Internet Mail Extension (MIME) Part One [46]. SMTP is inherently insecure because real mail security lies only in end-to-end methods involving the message bodies, such as digital signatures (Multipart/Signed and Multipart/Encrypted in RFC 1847, Pretty Good Privacy in RFC 4880 or Secure/Multipurpose Internet Mail Extensions (S/MIME) in RFC 3851) [45]. SMTP is used to transmit messages of CAIF, ARF and x-arf. The secure transmission and storage of CAIF, ARF or x-arf messages is outside the scope of these format specifications. When creating CAIF, ARF or x-arf messages or when interchanging the XML/JSON source of documents between sender and receiver, mechanisms such as canonical XML, digital signatures or methods for encryption have to be implemented, that assure the integrity and authenticity of the documents.

4) Syslog protocol: The syslog messages defined in RFC 3164 [31] are using the syslog protocol also published in RFC 3164. This syslog protocol uses the User Datagram Protocol (UDP), port 514, for communication. Therefore this syslog protocol does not guarantee a message delivery and is not encrypted. Further, the syslog protocol described in RFC 3164 is vulnerable to replay attacks. As the format of syslog was updated and the RFC 5424 was published, the syslog protocol was also updated. Unlike RFC 3164 the syslog protocol was published in a separate RFC, called RFC 5425 [47]: Syslog over UDP and RFC 5426 [48]: Syslog over TLS.

III. Evaluation

In this section, we evaluate the exchange formats and protocols. First, we describe the characteristics of the evaluation criteria. Furthermore, we introduce 10 evaluation criteria for the exchange formats and 7 evaluation criteria for the exchange protocols. Last, we present and summarize the results of the evaluation.

A. Evaluation methodology

The exchange formats are evaluated based on the following ten criteria: flow-based format interoperability, extensibility, scalability, aggregability, protocol independency, machine readability, human readability, integrity & authenticity, confidentiality and practical application. These criteria were chosen to provide a means of comparison between exchange formats and to measure if and how the use of this exchange format makes a large amount of security events manageable for network operators. In this evaluation, we assess the flow-based format interoperability by using a Cisco NetFlow v5 record. The findings are also valid for other flow-based accounting technologies, e.g., IPFIX.

The ability to use the exchange format in conjunction with flow-based data is described by the criterion flow-based format 'interoperability'. 'Extensibility' defines the aptitude of the exchange format to extend its functionality. This feature refers to the use of additional information or fields, and development of particular data structure in order to fit to one's needs. The ability to handle growing data volumes is described by the 'scalability' criterion. An exchange format is called scalable when it is capable of working properly with different volumes of data. 'Aggregability' defines the aptitude to aggregate single flows within the exchange formats. 'Protocol independency' describes the ability to transport the exchange format by using a protocol already existing in the infrastructure without adaption in case of an existing complementary protocol. 'Machine readability' describes whether a format is machine readable or not, whereas the 'human readability' criterion describes the human readability of each exchange format. The use of digital signatures within a message of an exchange format is described by the criterion 'Integrity & Authenticity'. The 'Confidentiality' criterion describes the ability to encrypt a message of each exchange format. The last criterion 'practical application' indicates if there are products available that make use of the exchange formats.

The exchange protocols are evaluated based on the following seven criteria: confidentiality, integrity, authenticity, reliable message transport, interoperability, scalablity and practical application. The criteria to evaluate an exchange protocol are part of the well-known model for security policy development CIA Triad.

The 'Confidentiality' criterion describes the capability to protect sensitive data from disclosure to unauthorized parties. The ability to determine whether the communication peers are, in fact who they declared to be and thus make use of authentication mechanisms is described by the 'Authenticity'

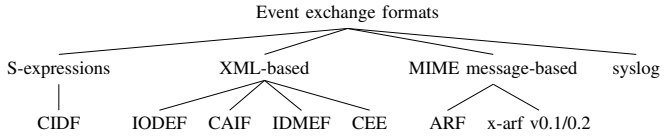

Fig. 2. Classification of the exchange formats

criterion.The ability to protect data from modification or deletion by unauthorized parties is described by the 'Integrity' criterion. An exchange protocol is using 'Reliable message transport' when the transmission of a security event uses a guaranteed delivery to avoid duplicate messages and message loss. The 'Interoperability' criterion describes the ability to transmit different data representations of security events. The ability to handle different network sizes is described by the 'scalability' criterion. The last criterion 'practical application' indicates if the exchange protocol is used within the network.

B. Evaluation results

In this section, we present and discuss the results of the evaluation of the security event exchange formats and protocols.

1) Evaluation of exchange formats: The security event exchange formats can be categorized into four groups: S-expressions, XML-based and MIMEMessage-based formats and syslog. This categorization facilitates the qualitative evaluation of the exchange formats based on common characteristics and is visualized in Figure 2.

Interoperability: CIDF uses nouns and verbs to encode security events, but there are no nouns and verb available to describe flow-based data. Due to this reason the interoperabilty of CIDF is low. To use Cisco NetFlow v5 in conjunction with XML-based formats the NetFlow record fields are passed to the structure of an appropriate XML-element. In the XML-based formats IODEF, CAIF and IDMEF, only a few NetFlow record fields fit into the defined data representation. Therefore, the flow-based format interoperability of the XML-based formats IODEF, CAIF and IDMEF in conjunction with NetFlow is low. The XML-based exchange format CEE, the MIMEMessage based formats ARF and x-arf, and both versions of syslog provide a free-form message part and thus result in a high interoperability.

Extensibility: The use of CIDF in conjunction with flow-based data requires a new definition of nouns and verbs and thus result in a high extensibility. The XML-based formats IODEF, CAIF and IDMEF use the `<AdditionalData>` structure to store all remaining NetFlow record fields. Another approach is to define an XML-schema to store flow-based data into the exchange formats IODEF, CAIF and IDMEF and thus results in a high extensibility. The MIME-based message formats, the XML-based format CEE and syslog do not have any additional fields to map NetFlow into their structure. Instead, the whole NetFlow record has to be attached as separate MIMEBodyPart or in a container of a free-form message. Thus the extensibility of the exchange formats ARF, x-arf and syslog is high.

Scalability: It is possible to map large input files into the structure of CIDF, the XML-based formats IODEF, CAIF, IDMEF and CEE and the MIMEbased formats ARF and x-arf. The main disadvantage of this approach is, that the output files will be as large as the input files or even larger. Thus the scalability of the exchange formats in general is low.

Aggregability: The exchange formats CIDF, IODEF, ARF, CEE, x-arfv0.1 and both versions of syslog do not provide any field or container to enter data of more than one security event. Therefore the aggregability is considered as low. CAIF and x-arfv0.2 address more than one security event within a single document and thus result in a high aggregability. IDMEF does not provide a field or container to enter more than one security event, but provides an XML element `CorrelationAlert` to group one or more previously sent security events together.

Protocol independency: CIDF uses CLT and a special architecture to transmit security events. Thus, the independency of CIDF is low. On the one hand the XML-based formats IODEF, IDMEF and CEE provide their own exchange protocol as a first choice and on the other hand these formats can be enveloped into a user-defined packet for transmission over network. Due to the reason that the designated exchange protocols are often not available in the network the independency of the XML-based formats result in a medium independency. CAIF does not define a specific protocol to exchange a CAIF message and thus can be enveloped into a user-defined packet for transmission. The protocol independency of CAIF is high. ARF and x-arf use a MIMEMessage format and thus must use SMTP for transmission over network. Besides ARF and x-arf, syslog also relies on an infrastructure that is available in most of the companies. The independency of the exchange formats ARF, x-arf and syslog results in a high independency.

Human readability: The exchange formats CIDF, IODEF, CAIF and IDMEF are hardly human readable. A parser is required to extract the requested values out of the underlying structure. Therefore the human readability of the exchange formats CIDF, IODEF, CAIF and IDMEF is low. The exchange formats ARF, x-arf, CEE and syslog consist of a human-readable container to store security events. Due to this container the human-readability of the exchange formats ARF, x-arf, CEE and syslog is high.

Machine readability: Except the exchange format syslog RFC 3164, all exchange formats are machine readable. A parser is used to extract the requested values out of the underlying structure. The exchange format syslog RFC 3164 can not be parsed unambiguously. One reason is that syslog RFC 3164 does not define a structure of a syslog message and does not define the character set or encoding for syslog messages. Therefore the machine readability of syslog RFC 3164 is low.

Confidentiality, Integrity and Authenticity: Except the exchange format x-arfv0.2, none of the exchange formats provide any security mechanisms to sign or encrypt a security event. x-arfv0.2 provides the ability to use Secure/Multipurpose Internet Mail Extensions (SMIME) for public key encryption and signing the security event.

Practical application: Due to the fact that the work on CIDF and CEE is presently stopped, no practical applications are available that make use of CIDF or CEE. The exchange formats IODEF, CAIF, IDEMF, ARF and x-arf are used in several applications (e. g. IODEF: DFLabs products; CAIF: RUS-CERT Stuttgart University, CERT-VW Volkswagen AG;

TABLE II. EVALUATION SUMMARY OF THE EXCHANGE FORMATS

Criterion	CIDF	IODEF	CAIF	IDMEF	ARF	CEE	X-ARF		Syslog	
							v0.1	v0.2	RFC 3164	RFC 5425
Interoperability	−	−	−	−	+	+	+	+	+	+
Extensibility	+	+	+	+	+	+	+	+	+	+
Scalability	−	−	−	−	−	−	−	−	−	−
Aggregability	−	−	+	0	−	−	−	+	−	−
Protocol independency	−	0	+	0	+	0	+	+	+	+
Human readability	−	−	−	−	+	+	+	+	+	+
Machine readability	+	+	+	+	+	+	+	+	−	+
Integrity & Authenticity	−	−	−	−	−	−	−	+	−	−
Confidentiality	−	−	−	−	−	−	−	+	−	−
Practical application	−	0	0	0	0	−	0	0	+	+

Legend: high (+), medium (0) and low (−)

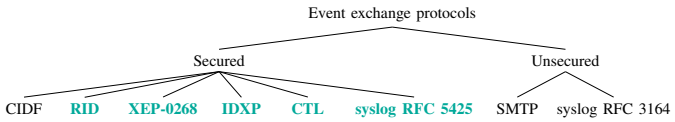

Fig. 3. Classification of the exchange protocols

IDEMF: Snort, Prelude; ARF: arffilter and x-arf: members of the Association of the German Internet Industry (eco)). Syslog is the most used exchange format.

2) Evaluation of exchange protocols: The security event exchange protocols can be categorized into the two groups: Secured and unsecured. This categorization facilitates the qualitative evaluation of the exchange protocols based on common characteristics. Figure 3 shows the categorization of the security event exchange protocols, whereas the colored and bold formatted protocols make use of TLS.

Confidentiality, Integrity and Authenticity: CIDF provides a matchmaking service through which CIDF components use authenticated and secured communications between CIDF components. The matchmakerservice also acts as a Certificate Authority (CA). The CIDF messages can include authentication headers and can be encrypted [49]. Therefore the authentication, the confidentiality and the integrity capabilities of the exchange protocol CIDF are high. RID uses digital signatures on a hash of the RID message with an X.509v3 certificate issued by a trusted party to authenticate the sender of an RID message. XEP-0268 is based on XMPP. As XMPP uses the Simple Authentication and Security Layer (SASL) for peer authentication, the authentication capabilities of XEP-0268 is considered as high. IDXP uses a mutual authentication of public-key certificates of the underlying BEEP security profile TLS. The syslog RFC 5425 uses mutually authenticated channels with widely deployed cryptographic algorithms and protocols. As CLT is basically syslog RFC 5424 it has the same characteristics. The authentication capability of the exchange protocol RID, IDXP, CLT and syslog RFC 5425 is high. The exchange protocols SMTP and syslog RFC 3164 do not provide the ability to verify that the declared sender is actually the peer who sent the message and thus the authentication capability of SMTP and syslog RFC 3164 is low.

RID uses XML encryption to provide confidentiality based on the `iodef:restriction attribute` in the IODEF and IODEF-RID schemes [36]. IDXP uses the underlying BEEP security profile of TLS with the TLS_DHE_DSS_WITH_3DES_EDE_CBC_SHA cipher suite to provide confidentiality and thus result in a high authentication capabilities. XMPP uses TLS to ensure the confidentiality and thus the confidentiality capability of XMPP and XEP-0268 is regarded as high. SMTP and syslog RFC 3164 do not make use of any mechanism to ensure confidentiality. Therefore the confidentiality capability is low. CLT and syslog RFC 5424 are using TLS to counter disclosure of the security events and thus the confidentiality capability is high.

The integrity of messages transmitted over the network is addressed through the use of authentication mechanisms. The TLS-based protocols RID, XEP-0268, IDXP, CTL and syslog RFC 5424 support message integrity through the use of message authentication of TLS, BEEP security profile TLS or members of the SASL profile. The capability to ensure integrity of the TLS-based exchange protocols RID, XEP-0368, IDXP, CTL and syslog RFC 5424 is high. In contrast, SMTP and syslog RFC 3164 do not provide any mechanisms to ensure integrity and thus the integrity capability is low.

Reliable-message transport: The default transport layer protocol for CIDF messages is reliable CIDF messaging over UDP. But CIDF also uses (i) message layer without acknowledgement and retransmission directly over UDP, (ii) the CIDF message layer with acknowledgement and retransmission over UDP, and (iii) the CIDF message layer directly over TCP. RID, XEP-0268 and IDXP require the use of connection-oriented protocol and thus make use of TCP. SMTP is independent of the particular transmission protocol but requires a reliable ordered data stream channel. Most companies are using SMTP over TCP. CLT and syslog RFC 5425 transmit syslog messages over a reliable channel over UDP. Therefore capability of a reliable message transport of the exchange protocols CIDF, RID, XEP-0268, IDXP, SMTP, CLT and syslog RFC 5425 is high. Syslog RFC 3164 relies on the simple UDP protocol and thus results in a low reliable message transport.

Interoperability: The exchange protocols RID, IDXP SMTP and both syslog versions provide the capability to transmit various exchange messages of different formats. The ability to transmit different data representations of security events of the exchange protocols RID, IDXP SMTP and both syslog version is high. The extension XEP-0268 was developed to transmit IODEF messages. Therefore only IODEF messages

can be transmitted via XEP-0268 and thus result in a low interoperability. CIDF and CLT can only be used to transmit CIDF messages and CEE messages and thus result in a low interoperability.

Scalability: All exchange protocols except RID were developed to interoperate and share data. Therefore the scalabilty of CIDF, IDXP, XEP-0268, SMTP, CLT and both syslog versions is high. Due to the fact that RID is a point-to-point exchange protocol it does not scale well. Therefore the ability to handle different network sizes is low. Similary to CIDF, XEP:0268

Practical application: CIDF was not intended as a standard that would influence the commercial marketplace. Due to the fact that CIDF was a research project and the work on CEE and XEP-0268 is presently stopped no practical applications are available that make use of CIDF, CTL and XEP-0268. Therefore the availability of the practical application is low. There are some implementations available that make use of RID and IDXP and thus the practical application is medium. Most companies make use of SMTP and syslog in both versions and therefore the practical application is high.

IV. Evaluation summary

In this section, we provide an aggregated overview of the key evaluation results to support network operators to identify an exchange format/protocol that can be used in their network. We have summarized the information presented in section III in Table II and III.

We found that the use of flow-based data within the XML-based exchange formats IODEF, CAIF and IDMEF requires a new XML scheme or the AdditionalData element needs to be used. The MIME-Message based formats transmit the same information multiple times without providing new knowledge.

Most of the exchange formats are machine readable. This ensures that security events can be handled automatically. In case the network operator focuses on machine readability, all exchange formats except syslog RFC 3164 are suitable. If a network operator focuses on an exchange format that is human readable and does not require an additional parser, the network operator should focus on the MIME based exchange formats ARF and x-arf. But also CEE and syslog might be of interest, as they provide an free form message part. However, CEE has never been finalized, so actions should first be taken into this direction before the exchange format can be used in practice.

With respect to the interoperability with flow-based data, the exchange formats ARF, CEE, x-arf and syslog are suitable to use in conjunction with flow-based data. To exchange sensitive data, however, the network operator might focus on mechanisms to sign or encrypt an security event. Except the exchange format x-arf v0.2, none of the exchange formats provide mechanisms to sign or encrypt an security event. Finally, we note that to establish a collaboration between exchange peers, a well-known and established format should be used. Even though, a lot of exchange formats were published in the last years, only the exchange format syslog provides a widespread use.

To transmit a security event, the network operator might focus on a high-security level. The exchange protocols CIDF, RID, XEP-0268, IDXP, CLT and syslog RFC 5425 provide a high-security level. SMTP and syslog RFC 3164 were not designed to ensure the four key aspects of information security (confidentiality, integrity, authenticity and non-repudiation). However, sysolg and SMTP have the advantage that they are widely spread. Therefore we identify here a tradeoff between security level and easiness to deploy. Even though SMTP has never been updated, the use of the S/MIME standard provides the ability to digitally sign messages and to encrypt message contents to overcome these missing security aspects. In case a network operator focuses on an exchange protocol that should be used in high-speed networks, all exchange protocols are suitable except RID. RID does not scale in high-speed networks because it was designed as point-to-point protocol.

TABLE III. Evaluation summary of the exchange protocols

Criterion	CIDF	RID	XEP-0268	IDXP	SMTP	CLT	Syslog	
							RFC 3164	RFC 5425
Confidentiality	+	+	+	+	−	+	−	+
Authenticity	+	+	+	+	−	+	−	+
Integrity	+	+	+	+	−	+	−	+
Reliable message transport	+	+	+	+	+	+	−	+
Interoperability	−	+	−	+	+	−	+	+
Scalability	+	−	+	+	+	+	+	+
Practical application	−	0	−	0	+	−	+	+

Legend: high (+), medium (0) and low (−)

V. Conclusion

In this paper we identified 10 exchange formats and 7 exchange protocols used in context of intrusion detection and incident management and assigned them to the different stages of the Operations Management Process of The Mitre Cooperation. We provided a structured overview of the exchange formats and protocols that can be used by network operators to decide what format and protocol should be used in their environment to exchange security events. We analyzed the interoperability of the exchange formats with flow-based data and evaluated and compared the formats and protocols in context of high-speed networks.

To conclude, it is still a challenge to find a standardized exchange format and protocol that is thoroughly validated and tested in full scale of industry trials. Future work should focus on the development of a standardized exchange format and protocol. However, a widespread use of these formats and protocols remain to be established in the community of network operators.

Acknowledgment

This work was partly supported by the German Federal Ministry of Education and Research under grant number 16BY1201F (iAID) and by CASED. In addition it is funded by FLAMINGO, a Network of Excellence project (318488) supported by the European Commission under its Seventh Framework Programme.

References

[1] Symantec Corporation, "Internet Security Threat Report," April 2013, http://www.symantec.com/threatreport/.

[2] S. Greengard, "The war against botnets," *Commun. ACM*, vol. 55, no. 2, Feb. 2012.

[3] S. S. Silva, R. M. Silva, R. C. Pinto, and R. M. Salles, "Botnets: A survey," *Elsevier B.V. Computer Networks*, vol. 57, no. 2, 2013.

[4] M. van Eeten, J. M. Bauer, H. Asghari, S. Tabatabaie, and D. Rand, "The Role of Internet Service Providers in Botnet Mitigation: An Empirical Analysis Based on Spam Data," in *The Tenth Workshop on the Economics of Information Security*, 2010.

[5] D. Anstee, D. Bussiere, G. Sockrider, and C. Morales, "Worldwide Infrastructure Security Report," Arbor Networks Inc., Tech. Rep. VIII, Jan. 2013, http://www.arbornetworks.com/research/infrastructure-security-report.

[6] J. François, I. Aib, and R. Boutaba, "Firecol: A collaborative protection network for the detection of flooding DDoS attacks," *Networking, IEEE/ACM Transactions on*, vol. 20, no. 6, Dec 2012.

[7] J. Steinberger, L. Schehlmann, S. Abt, and H. Baier, "Anomaly Detection and Mitigation at Internet Scale: A Survey," in *Emerging Management Mechanisms for the Future Internet*. Springer Berlin Heidelberg, 2013, vol. 7943.

[8] R. Koch, M. Golling, and G. Dreo, "Evaluation of state of the art ids message exchange protocols," *International Conference on Communication and Network Security (CNS 2013)*, 2013.

[9] P. Kampanakis, "Security automation and threat information-sharing options," *Security Privacy, IEEE*, vol. 12, no. 5, Sept 2014.

[10] Martin, Robert A., "Making Security Measurable and Manageable," Jul. 2013. [Online]. Available: http://msm.mitre.org/about/Making_Security_Measurable_and_Manageable.pdf

[11] J. Postel, "Internet Protocol," RFC 791 (Standard), IETF, Sep. 1981.

[12] M. Kerrisk, *The Linux Programming Interface* - . München: No Starch Press, 2010.

[13] "Information technology – Security techniques – Information security risk management," ISO/IEC 27005:2011, International Organization for Standardization, 2011.

[14] P. Cichonski, T. Millar, T. Grance, and K. Scarfone, "Computer Security Incident Handling Guide," Special Publication 800-61, NIST, Aug. 2012.

[15] B. Tung, "Common Intrusion Detection Framework," http://gost.isi.edu/cidf/, 1999.

[16] D. Schnackenberg, K. Djahandari, and D. Sterne, "Infrastructure for intrusion detection and response," in *DARPA Information Survivability Conference and Exposition, 2000. DISCEX '00. Proceedings*, vol. 2, 2000, pp. 3–11 vol.2.

[17] J. Betser, A. Walther, M. Erlinger, T. Buchheim, B. Feinstein, G. Matthews, R. Pollock, and K. Levitt, "GlobalGuard: creating the IETF-IDWG Intrusion Alert Protocol (IAP)," in *DARPA Information Survivability Conference Exposition II DISCEX Proceedings*, vol. 1, 2001.

[18] S. Staniford-Chen, "Intrusion Detection FAQ: What open standards exist for Intrusion Detection? ," Feb. 2008. [Online]. Available: http://www.sans.org/security-resources/idfaq/id_standards.php

[19] L. Shields and W. Heinbockel, "Common Event Expression," 2009, NIST's 5th Annual IT Security Automation Conference and Expo on October 27. [Online]. Available: http://cee.mitre.org/docs/NIST_SCAP_CEE_Briefing.pdf

[20] R. Danyliw, J. Meijer, and Y. Demchenko, "The Incident Object Description Exchange Format RFC 5070 (Proposed Standard)," IETF, Dec. 2007.

[21] E. W. Burger, M. D. Goodman, P. Kampanakis, and K. A. Zhu, "Taxonomy Model for Cyber Threat Intelligence Information Exchange Technologies," in *Proceedings of the 2014 ACM Workshop on Information Sharing & Collaborative Security*. New York, USA: ACM, 2014.

[22] A. Hefczyc, F. Jensen, M. Rémond, P. Saint-Andre, and M. Wild, "XEP-0268: Incident Handling," 2012.

[23] RUS-CERT University of Stuttgart, "CAIF - Common Announcement Interchange Format," http://www.caif.info/, 2004.

[24] H. Debar, D. Curry, and B. Feinstein, "The Intrusion Detection Message Exchange Format (IDMEF) RFC 4765 (Experimental)," IETF, Mar. 2007.

[25] The MITRE Corporation, "Common Event Expression," May 2013. [Online]. Available: https://cee.mitre.org/

[26] J. Falk and M. Kucherawy, "Battling Spam: The Evolution of Mail Feedback Loops," *Internet Computing, IEEE*, vol. 14, no. 6, nov 2010.

[27] J. Levine, "ARF is Now an IETF Standard," http://www.circleid.com/posts/20100901_arf_is_now_an_ietf_standard/, September 2010.

[28] Y. Shafranovich, J. Levine, and M. Kucherawy, "An Extensible Format for Email Feedback Reports RFC 5965 (Proposed Standard)," IETF, Aug. 2010.

[29] eco — Verband der deutschen Internetwirtschaft e.V., "Network Incident Reporting: Provider verstärken Zusammenarbeit," http://www.eco.de/2012/pressemeldungen/network-incident-reporting-provider-verstaerken-zusammenarbeit.html, 2012.

[30] J. Kohlrausch and S. Übelacker, "X-ARF: A Reporting and Exchange Format for the Data Exchange of Netflow and Honeypot Data," http://www.geant.net/Media_Centre/Media_Library/Media%20Library/xarf_geant_milestone2.pdf, 2011.

[31] C. Lonvick, "The BSD Syslog Protocol," RFC 3164 (Informational), IETF, Aug. 2001.

[32] R. Gerhards, "The Syslog Protocol," RFC 5424 (Proposed Standard), IETF, Mar. 2009.

[33] Assuria Ltd, "In Syslog we trust," Mar. 2012. [Online]. Available: http://www.assuria.com/uploads/In%20Syslog%20we%20trust.pdf

[34] D. Levi and J. Schoenwaelder, "Definitions of Managed Objects for the Delegation of Management Scripts," RFC 3165 (Proposed Standard), IETF, Aug. 2001.

[35] C. Kahn, D. Bolinger, and D. Schnackenberg, "Communication in the Common Intrusion Detection Framework," Jun. 1998. [Online]. Available: http://gost.isi.edu/cidf/drafts/communication.txt

[36] K. Moriarty, "Real-time Inter-network Defense (RID)," RFC 6545 (Proposed Standard), IETF, Apr. 2012.

[37] B. Trammell, "Transport of Real-time Inter-network Defense (RID) Messages over HTTP/TLS," RFC 6546 (Proposed Standard), IETF, Apr. 2012.

[38] M. Wood and M. Erlinger, "Intrusion Detection Message Exchange Requirements," RFC 4766 (Informational), IETF, Mar. 2007.

[39] D. Gupta, T. Buchheim, B. Feinstein, G. Matthews, and R. Pollock, "IAP: Intrusion Alert Protocol Internet-Draft," http://tools.ietf.org/html/draft-ietf-idwg-iap-05, IETF, Sep. 2001.

[40] T. Buchheim, M. Erlinger, B. Feinstein, G. Matthews, R. Pollock, J. Betser, and A. Walther, "Implementing the intrusion detection exchange protocol," in *Computer Security Applications Conference, 2001. ACSAC 2001. Proceedings 17th Annual*, Dec 2001.

[41] M. Rose, "The Blocks Extensible Exchange Protocol Core," RFC 3080 (Proposed Standard), IETF, Mar. 2001.

[42] E. Dumbill, "XML Watch: Bird's-eye BEEP," Dec. 2001. [Online]. Available: http://www.ibm.com/developerworks/webservices/library/x-beep/

[43] B. Feinstein and G. Matthews, "The Intrusion Detection Exchange Protocol (IDXP)," RFC 4767 (Experimental), IETF, Mar. 2007.

[44] The MITRE Corporation, "CEE Log Transport (CLT) Specification," Aug. 2012. [Online]. Available: https://cee.mitre.org/language/1.0-beta1/clt.html

[45] J. Klensin, "Simple Mail Transfer Protocol," RFC 5321 (Draft Standard), IETF, Oct. 2008.

[46] N. Freed and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies," RFC 2045 (Draft Standard), IETF, Nov. 1996.

[47] F. Miao, Y. Ma, and J. Salowey, "Transport Layer Security (TLS) Transport Mapping for Syslog," RFC 5425 (Proposed Standard), IETF, Mar. 2009.

[48] A. Okmianski, "Transmission of Syslog Messages over UDP," RFC 5426 (Proposed Standard), IETF, Mar. 2009.

[49] W. Lee, R. A. Nimbalkar, K. K. Yee, S. B. Patil, P. H. Desai, T. T. Tran, and S. J. Stolfo, "A Data Mining and CIDF Based Approach for Detecting Novel and Distributed Intrusions," in *Recent Advances in Intrusion Detection*. Springer Berlin Heidelberg, 2000, vol. 1907, pp. 49–65.