Collaborative Intrusion Detection Framework: Characteristics, Adversarial Opportunities and Countermeasures

Rainer Bye
DAI-Labor, TU Berlin, Germany

Seyit Ahmet Camtepe
DAI-Labor, TU Berlin, Germany

Sahin Albayrak
DAI-Labor, TU Berlin, Germany

Abstract

Complex Internet attacks may come from multiple sources, and target multiple networks and technologies. Nevertheless, Collaborative Intrusion Detection Systems (CIDS) emerges as a promising solution by using information from multiple sources to gain a better understanding of objective and impact of complex Internet attacks. CIDS also help to cope with classical problems of Intrusion Detection Systems (IDS) such as zero-day attacks, high false alarm rates and architectural challenges, e. g., centralized designs exposing the Single-Point-of-Failure. Improved complexity on the other hand gives raise to new exploitation opportunities for adversaries.

The contribution of this paper is twofold. We first investigate related research on CIDS to identify the common building blocks and to understand vulnerabilities of the Collaborative Intrusion Detection Framework (CIDF). Second, we focus on the problem of anonymity preservation in a decentralized intrusion detection related message exchange scheme. We use techniques from design theory to provide multi-path peer-to-peer communication scheme where the adversary can not perform better than guessing randomly the originator of an alert message.

1 Introduction

Information and Communication Technology permeates daily life in our modern society. We are increasingly dependant on IT infrastructures and services which promotes complex Internet attacks either because of commercial [31] or political reasons [16]. Intrusion Detection and Preventions systems (IDS/IPS) provide measures against attacks, but the evolution to more complex and pervasive networked IT also provides increasing opportunities for the adversary. In this regard, there has also been evolution in the research on intrusion detection towards collaborative IDS.

The first research in Intrusion Detection Systems (IDS) took place in the 1980s focussing on statistical anomaly detection and expert systems [4, 15] in a local fashion. In the 90s, IDS were started to be realized in a distributed manner. The first example was the DIDS, the Distributed Intrusion Detection System based on the centralized architecture that is still prevalent in many operationally used IDS [42]. Distributed sensors, so called LAN-managers, provide the data to a central analysis component, the DIDS-director. In this fashion, the director aggregates the data and creates a centralized view of the decentralized system. This architecture exhibits a Single-Point-of-Failure to attackers. The NIDES (Next-Generation IDES), a further development of IDES, uses a similar architecture as DIDS while not considering data reduction via aggregation [2]. In this regard, the central component may become in addition a bottleneck for sensor data processing in the overall system.

The next step in IDS development was the introduction of hierarchical, multi-layered approaches: EMERALD (Event Monitoring Enabling Responses to Anomalous Live Disturbances)[37]. In this manner, scalability issues of prior works were targeted. GrIDS, (Graph-based IDS) [14] went a step further: here, graphs based on network activity were created.

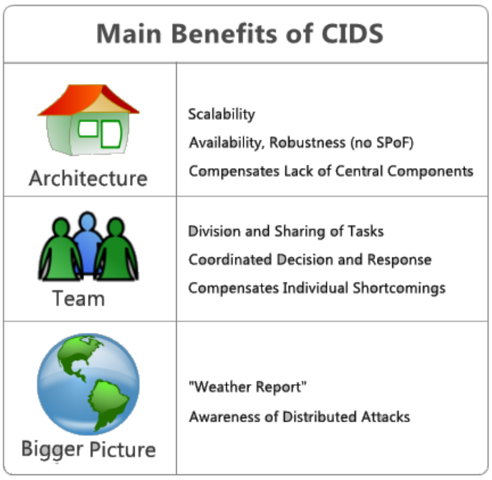
Since 2000, there have appeared many works discussing and incorporating collaborative aspects in two different ways: first, the advent of new distributed system paradigms such as agents, Peer-to-Peer or to smaller extent grid-based systems enabled the IDS to solve existing problems, e.g. Single-Point-of-Failure or scalability. Next, new application domains for IDS evolved such as internet-wide working IDS or IDS for mobile (ad-hoc) networks. Recently, collaborative IDS have been discussed in the context of Future Internet activities [55].

The main advantages of collaboration for intrusion detection and prevention are depicted in Figure 1. The architectural benefits include the scalability of solutions as well as robustness and availability, e.g. by means of the absence of the Single-Point-of-Failure. CIDS are also able to compensate lack of central components, e.g. in the case of MANETs (Mobile ad hoc networks) there are no centralized entities to rely on.

The second main advantage is teamwork: known from organizational studies, the team represents nowadays the atomic unit to solve complex problems. Tasks can be shared, e.g. load balancing, and can be solved by directing them to the most capable member. This approach can also compensate the shortcomings of individuals, e.g. an agent capable of misuse detection collaborates with an agent capable of anomaly detection. In addition, coordinated decision, e.g. voting or a joint detection status, as well as coordinated response for fast containment of malicious activity is enabled.

Third, the “Bigger Picture” is realized by collaborating monitors. This allows the awareness for distributed attacks. Also, a “Weather Report” can be derived, i.e. the state of the network/Internet.

The research on Intrusion Detection Systems (IDS) can be decomposed in two main areas, the *detection aspect* and the *system* respectively *framework aspect*. Existing works can be organized contributing exclusively to one or both areas. In this research, we consider *CIDF* (Collaborative Intrusion Detection Framework), the system aspect of the Collaborative Intrusion Detection System (CIDS). Recently, there have shown up many works on IDS discussing collaborative respectively cooperative aspects. However, this property of an intrusion detection approach has to the best of our knowledge not properly been defined yet. However, this is a necessary prerequisite defining CIDF in an appropriate manner. The application of collaborative techniques also enables new opportunities for the adversary. These need further investigation in the scope of a CIDF. The challenges are to be discussed intensively to understand the underlying research problems that may be solved with existing approaches or require new solutions to confront the adversary.



Main Benefits of CIDS	
Architecture	Scalability Availability, Robustness (no SPoF) Compensates Lack of Central Components
Team	Division and Sharing of Tasks Coordinated Decision and Response Compensates Individual Shortcomings
Bigger Picture	"Weather Report" Awareness of Distributed Attacks

Figure 1: CIDS provide advantages in three main areas

1.1 Contribution and Organization

The contribution of this paper is twofold. First, we investigate related research on CIDS to identify the common building blocks creating the Collaborative Intrusion Detection Framework (CIDF) and to understand its vulnerabilities. We investigate key aspects for collaborative intrusion detection framework (*Communication Scheme*, *Organizational Structure*, *Group Formation*, *Information Sharing and Interoperability*, *System Security*) and discuss relevant works in the field with respect to them.

Second, we focus on the problem of anonymity preservation in a decentralized intrusion detection related message exchange scheme where we try to hide the ID of the alert-originating IDS from compromised IDS. We use techniques, namely Symmetric BIBD and Generalized Quadrangles, from design theory to provide multipath peer-to-peer communication scheme where adversary can not perform better than trying to randomly guess the originating IDS. We perform overhead analysis for the scheme using the metrics maximum hop count and total message count to communicate an alert messages to all IDS nodes in a collaborating group. We quantify the trade-off between multi-path design scheme, hop-count and number of the messages.

Organization of the paper is as follows: in Section 1.2, we introduce the relevant terminology. In Section 2, the Collaborative Intrusion Detection Framework model is presented with the relevant works in the field. In Section 3, vulnerabilities of CIDF and adversarial opportunities are discussed. In Section 4, we present our scheme for anonymity preservation in a decentralized intrusion detection related message exchange. We conclude in Section 5. In the appendix, we provid background information on the techniques we employed from the design theory.

1.2 Terms and Definitions

Technical terms in a research field need to be clearly defined; but, to the best of our knowledge, a clear definition of *Collaborative* respectively *Cooperative Intrusion Detection Systems* does not yet exist and in most cases the terms are even used synonymously.

Collaboration and cooperation have been investigated in different research areas such as Community Research [24], Business Logistics ([29]) or Organizational Studies [46]. Winer and Ray define work coupling activities as cooperative if the individuals retain their authority, resources are not shared and the relationship is informal. In contrast, collaboration forms a new entity having the authority, characterized by commitment to common mission, comprehensive planning, well-established communication channels and a considerable investment of resources. The reward of collaboration is to reach "solutions that go beyond their [remark of the author: the people's] own limited vision of what is possible".

Hence, we define *Cooperative Intrusion Detection System* as a distributed system, where participants exchange intrusion detection related information for intrusion detection and prevention. In such a system, the individual entities can work on their own, and the outcome of cooperation is creating additional benefits, but not completely new opportunities.

In contrast, a *Collaborative Intrusion Detection System* is a dynamic, distributed system where participants form new organizational structures such as teams and can adapt to different roles to fulfill a common task not solvable by a participant on its own, i.e. the result must substantially differ from the individual functionality.

We define a CIDF, *Collaborative Intrusion Detection Framework*, as the set of mechanisms, to enable Collaborative Intrusion Detection for a given detection/correlation algorithm. Throughout the paper, we use the term agent for the participant of a CIDF, and a (detection) group as a subset of agents administrated by a CIDF following the same objective, e.g. "misuse detection" or "anomaly detection". These agents can share (detection) messages to collaborate for intrusion detection or prevention.

2 Framework for Collaboration

Research and development on CIDS requires clear understanding of the related works and building blocks of a CIDS approach which we call "pillars". In this regard, we present the building blocks of CIDF in Section 2.1. In Section 2.2 we define a CIDF system as a detection group and give an example on CIDF based on a priori work in Section 2.3.

2.1 Pillars of CIDF

Zhou *et al.* made a first attempt to categorize CIDS according to the system topology, either decentralized, hierarchical or centralized [54]. In contrast, we consider the topic to be more complex and deem the following differentiation appropriate. According to the requirements by Winer and Ray, we define key aspects for the CIDF denoted in Figure 2: Communication Scheme (well-established communication channels), Organizational Structure and Group Formation (comprehensive planning and commitment to common mission) and Information Sharing and Interoperability (resource investment and sharing). In addition, the System Security is an important characteristic owed to the application domain. We focus on the framework aspect, however for further reading on the topic of distributed detection we refer to the dedicated work of Xu and Ling on alert correlation [49], or surveys dedicated to detection mechanisms such as anomaly detection [13] or the application of computational intelligence to intrusion detection [47].

2.1.1 Communication Scheme (A)

Since components of a CIDS are distributed across a network, they need to communicate with each other. In this regard, the fundamental requirement for "well-established communication channels" is dealt with in various ways: DOMINO system uses an unspecified Peer-to-Peer overlay architecture of IDS "axis" agents to exchange black lists of IP addresses [51]. Each axis agent is the root agent of a hierarchical IDS. The INDRA system, uses the structured Peer-to-Peer protocol Pastry to exchange intrusion information [25]. In this regard, the application-level multi-cast mechanism SCRIBE is used [12]. LarSID also uses a Publish/Suscribe mechanism for evidence exchange, in particular suspicious source IP addresses [53]. In previous works, we employed a custom Peer-to-Peer protocol to realize a cooperative AIS (Artificial Immune System) approach [34].

Gorodetsky *et al.* discuss the combination of the Peer-to-Peer and the agent paradigm. This is realized via a three layer architecture comprised of a Peer-to-Peer provider, a Peer-to-Peer agent platform and agent services on top, e.g. negotiation or service matcher [23]. Gopalakrishna and Spafford present an agent-based approach, where agents propagate and escalate intrusion related information to other interested agents [22] in a Publish/Suscribe fashion. A subset of agent-based IDS are using mobile agents capable of migration, e.g. Xiao *et al.* [48], where mobile agents realize a collaborative voting mechanism for coordinated reaction. Agent-based IDS are found frequently in the literature targeting MANETs such as surveyed by Anantvalee *et al.* [44].

There also exist specialized middlewares, e.g. xmlBlaster[1] or specialized algorithms for communication purposes. Garcia-Alfaro *et al.* propose application of RSS- and xmlBlaster for distributed exchange of alerts [21]. Gamer *et al.* [18] discuss the application of expanding ring search or path-coupled mechanisms for neighborhood discovery. Publish/Suscribe is the most prevalent mechanism, either be realized by Peer-to-Peer , agents or (other) specialized middleware. Its application is further discussed in the section on group formation.

[1] http://www.xmlblaster.org

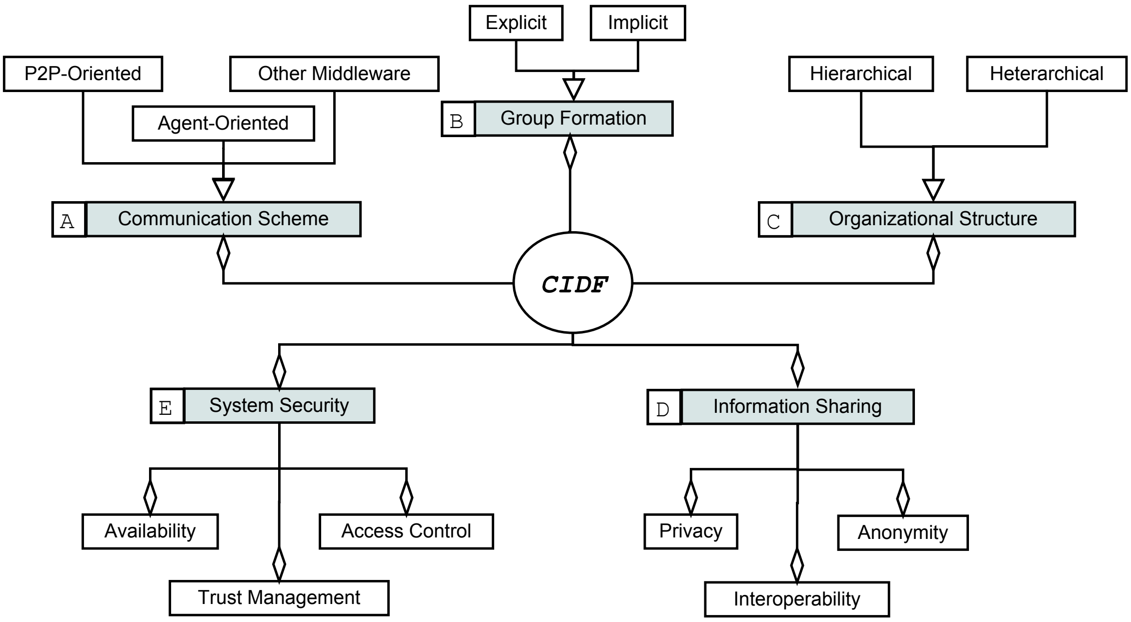


Figure 2: Pillars of CIDF: boxes connected to CIDF represent the building blocks of Collaborative Intrusion Detection Framework. The semantics of the connectors are as following: a triangle represents exclusive (sub)-classification, whereas the rhombus denotes non-exclusive aspects, constituting the high-level concepts.

2.1.2 Group Formation (B)

The creation of a team plays an essential role in the process of collaboration. To pursue a joint objective, potential collaborators need to define and agree on a "common mission". The grouping enables the exchange of relevant objective data between only the interested parties. This reduces communication overhead, but also allows to focus on particular objectives by creating specialized teams. To this end, the specification of a grouping strategy garners the benefit from both special abilities and particular knowledge of the group members. Katti *et al.* demonstrated the value of groups in CIDS by evaluating real data from multiple sources such as DSHIELD, as well as logs from universities and network providers [28]. The creation of small *correlation groups*, i.e. four to six IDS, enable similar results as collaborating with 1700 agents. In the literature, group formation is either addressed explicitly or implicitly in a used scheme.

Grouping concepts may be inherent to a used communication protocol, e.g. multicast, Publish/Suscribe, Peer-to-Peer etc. Basicevic *et al.* discuss the application of Publish/Suscribe for signature base updates and alert notification [6]. The characteristics of the participating IDS agents are taken into account, as signature bases on agents are only updated if an installed application is subject to a corresponding vulnerability. Another example for implicit grouping is carried out by Zhou *et al.* [53]; in this work, the number of subscriptions for a suspicious activity is used for detection, i.e. groups are created for a particular activity.

Explicitly, Janakiraman *et al.* discuss agents following a common interest, e.g. "failed log-in attempts" [25]. Vlachos *et al.* use JXTA [2] for peer groups exchanging incident information. Luther *et al.* discuss the creation of homogeneous groups for cooperative anomaly detection to reduce the false-positive rate[34]. The CIMD approach (c.f. Section 2.3) provides sematic group formation taking IDS characteristics, objectives and associated interests into account.

2.1.3 Organizational Structure (C)

The aforementioned "Comprehensive Planning" as well as the operation of collaborative IDS require an organizational structure, either hierarchical or heterarchical. In a heterarchical system, all participating agents do have the same competence, but it remains possible that agents are given special authority for a limited time frame to fulfill finite tasks [25, 53, 34, 21, 23].

In contrast, hierarchical systems rely continuously on agents with more competencies; typically there exist one topmost agent, and for the agents of each layer of hierarchy, there exist at least one agent superior in authority. Special forms are centralized systems, with only one root and a two level hierarchy, or polyhierarchical sys-

[2] https://jxta.dev.java.net/

tems, where a agent, except the topmost, is subordinate to more than one parent agent [19, 51, 6, 10, 1].

2.1.4 Information Sharing (D)

Information sharing can take place on different levels. On the most basic level, raw sensor data can be shared. On a second level, components for detection and analysis provide processed data to be exchanged such as a statistical analysis or detection status. Finally, knowledge, e.g. signatures, IP addresses of suspicious sources or even ontological domain models are exchanged. There are two types of publishing architectures discussed in the literature: the data contributor/ data repository model or a completely decentralized CIDS; this is similar to the discussion in Section 2.1.3 of organization. We consider three important aspects of Information Sharing: *Privacy*, *Anonymity* and *Interoperability*.

Georgios *et al.* discuss *privacy*-preservation by means of a rule-based access control scheme, using an ontological model to decide on the fly whether access to data is granted or not [33]. Xu *et al.* present a technique for consistent, prefix-preserving IP address anonymization and also discuss attacks on such schemes in general [50]. Koukis *et al.* present a tool for the flexible anonymization of packet traces [30]. Lincoln *et al.* discuss privacy-preserving techniques for exchange and correlation of security alerts[32]. For the sake of *anonymity*, Lincoln *et al.* also discuss a randomized alert routing following the work of Reiter *et al.* for anonymity in web transactions [32, 38]. Here, source anonymity is preserved to some degree, but this approach differs from the CIDF scenario, where all agents in a detection group are supposed to receive the detection messages.

The interaction of collaborators from distinct IDS poses various requirements. Dependent on the level of collaboration, these include a common language, protocol or even a complete frameworks. There exist a variety of exchange formats, prominent examples include IDMEF and IODEF. The main intention of IDMEF *Intrusion Detection Message Exchange Format* (RFC 4765) is to provide a communication standard enabling different intrusion detection analyzers from different origin (commercial, open source and research systems) to report to a managing entity in one administrative domain. The IDMEF is mentioned in several works as either be used or considered for future works [51, 20, 52]. In contrast, the objective of IODEF, *Incident Object Description Exchange Format* (RFC 5070), also an XML-based format, is the exchange of incident reports between different CSERT (*Computer Security Emergency Response Teams*) in different administrative domains. In contrast to IDMEF, IOEDF provides extension strategies to prevent changes of the XML schema for the sake of interoperability.

2.1.5 System Security (E)

Owed to the application domain, security of the CIDS itself plays an important role. For CIDF we consider *Trust Management*, *Access Control* and *Availability* as relevant topics. The other to classic properties of information security, confidentiality and integrity manifest in different aspects such as encryption protocols in the underlying communication scheme or access control mechanisms.

The dynamics of a system, e.g. benign agents becoming malicious during runtime, can be dealt with *collaborative trust management* in contrast to a-priori fixed trust assignments with a PKI. The collaborative trust management has emerged as an important research topic on its own and is studied for various application domains. We refer to Artz and Gil [5] and the Trust Management works in [39] for further details. Exemplarily, a dynamic component based on feedback, as presented by Kamvar *et al.* [26], can be used to reflect the dynamics of a system.

In the CIDS literature, most schemes try to counter the adversary by rigorous *access control* mechanisms. The prototypical *Indra* version uses central key servers, but in the opinion of the author, the *Web of Trust*- approach is better suited for a decentralized peer-to-peer system [25]. The DOMINO system uses public-key cryptography for the authentication of the exchanged messages. Zhou *et al.* propose to use PKI as well [53], so that participants of LarSID (Large Scale Intrusion Detection) framework report about intrusion intelligence authenticated. A mutual authentication scheme is used and data is transmitted via SSL to assure integrity. Ganame *et al.* in the DSOC approach use certificates to protect communications between DSOC components, but do not discus the issuing entity [19]. There are schemes discussing access control to be realized by means of an underlying agent framework or the integration of group key management protocols such as presented in [10]. Basicevic *et al.* propose to integrate their proposed Publish/Suscribe mechanism directly into the IDXP protocol that itself used BEEP (c.f. Section 2.1.2) profiles to benefit from the inherent security functionality [6].

The *availability* of the system is highly affected by attacks, such as DDoS, or exploitation of protocol flaws. In overlay networks, the selective disabling of agents or compromise of routing schemes are important problems. Fiat and Saia present a censor-resistant peer-to-peer network that sustains the breakdown of up to 50% of the participating nodes [17]. Kapadia *et al.* discuss reliable resource look-up in structured Peer-to-Peer networks by means of redundant searches. The authors provide a scheme reducing substantially the failure rate in the look-up process [27]. We further elaborate on the system security in Section 3.

2.2 CIDF System Definition

We consider a Collaborative Intrusion Detection Framework as an environment, where each participating agent $a \in A$ is given the opportunity to follow one or more objectives $O_i \in \mathcal{P}(O)$, where O is the set of objectives, for intrusion detection, e.g. "IP Blacklist Exchange", "Anomaly Detection" or "Signature Exchange". Hence, a group with a given maximum membership size k, a common objective o and additional restrictions R is defined as a 4-tuple $G(k, o, R) = (A_G, o, E_G, R)$, with $A_G \in \mathcal{P}_k(A)$ and E_G a set of overlay links between the group members. The additional constraint set respectively properties of groups is used, e.g. by the CIMD approach in the following. The consequence of a maximum group size and further constraints are that for one objective there may exist more than one group. Exchanged messages are either asynchronous, i.e. event-based messages, e.g. "Attack Detected from x" or synchronous, i.e. periodic messages, e.g. "Detection Status y".

2.3 CIMD: A Realization of CIDF

In prior work, we introduced the CIMD (Collaborative Intrusion & Malware Detection) approach [8, 9], a realization of CIDF. CIMD offers a scheme for the formation of detection groups based on an overlay network, including a collaboration model and a decentralized group formation algorithm.

Every agent in an overlay network should be able to express its interest regarding collaboration partners. In the CIMD architecture, those interests are expressed using a collaboration ontology for the specification of potential collaboration partners in the look-up phase, but also for the description of the agents themselves. This model reflects security relevant characteristics such as the operating system and supported applications, the network and hardware configuration as well as detection capabilities.

The algorithm for the group formation assumes to have an overlay network providing search capabilities. The algorithm performs the grouping of devices connected to an overlay structure such as discussed in Section 2.1.1. The CIMD approach enables the semantic group formation for agents to exchange intrusion detection related information according to a common objective in combination with the definition of associated interests.

In detail, CIMD adds an additional concept to the model for CIDF introduced in Section 2.2. Each agent $a \in A$ provides a property model p^a, and can associate to each followed objective $o \in O$ a set of interests $c^a_{o1}..c^a_{om}$, i.e. an agent can have more than one interest associated with one objective. The device property model and the interests are instances of the aforementioned collaboration ontology. This enables the matching of the property source against interests to select appropriate groups. This results in $G(k, o, R) = (A_G, o, E_G, R)$ with $R = \{\{p^i, c^j_{ok}\}\}$ for $i, j \in \{1..n\}, |A_G| = n$ and $c^j_{ok} \in \{c^j_{o1}..c^j_{om}\}$.

As an illustrative example, we outline a case study presented in [9]: Three different IDS manufacturers are selling NIDS appliances. These systems are capable of detecting a known malware by stored signatures provided centrally by their corresponding manufacturers. Accordingly, exclusively detecting known threats leaves the customer vulnerable to zero-day attacks and other unknown threats. As a result, the vulnerability window needs to be minimized. The companies provide updates about new attacks independent from each other.

In a non-collaborative scenario, the NIDS are working on their own, whereas in the collaborative scenario the heterogeneous detection groups are built with the purpose to mediate upcoming signatures between the appliances to benefit from the updates of other manufacturers. Hence, groups can be created following the common objective of "signature exchange". However, the property base contains information about the manufacturer, device ID and type, and the associated interests to join groups with a different manufacturer enables the creation of a heterogeneous detection group. In an additional scenario, we discussed the incorporation of dedicated signature generators into a group.

3 Challenges and Adversarial Opportunities in CIDF

At this point, we would like to highlight three adversarial opportunities in the scope of CIDF based on the literature we examined so far.

First, the adversary may gather critical information about the IDS agents in a CIDF using active or passive attack techniques, e.g. by means of the *probe-response attack*. Probe-response attacks are based on crafted packets containing a unique mark, e.g. seldom used port number, and are supposed to be detected [32]. The defenders report reveals different opportunities for exploitation: first, it indicates that an IP address is monitored. In addition, further analysis can reveal the detection capabilities and consequently type of detection system used or running services. In the same context, Shinoda *et al.* discuss vulnerabilities of threat monitors [40] and exploitation of the feedback of detecting attacks visible in public data repositories. They also present exploitation algorithm to determine sensor locations and discuss further analysis such as fingerprinting devices types or inferring network topologies. Similarly, Bethencourt *et al.* propose an algorithm for probe-response attack to determine monitored IP address regions locations [7].

The exploited knowledge is of high value when devising further attack strategies, e.g avoiding well protected networks, finding vulnerable unprotected systems etc. Porras and Shmatikov discussed research challenges of large-scale collection and sanitization from data provider and data repository point of view. In this work, the authors emphasize the importance of anonymous data delivery, because it makes fingerprinting and probe-response attacks more difficult [36]. In this regard, anonymity and privacy-preserving data exchange help to hide such information: (i) if adversary manages to bypass the access control mechanisms in CIDF, or (ii) if the data is publicly available, e.g. a public threat repository such as DSHIELD[3] or CAIDA network telescope research [4].

Next, the adversary may want to compromise the detection scheme if she can collect aforementioned critical information about the CIDF. The benefits for the adversary are that she can hide her own activities or even use a detection scheme itself as an attack tool. As an example, we refer to the illustrative example of Section 2.3. If a malicious signature generator injects "HTTP/1.1 200 OK" for misuse detection, this would block every web server response resulting in a Denial-of-Service (DoS). Hence, an adversary may be capable of compromising a detection scheme with the help of fake messages. Finally, the adversary may want to disable the overall CIDF, preventing it from realizing the entire collaborative intrusion detection approach. This can be achieved by the means of a Distributed Denial-of-Service attack (DDoS) on the underlying overlay network.

Other attacks on overlay networks, supporting the adversarys objectives include the Eclipse and the Sybil attack. In the Eclipse attack, a small sub set of compromised nodes is capable of modifying the routes in the overlay network, so that the benign nodes are eclipsed. In this regard, the overlay links may be manipulated and the adversary may become a direct contact for all intrusion detection related message exchange. With the help of the Sybil attack, the adversary introduces multiple pseudonym identities of herself into the system. In the scope of a CIDF, such an attack may be used to compromise the detection approaches by sending false detection messages from not just one, but multiple virtual nodes. We do not consider these attacks in this work, but we refer to (i) Peng *et al.* surveying detection of and countermeasures against DDoS attacks and Urdaneta *et al.* surveying countermeasures to attacks on DHT-based Peer-to-Peer systems such as Eclipse or Sybil attack [35, 45].

4 Design of Resilient CIMD

4.1 Problem Specification

We consider a collaborative detection group, that has been created by a CIDF such as CIMD. The resulting collaborating group $G(k, o, \emptyset) = (A_G, o, E_G, \emptyset)$ has $|A_G| = N$ IDS nodes (a.k.a. agents) with a common objective o and $|E_G| = m$ peer-to-peer overlay links connecting them. Each IDS can be linked with a subset of IDS (due to source anonymity discussed below). IDS nodes want to exchange intrusion detection related messages by the means of the overlay links, either (i) event-based or (ii) interval-based.

There may exist one or more adversaries in the group, capable of listening to all the exchanged messages. These adversarial agents can store and analyze every message in a group with the corresponding sender and collaborate among each other. In this regard, the adversary can create profiles of IDS, i.e. determining the capabilities or locations in terms of the subnets that they are responsible for. The adversaries have access to this information either because it is a public group or an agent becomes compromised through a trojan or similar attacks.

The general problem can be defined as preserving the source anonymity by hiding the source information for the initiator of a message in a group, while still making sure that all other group members are informed about the message timely. In this work, we focus on the special problem of probe-response attacks (c.f. Section 3).

4.2 Adversarial Models

We assume that the adversary can compromise IDS nodes in a group of collaborating peers. In this regard, we consider adversarial agents capable of eavesdropping a subset of compromised agents $A_{Gc} \subset A_G$, participating in the detection group. However, the adversaries can not inject messages by themselves into the group.

The adversarial objective is to locate the initiator of the alarm messages related to the probe-response attack. In this regard, a simple heuristic is followed: the adversary assumes, that the first node she receives an alert from, related to the probe-response attack, is the initiator of the detection alert.

This heuristic can work, (i) if one of the adversaries is located within one-hop-neighborhood of the initiator and (ii) if the initiator selects the adversary to be contacted at first rather than picking another node from its neighborhood.

[3] www.dshield.org

[4] http://www.caida.org/research/security/telescope/

4.3 Combinatorial Design of Resilient CIMD

In this work we are interested in preserving the source anonymity by hiding the source information for the initiator of a message in a group. We use techniques, namely symmetric BIBD and generalized quadrangle, from design theory to decide which IDS nodes is to be linked which other IDS nodes so that an adversary can get limited information regarding the source of an alert message. Thus, we target the stage after selection of group $G(k, o, \emptyset) = (A_G, o, \emptyset, \emptyset)$ as introduced in Section 2.3. We define a peer-to-peer link design scheme ($E_G = \{e_1, ..e_m\}$), randomized and delayed alert initiation scheme, and alert propagation scheme for preserving the source anonymity.

```
Input  : $N$ Total number of IDS agents
         $IDS_i$ ID of the IDS agent
         $Algorithm$ SBIBD, GQ
Output: Initial Block Assignment $B_i$ for all $IDS_i$
         participating in the group
begin
    if (Algorithm == SBIBD) then
        Generate $(v, k, \lambda)$-Design where
        $v = N = q^2 + q + 1$ and $k = q + 1$
        $IDS_i$ selects the block $B_i$ ($|B_i| = k$)
    else if (Algorithm = GQ) then
        Generate $GQ(q, q)$-Design where
        $N = q^3 + q^2 + q + 1$ and $k = q + 1$
        $IDS_i$ selects the block $B_i$ ($|B_i| = k$)

end
```

Algorithm 1: Initial Block assignment reflecting the overlay link destinations for each IDS.

Shmatikov *et al.* [41] presented an approach for alert propagation by using a DHT-based Peer-to-Peer mechanism together with a probabilistic propagation scheme. The DHT has the advantage of logarithmic and deterministic look-up operation. We also discussed the application of DHT as being valuable for the group formation process in [9]. Probabilistic propagation in [41] assumes that adversaries can not compromise group members but can eavesdrop a link. The presented scheme provides two distinct paths between alert initiator and each other IDS nodes. Each sent message is encoded with a key so that the adversary can not obtain the message unless she is monitoring both paths.

For alert propagation, we assume that the adversary can compromise number of IDS nodes using techniques such as trojan attacks. We provide an overlay structure between the group members which (i) provides multiple paths between any pair of IDS nodes, (ii) has a good connectivity, (iii) helps preserving anonymity to a certain degree. We are using Combinatorial Design techniques in determining the overlay links between the members of the group, E_G. The Combinatorial Design enables the construction of sets whose intersections have beneficial properties. Camtepe and Yener [11] applied these techniques successfully to key distribution schemes in wireless sensor networks. We apply similar techniques, Symmetric Balance Incomplete Block Design (SBIBD) and Generalized Quadrangles (GQ) as detailed in the Appendix, to generate sets (a.k.a. blocks) of IDS node identities. Each IDS receives one such set which simply tells which other IDS nodes to be contacted when there is an alert to be initiated or propagated. We consider three stages as described below.

Group Organization: In the first phase, the overlay links between the members of the group need to be defined and established. Therefore, Algorithm 1 is applied for each individual peer. Here, each IDS_i decides on a list of node ID B_i for $|B_i| = k$ by using either SBIBD or GQ approach. The approach to be used is determined in the group formation phase. The outcome of the algorithm for all IDS_i is the initial Block assignment for all IDS_i participating in the group. Both SBIBD and GQ design assign v objects, i.e. agents in the overlay, into b blocks, i.e. contact lists, ($v = b = q^2 + q + 1$ in SBIBD and $v = b = q^3 + q^2 + q + 1$ in $GQ(q, q)$) so that every pair of blocks has exactly one common object in SBIBD, and at most one common object in the GQ design. The mapping of the node IDs in the blocks to IP addresses can be realized in a public fashion, e.g. store them in a DHT with a common group key or requesting them from the group members in the group formation.

```
Input: $B$ Assigned Block
       $e$ Attack Event
       $t$ Time Span to wait
       $IDList$ List of Alert IDs
$targetList$ = B
$id_e$ = GeneratePseudoID($e$)
Add($id_e$, $IDList$)
foreach $i = 1$ to $|B|$ do
    target = Random(ReceiverList)
    Remove(target, targetList)
    SendMessage($id_e$, $m_e$, target)
    Sleep($t$)
end
```

Algorithm 2: Initial propagation of alerts

Alert Generation: In case of an alert, Algorithm 2 is used. At first, a pseudo ID is generated for an alert message, e.g. by a hash function taking message and time stamp as parameters. This ID is necessary for the termination of the message propagation and stored in the variable $IDList$. In addition, a temporary variable $targetList$ is initialized containing all the block entries. Then, an agent from $targetList$ is chosen randomly, and the alert as well as the ID sent to it. To ensure the agent is not contacted again, it is removed from the temporary list and after a specific time span, the next agent from the list is contacted. The length of time span can be decided using maximum round trip time in the network.

Alert Propagation: Similarly, messages are forwarded upon receipt. This is shown in Algorithm 3. For the sake of terminating the message propagation, it is first checked whether this message has already been send based on the stored ids. If not, the message is propagated in the same manner as in Algorithm 2.

```
Input: $B$ Assigned Block
       $e$ Attack Event
       $t$ Time Span to wait
       $IDList$ List of Alert IDs
       $m$ Received Message
m = Receive(message)
if GetID($m$) $\notin$ IDList then
    Add($id_e$, $IDList$)
    $targetList$ = B
    foreach $i = 1$ to $|B|$ do
        target = Random(ReceiverList)
        Remove(target, targetList)
        SendMessage($id_e$,$m_e$, target)
        Sleep($t$)
    end
end
```

Algorithm 3: Further propagation of alerts on message receipt

4.4 Evaluation

We consider three different criteria: (i) anonymity, (ii) reliability of message exchange and (iii) communication overhead. In both approaches, SBIBD and GQ, the ID list has exactly $q+1$ entries, where in SBIBD the overall number of nodes is $N = q^2 + q + 1$ and in GQ $N = q^3 + q^2 + q + 1$. In this regard, the ratio between number of nodes in the ID list and the overall number of nodes is $\frac{q+1}{q^2+q+1} \approx \frac{1}{\sqrt{N}}$ for SBIBD and $\frac{q+1}{q^3+q^2+q+1} \approx \frac{1}{\sqrt[3]{N}}$ for GQ.

Anonymity: Reiter *et al.* introduced in their work anonymity concepts of "beyond suspicion" and "probable innocence" [38]. Beyond suspicion denotes that a sender "is no more likely to be the originator of that message than any other potential sender in the system". In contrast, probable innocence is a weaker concept, that considered from the attackers perspective "the sender appears no more likely to be the originator than to not be the originator".

Theorem 1 *Resilient CIMD scheme described in Section 4 provides anonymity "beyond suspicion".*

We look at the probability that: (i) at least one of the compromised IDS is in the first hop neighborhood of the source (alert initiating) IDS (a.k.a. in the ID list of the source IDS) and (ii) one of the compromised IDS is picked as the first node to receive the alert message. Only under these conditions, the adversary can correctly estimate the source IDS. In all other cases, due to random and delayed alert propagation, the adversary fails to correctly estimate source. Thus, we will show that the above mentioned probability is equivalent to the probability that an IDS can be the source ($\frac{1}{N}$). That means, an adversary can not perform better than guessing an IDS to be the source, or a sender "is no more likely to be the originator of that message than any other potential sender in the system".

Single Adversary: Each node is listed in $q+1$ ID lists and hence the probability for an source IDS to have a block with a particular node is $\frac{q+1}{N}$. In addition, there exist $q+1$ opportunities a node can be selected as an initial message receiver. Thus, probability that a compromised IDS is in the first hop neighborhood of the source IDS, and it is picked as the first node is $\frac{(q+1)}{N}\frac{1}{(q+1)} = \frac{1}{N}$.

Two Adversaries: Compromised nodes can be in distinct ID lists or in the same ID list. Due to the characteristics of SBIBD, each compromised node appears in $q+1$ ID lists: there exists exactly one ID list containing both adversaries, and each compromised node appears in q ID lists alone. Thus, probability that a compromised IDS is in the first hop neighborhood of the source IDS, and it is picked as the first node is $\frac{(2q)}{N}\frac{1}{(q+1)} + \frac{1}{N}\frac{2}{(q+1)} = \frac{2}{N}$. This probability is equivalent to the sum of independent guesses of two adversaries.

Consider the following example of SBIBD: Let $N = \{1, 2, 3, 4, 5, 6, 7\}$, There also exist 7 ID lists of size $q + 1 = 3$: $\{1, 2, 3\}$, $\{1, 4, 5\}$, $\{1, 6, 7\}$, $\{2, 4, 6\}$, $\{2, 5, 7\}$, $\{3, 4, 7\}$, $\{3, 5, 6\}$. Let 1 and 4 be the adversaries, there exists one ID list containing both entries, $\{1, 4, 5\}$ and for each adversary, there exist two more lists each adversary appears alone:$\{1, 2, 3\}$ $\{1, 6, 7\}$ and $\{2, 4, 6\}$ $\{3, 4, 7\}$.

m Adversaries: In the general case, when there exist m adversaries, we have $\frac{m(q-m+2)}{N}$ disjoint contact lists. There exist at most $\binom{m}{2}$ different contact lists with two entries. Hence, the generalized formula for the probability to contact an adversary in the first hop is:

$$\frac{(q-m+2)}{N(q+1)} + \binom{m}{2}\frac{2}{N(q+1)}. \quad (1)$$

$$\frac{m(q-m+2)+(m(m-1)/2*2)}{N(q+1)} = \frac{m}{N}. \quad (2)$$

Hence, the probability that a compromised IDS is in the first hop neighborhood of the source IDS, and it is picked as the first node is $\frac{m}{N}$. This probability is equivalent to the sum of independent guesses of m adversaries. Similar discussions follows for GQ-design.

Message Exchange: Every node in our scheme has overlay links to $q+1$ other nodes, where in SBIBD $N = q^2+q+1$ and in GQ $N = q^3+q^2+q+1$. We consider the worst case where each IDS node has its own ID in its ID list. Thus, source IDS can send alert message to q other IDS. Each of these IDS can propagate the alert messages at most q others. Thus, maximum number hops required for an alert message to reach all IDS nodes will be:

$$\lceil \log_q N \rceil. \quad (3)$$

$$\lceil \log_q q^2+q+1 \rceil = 3 \; for \; SBIBD. \quad (4)$$

$$\lceil \log_q q^3+q^2+q+1 \rceil = 4 \; for \; GQ. \quad (5)$$

Communication and Computational Costs: Each IDS_i generates the blocks in SBIBD or GQ. SBIBD can be constructed in $O(N^{3/2})$ and GQ designs can be constructed in $O(N^2)$ (c.f. Section 5). In both SBIBD and GQ scheme, size of ID list is $(q+1)$. Thus, $(q+1)$ messages are sent in the overlay, resulting maximum of $(N-1)*(q+1)$ messages ($O(q^3)$ messages for SBIBD and $O(q^4)$ messages for GQ). The termination condition prevent the agents from relaying the same alert more than once to the same receiver. In comparison, in a simple not anonymous broadcast, $N-1$ messages would have to be sent by the initial propagator. The approach discussed by Shmatikov *et al.* considers anonymization to be realized via disjoint routing on a DHT [41]. The authors do not target exactly the same problem as we do, but in order to contact all nodes in a DHT, every message would need to be routed over $log(n)$ hops to each destination. However, the authors also introduce a probabilistic component which does not guarantee reliable message exchange. The minimum number of messages to be exchanged would be $2*(N-1)log(n)$. The factor 2 applies as the authors consider two disjoint routes for every message.

Evaluation results identifies a trade-off between design technique, maximum hop count and communication overhead. For the same group size of $N = O(q^3)$, BIBD provides better connectivity and lower maximum hop count of 3 at the cost of increased communication overhead of $O(q^{9/2})$ messages. For the same group size, GQ has maximum hop count of 4 at the cost of communication overhead of $O(q^4)$ messages.

The drawback of the design scheme, is that design parameter q has to be a prime power. Also after design phase, the overall number of nodes is fixed or a new design has to be applied. Our solution is unaffected, as we consider the design to be used after group has been created. In this regard, the number of members does not fluctuate significantly.

5 Conclusion

In this work, we have investigated the common building blocks defining the Collaborative Intrusion Detection Framework (CIDF). In this regard, we identified five relevant building blocks that need to be covered when doing research on CIDS. Motivated by the adversarial opportunities in a CIDF, we have realized a scheme for anonymity preservation in a decentralized intrusion detection related message exchange scheme. We used SBIBD and GQ from design theory to guarantee that alert messages are send to all other nodes with a fixed hop count. Moreover, the solution ensures the probability that the single adversary correctly estimates the source of the alert is not more ID than $\frac{1}{N}$.

We show that our scheme provides anonymity "beyond suspicion". This implies that the adversary can not perform better than guessing an IDS to be the source, or a sender "is no more likely to be the originator of that message than any other potential sender in the system". Finally, evaluation results show us the trade-off between design technique, maximum hop count and communication overhead.

References

[1] Ajith Abraham, Ravi Jain, Johnson Thomas, and Sang Yong Han. D-SCIDS: Distributed Soft Computing Intrusion Detection System. *J. Netw. Comput. Appl.*, 30(1):81–98, Oct 2007.

[2] Debra Anderson, Teresa Lunt, Harold Javitz, Ann Tamaru, and Alfonso Valdes. Next-generation Intrusion Detection Expert System (NIDES): A summary. Technical report, Computer Science Laboratory, 1995.

[3] I. Anderson. *Combinatorial designs: construction methods*. Ellis Horwood Limited, 1990.

[4] James P. Anderson. Computer security threat - monitoring and surveillance. Technical report, Fort Washington, 1980.

[5] Donovan Artz and Yolanda Gil. A survey of trust in computer science and the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2):58–71, 2007.

[6] I. Basicevic, M. Popovic, and V. Kovacevic. Use of publisher-subscriber design pattern in infrastructure of distributed ids systems. In *Third International Conference on Networking and Services, ICNS.*, pages 56–56, June 2007.

[7] John Bethencourt, Jason Franklin, and Mary Vernon. Mapping internet sensors with probe response attacks. In *SSYM'05: Proceedings of the 14th conference on USENIX Security Symposium*, Berkeley, CA, USA, 2005. USENIX Association.

[8] Rainer Bye, Ahmet Camtepe, and Sahin Albayrak. *Collaborative Computer Security and Trust Management*, chapter Teamworking for Security: The Collaborative Approach, pages 12–33. Reference. Information Science Reference, 1 edition, November 2009.

[9] Rainer Bye, Seyit A. Camtepe, and Sahin Albayrak. Design and modeling of collaboration architecture for security. *International Symposium on Collaborative Technologies and Systems*, pages 330–341, 2009.

[10] Yu Cai. A distributed autonomous intrusion detection framework. In *IEEE Globecom Workshops*, pages 1–5, Nov. 2007.

[11] S. A. Camtepe and B. Yener. Combinatorial design of key distribution mechanisms for wireless sensor networks. *IEEE/ACM Transactions on Networking*, 15(2):346–358, April 2007.

[12] M. Castro, P. Druschel, A. M. Kermarrec, and A. I. T. Rowstron. Scribe: a large-scale and decentralized application-level multicast infrastructure. *Selected Areas in Communications, IEEE Journal on*, 20(8):1489–1499, 2002.

[13] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3):1–58, 2009.

[14] Steven Cheung, Rick Crawford, Mark Dilger, Jeremy Frank, Jim Hoagland, Karl Levitt, Jeff Rowe, Stuart Staniford-Chen, Raymond Yip, and Dan Zerkle. The design of GrIDS:a graph-based intrusion detection system. Technical report, Department of Computer Science, University of California at Davis, 1999.

[15] Dorothy E. Denning. An intrusion-detection model. *IEEE Transactions On Software Engineering*, 13(2):222–232, 1987.

[16] Gadi Evron. Battling botnets and online mobs. *Georgetown Journal of International Affairs*, 9(1):121–126, 2008.

[17] Amos Fiat and Jared Saia. Censorship resistant peer-to-peer content addressable networks. In *SODA '02: Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 94–103, Philadelphia, PA, USA, January 2002. Society for Industrial and Applied Mathematics.

[18] Thomas Gamer, Michael Scharf, and Marcus Schöller. Collaborative anomaly-based attack detection. In *Proceedings of 2nd International Workshop on Self-Organizing Systems (IWSOS 2007)*, Lecture Notes in Computer Science, pages 280–287, English Lake District, UK, September 2007. Springer.

[19] Abdoul Karim Ganame, Julien Bourgeois, Renaud Bidou, and François Spies. A global security architecture for intrusion detection on computer networks. In *IPDPS'07, Proc. of the ACM/IEEE Int. Parallel and Distributed Processing Symposium*, pages 1–8, Long Beach, California USA, March 2007. IEEE computer society press.

[20] Joaquin García, Michael A. Jaeger, Gero Muehl, and Joan Borrell. Decoupling components of an attack prevention system using publish/subscribe. In *Intelligence in Communication Systems*, volume 190 of *IFIP International Federation for Information Processing*, pages 87–97. Springer Boston, 2005.

[21] J. Garcia-Alfaro, M. A. Jaeger, G. Muehl, I. Barrera, and J. Borrell. Distributed exchange of alerts for the detection of coordinated attacks. *Communication Networks and Services Research, Annual Conference on*, pages 96–103, 2008.

[22] Rajeev Gopalakrishna and Eugene H. Spafford. A framework for distributed intrusion detection using interest driven cooperating agents. In *International Symposium on Recent Advances in Intrusion Detection (RAID)*, Purdue University, West Lafayette, USA, 2001.

[23] Vladimir Gorodetsky, Oleg Karsaev, Vladimir Samoylov, and Sergey Serebryakov. Multi-agent peer-to-peer intrusion detection. In *Communications in Computer and Information Science*, 2009.

[24] Teresa Hogue. Community based collaborations wellness multiplied. Technical report, Oregon Center for Community Leadership, 1994.

[25] Ramaprabhu Janakiraman, Marcel Waldvogel, and Qi Zhang. Indra: A peer-to-peer approach to network intrusion detection and prevention. In *WETICE '03: Proceedings of the Twelfth International Workshop on Enabling Technologies*, Washington, DC, USA, 2003. IEEE Computer Society.

[26] Sepandar D. Kamvar, Mario T. Schlosser, and Hector Garcia-Molina. The eigentrust algorithm for reputation management in P2P networks. In *WWW '03: Proceedings of the 12th international conference on World Wide Web*, pages 640–651, New York, NY, USA, 2003. ACM.

[27] Apu Kapadia and Nikos Triandopoulos. Halo: High-Assurance Locate for Distributed Hash Tables. In *Proceedings of the 15th Annual Network and Distributed System Security Symposium (NDSS)*, pages 61–79, February 2008.

[28] Sachin Katti, Balachander Krishnamurthy, and Dina Katabi. Collaborating against common enemies. In *IMC '05: Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*, pages 34–34, Berkeley, CA, USA, 2005. USENIX Association.

[29] E.-M Kern and W Kersten. Framework for internet-supported inter-organizational product development collaboration. *Journal of Enterprise Information Management*, 20:562–577, 2007.

[30] D. Koukis, S. Antonatos, D. Antoniades, E. P. Markatos, and P. Trimintzios. A generic anonymization framework for network traffic. In *Proceedings of the IEEE International Conference on Communications (ICC 2006*, 2006.

[31] J Leyden. Cybercrime losses almost double - FBI figures show huge rise in online miscreantage, March 2010. Accessed 30th of Aril 2010.

[32] Patrick Lincoln, Phillip Porras, and Vitally Shmatikov. Privacy-preserving sharing and correction of security alerts. In *SSYM'04: Proceedings of the 13th conference on USENIX Security Symposium*, Berkeley, CA, USA, 2004. USENIX Association.

[33] Georgios V. Lioudakis, Fotios Gogoulos, Anna Antonakopoulou, Dimitra I. Kaklamani, and Iakovos S. Venieris. Privacy protection in passive network monitoring: an access control approach. In *Proceedings of the 23rd IEEE International Conference on Advanced Information Networking and Applications (IEEE AINA-09)*, pages 280–289, 2009.

[34] Katja Luther, Rainer Bye, Tansu Alpcan, Sahin Albayrak, and Achim Mueller. A cooperative approach for intrusion detection. In *IEEE International Conference on Communications (ICC 2007)*. IEEE, 2007.

[35] Tao Peng, Christopher Leckie, and Kotagiri Ramamohanarao. Survey of network-based defense mechanisms countering the dos and ddos problems. *ACM Comput. Surv.*, 39(1):3, 2007.

[36] Phillip Porras and Vitaly Shmatikov. Large-scale collection and sanitization of network security data: risks and challenges. In *NSPW '06: Proceedings of the 2006 workshop on New security paradigms*, pages 57–64, New York, NY, USA, 2007. ACM.

[37] Phillip A. Porras and Peter G. Neumann. EMERALD: Event monitoring enabling responses to anomalous live disturbances, 1997.

[38] Michael K. Reiter and Aviel D. Rubin. Crowds: Anonymity for web transactions. *ACM Transactions on Information and System Security*, 1:66–92, 1998.

[39] Jean-Marc Seigneur and Adam Slagell, editors. *Collaborative Computer Security and Trust Management*. Reference. Information Science Reference, 1 edition, November 2009.

[40] Yoichi Shinoda, Ko Ikai, and Motomu Itoh. Vulnerabilities of passive internet threat monitors. In *SSYM'05: Proceedings of the 14th conference on USENIX Security Symposium*, pages 14–14, Berkeley, CA, USA, 2005. USENIX Association.

[41] Vitaly Shmatikov and Ming-Hsiu Wang. Security against probe-response attacks in collaborative intrusion detection. In *LSAD '07: Proceedings of the 2007 workshop on Large scale attack defense*, pages 129–136, New York, NY, USA, 2007. ACM.

[42] Steven R. Snapp, James Brentano, Gihan V. Dias, Terrance L. Goan, L. Todd Heberlein, Che lin Ho, Karl N. Levitt, Biswanath Mukherjee, Stephen E. Smaha, Tim Grance, Daniel M. Teal, and Doug Mansur. DIDS (distributed intrusion detection system) - motivation, architecture, and an early prototype. In *In Proceedings of the 14th National Computer Security Conference*, pages 167–176, 1991.

[43] D. R. Stinson. *Combinatorial designs: construction and analysis*. Springer-Verlag, 2004.

[44] Jie Wu Tiranuch Anantvalee. *Wireless Network Security*, chapter A Survey on Intrusion Detection in Mobile Ad Hoc Networks, pages 159–180. Springer, 2007.

[45] Guido Urdaneta, Guillaume Pierre, and Maarten van Steen. A survey of DHT security techniques. *ACM Computing Surveys*, 2009. `http://www.globule.org/publi/SDST_acmcs2009.html`, to appear, preprint available.

[46] Michael Barry Winer and Karen Louise Ray. *Collaboration handbook: Creating, sustaining, and enjoying the journey*. Fieldstone Alliance, 1994.

[47] Shelly Xiaonan Wu and Wolfgang Banzhaf. The use of computational intelligence in intrusion detection systems: A review. *Applied Soft Computing*, 10(1):1 – 35, 2010.

[48] Kun Xiao, Ji Zheng, Xin Wang, and Xiangyang Xue. A novel peer-to-peer intrusion detection system. In *PDCAT '05: Proceedings of the Sixth International Conference on Parallel and Distributed Computing Applications and Technologies*, pages 441–445, Washington, DC, USA, 2005. IEEE Computer Society.

[49] Dengbang Xu and Peng Ling. *Intrusion Detection Systems*, volume 38 of *Advances in Information Security*, chapter Correlation Analysis of Intrusion Alerts, pages 65–92. Springer US, 2008.

[50] Jun Xu, Jinliang Fan, M.H. Ammar, and S.B. Moon. Prefix-preserving ip address anonymization: measurement-based security evaluation and a new cryptography-based scheme. In *Proceedings of 10th IEEE International Conference on Network Protocols*, pages 280 – 289, 12-15 2002.

[51] Vinod Yegneswaran, Paul Barford, and Somesh Jha. Global intrusion detection in the DOMINO overlay system. In *Proceedings of the Network and Distributed System Security Symposium, NDSS 2004*, 2004.

[52] Chenfeng Vincent Zhou, Shanika Karunasekera, and Christopher Leckie. A peer-to-peer collaborative intrusion detection system. In *Proceedings of the IEEE International Conference on Networks (ICON)*, pages 118–123, November 2005.

[53] Chenfeng Vincent Zhou, Shanika Karunasekera, and Christopher Leckie. Evaluation of a decentralized architecture for large scale collaborative intrusion detection. In *The Tenth IFIP/IEEE International Symposium on Integrated Network Management (IM 2007)*, pages 80–89, May 2007.

[54] Chenfeng Vincent Zhou, Christopher Leckie, and Shanika Karunasekera. A survey of coordinated attacks and collaborative intrusion detection. *Computers & Security*, 29(1):124 – 140, 2010.

[55] Tanja Zseby, Thomas Hirsch, Michael Kleis, and Radu Popescu-Zeletin. *Towards the Future Internet - A European Research Perspective*, chapter Towards a Future Internet: Node Collaboration for Autonomic Communication, pages 123–135. IOS Press, 2009.

Appendix

Balanced Incomplete Block Designs (BIBD)

A *BIBD* is an arrangement of *v* distinct objects into *b* blocks such that: (i) each object is in exactly *r* distinct blocks, (ii) each block contains exactly *k* distinct objects, (iii) every pair of distinct objects is in exactly λ blocks. The design is expressed as (v, b, r, k, λ) (a.k.a., (v, k, λ)) where: $b \cdot k = v \cdot r$ and $\lambda \cdot (v-1) = r \cdot (k-1)$. It is called *Symmetric BIBD* (a.k.a., *Symmetric Design* or SBIBD) when $b = v$ and $r = k$ [3] meaning that not only every pair of objects occurs in λ blocks but also every pair of blocks intersects on λ objects.

In this paper, we are interested in the *Finite Projective Plane* which is a subset of *Symmetric BIBD*. The *Finite Projective Plane* consists of points (a finite set P of points) and lines (a set of subsets of P) of the *projective space* $PG(2, q)$ of dimension 2 and order q. For each prime power q where $q \geq 2$, there exists a *Finite Projective Plane* of order q [43, Theorem 2.10] with following four properties: (i) every line contains exactly $k = q + 1$ points, (ii) every point occurs on exactly $r = q + 1$ lines, (iii) there are exactly $v = q^2 + q + 1$ points, and (iv) there are exactly $b = q^2 + q + 1$ lines. Thus, a *Finite Projective Plane* of order q is a *SBIBD* with parameters $(q^2 + q + 1, q + 1, 1)$ [3]. Symmetric designs can be constructed in $O(v^{3/2})$ time as described in [11] and references there in.

Consider $(v, k, \lambda) = (7, 3, 1)$ *Symmetric Design* as an example. Let $S = \{1, 2, 3, 4, 5, 6, 7\}$ be a set of $|S| = v = 7$ objects. There are $b = 7$ blocks: $\{1, 2, 3\}$, $\{1, 4, 5\}$, $\{1, 6, 7\}$, $\{2, 4, 6\}$, $\{2, 5, 7\}$, $\{3, 4, 7\}$, $\{3, 5, 6\}$. Each block contains $k = 3$ objects, every object is in $r = 3$ blocks, every pair of distinct objects is in $\lambda = 1$ block and every pair of blocks intersects in $\lambda = 1$ object.

Finite Generalized Quadrangle (GQ)

A *Finite Generalized Quadrangle* $GQ(s, t)$ is a point-line incidence relation with the following properties: (i) each point is incident with $t + 1$ lines ($t \geq 1$) and two distinct points are incident with at most one line, (ii) each line is incident with $s + 1$ points ($s \geq 1$) and two distinct lines are incident with (a.k.a., intersect on) at most one point, and (iii) if x is a point and L is a line not incident (I) with x, then there is a unique pair $(y, M) \in Points \times Lines$ for which $x\ I\ M\ I\ y\ I\ L$. In a $GQ(s, t)$, there are $v = (s + 1)(st + 1)$ points and $b = (t + 1)(st + 1)$ lines where each line includes $s + 1$ points and each point is incident with $t + 1$ lines. In this work, we are interested in $GQ(q, q)$ from projective space $PG(4, q)$. Probability that two lines intersect in $GQ(q, q)$ is given by the Equation 6.

$$P_{GQ} = \frac{t(s+1)}{(t+1)(st+1)} = \frac{q(q+1)}{(q+1)(q^2+1)} \approx \frac{1}{q}. \tag{6}$$

In $GQ(s, t) = GQ(q, q)$, there are $v = b = q^3 + q^2 + q + 1$ lines and points. Each line contains $s + 1 = q + 1$ points, and each point is incident with $t + 1 = q + 1$ lines. $GQ(q, q)$ can be constructed in $O(v^2)$ time as described in [11] and references there in.

Consider $GQ(s, t) = GQ(2, 2)$ for $q = 2$ as an example: There are 15 points $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}$ and 15 lines $\{1, 8, 9\}$ $\{1, 12, 13\}$ $\{1, 4, 5\}$ $\{3, 12, 15\}$ $\{2, 8, 10\}$ $\{2, 12, 14\}$ $\{2, 4, 6\}$ $\{5, 11, 14\}$ $\{3, 4, 7\}$ $\{6, 11, 13\}$ $\{5, 10, 15\}$ $\{3, 8, 11\}$ $\{7, 9, 14\}$ $\{7, 10, 13\}$ and $\{6, 9, 15\}$ where each line contains $s + 1 = 3$ points and each point is incident with $t + 1 = 3$ lines. Note that lines $\{1, 8, 9\}$ and $\{3, 12, 15\}$ do not intersect but GQ provides three other lines intersecting with both: $\{1, 12, 13\}$, $\{3, 8, 11\}$ and $\{6, 9, 15\}$.