Detecting VoIP based DoS attacks at the Public Safety Answering Point

Christoph Fuchs, Nils Aschenbruck, Felix Leder, Peter Martini
University of Bonn - Institute of Computer Science IV, Roemerstr. 164, 53117 Bonn, Germany
{cf, na, leder, martini}@cs.uni-bonn.de

ABSTRACT

In the recent years Voice over IP (VoIP) telephony started to migrate from research to the market. In the future, All-IP networks will substitute the classical Public Switched Telephone Networks (PSTNs). Nowadays, there is no All-IP network yet, but many VoIP-providers already enable calls from VoIP to a PSTN and vice versa. Thus, critical infrastructures within the PSTN like the emergency call service, are accessible from the VoIP network (e.g. the Internet) and get exposed to new security threats. In particular, there is the risk of Denial of Service (DoS) attacks originating from the VoIP network. An attacker could jam the emergency call service by generating a massive load of faked emergency calls, which could lead to the loss of lives in the worst case. For us, this was the motivation to analyse the applicability of the concept of Intrusion Detection (ID) in the emergency call context and develop an adapted ID-architecture including its implementation. In an evaluation of the ID-architecture, using real emergency call traces from the fire department of Cologne, we show that the developed concept can reliably detect emerging DoS attacks from VoIP networks up to a certain VoIP diffusion rate.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General—*Security and protection*

General Terms

Security, Reliability

Keywords

Intrusion Detection, Emergency Calls, Denial of Service, PSAP, VoIP, 911

1. INTRODUCTION

Over the last years, numerous improvements and standardisations were achieved concerning the Voice over IP (VoIP) technique, leading to a migration to VoIP-based telephony as an alternative to the conventional telephone system.

The integration of VoIP telephony into the existing telecommunications infrastructure that consists mainly of the Public Switched Telephone Networks (PSTNs) and mobile networks, is realised by VoIP gateways which are run by VoIP service providers. These gateways handle the translation of signalling information and provide mechanisms for addressing and routing of VoIP calls to the PSTN or mobile networks and vice versa.

This interconnection of packet-based IP networks and the circuit-switched PSTN, transfers several security risks which are present in the IP networks to the telephone network. Besides the issues of authentification and privacy, the danger of Denial of Service (DoS) attacks caused by a large amount of faked telephone calls imposes a serious threat on the up to now comparatively secure telephone network. The reason for the constitution of this new peril lies within the VoIP gateways which combine two different kinds of resources namely data packets on the IP side and (more limited) circuits on the switched network side. To conduct an effective DoS attack with the means of the conventional PSTN only, an attacker would have to dispose of a large number of fixed lines and furthermore expose himself to a high risk of being backtracked. The deployment of VoIP, however, enables an attacker to generate a large number of parallel calls from a single terminal, all of which are transformed into switched circuits at the VoIP gateway. The risk of being backtracked depends on the security mechanisms applied by the VoIP provider. It can be circumvented by gaining access to a legion of foreign VoIP-enabled PCs, e.g. in the form of a botnet.

Critical telecommunication infrastructures like the emergency call service (e.g. 911 in the US and 112 in Europe) are put at risk by the potential of an overload caused by a DoS attack. Emergency calls are routed to their regional emergency call centre where call agents process the incoming requests. These call centres are also referred to as Public Safety Answering Points (PSAPs). As a PSAP's resources of incoming lines and agents are limited, emergency calls get dropped if there is an overload of incoming calls. In this case, real emergency calls can not be answered. This may in the worst case even lead to the loss of lives.

The rest of the paper is organised as follows: In section 2 related work and possible solutions to handle the threat of DoS attacks on a PSAP are considered leading to the approach of an adapted type of intrusion detection within the

ASIACCS '08, March 18-20, Tokyo, Japan
Copyright 2008 ACM 978-1-59593-979-1/08/0003 ...$5.00.

PSAP. Section 3 describes the realisation of this approach in terms of an intrusion detection module which can be deployed at the PSAP. Finally, section 4 evaluates the attack detection capabilities of the proposed method and presents the results achieved followed by a concluding section.

2. RELATED WORK

The potential risk of DoS attacks on a PSAP initiated from the VoIP domain has already been perceived by the IETF working group ECRIT [17, 16] and NENA [12], which are actively involved in the subject of VoIP emergency calls. First approaches to face and counteract this threat included the following possibilities [17]:

- *Turing Test*: Before an emergency call is put through, an automated procedure in analogy to a turing test could be deployed to authenticate a human caller. In addition to the general complexity of this task [15], it can not be assumed that an emergency caller being in shock or injured will be able to pass such a test.
- *Inspection of IP address*: In order to suppress repetitive calls from the same caller, the source IP address of VoIP emergency calls could be inspected. This measure will not prove effective when the attacker is able to spoof his IP address or when he uses numbers of foreign machines brought under his control. Apart from this, multiple calls from the same IP address can constitue a valid condition, e.g. from enterprise networks using address translation.
- *Verification*: Assuming that eligible methods existed to verify the identity of a VoIP caller, his location or the VoIP service provider used, this information could serve as a starting point to prevent or backtrack DoS attacks. It is questionable though, to which extent the procedure of verification would restrict the public availability of the emergency service. According to ETSI, the emergency call service should be accessible without prior authentication [8].

In conclusion it can be stated that no feasible solution to the risk of VoIP initiated DoS attacks on the emergency call service has been proposed or realised so far.

Generally, it must be pointed out that the protection of a PSAP can not be achieved by means of preventive measures like authentification or up-front rate limiting. These measures would restrain the public and unrestricted accessibility of the emergency service which is also enforced by law in many countries, e.g. in Germany [1]. Due to these specific regulations for the handling of emergency calls, only reactive measures are eligible in this context. Reactive measures come into action only when a DoS is already emerging and then try to prohibit the total breakdown of the PSAP. The effectiveness of these measures depends on the ability of detecting an attack as early as possible.

In the field of IP networks, the concept of Intrusion Detection (ID) [7, 11] is widely deployed for the early detection of upcoming DoS attacks [6]. Here, selected aspects of the network traffic are observed and analysed. Two major approaches exist to detect possible intrusions: Anomaly detection and misuse detection. Anomaly detection aims at discovering deviations from a previously learned "normal" behaviour that exceed a certain threshold. Misuse detection usually uses a database of known attacks and scans the incoming traffic for matching patterns. Both methods can generate false alarms of two categories: Attacks that pass unrecognised (*false negatives*) are mainly a problem in an IDS using misuse detection as only known attacks can be detected. Alarms which get triggered even though no attack is taking place (*false positives*) occur more frequently in anomaly-based IDSs due to irregularities in the ordinary network traffic.

Applied to the emergency service, an ID-system deployed within the PSAP would have to analyse the total load and the characteristics of the stream of incoming emergency calls and — on this basis — realise a reliable attack detection. However, it may be hard to distinguish between DoS attacks and events triggered by real disasters, such as severe storms or major fires: In both cases, a strong increase of incoming emergency calls will be observed, leading to a high load at the PSAP. As the availability of the emergency service is crucial especially during events of disaster, it is important not to classify these events as attacks (false positives) which could result in mistaken countermeasures. This observation leads to the constraint that an intrusion detection method deployed in the emergency context must eliminate false positives or keep them at an absolute minimum.

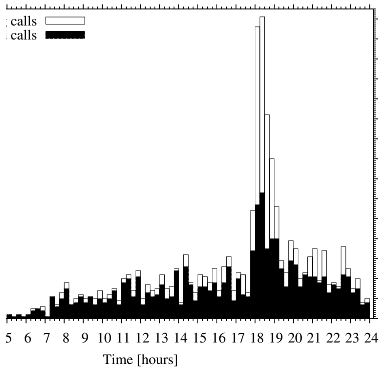

Figure 1: Emergency call load at PSAP Cologne on July 27, 2005.
(calls within 15 minute intervals)

Fig.1 shows a heavy load of emergency calls measured at the PSAP Cologne during July, 2005 which was caused by a thunderstorm . It can be seen that at about 5:45 p.m. the number of incoming emergency calls increases drastically. Due to the heavy call load, the resources are reaching their limits and a large fraction (sometimes over 50%) of the calls get lost.

In [3], a possible method to distinguish between attacks and disasters is proposed. Here, the incoming emergency calls are examined separated by their originating network domain that is PSTN, mobile networks or VoIP. The load characteristics for the calls of each originating domain are examined separately and compared to each other afterwards. During a disaster one expects that all domains show an increased or anomalous load of calls, whereas in the case of an attack only the VoIP domain would register a suspicious call behaviour. The outlined concept is similar to the principle of cooperative intrusion detection where multiple indepen-

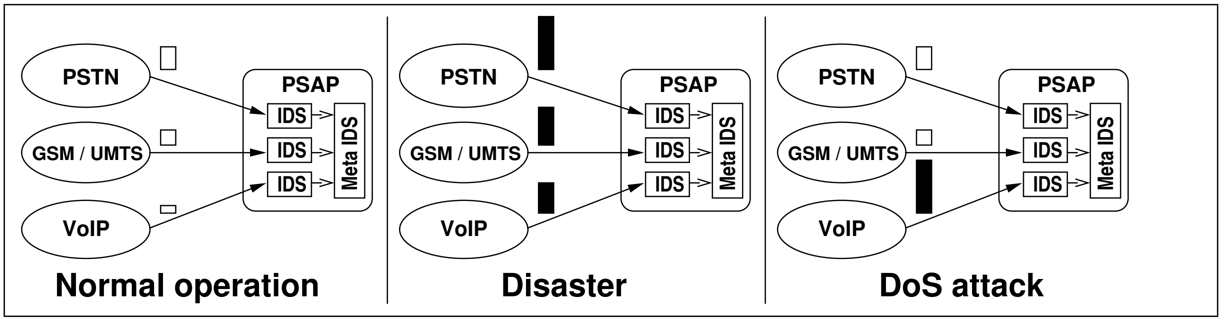

Figure 2: Expected call load of originating networks at the PSAP during different possible scenarios.

dent ID-systems located in different network domains submit their data to a superior meta ID-system. The meta IDS can then achieve a better detection rate by taking advantage of the broader data basis or by correlation of the individual datasets from each domain [9].

Fig. 2 visualises the described method including the expected call load from each originating domain during the different possible scenarios within the PSAP. The emergency call load from a particular domain is symbolised by the height of the respective bar icon. A filled bar denotes an especially high or abnormal call load.

The proposed partitioning of calls according to their originating domain also enables new possibilities for intrusion response measures. A possible intrusion response measure in case of a detected attack would be to limit the resources for each domain in order to prevent the complete allocation of resources by calls from the VoIP domain. The drawback of such fixed resource shares is that the available resources may be used in a suboptimal way. A less drastic intrusion response option would be to prioritise specific domains, e.g. the PSTN and the mobile networks. Although this would also penalize authentic VoIP emergency callers, this restriction would exist only in times of a detected attack. The described disadvantages of fixed resource partitioning and priorisation make clear, that these measures should not be installed up front but only in case of a DoS attack. Therefore, a fast and reliable method for attack detection is necessary.

3. REALISATION

We implemented the ID concept described in the preceding section in the form of an ID module that extends the Private Branch Exchange (PBX) within the PSAP. The PBX terminates the incoming mainlines (e.g. several ISDN lines) and distributes incoming emergency calls to the available call agents or buffers pending calls in a queue. To monitor the call traffic, it is necessary that the ID module is connected to the PBX via an appropriate interface. The comparison of different open-source software-based PBX systems [14, 13, 4] led to the choice of the widely-used software PBX Asterisk [4]. Asterisk is able to handle high call loads, which we confirmed by conducting stress tests with 4800 calls/minute. Furthermore, Asterisk features a convenient event-based management interface allowing for the detailed monitoring of the processed calls within the PBX.

The realisation of the ID module can be divided into several tasks which are described in the following:

3.1 Classification of calls by originating network

To allow for a separation of the calls from each originating network class, a procedure is required to assign a call to its originating domain. The distinction between PSTN and mobile networks could be accomplished by examining the transmitted telephone number of the caller as mobile networks have exclusively assigned number prefixes. Unfortunately, this does not hold true for VoIP. VoIP numbers are prone to be spoofed [2] and can also be identical to a regular PSTN number.

Thus, the classification of calls must take place within the telephone network routing the emergency calls to the PSAP. As the entry point of the call is known to the network, e.g. VoIP gateway, PSTN exchange or mobile switching centre, the mapping to the originating domain is straightforward. One option to relay this information to the PSAP would be to encode it into the protocol signalling information during the call establishment. A simpler alternative without changes to the routing protocol is the following: The uniform emergency call number (112) is translated to the internal routing number of the destination PSAP in the first exchange the call passes. Similar to multiple subscriber numbers of a single ISDN line, the PSAP can be equipped with multiple routing numbers where each of these numbers is dedicated to an originating domain [10]. Hence, the originating network domain can be determined by the destination number the emergency call is routed to.

3.2 Identification of suitable call load characteristics

Each intrusion detection is based on the collection of measurable characteristics (metrics) of the monitored system. For attack detection, the value or development of the surveyed metrics is interpreted to decide whether an attack is in progress. For a high detection rate, it is important to identify those properties of the emergency call load which are significantly and characteristically influenced by a DoS attack.

As we use the concept of cooperative intrusion detection, the communication architecture of the PSAP has to be divided into adequate functional areas, each constituting an autonomous intrusion detection subdomain. Figure 3 shows the communication architecture within a PSAP along with five identified functional areas:

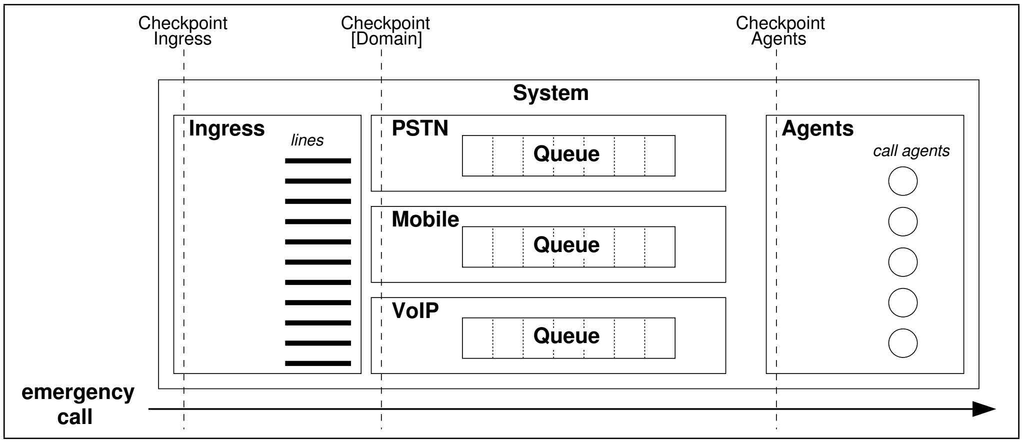

Figure 3: Communication architecture within a PSAP, showing functional areas and emergency call flow.

- The INGRESS area, representing the incoming telephone lines.
- An area for each originating domain consisting of its respective call queue.
  (in the following labelled as QUEUE[DOMAIN])
- The AGENTS area with the available call agents.

Each incoming and processed emergency call traverses these functional areas from left to right, starting at the INGRESS, then entering the QUEUE according to its origin and finally being answered by an agent. Next, it can be stated that each functional area has a limited capacity. The corresponding resources are the lines for the INGRESS, the number of slots for each QUEUE, and the number of agents. By establishing a checkpoint at the entry point of each functional area which registers each passing or rejected call, the first two metrics can be captured as the *interarrival times* of calls and the number of *rejections* at each area. Adding a checkpoint at the exits of the areas enables us to measure the *occupancy* of resources and the *residence* time of a call within an area. Thus, we obtain a set of four metrics for each intrusion detection subdomain. Each set represents the internal state of the corresponding area.

The metrics chosen are significantly influenced during a DoS attack: An increase of the arrival rate is typical for any DoS attack as otherwise no critical resource consumption could be achieved. Since it is the primary target of a DoS attack to exhaust available resources, monitoring the occupancy of each area's resources is crucial. Rejections occur when resources are already at their limit. Thus, they should be treated as a high priority indication of overload. The residence time within a QUEUE represents the waiting time of calls in that queue whereas in the AGENTS area the residence time denotes the call duration. Faked emergency calls related to a DoS attack are hung up by the agents after a short time. Thus, the mean call duration is affected by a DoS attack.

With respect to the metrics interarrival time and residence time, we are not interested in the single values for each call but in the mean values. However, the metrics shall only reflect the current situation. Therefore, the mean values are computed as a floating mean over a window of the last n calls. The value n can thus be thought of as the size of the observed event horizon. For the occupancy and rejections we use absolute values, limiting the rejection counts also to the last n calls.

3.3 Design of a structured architecture of the ID module

The acquisition of the metrics is followed by the analysis and interpretation as an alarm state of each functional area. Corresponding to the principle of cooperative intrusion detection, each area is associated with its own instance of an ID-subsystem. These ID-subsystems submit their results to a superior meta ID-component where the individual results are consolidated. This leads to a hierarchical organisation of the overall ID module, shown in figure 4. Each of the four metrics is captured and calculated by a dedicated sensor, processing the received events via the Asterisk management interface. When a metric changes, its value is propagated to the attached ID-subsystem by way of a *sensor update* message.

The frequency in which the values of the metrics change is comparatively high, especially during a high call load. Therefore, it is necessary to preprocess the received data in the ID-subsystems and abstract to a higher level in order to reduce the amount of communicated information to the meta IDS. This is achieved by setting alarm thresholds for each metric in each subsystem. Whenever the value of a metric exceeds its threshold, the meta IDS is informed by an alarm message containing the orginating subsystem and name of the metric that caused the alarm. As soon as the metric returns to normal, an alarm release is signalled to the meta IDS. Thus, the meta IDS' perspective on the overall system is abstracted to the accumulated alarm states of the observed metrics of all functional areas. Regarding the set of activated alarms as the system state, we obtain a finite state machine with 2^{20} = 1.048.576 possible states, where each alarm message or release triggers a state transition.

Regarding the alarm thresholds for each metric, it is possible to either use constant values or dynamic thresholds. The dynamic thresholds float around the actual mean value by a multiple of the deviation, thus adapting to the long term variability of the metric. We chose constant thresholds which is motivated by two reasons: First, dynamic thresh-

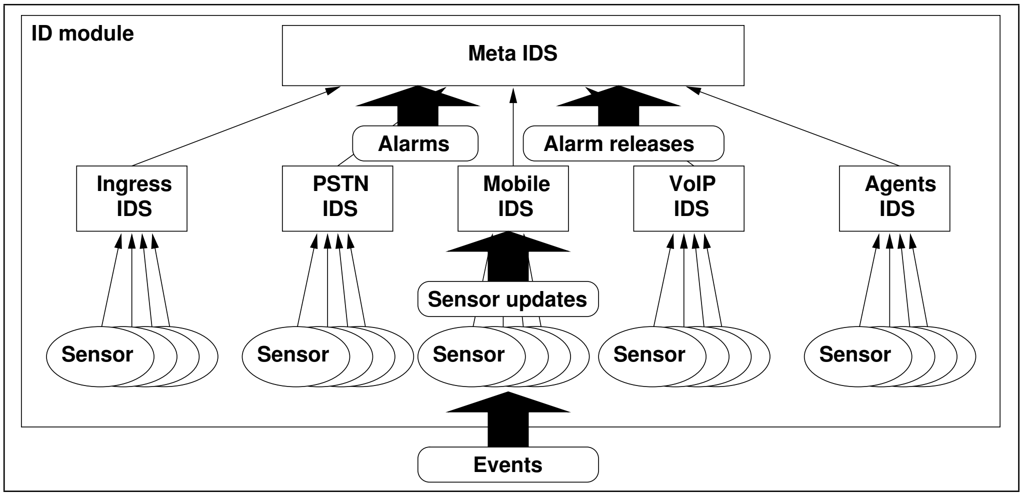

Figure 4: Hierarchical organisation of the ID module with flow of events.

olds are useful for parameters subject to seasonal or economic fluctuations with a small local variance. Alarms are issued when abrupt deviations from the mean value occur. However, the nature of emergency call traffic is highly variable also during normal situations, which would result in many spurious alarms for individual metrics. Second, as DoS attacks aim to exploit available resources, it is advisable to deduce constant alarm thresholds from the given resource limits. Thresholds for occupancy and interarrival times can be adapted to the amount of resources and to the mean service rate, respectively in each functional area. Residence times can be limited to an interval experienced in normal operation. Rejections are critical, so that already one rejection within the last call window should cause an alarm.

3.4 Determination of attack criteria

The actual attack detection is conducted on the meta IDS level, based on the propagated alarm states of all functional areas which jointly form a system state. Therefore, those system states had to be defined which are in all probability indicating a DoS attack and thus trigger an attack alarm.

As stated in section 2, it is essential to minimize the false positive rate in the emergency context. Therefore, we want to report only DoS attacks with the potential to lead to a severe restriction of the emergency call service. To achieve this, we use a two-stage alarm criterion, symplifying the identification of attack states within the comparatively large space of system states.

In the first stage, the two areas INGRESS and AGENTS, which are indicators for the allocation of incoming telephone lines and available call agents form the decisive factor representing the utilisation of critical resources of the overall system. Only at times when one of these areas reports an impending shortage of resources due to a high call load, the availability of the emergency service is endangered.

The second stage of the attack detection is only executed when the alarm criterion of the first stage is met. This implies that resources are already running short, leaving the decision whether the high load is caused by a DoS attack or by a disaster situation. The correlation of the different originating domains is applied by comparing the state of each domain on the abstract level of simultaneously active alarms across the domains.

In detail, we chose the alarm criterion of the first stage to be fulfilled when either one of the occupancy alarms of the areas INGRESS or AGENTS is set. A single interarrival alarm occurs frequently, caused by short fluctuations in the call rate. For this reason the interarrival alarm is not adequate to serve as criterion. As we want to reflect the consumption of resources within the first stage, the occupancy alarm is an appropriate measure. This partition of the attack detection also allows for a simple adjustment of the overall sensitivity of the ID module by chosing the alarm thresholds of the areas INGRESS and AGENTS accordingly.

To realise the correlation of the originating domains and to derive a proper alarm criterion for the second stage, a scheme is required that rates and compares the severity and the probable cause of the accumulated alarms in different domains. Comparing the alarm states instead of each metric's absolute value already has the advantage that fluctuations within a metric's course are attenuated due to the floating average determination in the sensors. Possible correlation methods would be the one-to-one matching of equal alarms in each domain or the comparison of the sum of active alarms as a measure for severity. The latter method lacks significance concerning the *cause* of a high alarm count. The drawback of the first method is the missing cross-relation of different types of alarms.

A further observation is that the sequence of reported alarms can reflect the causal chain of a domain's alarm evolution. For example, an occupancy alarm within a domain is not necessarily the consequence of a high call rate for that domain which would have been indicated by a preceding interarrival alarm. It can also be caused by slow agents or by a high load in another domain, leading to an accumulation of pending calls within the other queues. These cases should be identified and be considered when rating the severity of a domain's alarm state.

To achieve this, we introduced a concept of four alarm levels, based on the evolution of the observed metrics during a DoS attack. As a prerequisite, a DoS attack starts with an increased call load, indicated by an interarrival alarm within the domain's functional area. When calls accumulate in the QUEUE, an occupancy alarm is triggered and with sustained call load eventually followed by a rejection alarm. A chronological indexing of the residence alarm is inapplicable.

The following list shows the definition of the four alarm levels. For each domain, the corresponding alarm level is determined accordingly. Note, that a triggered interarrival alarm within a domain constitutes a fundamental precondition for an alarm level higher than zero. As the exact order of incoming alarm messages can differ slightly even during an attack, we defined the alarm levels based on the simultaneously active alarms, thus consolidating different permutations on the basis of their resulting alarm set.

- **Alarm level 0:** *interarrival* alarm of domain is **not** set
- **Alarm level 1:** *interarrival* alarm of domain is set
- **Alarm level 2:** *interarrival* alarm **and** *occupancy* alarm of domain are set
- **Alarm level 3:** *interarrival*, *occupancy* **and** *rejection* alarms of domain are set

The actual alarm criterion of the second stage, depending on the present alarm levels of the three domains PSTN, Mobile and VoIP is defined in Table 1. As we want to detect DoS attacks originating from the VoIP domain, it is reasonable to assume that during an attack the alarm level of the VoIP domain is at least 1, indicating an increased call rate. As short bursts in the call rate are normal and frequent, we require the VoIP domain to be at a minimum alarm level of 2 to possibly assume an impending attack. Only a longer lasting high call rate leads to an occupancy alarm resulting in an alarm level of 2. If the alarm level is 0 in one or both other domains, we assume an attack and an alarm is triggered. An alarm level of 1 (or higher) in *both* other domains is considered as a disaster where the transition from level 1 to 2 happened first in the VoIP domain. When alarm level 3 is reached within the VoIP domain, the capacitiy of the queue is already exhausted. If a DoS attack was in progress, it would already be in an advanced state. Therefore, in this case an alarm is also triggered if both remaining domains show an alarm level of 1. All other combinations not listed in Table 1 do not result in an attack alarm.

The residency alarm was not included in the definition of alarm levels, as its temporal occurrence during a heavy load situation is not determined. As an option, the residency alarm could be used aside from the previously defined alarm criteria, so that an attack is not indicated unless a residency alarm of the PSTN or Mobile domains is triggered. In this case, an attack alarm would not be signaled until the quality of the emergency service is affected in the way that the holding time of calls from the PSTN or mobile devices exceeds a predefined threshold.

4. EVALUATION AND RESULTS

To evaluate the quality of the attack detection, two aspects are of major interest. First, the response time of the attack detection constitutes a critical factor, as countermeasures in response to a detected attack are only effective if the attack is signaled during an early stage. Second, the reliability of the detection is important, as it is useless to have a short response time at the cost of a high false positive rate.

VoIP	PSTN	Mobil	Attack Alarm
0	*	*	No
1	*	*	No
2	0	0	Yes
2	0	1	Yes
2	1	0	Yes
2	1	1	No
3	0	0	Yes
3	0	1	Yes
3	1	0	Yes
3	1	1	Yes

Table 1: Second stage alarm criterion based on the domain alarm levels.
***: alarm level 0, 1, 2 or 3**

Unfortunately, it was not possible to test our approach in a real PSAP. However, the fire department of Cologne (one of the largest fire departments in Germany) provided traces of real emergency call traffic. Thus, we were able to emulate realistic regular and disaster call traffic.

The parameters to adjust the sensitivity of the intrusion detection module are the alarm thresholds for the metrics in the different functional areas. It is obvious that appropriate values for these parameters depend on the configuration of the PSAP, defined by the number of incoming lines, agents, average call load, etc. As mentioned above, we based our evaluation on a setup motivated by the PSAP of Cologne, Germany. This results in 30 incoming lines and 10 available agents. In the course of analysing the response time and reliability we also optimised some key thresholds for the deployed configuration.

For the evaluation of the response time, a DoS attack from the VoIP domain is emulated as a quick succession of calls at a certain rate with uniformly distributed interarrival times. The coexistent regular call traffic from the PSTN and mobile networks is generated using real tracefiles measured at the PSAP Cologne during 2002 to 2006. If the DoS attack is executed with a high call rate, the response time of the ID module must be shorter than in the case of an attack with a lower rate. Therefore, we analysed the response time against the call rate of the DoS attack. Fig. 5 shows the results for different parameters for the threshold of the agent occupancy alarm. The upper solid curve marks the calculational time until all agents are busy at the given call rate. The graphs beneath depict the mean latency until the alarm is triggered for 25, 30, 50, 100, 400, 800 and 1200 calls per minute (with the 95% confidence intervals). For both thresholds it can be stated that an alarm is released well prior to the full utilisation of the agent resources. Furthermore, an occupancy threshold at 4 agents results in a significantly faster response time at lower call rates.

The evaluation for the INGRESS occupancy threshold resulted in the curves shown in Fig. 6. Here, the upper solid curve shows the time until all incoming lines are busy. Again, the DoS attack is indicated before the available lines are fully utilised. The alarm threshold at 8 lines for the INGRESS occupancy alarm results in significantly faster response times at high call rates, which is shown in the close-

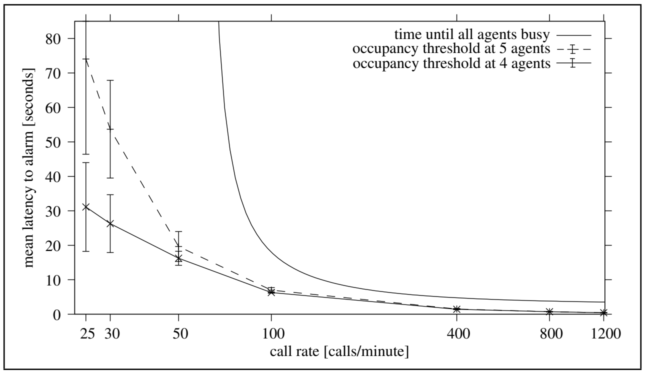

Figure 5: Response times for different agent occupancy thresholds.
(Mean values are shown with 95% confidence intervals)

up in Fig. 7. This effect is due to the fact that at high call rates the occupancy threshold of the INGRESS is reached prior to the threshold of the AGENTS. Thus, the first stage's alarm criterion is met early. For the setup with 30 incoming lines and 10 available agents, we did not further decrease the thresholds as this would no longer justify an occupancy alarm and only increase the risk of false positives.

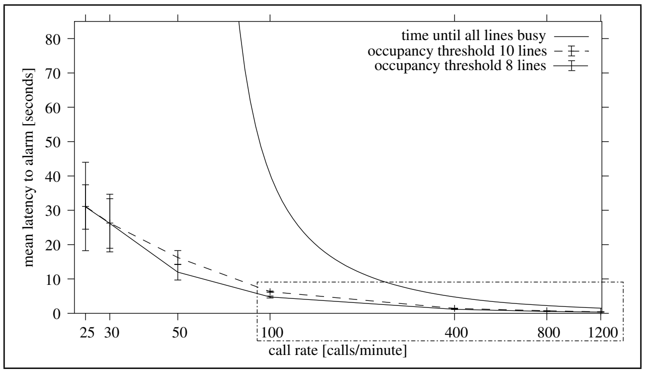

Figure 6: Response times for different INGRESS occupancy thresholds.
(Mean values are shown with 95% confidence intervals)

A second evaluation aspect is, whether the determined parameters that delivered good response times do also provide a good reliability, i.e. a low false positive rate during disaster scenarios. Since the distinction between an attack and a disaster is based on the correlation of emergency calls separated by their originating domains, the share of VoIP calls in the overall load constitutes a decisive factor. Therefore, we tested up to which share of VoIP emergency calls our technique can be deployed without leading to false positives in case of heavy load caused by disaster scenarios. We identified several disaster situations within the tracefiles from the PSAP Cologne and added a VoIP call load corresponding to VoIP shares of 11, 20, 27 and 33% of the overall traffic.

Fig. 8 shows the results for the thunderstorm scenario described in section 2. The bars depict the number of false positives whereas the curves mark the fraction of time during which the alarm criteria of the first and second stage were fulfilled. For all analysed disaster scenarios, false positives did not occur below a VoIP call share of 27%. This conforms to the observation that at this VoIP share the criterion of the second stage starts to be met during a small fraction of time. In contrast, the alarm criterion of the first stage is met during 40% of the tested period in Fig. 8 which corresponds to the heavy load situation present in a disaster scenario.

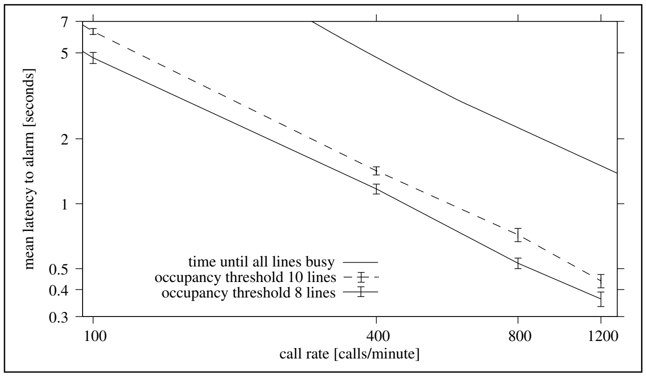

Figure 7: Response time for different INGRESS occupancy thresholds.
(Close-up on high call rates, x- and y-axis in log scale)
(Mean values are shown with 95% confidence intervals)

To sum it up: The evaluation showed that a fast and reliable attack detection is achieved up to a share of VoIP emergency calls of at least 20%. At the time of writing, the fraction of VoIP customers compared to PSTN and mobile networks constitutes approximately 0.4% in Germany [5].

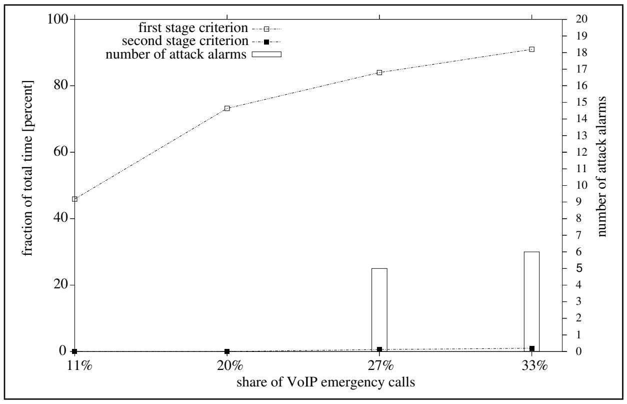

Figure 8: Evaluation of the reliability of the attack recognition.
(Thunderstorm on July 27, 2005)

5. CONCLUSION AND FUTURE WORK

We designed and implemented cooperative intrusion detection at the PSAP to detect VoIP-initiated DoS attacks as an extension ID module for the popular software-based PBX Asterisk.

To map the concept of multiple, cooperative intrusion detection domains on the PSAP, we identified separate functional areas within the telecommunication architecture of a PSAP. This partitioning led to a hierarchical design of the ID module consisting of ID-subsystems for each functional area, reporting to a meta IDS. We defined a two-stage alarm detection process, separating the estimation of resource criticality from the discrimination between attack and disaster situation. This method reduces the risk of false positives by limiting attention to attacks with sufficient potential to limit the resources of the PSAP and to degrade or interrupt service availability.

We specified metrics that lead to a characterisation of the traffic profiles for each originating network domain, particularly with respect to the allocation of valuable resources, e.g. incoming lines or call agents. By preprocessing the collected data and by abstraction within the IDS hierarchy, the number of events during a high call load could be reduced, allowing for an alarm decision based on discrete system states within the meta IDS. The response time of the attack detection proved to be well below the time needed for a full utilisation of resources and dynamically adapts to the call rate of a conducted attack. The analysis of the reliability of the attack detection showed that a reliable differentiation between attack and disaster is attained up to a VoIP share of approximately 20% on the overall emergency calls, which is far above the estimated 0,4% VoIP share in Germany today.

If the VoIP share exceeds 20% further properties of the incoming alarms could be considered, thus extending the state-based approach. For example, temporal properties like the duration that each alarm stays activated could be taken into account. This consideration of alarm persistence could serve to ignore short alarms and reduce false positives.

Another application area for the presented method are commercial call centers, which would suffer financial damage from a DoS attack. The different call load characteristics in a commercial call center would require an adaptation of the ID module's parameters.

Finally, the question of adequate intrusion response in case of a detected attack remains open. Here different possibilities of intrusion response, e.g. those proposed in section 2, must be studied with respect to their response time, effectiveness and applicability in the emergency context.

6. ACKNOWLEDGMENTS

We wish to thank the fire department Cologne, in particular Roland Legat for providing anonymised emergency call traces and for fruitful discussions on the operating principles of a PSAP.

7. REFERENCES

[1] Telekommunikationsgesetz (TKG), Juni 2004. BGBl I 2004, 1190.

[2] M. Arora and S. Chakravarty. VoIP security: Scenarios, challenges, and counter measures–Part I, January 2007.

[3] N. Aschenbruck, M. Frank, P. Martini, J. Tölle, R. Legat, and H.-D. Richmann. Present and Future Challenges Concerning DoS-attacks against PSAPs in VoIP Networks. In *Proceedings of the Fourth IEEE International Workshop on Information Assurance*, pages 103–108, April 13 - 14, 2006.

[4] Asterisk Homepage, May 2007. `http://www.asterisk.org`.

[5] Jahresbericht der Bundesnetzagentur 2005, Februar 2006.

[6] J. Branch, A. Bivens, C.-Y. Chan, T.-K. Lee, and B. Szymanski. Denial of service intrusion detection using time dependent deterministic finite automata. In *Proc. Graduate Research Conference, RPI, Troy, NY*, October 2002.

[7] D. E. Denning. An Intrusion Detection Model. In *IEEE Transactions on Software Engineering*, volume SE-13 No.2, pages 222–232. Februar 1987.

[8] Requirements for communication of citizens with authorities/organizations in case of distress (emergency call handling). Special Report 002 180, ETSI, 2003.

[9] N. gentschen Felde. Einsatz der graphbasierten Meldungsstrukturanalyse in domänenübergreifenden Meta-IDS. In A. B. Cremers, R. Manthey, P. Martini, and V. Steinhage, editors, *GI Jahrestagung (2)*, volume 68 of *LNI*, pages 653–657. GI, 2005.

[10] Erarbeitung landesweiter Standards für die Errichtung Integrierter Leitstellen in Bayern, August 2001. `https://www.bayern-ils.de/ILSWebseite/downloads/extern/Erarbeitung%20landesweiter%20Standards/Teil3.pdf`.

[11] B. Mukherjee, L. Heberlein, and K. Levitt. Network intrusion detection. *IEEE Network*, 8(3):26–41, May/June 1994.

[12] PSTN Risks & Challenges, 2005. (Results of a fast track working group) `http://nena.org/VoIP_IP/PSTN%20Risks%20&%20Challenges%20FINAL.doc`.

[13] Homepage von Voicetronix, November 2005. `http://www.voicetronix.com.au`.

[14] Homepage von PBX4Linux, November 2005. `http://isdn.jolly.de`.

[15] A. P. Saygin, I. Cicekli, and V. Akman. Turing Test: 50 Years Later. In *Minds and Machines*, volume 10, pages 463 – 518. Kluwer Academic Publishers, 2001.

[16] H. Schulzrinne and R. Marshall. Requirements for Emergency Context Resolution with Internet Technologies, Oktober 2005. IETF ECRIT Draft draft-schulzrinne-ecrit-requirements-01.txt.

[17] H. Schulzrinne, M. Shanmugam, P. Taylor, and H. Tschofenig. Security Threats and Requirements for Emergency Calling, 2005. IETF ECRIT Draft draft-taylor-ecrit-security-threats-00.txt.