Efficient Lawful Interception in Mobile IPv6 Networks

Andres Rojas, Philip Branch and Grenville Armitage
Centre for Advanced Internet Architectures
Swinburne University of Technology
Melbourne, Australia
Email: (anrojas, pbranch, garmitage)@swin.edu.au

Abstract— The Lawful Interception (LI) of communications is necessary in modern telecommunications networks in order to help law enforcement agencies with the investigation and prosecution of criminal activities.

For IP networks that support user mobility, such as Mobile IP, the LI solution must be distributed throughout the network in order to satisfy the requirement that it capture 100% of a target user's communications. This distribution is in conflict with another requirement which requires that the identity of the interception target remain unknown outside of the LI system : increasing the distribution of the LI solution to more nodes, as well as being inefficient, means that the identity of the target of the interception is more susceptible to being known via attacks, social engineering or unintentional misconfiguration.

In this paper, we analyse LI solutions for Mobile IPv6 networks in terms their efficiency. We define an efficiency ratio which measures the balance between the requirement to capture 100% of target traffic and how distributed the efforts to intercept are. We examine the effect of alternative LI solutions on this measure through simulation.

Our results show that a client driven approach which incorporates appropriate caching mechanisms provides a significant improvement in efficiency when compared to a traditional client–server approach.

I. INTRODUCTION

Lawful Interception (LI) is the process whereby a Lawful Enforcement Agency (LEA) is legally allowed to intercept a target's communications for the purpose of law enforcement.

One of the requirements of a LI solution is that it must capture as much of a target individual's communications as possible, as specified under the terms of a legally authorised warrant. This is a critical requirement that directly impacts on the ability of law enforcement agencies to prosecute and convict criminals.

Today, to meet this requirement in mobile telephony networks, the LI function is distributed throughout the network so that a mobile user can be targetted wherever the user moves. A client–server model is used whereby the server controls the interception start and stop times (as well as other parameters). Each client is responsible for the interception of traffic within its own coverage area. The server effectively activates the interception of each target on every client. This model works well for mobile telephony because, although the coverage area is large, the number of clients is small: telephony exchanges are large machines that handle many thousands of subscriber calls simultaneously [1].

For IP networks that support user mobility, such as Mobile IP, applying the same client-server model leads to two problems as a result of the inherent decentralization of IP networks and the distribution of the LI solution. Firstly, the solution is inefficient in that the server activates the interception on all clients. This leads to wasted resources on those clients that are responsible for the coverage areas that the interception target does not visit. Secondly, effectively broadcasting the identity of the interception target, and maintaining that information on all clients conflicts with another LI requirement that mandates that the identity of the interception target not be divulged.

In this paper we define a measure of the balance between these opposing requirements. The aim of the defined ratio is to measure how efficient an LI solution is at capturing a target's traffic when considering how distributed the interception is (ie. how many clients know about the interception target, and for how long). Further to this we describe a more efficient LI solution for IP networks that support user mobility and evaluate it against a naive solution using simulation.

II. RELATED WORK

As far as the authors can ascertain, there are no studies which examine the efficiency of LI solutions in any networking context.

Rojas and Branch explored the problems associated with using traffic sniffers, which are an integral part of current LI solutions, in the interception of future networks [2]. Although the present paper assumes the same use of traffic sniffers, [2] defines some alternatives which could be used as replacements.

In [3], Sherr et.al. describe exploitable vulnerabilities in wiretap systems that employ loop extenders frequently used in telephony networks. In the most serious vulnerability found, a target could cause the wiretapping system to suspend audio recording. Although the vulnerabilities, for the most part, only apply to interception of telephony networks using physical devices, the paper highlights the fact that LI systems are seldom explored within the research community.

Aside from technical discussions, a number of authors have explored broader topics: implications of LI on network design, definition of LI legislation in technologically neutral policy, and the advantages and disadvantages of different approaches to LI, [4]–[6].

III. A SIMPLE CLIENT/SERVER SOLUTION

A. The Lawful Interception Process

Interception of a targets communications, in most western countries, begins with a warrant that is issued by a judicial officer to allow a law enforcement agency to intercept the communications of a person who is under criminal investigation.

Each warrant typically has a commencement and ending date and time. Interception of communications outside of this time is unlawful - as is any interception that is not subject to a warrant. The warrant also specifies the identity of the target (eg. telephone no., email address, IP address) and may also specify whether the content of the communications and/or meta–data associated with the communications is subject to interception [7].

All Australian Telecommunications carriers and Internet Service Providers (ISP) are required to be able to intercept any communications that pass over their network, [7], [8].

B. Server Driven Broadcast Activation (SDBA)

A rudimentary example of an LI architecture that could be employed for Mobile IPv6 networks is one which is used today for mobile telephony networks such as GSM and GPRS. It uses a simple client–server model, with one central LI server and many LI clients. Each LI client is responsible for the interception of traffic for a certain area of wireless coverage.

We name this architecture Server Driven Broadcast Activation (SDBA) because the activation of interception for a target is broadcast from the central server. A depiction of how SDBA works is shown in figure 1, which shows the sequence of LI messages involved.

In the description of SDBA below, we assume that each LI client has access to all traffic transmitted within the coverage area that it is responsible for. This assumption is consistent with how commercial LI equipment operates today. This is achieved through port mirroring from a local switch or router.

At the interception commencement time, the server sends an `activate` message to all clients to activate interception for the target's IP address. Before this time, any traffic that is mirrored to a client for which it does not have interception activated is discarded by that client. After this time, any IP datagram that is mirrored to a client for which it has interception activated (a match to either the source or destination address of the datagram), is forwarded to the server for processing and delivery to the law enforcement agency.

At the interception end time, to deactivate interception for a target, the server sends a `deactivate` message to all clients. After processing this message, clients no longer have any knowledge of that target.

C. Balancing Lawful Interception Requirement

An efficient LI solution for IP networks that support user mobility, both tries to capture as much target traffic as possible and only activates interception at the clients that are responsible for the areas where the target actually moves through. An efficient LI solution limits the amount of time that the LI system is susceptible to leaking information about who is being intercepted. This threat can be from remote attacks and social engineering, through to unintentional misconfiguration. Also, limiting the scope of interception activation affects the resource usage on network equipment and network utilisation.

Therefore, we define

$$\tau_{captured} = \left(\frac{\tau_{sent} + \tau_{received}}{\mu_{sent} + \mu_{received}} \right) \cdot 100$$

where $\tau_{captured}$ is the percentage of target traffic captured, $\tau_{sent}, \tau_{received}$ is the amount of captured target traffic sent and received, respectively, and, $\mu_{sent}, \mu_{received}$ is the amount of traffic the target sent and received.

We also define β to be the mean time that an LI client has interception activated:

$$\beta = \frac{\sum_{i=1}^{n} t_{deactivation} - t_{activation}}{n}$$

where n is the total number of LI clients in the network, and $t_{deactivation}, t_{activation}$ are the deactivation and activation times for each client. The mean client activation time, as a percentage of the maximum activation time (ie. the duration of the interception period as specified in a warrant) is then given by

$$\delta = \left(\frac{\beta}{t_{warrantEnd} - t_{warrantCommence}} \right) \cdot 100$$

Finally, we define the efficiency ratio of an LI solution to be

$$\eta = \frac{\tau_{captured}}{\delta}$$

Ideally, SDBA should have an η of 1 as 100% of target traffic is captured and each client has interception activated for 100% of the maximum interception period, for one target.

IV. CLIENT DRIVEN TARGETTED ACTIVATIONS

In this section we describe an LI architecture for IP networks that support user mobility that aims to reduce the amount of time that each LI client has interception activated.

We name this architecture Client Driven Targeted Activation (CDTA) because the LI client initiates the activation of
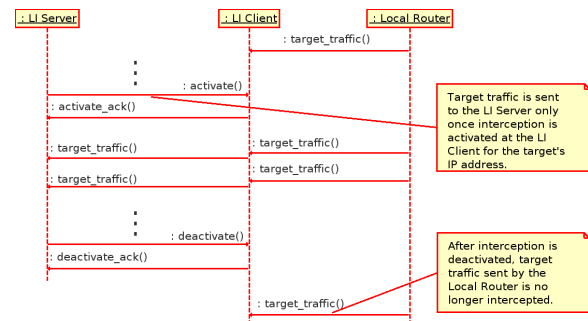

Fig. 1.   Sequence diagram for Server Driven Broadcast Activation.
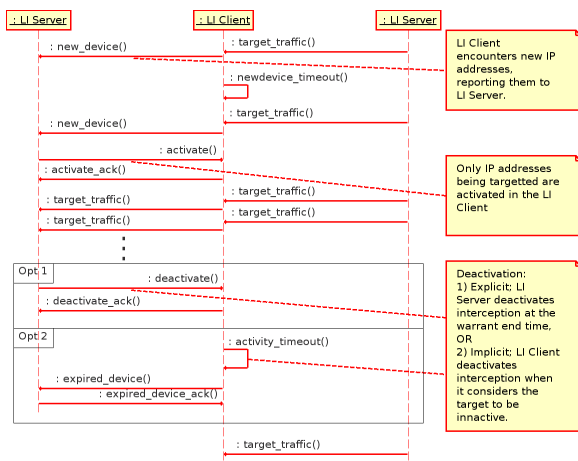
Fig. 2. Sequence diagram for Client Driven Targetted Activation.

interception for a target, and because the interception is only activated at the client when the target moves into the area for which the client is responsible.

In CDTA, as shown in figure 2, clients report any IP addresses they encounter to the server (`newdevice` message). Only IP addresses that are being targetted are activated at the client by the server. This ensures that interception is activated at a client only when necessary.

Deactivation of interception at clients can occur in two ways. The client can simply wait for an explicit `deactivate` message from the server at the interception end time, or, the client can act in a more proactive manner and deactivate interception of the target when it considers the target to be inactive. In the second case, the client uses a timeout message to determine innactivity (`activity_timeout` message). If the client does not see traffic from the target IP address within the timeout period, then the target is considered inactive. This activity timeout is designed to ensure that interception is deactivated at a client as soon as the target moves away from the area for which the client is responsible.

Naturally, reporting every single IP address that a client encounters to the server leads to a large amount of traffic, especially as the network's size increases. To counter this volume of messages, we introduce a caching mechanism so that the `newdevice` message is only sent to the server when an IP address has not been seen for a period of time. This can be implemented using a timeout, as shown in figure 2 (`newdevice_timeout` message).

V. SIMULATION

A. Simulation Setup

To compare the efficiency of the LI architectures described in general and to analyse the effect of the network size on this efficiency, we ran a number of simulations. We used the Omnet++ simulator together with the IPv6SuiteWithINET model to simulate Mobile IPv6 networks, [9], [10].

In our simulations, we used a standardised local network in order to vary the size of the simulated network. Each local

network consisted of four IEEE 802.11b access points (AP) configured to provide wireless coverage for an area 400m x 400m in size. These APs were connected to a local router, which also connected to a local server. The local networks were interconnected to each other via core routers.

In terms of LI, each local network was serviced by one LI client, the local router configured to mirror all traffic transmitted on its AP links to the LI client. The LI server was connected to one of the core routers.

We simulated one mobile user travelling throughout a square network coverage area which varied from 800m x 800m to 2400m x 2400m. The user travelled according to a Random Waypoint model (RWP) for one hour, with a random speed chosen uniformly between 5 and 15 m/s, pausing at each destination for exactly 10s; a model designed to be consistent with vehicular travel for the network sizes simulated. This mobility model was implemented taking into account the stationary distribution demonstrated by the RWP [11].

The mobile user pinged one of the local servers with a rate of 1 ping/sec. Consequently, barring any routing problems, a ping–reply packet was sent by the server as a response.

To test the efficiency of the following different LI architectures we ran 50 simulation runs for each network size,

- SDBA
- CDTA with server deactivations
- CDTA with activity timeout = 60s

To analyse the effect of different values of the activity timeout period on the mean client activation time and the percentage of target traffic captured (δ and $\tau_{captured}$, respectively), we also ran 35 simulation runs for each activity timeout period value varying from 0.2s to 60s.

To analyse the effect of different values of the new–device timeout period on the number of messages sent to the server and on the mean client activation time (δ), we also ran 35 simulation runs for each new–device timeout period value varying from 0.2s to 60s.

B. Results

Figure 3 shows the performance of the 3 LI architectures in terms of the mean LI client activation time for different network sizes. We see that CDTA with an activity timeout of 60s gives the smallest mean LI client activation time, almost reaching the theoretical minimum which is also shown.

From our simulations, the 2 CDTA architectures tested had means of 99.7% of target traffic captured for all network sizes. The SDBA architecture had a mean captured target traffic of 99.9%. This was not quite 100% because the server discarded any target traffic that was sent by a client before the interception end time but arrived at the server after this time.

Together, these results influence the efficiency ratio, η, as shown in figure 4. We see that for a CDTA architecture that employs an activity timeout, η increases with network size. This is explained by the fact that the mobile user moves through less coverage area when the network size is larger.

Figure 5 shows how the value of the activity timeout affects the mean LI client informed time, and the percentage of total
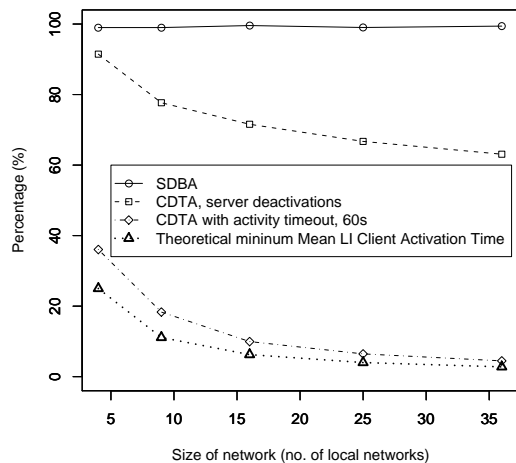
Fig. 3. Mean LI client activation time as percentage of the maximum activation time.
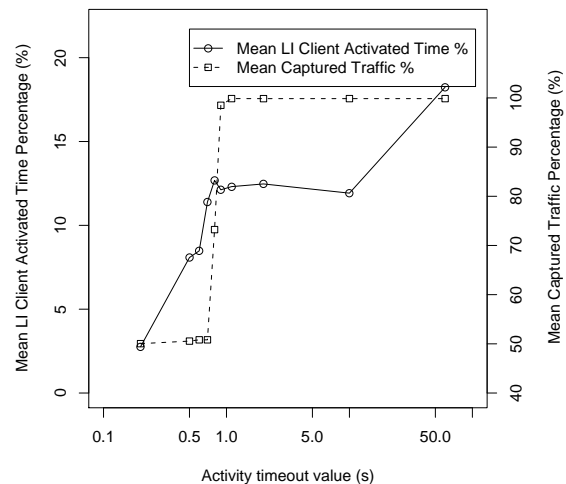

Fig. 4. Mean efficiency ratio (Captured traffic percentage : Mean LI client activation time percentage).

Fig. 5. Effect of activity timeout on mean LI client activated time and mean captured traffic. (CDTA with activity timeout, 9 local networks).
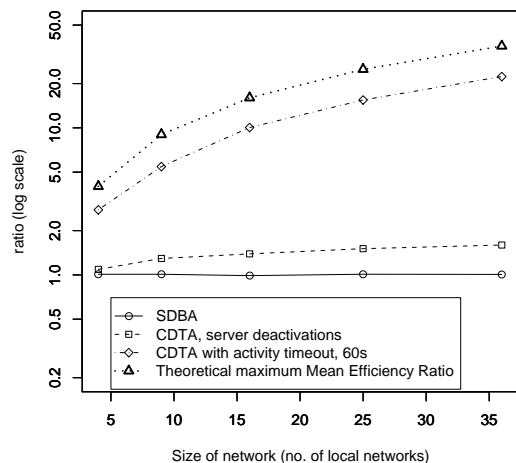
captured traffic. We see that trying to reduce the mean LI client activation time by reducing the activity timeout value has a tradeoff in that the mean captured traffic is severely affected, in this case, when the timeout value is less than 1s.

Although both CDTA architectures tested improve the efficiency ratio, the one drawback is the amount of network traffic that is generated from LI client to server. Figure 6 shows how the value of the new–device timeout affects the mean number of total LI messages, and the mean LI client informed time. We see that trying to reduce the mean number of messages by increasing the new–device timeout value also increases the mean LI client activation time. In our simulations, increasing the new–device timeout value past approximately 1.5s reduces the mean number of messages to an acceptable level without a substantial effect on the mean LI client activation time.

VI. DISCUSSION

We have seen, through the results presented in the previous section, that our proposed CDTA architecture is a more efficient LI solution than an SDBA architecture for all network sizes, and approaches the maximum theoretical efficiency. Also, the efficiency of our proposed CDTA architecture can be optimised by using an appropriate activity timeout value, and the high number of messages associated with CDTA can be controlled with an appropriate new–device timeout value.

Although these results show that CDTA can provide law enforcement agencies with the traffic that they are lawfully entitled to have, while also minimizing the extent of the distribution of interception information, we make the reader aware that there are a number of assumptions, listed below, that affect our simulations and the interpretation of the results.

It is unclear how Mobile IPv6 will be deployed. If a deployed network uses authentication protocols, it seems logical that an LI solution could be aided in it's detection of where a mobile user is by authentication.

In the LI architectures described in this paper, we have relied on the use of port–mirroring to deliver the communications of all users of the network to LI clients. Although this technique is in use today with commercially available LI solutions, it is not clear whether it is an effective, robust, scalable, or viable way of getting access to all network traffic in practice.

A technologically astute criminal would likely communicate using end to end encryption. Anecdotal evidence suggests that law enforcement agencies are still interested in communication meta–data such as the destination IP address of the traffic. The
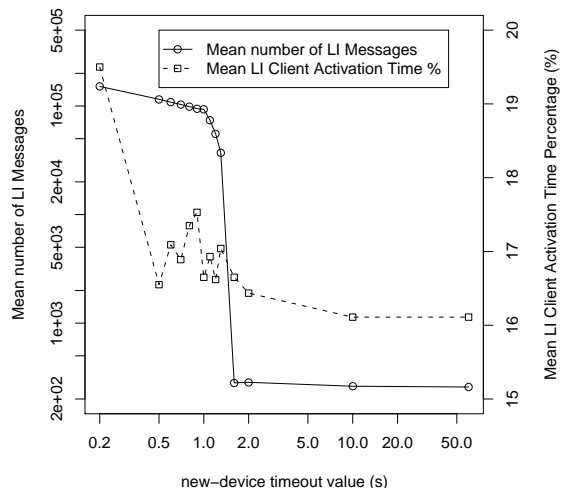
Fig. 6.   Effect of new–device timeout on mean number of messages and on mean LI client activation time. (CDTA with new–device timeout, 9 local networks).

target could use other measures such as using encrypted and/or anonymous tunelling, or anonymous IPv6 address configuration to avoid detection. These issues are not considered.

There were a number of limitations in our simulations which we view as opportunities to develop the research further.

First, we used a mobility model that, although simplistic, was aggressive. A conservative model would probably result in the CDTA architectures tested exhibiting even better η values.

Also, we used a very simple traffic model that modelled an application with a constant packet transmission rate. The results shown in figures 5 and 6 are highly dependent on this simple traffic model which used a relatively slow rate of 1 packet/sec. Notwithstanding, our results are indicative for other traffic sources which use a constant packet transmission rate, such as VoIP or video streaming. We leave it as future work to test the architectures with traffic models that represent applications with non–constant packet transmission rates such as www or e–mail, [12], [13].

One interesting side effect of how CDTA currently works is that, when it receives mirrored traffic and sends the `new_device` message to the server, it effectively discards this traffic until it receives an `activate` message from the server. E–mail is an example of a type of communication where capturing the first packet in the stream is absolutely critical to law enforcement. This means that we need to explore ways of capturing no less than 100% of the target traffic. We also note that it is hard to know what exact applications a target of interception (ie. a criminal) would likely use.

Finally, in our simulations the distribution of APs was uniform across the coverage area. In reality, APs are more densely distributed among areas of high wireless device use such as along roads, in pedestrian thoroughfares, and in places

were there is a high density of people. An interesting extension to this work is to investigate how a more realistic distribution of APs affects the efficiency of the LI architectures described.

VII. Conclusion

In this paper we have defined a ratio which can be used to measure how efficient an LI solution is at performing its primary goal of capturing 100% of a target's traffic. We define efficiency to mean that a distributed LI solution should optimise the amount of time that it activates interception in each of its clients when seeking to capture target traffic.

We used the ratio to compare the efficiency of a simple distributed LI solution, which uses a Server Driven Broadcast Activation (SDBA) architecture, to that of an architecture that we propose, Client Driven Targetted Activation (CDTA). Through simulation, we have found that CDTA offers an efficiency advantage, especially when appropriate caching mechanisms are used to control the number of messages transmitted and the activation time of each LI component.

LI efficiency is an important topic to study as an efficient LI solution offers a number of advantages. Firstly, savings in network bandwidth and network equipment resource consumption such as CPU cycles, router and switch memory usage. Secondly, an efficient LI solution also limits the amount of time that the LI system is susceptible to divulging the identity of the interception target. This latter advantage, not only satisfies one of the requirements on LI, it also indirectly increases the general public's security from criminal activity by making the knowledge about who is being intercepted less likely to be compromised.

References

[1] M. Rahnema, "Overview of the GSM system and protocol architecture," *IEEE Communications Magazine*, Apr. 1993.

[2] A. Rojas and P. Branch, "Lawful Interception based on Sniffers in Next Generation Networks," in *Proceedings of the Australian Telecommunications Networks and Applications Conference 2004*, Dec. 2004.

[3] M. Sherr, E. Cronin, S. Clark, and M. Blaze, "Signaling Vulnerabilities in Wiretapping Systems," *IEEE Security and Privacy*, pp. 13–25, Nov. 2005.

[4] D. Denning, "To Tap or not To Tap," *Communications of the ACM*, vol. 36, no. 3, pp. 24–33, Mar. 1993.

[5] S. Chan and L. Camp, "Law Enforcement Surveillance in the Network Society," *IEEE Technology and Society Magazine*, vol. 21, no. 2, pp. 22–30, 2002.

[6] A. Escudero-Pascual and I. Hosein, "Questioning Lawful Access to Traffic Data," *Communications of the ACM*, Mar. 2004.

[7] "Telecommunications (Interception) Act 1979," Parliament of Australia.

[8] "Internet Service Providers Interception Obligations," Australian Communications Authority.

[9] (2006) Omnet++ discrete event simulation system. [Online]. Available: http://www.omnetpp.org

[10] (2005) IPv6 Suite with INET Framework. [Online]. Available: http://ctieware.eng.monash.edu.au/twiki/bin/view/Simulation/IPv6Suite

[11] W. Navidi and T. Camp, "Stationary Distribution for the Random Waypoint Mobility Model," *IEEE Transactions on Mobile Computing*, vol. 3, no. 1, pp. 99–108, 2004.

[12] L. H. Gomes, C. Cazita, J. M. Almeida, V. Almeida, and J. Wagner Meira, "Characterizing a spam traffic," in *IMC '04: Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*. New York, NY, USA: ACM Press, 2004, pp. 356–369.

[13] M. Crovella and A. Bestavros, "Explaining world wide web traffic self–similarity, Tech. Rep. 1995-015, 29, 1995. [Online]. Available: citeseer.ist.psu.edu/crovella95explaining.html