Chapter 15

USING SIGNALING INFORMATION IN TELECOM NETWORK FORENSICS

T. Moore, A. Meehan, G. Manes and S. Shenoi

Abstract Telephones are often used to facilitate criminal and terrorist acts. The signaling core of public telephone networks generates valuable data about phone calls and calling patterns, which may be used in criminal investigations. However, much of this data is not maintained by service providers and is, therefore, unavailable to law enforcement agencies. This paper presents techniques for collecting and storing important data pertaining to phone calls and calling patterns. The techniques leverage existing telecommunications network resources, and address the long-term storage issues of massive amounts of call record data.

Keywords: Telecommunications networks, signaling messages, call detail records

1. Introduction

Detailed information about telephone calls and the calling patterns of suspects can be very useful in criminal and terrorism investigations. But the call detail records currently collected by service providers for billing purposes are not comprehensive enough, nor are they retained for more than a few months.

A modern telecommunications network incorporates a transport network that carries voice and data, and a vital signaling network core that controls and manages voice and data circuits in the transport network. Call setup and other messages that traverse the signaling network are routinely examined by service providers to maintain quality of service, debug network problems and generate billing records. These messages provide a wealth of forensic information about phone calls and calling patterns. Since signaling messages provide data about phone calls – not the content of phone conversations – collecting and analyzing these

messages may not be subject to the same legal restrictions as recording voice conversations.

This paper describes techniques for collecting detailed information about phone calls and calling patterns. The techniques can be implemented using current telecommunications network surveillance equipment (e.g., [9]), and the collected data can be analyzed and stored at little additional cost.

The following sections describe signaling networks and techniques for collecting signaling data pertaining to telephone calls and calling patterns. The storage requirements for the techniques are analyzed, and a post-capture data processing technique that addresses the long-term storage issues of massive amounts of call record data is proposed.

2. Signaling Networks

Public telephone networks incorporate a transport network that carries voice and data, and a vital (out-of-band) signaling network that controls voice and data circuits in the transport network. The Signaling System 7 (SS7) protocol and its variations are used worldwide in signaling networks [7, 10, 11]. SS7 is responsible for setting up calls and implementing advanced features such as calling cards and toll-free service. VoIP and wireless networks use different protocols, but interface to public telephone networks using SS7. The following subsections describe the SS7 network architecture, SS7 messages that provide forensic information, and strategies for collecting SS7-related call data.

2.1 SS7 Overview

SS7 networks have three types of components (signaling points): service switching points (SSPs), signal transfer points (STPs) and service control points (SCPs) (see Figure 1). Each signaling point has a unique point code for routing SS7 messages. SSPs are central office switches that connect phones to voice trunks (multiple solid lines in Figure 1); they also generate SS7 messages for call setup and database queries.

STPs receive and route SS7 messages between signaling points using out-of-band signaling links (dashed lines in Figure 1). In U.S. telephone networks, each SSP is connected to at least one mated pair of STPs [12, 13]. SCPs (not shown in Figure 1) mainly provide database access for advanced services, e.g., call forwarding and toll-free numbers; like SSPs, they connect to STPs via signaling links.

SS7 messages contain important control data about telephone calls, e.g., calling and called party numbers, as well as the time and duration of calls. Collecting this data does not require an inordinate amount

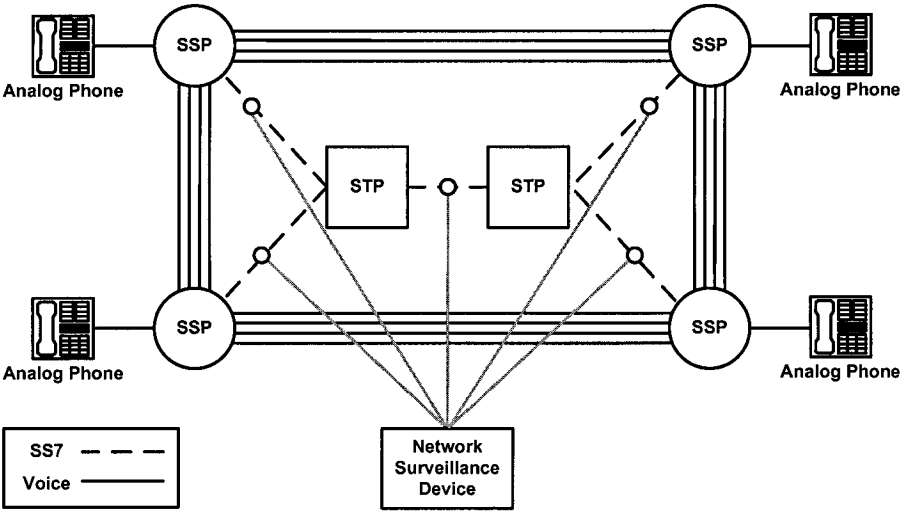

Figure 1. Signaling System 7 network.

of resources – many U.S. providers already employ surveillance devices [8, 9] to monitor SS7 links (see Figure 1). In U.S. networks, these devices are co-located at STPs as all signaling links pass through STPs. Some countries (e.g., U.K.) employ "fully-associated" networks, where SSPs are connected directly by signaling links (not via STPs as in U.S. networks). Such a topology makes it impractical to monitor every SS7 link.

2.2 SS7 Messages

SS7 messages contain valuable forensic information. Call setup is governed by the ISUP protocol [1]. One of the most important ISUP messages is the initial address message (IAM), which initiates phone calls. Other messages for setting up calls – address complete (ACM), answer (ANM), release (REL) and release complete (RLC) messages – also contain useful data. Some ISUP messages, e.g., block (BLO) messages that remove voice circuits for maintenance, are very powerful but are used rarely.

Database queries and responses are implemented by TCAP messages [4]. These general-purpose messages translate toll-free and ported phone numbers. Note that the formats of the message payloads vary considerably, so any analysis software must be tailored to specific applications. The underlying SCCP header in TCAP protocol messages, however, does provide useful routing and service information [3].

MTP messages used for network management provide data about network behavior and response [2]. They include transfer prohibited (TFP) messages sent by STPs during service outages, and transfer controlled (TFC) messages used to deal with network congestion.

Emergency communications systems – 911 and GETS in the United States – use the SS7 infrastructure. When a 911 call is placed, SS7 messages are sent to a dedicated 911 tandem switch to set up the voice path. From the signaling point of view, the only difference is that 911 call setup messages have higher MTP priority values than those for normal calls. Storing SS7 message traffic destined for a 911 tandem is a simple and effective way to record activity, especially for *post mortem* analysis in the event of accidental service outages or outright attacks.

2.3 Message Collection Strategies

SS7 networks generate much less traffic than IP networks; even so, discretion must be applied when capturing traffic for forensic purposes. We consider three strategies: collecting all messages, collecting messages based on protocol type, and collecting messages based on message type.

The simplest approach is to capture all SS7 message traffic. SS7 links have low bandwidth (56 Kbps); nevertheless, this strategy is too simplistic. First, large providers may operate tens of thousands of links, making the storage of all SS7 traffic prohibitive. Second, most SS7 messages are irrelevant for forensic purposes. For example, large numbers of special fill-in signal units (FISUs) are transmitted to maintain proper timing.

A better approach is to capture traffic for specific protocols, e.g., ISUP and MTP. This is achieved by filtering messages based on their service indicator octet (SIO) field. This approach is simple and computationally efficient. The vast majority of ISUP traffic deals with call setup, so little extraneous information is stored when generating call records. Moreover, most network surveillance devices can capture protocol-specific traffic [8, 9].

Since not all messages contain useful forensic information, it is prudent to capture only specific types of messages. The corresponding message capture algorithm must examine the SIO to verify the protocol and the first byte after the routing label to determine the message type. Only messages whose protocols and types are of interest are collected.

Many network surveillance systems support network-wide message correlation and storage for *post mortem* analysis [8, 9]. However, they filter traffic based on protocol type, not message type. Nevertheless, adding this capability is straightforward. While the storage requirements are reduced, more processing is needed to filter messages in real

Table 1. Basic call record data from call setup messages.

Call Record Data	ISUP Message
Incoming Calls	IAM, ANM
Outgoing Calls	IAM, ANM, REL
Call Duration (Hold Time)	IAM, ANM
Call Duration (Conversation Time)	IAM, ANM, REL

time. Applications that require only one or two specific types of messages to be collected would benefit from this strategy. On the other hand, applications that collect practically all the traffic for a given protocol, e.g., ISUP-based call record data for forensic investigations, need not use this approach.

3. Call Record Signatures

Law enforcement agencies are primarily interested in calls involving specific phone numbers. This information can be obtained from three ISUP call setup messages: IAMs, ANMs and RELs (see Table 1). Note that SS7 is only used to connect voice circuits of different switches (SSPs). Therefore, no SS7 messages are generated for a call to a number serviced by the same switch. Still, the benefits of obtaining historical data, even if only for inter-switch calls, are quite significant. See [1, 7] for details of the call setup process.

Initial address messages (IAMs) contain three key parameters: called party number, calling party number and circuit identification code (CIC), which indicates the voice circuit used for call setup. Only IAMs contain the called and calling party numbers; other call setup messages include just the CIC. The CICs in these other messages must match with the IAM's CIC for the messages to be correlated with the IAM.

The original called number parameter in an IAM also conveys useful information. When a ported or toll-free number is dialed, the called party number is set to the translated number. The original called number is the actual dialed number.

Figure 2 presents three call record signatures. To identify an outgoing call from A, it is necessary to observe the IAM sent from A's SSP and note its CIC value. Next, it is necessary to observe an answer message (ANM) sent to A's home SSP with the same CIC value. This IAM-ANM sequence reveals that a call is made from A to B (Signature 1).

A similar IAM-ANM message sequence identifies a call from B to A (Signature 2). An IAM is sent to A's SSP with the called party number set to A. An ANM is then returned from the SSP with the same CIC

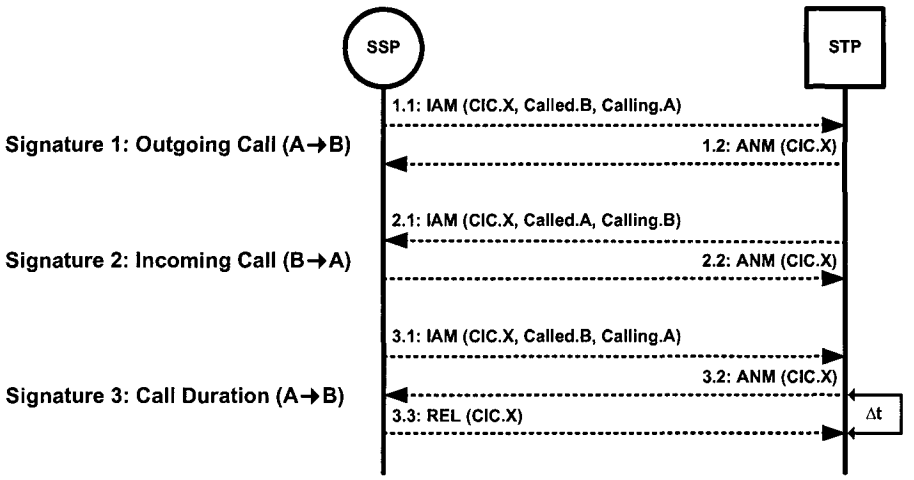

Figure 2. Basic call record signatures.

value. Since SS7 links are bidirectional, the direction of a message is inferred from its originating and destination point code parameters.

Call duration may be computed using IAMs and RELs or correlating IAMs, ANMs and RELs. The first method observes the release (REL) message and correlates it with the appropriate IAM. The call duration is then estimated based on the time difference between the generation of the IAM and the receipt of the REL. Because an IAM is generated when a number is dialed, this time difference overestimates the call duration (it includes the time spent waiting for the phone to be answered). To address this discrepancy, the second method also looks for ANMs with matching CICs. The time difference is calculated only after the called party answers and an ANM message is returned (Signature 3).

Table 2. Advanced call record data from call setup messages.

Call Record Data	ISUP Message
Unanswered Calls	IAM, ANM, REL
User Busy Failure	IAM, REL
User Release Direction	IAM, REL
Preemptive Release	IAM, ANM, REL

Analysis of signaling messages yields information about unanswered calls, calls terminated due to busy signals, preemptive hang ups by the caller, and preemptive hang ups by the receiver. Table 2 lists the ISUP messages that provide this data. Note that ISUP data are not collected by service providers and are, therefore, not available to law enforcement

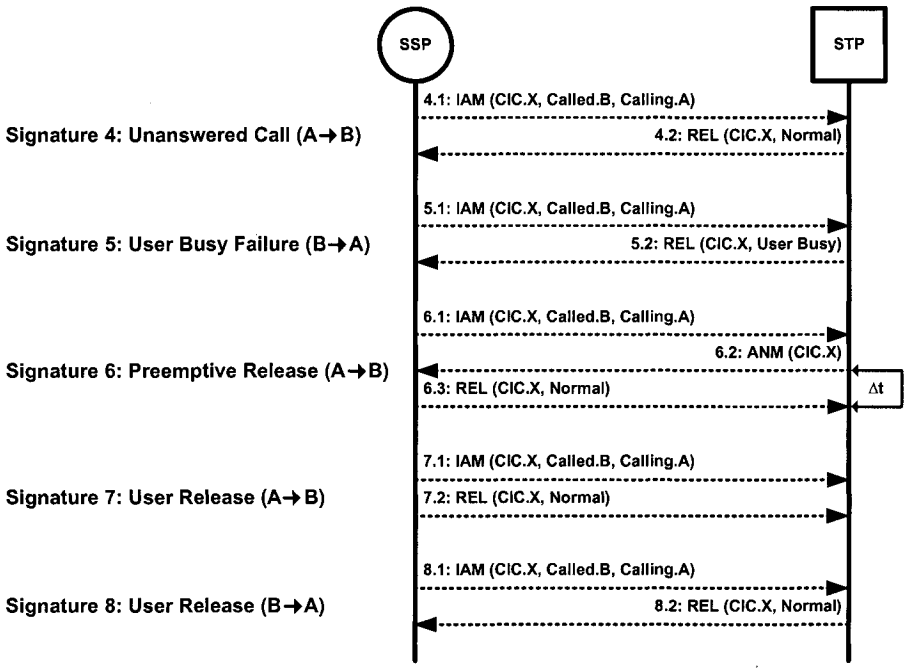

Figure 3. Advanced call record signatures.

agencies. For example, if an investigator has call records that contain only completed calls, and there is a record of a single call to an implicated phone number, the defendant could argue that he dialed the incorrect number. On the other hand, if the investigator has a record of all attempted calls, including numerous failed attempts to the implicated number before and after the completed call, it would be difficult to argue that a dialing error occurred.

Figure 3 presents five advanced call record signatures involving REL messages, whose cause code parameter indicates the reason for call termination. This parameter specifies whether a call completes normally, fails due to a busy signal, or ends because no one answers.

An unanswered call occurs when a caller hangs up before the phone is answered. Unanswered calls are detected by checking for IAMs, ANMs and RELs, even though the signature only involves an IAM and a REL (Signature 4). The presence of an ANM means that the call has been answered; therefore, an unanswered call does not have an ANM. The REL cause code is set to normal clearing because the caller hangs up and terminates the call normally.

Table 3. Storage requirements (full capture).

Links	Storage/Day
1	4.84 GB
100	483.84 GB
1,000	4.84 TB
10,000	48.38 TB

When a call is placed and a REL is received in response, the call does not complete. The attempted call is detected by correlating an IAM and REL message. The signature for a busy failure is an IAM followed by a REL with cause code set to user busy (Signature 5).

A preemptive release occurs when the caller hangs up immediately after the receiver answers – common behavior for pranksters and stalkers. The signature (Signature 6) is an IAM-ANM pair, followed by a REL within a short time period (1 to 3 seconds). Preemptive releases are also associated with call fraud. Many network surveillance devices record long distance calls with durations of a few seconds or less [9]. However, preemptive releases are distinct because hang ups occur in the initiating direction.

To determine who hangs up first, it is necessary to check the originating point code (OPC) and destination point code (DPC) of the REL message on the originating switch's SS7 link. If the OPC matches the originating switch's point code, then the caller hung up first (Signature 7). On the other hand, if the DPC of the REL matches the originating switch's point code, then the receiver hung up first (Signature 8). In both cases, the REL cause code is normal clearing.

4. Data Storage Requirements

As described in Section 2.3, the three options for collecting signaling traffic are: (i) full capture, (ii) protocol-specific capture, and (iii) message-specific capture. This section quantifies the storage requirements for each technique and presents a post-capture data processing technique that minimizes long-term storage costs.

Each SS7 link has a 56Kbps capacity. The storage requirement per link is also 56 Kbps (because fill-in messages are sent when regular messages are not transmitted). Table 3 lists the storage requirements for various numbers of links. The full capture technique is not feasible because most phone companies operate between 1,000 and 10,000 links.

Capturing only ISUP messages is a more reasonable approach because ISUP messages are mainly used for call setup. However, estimating the

storage required is difficult for three reasons. First, most – but not all – ISUP traffic is used for call setup. Second, all calls do not use the same number of messages. The number of messages depends upon the number of intermediate switches (nodes) along the voice path; typically, no more than six nodes are involved. Third, ISUP messages contain many optional parameters, most of which are never used. Therefore, the storage computation assumes that all ISUP messages are used for call setup and that these messages have only the optional parameters that are most commonly used.

Table 4. Storage requirements (ISUP message capture).

Number of Calls	2 Nodes	3 Nodes	4 Nodes	5 Nodes	6 Nodes
1 Call	256 B	512 B	768 B	1.024 KB	1.28 KB
1 Million Calls	256 MB	512 MB	768 MB	1.024 GB	1.28 GB
100 Million Calls	25.6 GB	51.2 GB	76.8 MB	102.4 GB	128 GB
1 Billion Calls	256 GB	512 GB	768 GB	1.024 TB	1.28 TB

In the simplest case, only two nodes are required to set up a call. Ten ISUP call setup messages are involved: an IAM (43 Bytes), ACM (17 Bytes), ANM (15 Bytes), REL (19 Bytes) and RLC (14 Bytes); each message is sent on two links. Table 4 indicates the storage requirements for various numbers of calls depending on the number of nodes involved in call setup. These figures incorporate 4-byte message timestamps. But they also assume that (aside from call setup) no other ISUP messages are sent. Therefore, the actual storage requirements could be 5% to 10% higher than specified in Table 4.

Table 5. Storage requirements (IAM-REL-ANM message capture).

Number of Calls	2 Nodes	3 Nodes	4 Nodes	5 Nodes	6 Nodes
1 Call	178 B	356 B	534 B	712 B	890 B
1 Million Calls	178 MB	356 MB	534 MB	712 MB	890 MB
100 Million Calls	17.8 GB	35.6 GB	53.4 MB	71.2 GB	89 GB
1 Billion Calls	178 GB	356 GB	534 GB	712 GB	890 GB

Since some ISUP messages, e.g., REL and RLC, contain redundant data, the storage requirements can be reduced by only capturing messages with useful data. Table 1 indicates that valuable call record data is contained in IAM, ANM and REL messages. The storage requirements for capturing these three types of messages are presented in Table 5.

Note that this technique has some inefficiencies. Often, a message traverses several links before it reaches its destination. As a result, identical messages are stored multiple times. It might seem reasonable to modify surveillance devices to store only single copies of a message. Unfortunately, the computational overhead outweighs any storage benefits.

However, post-capture data processing can significantly reduce the storage requirements without adversely impacting computational costs. The first step is to use the full ISUP capture or multiple-message capture technique. Next, the messages are analyzed and only the relevant fields from message sequences are retained.

Table 6. Post-capture call record attributes.

Attribute	Message	Size
OPC	IAM	3 B
DPC	IAM	3 B
Called Number	IAM	5 B
Calling Number	IAM	5 B
Call Start Time	IAM	2 B
Call Duration	IAM, ANM, REL	2 B
Received Answer Flag	IAM, ANM	1 B
Busy Failure Flag	IAM, REL	1 B
Preemptive Release Flag	IAM, ANM, REL	1 B
User Release Direction Flag	IAM, REL	1 B

Table 6 lists the important attributes contained in ISUP post-capture call records along with their size. Originating and destination point codes (OPCs and DPCs) are collected to assist in searching records. To track all calls to a particular number, one could naively search through all the records for the called and calling party numbers. However, the search effort can be significantly reduced by only looking for OPCs or DPCs with the point code of the home switch of the target number.

Table 7. Storage requirements for 1 billion calls.

Data Collection Technique	Storage
Full ISUP Capture	768 GB
IAM-REL-ANM Message Capture	534 GB
Post-Capture Data Processing	26 GB

Table 7 compares the storage requirements for the three techniques based on one billion phone calls. The average of each of the node costs is used for computing the storage requirements for the full ISUP capture

and the IAM-REL-ANM capture techniques. Note that the IAM-REL-ANM message capture technique requires significantly more storage than the post-capture data processing technique. This is because ISUP messages contain only a few parameters that are required for creating useful call records, and the same messages are repeated several times along the call setup path. The post-capture data processing technique is promising because it eliminates extraneous and redundant data.

5. Conclusions

Call records generated from signaling messages can be very valuable in criminal investigations. Signaling messages provide data about phone calls – not the content of phone conversations. Therefore, collecting and analyzing signaling messages may not be subject to the same legal restrictions as recording voice conversations. Signaling messages are routinely examined by service providers to maintain quality of service, debug network problems and generate billing records. Service providers could use these same messages to obtain data about user calling patterns that could be provided to law enforcement when authorized. This data can be collected using existing surveillance equipment [8, 9], and it can be analyzed and stored at little additional cost.

Collecting and archiving massive quantities of call setup traffic raises security and privacy concerns. Obviously, legal and technical safeguards must be implemented to prevent abuses by service providers and law enforcement personnel. Still, storing all call records for prolonged periods of time is almost Orwellian.

Hashing techniques can be employed to allay these concerns. Instead of storing call records, their hash values are computed and saved. Then a law enforcement agent could query the hash values, for example, to check if a certain phone number was called from a target phone number. Bloom filters [5] permit (precise) negative responses to queries, but affirmative responses only have an associated probability. These techniques provide numerical confidence levels for affirmative responses. Such information could assist investigations by ruling out certain theories or corroborating other evidence. The hashing techniques would not provide the same detail and level of confidence as storing (and querying) all relevant data – but they would certainly protect privacy.

A major technical challenge is introduced by the convergence of the public telephone infrastructure with VoIP networks. This convergence introduces new protocols and signaling points to the infrastructure [6]. Because these networks still interface using the SS7 protocol, monitoring the signaling links that connect VoIP systems to public telephone

networks can cover a significant amount of call traffic. However, VoIP calls routed exclusively over the Internet are difficult – if not impossible – to monitor because of the decentralized nature of pure VoIP networks.

References

[1] American National Standards Institute (ANSI), *T1.113-1995: SS7 Integrated Services Digital Network (ISDN) User Part*, New York, 1995.

[2] American National Standards Institute (ANSI), *T1.111-1996: SS7 Message Transfer Part (MTP)*, New York, 1996.

[3] American National Standards Institute (ANSI), *T1.112-1996: SS7 Signaling Connection Control Part (SCCP)*, Institute, New York, 1996.

[4] American National Standards Institute (ANSI), *T1.114-1996: SS7 Transaction Capabilities Application Part (TCAP)*, New York, 1996.

[5] B. Bloom, Space/time trade-offs in hash coding with allowable errors, *Communications of the ACM*, vol. 13(7), pp. 422-426, 1970.

[6] O. Ibe, *Converged Network Architectures*, John Wiley, New York, 2002.

[7] T. Russell, *Signaling System #7*, McGraw-Hill, New York, 2000.

[8] Tekelec, Integrated Application Solutions (www.tekelec.com).

[9] Tektronix, GeoProbe (www.tek.com).

[10] Telcordia, *GR-82: Signaling Transfer Point (STP) Generic Requirements*, Morristown, New Jersey, 2001.

[11] Telcordia, *GR-1241: Supplemental Service Control Point (SCP) Generic Requirements*, Morristown, New Jersey, 2001.

[12] R. Thompson, *Telephone Switching Systems*, Artech, Norwood, Massachusetts, 2000.

[13] J. Van Bosse, *Signaling in Telecommunication Networks*, John Wiley, New York, 1997.