The Quarterly Magazine for Digital Forensics Practitioners

DIGITAL FORENSICS / MAGAZINE

ISSUE 02

COMPETITION!
Win 3 State of the Art Sony Dictaphones

INSIDE

ANDROID ON THE LOOSE

Andrew Hoog unveils Google's new mobile operating system, showings us exactly what's important for forensic investigators

9 772042 061103 01
Issue 2 / £17.50 TR Media

SETTING STORE ON NEW DATA RULES

WHAT SECURITY PROFESSIONALS NEED TO KNOW ABOUT EU DATA RETENTION

***Mark Osborne** takes a personal look at the challenges posed to UK & European organizations by the 2006 directive*

/ INTERMEDIATE

EU Directives always cause debate. Think of the problems caused by the working-hours directive or the data protection directive. But the EU data retention directive has caused more than most. But there are reasons for transposing the Directive into UK law. As network complexity increases through such as BT 21CN it becomes harder to monitor communications traffic in the reasonably practicable manner allowed by the Regulation of Investigatory Powers Act 2000. 'Traditional' PSTN calls admit taping anywhere along the line but Next Generation Network packets travel myriad, frequently impenetrable paths. The Directive, in accordance with the Home Office's Interception Modernisation Programme, ensures lawful interception on demand, even in complex environments.

The title and reference is: 'Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC'1. What a mouthful! From here on we shall refer to it as the Directive.

Its' purpose stated at Article 1 is to harmonise Member States' provisions concerning the obligations of providers of public communications services or networks regarding data generated or processed by them so that it is available for the purposes of investigation, detection and prosecution of serious crime".

It sounds clear enough. And not unreasonable... at least, apparently not. Just look around at the many lucrative conference, online forums and journals that have spent some time conjecturing, discussing and arguing about the true meaning of it all.

There are many reasons cited for, but the main points of contention aree:

• Some people believe that the Directive is an infringement of their rights, and that it enables the authorities to bug all conversations and transactions on the Net. Add to the mix, the various reported comments from a number of authorities that the Directive's stated requirements fall short of what they really wanted, and that some of guidance documents explicitly ask for items that could not be justified by the directive, and you can certainly see a level of discord.

• People that wrote the directive claim that all the data required to be retained would be collected as part of normal business – yet much of the non-voice data is far from what is routinely collected. Indeed, it is expensive and difficult to obtain, especially when you consider the abundance of traffic on today's Internet. Such a massive change will inevitably cause confusion. Many engineers across EU are saying to themselves, "How in the Hell am I going to get that data – I must have read it wrong!".
• The Directive doesn't clearly identify what is meant by "the providers of publicly available electronic communications services". Many people that should comply are avoiding doing so because "it only applies to ISPs."

The original purpose of this article was to dispel these causes of confusion, particularly the second and third points, in the light of supporting documentation and the Directive itself. Initially, I had some sympathy for the arguments that the Directive was confusing.

However, on revisiting the directive to construct this article, I found myself believing the requirements to be relatively clear (by the standards set by other computer laws). Could it be that our industry relies on its legal requirements being spoon-fed to it by the very magazines, conferences and journals mentioned previously, and they, plus the activists from the first point above, could be engaged in successful campaign of obfuscation?

Perhaps too many years spent solving security problems, caused by "events" that everybody said would never happen, have left me paranoid. I will let you decide the reason for confusion – but confusion there is. Lots of time has been spent on the Civil liberties issue, and as much as I sympathise, I am a bloke with a job to do and these guys aren't really helping.

SO, LET'S GO STRAIGHT TO POINT 2: WHAT DO WE NEED TO STORE?

Data retention requirements are described in Article 5 of the Directive. As I have said, on my first reading, I found the requirements complex. However, I must have been suffering from some kind of mid-life crisis, as, on later reflection, the data requirements are actually not very difficult to understand.

The data requirements are sub-divided into two general types as they are specified. These are:

• Fixed network telephony and mobile telephony
• Internet access, Internet e-mail and Internet telephony

From this we can plainly see, with no hypothesis or extrapolation (don't worry, that will come!), that we are supposed to record information about four categories of network traffic:

SOME PEOPLE BELIEVE THAT THE DIRECTIVE IS AN INFRINGEMENT OF THEIR RIGHTS, AND THAT IT ENABLES THE AUTHORITIES TO BUG ALL CONVERSATIONS AND TRANSACTIONS ON THE NET

• Fixed network telephony and mobile telephony
• Internet access
• Internet e-mail
• Internet telephone

Fixed network telephony and mobile telephony

My expertise in this area derives from my sitting next to many experts in this field, which is a dubious qualification at best. That said, the stated requirements appear succinct and with some experience in the area, most telecom security pros will come up with a valid implementation. The stated retention requirements for fixed network telephony and mobile telephony cover:

• (A1i) the calling telephone number – as this will be a customer, the name and address of the subscriber or registered user;
• (B1i) The telephone number(s) called or the number end-point if it has been forwarded or transferred, routed; If it is your customer you should retain the name(s) and address(es) of the subscriber;
• (C1) the date and time of the start and end of the communication;
• (D1)the telephone service used;
• (E1) Where a mobile is concerned, data to identify the handset and its data necessary to identify the location of mobile communication equipment:
• The calling and called telephone numbers – surely a careless repetition in the document as these requirements are already stated
• The International Mobile Subscriber Identity (IMSI) of the calling party; (mobile only)
• The International Mobile Equipment Identity (IMEI) of the calling party; (mobile)
• The IMSI of the called party; (mobile only)
• The IMEI of the called party; (mobile only)
• In the case of pre-paid anonymous services, the date and time of the initial activation of the service and the location label (Cell ID) from which the service was activated; (mobile only)
• (f1) the location label (Cell ID) at the start of the communication;
• (f2) data identifying the geographic location of cells by reference to their location
• Labels (Cell ID) during the period for which communications data are retained

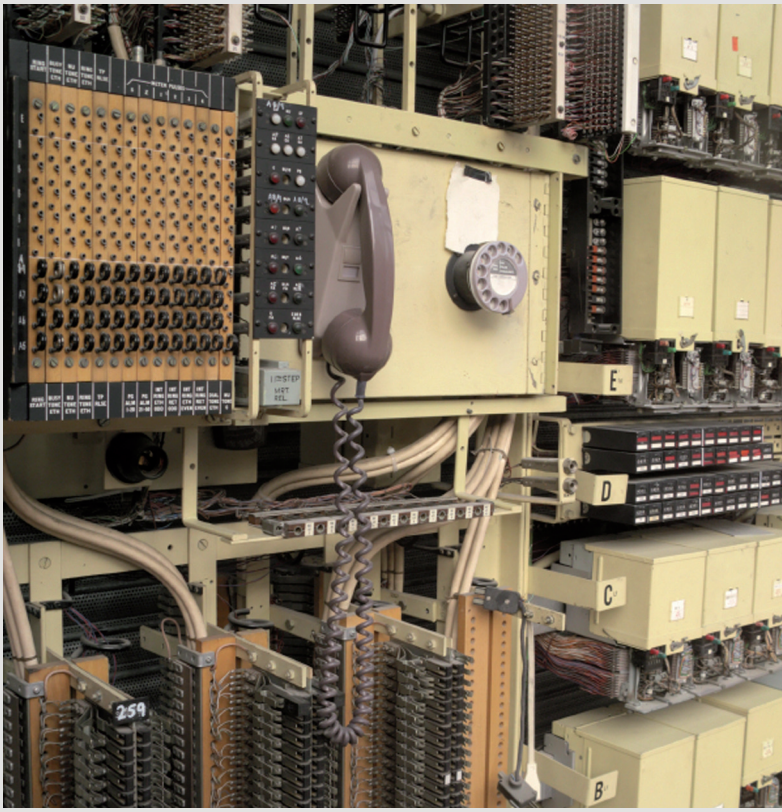
MOST AUTHORITIES HAVE NEARLY A HUNDRED YEARS OF EXPERIENCE IN TELEPHONE-RELATED LAW ENFORCEMENT

Having broken it down like this, it doesn't look too outrageous. In fact, anyone who runs a public voice system (please note this phrase), will know they are able to source nearly or all of this information from CDRs (call data records) produced by the switch manufacturers to allow call billing. In fact, they will probably be storing these for 6 to 12 months for billing purposes and will, no doubt, be used to providing this information for warrants.

As I said, these requirements are not outrageous at all. The truth is most authorities have nearly a hundred years of experience in telephone-related law enforcement. They understand what is available to a carrier and how it might be used. This isn't to say there isn't a whole host of problems that need to be dealt with, including those of jurisdiction, lawful requests and proportionality, and International number rationalization. But the major issues of this nut are well and truly cracked.

This is good, because you should have been compliant some time ago. So let's not spend any more time on it.

The aforementioned breadth of knowledge of the law-makers is not apparent when dealing with the authorities and their understanding of data networks, and particularly IP networks. So let's dissect the requirements for "Internet access."



/ INTERNET ACCESS

"Internet access" generally refers to the process of "How the connection from the subscriber to the Internet is performed". So here we should be looking at how a user will acquire an IP address, replete with routes, and how that user will talk to the Internet.

Under the section "a2) data necessary to trace and identify the source of a Communication" we can see the following data must be kept:

• The user ID
• The user ID and telephone number allocated to any communication entering the public telephone network;
• The name and address of the subscriber or registered user to whom an Internet Protocol (IP) address, user ID or telephone number was allocated at the time of the communication;

Most people reading this (1 and 2) will spot the repetition. Point Number 1 - they want me store the userid! Point Number 2 – they want me to store the userid!! – Thanks but I got it the first time.

Legacy providers will instantly think of SLIP and PPP used in the bad old days of dial-in access. If you are a 21st century access provider via WiFi or GSM or even provide Ethernet connections in public places like conference centres, this section will be very pertinent. But this information is easily and readily available your access servers which will record CHAP or PAP authentication information in a RADIUS or TACACCS+ system, even if some augmentation of the information from your RAS or DHCP system is needed. Traditionally, operators throw away the information after the transaction has expired – but not any more.

Paragraph a2 (iii) has a meaning that is clear: You need to be able to correlate the IP address, telephone number and

user id to a subscriber name and postal address. For a backbone provider, a provider that supplies fixed links with a fixed IP range, this is all you need to store.

There is no data relevant to how the internet access is gained in B2 so let's look at C2, which is mainly about time of internet access. It contains the following jumble of information:

"the date and time of the log-in and log-off of the Internet access service, based on a certain time zone, together with the IP address, whether dynamic or static, allocated by the Internet access service provider to a communication, and the user ID of the subscriber or registered user".

What this actually means is that when you login or log out, you need to record the IP and the user ID allocated plus the Date/time. For those of you seeking guidance, I suggest that you synchronise to a common time source and use UTC.

Section (D) is entitled "Data necessary to identify the type of communication" and for just Internet Access I would suggest we record this data as either a login event or logoff event.

Section E seeks to record the line and hardware used for access:

- The calling telephone number for dial-up access;
- The digital subscriber line (DSL) (or other end point) of the originator of the communication;

This is self explanatory.

Internet Access: summary

For each login or logout, we need to store:

- User ID, if available or Telephone Number, If available
- A link to the bill payers name and address
- The ip address
- The date-time
- xDSL line, Telephone number or other access media if no fixed access.

INTERNET E-MAIL AND INTERNET TELEPHONY

As before, if we re-analyze the storage requirements for section "a2" in terms of VOIP & Email, we see:

- The user ID
- The user ID and telephone number allocated to any communication entering the public telephone network;
- The name and address of the subscriber or registered user to whom an Internet Protocol (IP) address, user ID or telephone number was allocated at the time of the communication;

Items (i) and (ii) are straightforward:

- For each VOIP communication, they want you to store a record containing SIP URL, userid (if different but not usually) and a PSTN telephone number (if one is allocated/used (e.g.,
- For a skype-out type applications);
- For each Email sent, they want you to store a record containing Email address and userid (if different).

In Section B, the document simply refers to the destination(s) of a given VoIP call or any sent email. They require you to store:

- The user ID or telephone number of the intended recipient(s) of an Internet telephony call;
- The name(s) and address(es) of the subscriber(s) or registered user(s) and user ID of the intended recipient of the communication.

So:

- For each VoIP communication, they want you to store a record with a field-containing destination SIP URL, a user ID (if different but not usually) and a PSTN telephone number if one is allocated/used (e.g., for a Skype-out type applications).
- For each Email sent, they want you to store a record with a field containing destination Email addresses (if there is more than 1) and a user ID (if different).

Para (ii) is only relevant if you host source and destination address.

Section C contains the relevant paragraph:

(ii) the date and time of the log-in and log-off of the Internet e-mail service or Internet telephony service, based on a certain time zone;

YOU NEED TO BE ABLE TO CORRELATE THE IP ADDRESS, TELEPHONE NUMBER AND USER ID TO A SUBSCRIBER NAME AND POSTAL ADDRESS

So:

- For a SIP based VoIP service, they want you to store a record showing when a user registers with a SIP server, which is the equivalent to a login. SIP will treat a Register with an expiry=0 as a logout, but most servers rely on the registration timing out;
- For each Email post office service like POP3 or IMAP, you need to record when you login or logout.

"Section D - data necessary to identify the type of communication". As before, to satisfy this requirement we need record the protocol. This would be:

- SIP/RTP
- SMTP
- POP3
- IMAP

Please note for the purposes of simplicity, I have ignored other VoIP protocols such as H323, SKYPE and have concentrated on SIP/RTP.

DATA RETENTION STORAGE SUMMARY

	EVENT	NAME / ADDRESS	TIME	USERID	ALLOCATED / SRC IP	PSTN BREAKOUT	SRC ADDR	DEST ADDR(S)	OTHER INFO	COMMENT
INTERNET ACCESS										
DAIL-IN OR BROADBAND	LOGIN	X	UTC – at record time	X	X				The CLI or DSL identifier	
	LOGOUT	X	As above	X	X					
WIFI	ASSOCIATE	X	As above	X	X					
	DISASSOCIATE	X	As above	X	X					
INTERNET TELEPHONY										
SIP/RTP	REGISTER	X	As above	X	ADVISABLE	X	SIP URL (userid)			SIP has no LOGIN or LOGOUT
	INVITE	X	As above	X	ADVISABLE		SIP URL (userid)	SIP URL	Dst IP/port	
	BYE/ CANCEL	X	As above	X	ADVISABLE		SIP URL (userid)	SIP URL		
INTERNET E-MAIL										
SMTP	RCTP TO / MAIL FROM	X	As above		ADVISABLE		E-MAIL ADDRESS	E-MAIL ADDRESS	Dst IP/port	SMTP has no LOGIN or LOGOUT
POP3	USER	X	As above	X	ADVISABLE		E-MAIL ADDRESS			LOGIN
	QUIT	X	As above	X	ADVISABLE		E-MAIL ADDRESS			Must be carried forward from the login
IMAP	LOGIN	X	As above	X	ADVISABLE		E-MAIL ADDRESS			
	BYE	X	As above	X	ADVISABLE		E-MAIL ADDRESS			LOGOUT

The table above summaries fields within a packet or protocol command verbs that can be used to satisfy the "non-technical" terminology used in the directive

/ SO WHERE ARE THE PROBLEMS?

Q If I just provide Internet Access and Colo, How can I get access to the customers Email logs to record the information?
A Good news: you don't have to. The Directive states in para 13 that:

"Data generated or processed when supplying the communications services concerned refers to data that are accessible. In particular, as regards the retention of data relating to Internet e-mail and Internet telephony, the obligation to retain data may apply only in respect of data from the providers' or the network providers' own services."

So if you are an ISP that provides a user an Internet connection, you don't have to raid your customer machines to get their VoIP logs.

Q I am a backbone provider but I provide an SMTP relay for use by my Internet access customers?
A You are providing an Email service to a Public Network so the Directive applies to you. You cannot provide Login/Logout events but you must provide details of your send events.

Q I have seen documents that refer to Web traffic or IM.
A There are a number of current UK Government documents predating the Directive that specifically asked for more. For example, the document called "Retention Of Communications Data Under Part 11: Anti-Terrorism, Crime & Security Act 2001 Voluntary Code Of Practice" 2 asks for the retention of Web and IM traffic. Clearly omission of IM from the Directive was a mistake. Some parties try to rectify that mistake by claiming that IM is a form of Email.

Q For how long should I keep the data?
A 12 Months is a good starting place. I suspect that it is unlikely that the "prescriptive online" requirement would apply to data after this period so archive it to secondary storage to be on safe side – it will only cost you one tape cycle.

Q To whom does it apply?
A The Directive states that it applies to "the providers of publicly available electronic communications services or of public communications networks". In general, most people assume that it refers only to ISPs and Telcos. But think: the Directive is designed to provide evidence against serious crime – Terrorists or MafiA These are people who are mobile and technically savvy – they are going to be on the move.

The authorities are well aware of this, so the Directive is worded to include anyone who provides an Internet connection to the Public. I know this broader definition hasn't been widely considered. However, it must cover more than Telcos and ISPs, if the community at large are to derive any benefit. This means it should cover access providers in the form of:

• Hosting companies;
• Internet Cafes & Wireless Hot Spots;
• Ethernet providers in conference centres.

Most of the people committing serious criminal activities and crimes against humanity will use these in preference to a registered fixed line to Dr Evil's HQ There is some evidence to support this assertion. France and Italy have already issued guidance or legislation that incorporates this broader scope. In fact, Italy has already passed legislation that requires Internet Cafes with more than 3 terminals to comply.

Additionally, many leading HotSpot software providers now advertise EU data retention features as standard. These commercial companies would not have developed these features if it hadn't been necessary. Typically, the UK authorities have not issued anything since the Act referred to above, so watch-out for some future statement.

Q Does the Directive impact civil liberties?
A Only time will tell – but my best bet is it may, but not in the way predicted by the campaigners. Some campaigners have suggested that the Directive endorses eavesdropping. It does not. The last paragraph of Article 5 of the Directive strictly forbids it:

"No data revealing the content of the communication may be retained pursuant to this Directive."

Article 7 goes on to re-enforce the security requirements. The data must be stored with better security than that prevailing when it was transmitted and with due consideration of data-protection requirements. So our rights to private communication are not damaged in that respect.

If there is a risk, it is that the data will become used to prosecute more mundane crimes. Currently, this is the case for RIPA3 in the UK where the legislation is often not used only to protect society from so-called High-Crimes but for the pursuance of relative misdemeanours – recently publicised cases show that local authorities have been using the legislation for very minor cases like dogs fouling public parks (which is disgusting but hardly a serious crime).

Most CISO or CIOs that work in service providers will have assisted the authorities with enquiries. Personally I am happy to do so when presented with the correct documentation. Unfortunately but frequently, this is not presented as necessary with the initial requests. It is my duty under law and as a member of a number of profession associations not to provide the information until the correct warrants etc are in place. It is my belief that sufficient controls are in place but the government representatives requesting the information AND the corporate officers providing the data needed to be reminded of the gravity of the process. This opinion was reinforced by an encounter with a bunch of government & police analysts at an "executive briefing" at the beginning of the year. Some fairly capable Data Retention software that had been developed for a large mobile provider was being demonstrated to a dozen provider representatives. And all was going well until the sponsor declared his intention to allow certain agencies to retrieve the data directly. I represented that with out dual-authorisation from both LEA management and service provider management, the service provider would be failing to meet their obligations (legal & moral) to the customer. One of the analysts suggested in a far too glib manner that should wrong information be retrieved a simple "notify" under the data protection act would make everything better. I hope the inevitable but unfortunate victim feels the same way!

Fortunately others in authority understand the subject better and take matters more seriously. Sir Paul Kennedy, in his recent report4 notes that most of the reported 55 errors in data intercepts resulted from simple typos and don't have a damaging impact. However, he acknowledges the potential for damage through control failure and in his words describes the impact as potentially "Catastrophic".

CONCLUSION

The directive on data retention could be a powerful tool to protect us all against serious crimes – and have a minimal impact on our freedoms as long as we concentrate on the obvious flaws and don't just jump on the bandwagon. However, like most computer and security law the people writing the directive would have benefitted with a better knowledge of the protocols and telecoms operations. This produces confusion and correspondingly too much idiosyncratic interpretation in the areas of the Directive that impact newer Internet technologies.

A draft schema for a storage model can be derived from the information in this article that could serve an operator well or at least provide a good starting comparison when considering commercial offerings.

References:
1. Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC. Available at http://eur-lex.europAeu/LexUriServ/site/en/oj/2006/l_105/l_10520060413en00540063.pdf
2. UK Acquisition and Disclosure of Communications Data - a Code of Practice, HMSO
3. UK Regulation of Investigatory Powers Act 2000, Office of Public Sector Information http://www.opsi.gov.uk/acts/acts2000/ukpga_20000023_en_1
4. Report of the Interception of UK Communications Commissioner for 2008

AUTHOR BIO

Mark Osborne ran the KPMG security practice for many years (1993-2003). He has published several Zero-Day security vulnerabilities (e.g. Fatajack), and has also been an expert witness in the "cash-for-rides" case. Mark has designed the popular open-source wireless IDS/IPS (WIDZ), as well as the largest Cyber Security System in Europe. He is the author of "How To Cheat at Managing Information Security", which reached the Amazon.com Top-500.