The Investigatory Powers Act 2016 and Internet Connections Records:
some surprising truths?

Authors
Daniel Coats, MSc (Royal Holloway, 2016)
Peter Komisarczuk, ISG, Royal Holloway

Abstract
Internet Connection Records were one of the key aspects of the recent Investigatory Powers Act 2016. This was a new term for what had been previously referred to as “communications data”, that is, data sent along with a message but not the message itself. Access to this data, which is already allowed and undertaken, by law enforcement is seen as a crucial part of their evidence and investigatory capabilities. Civil libertarians see this as further encroachment of an individual's privacy. Close examination of the Act revealed some surprising truths between these two opposing viewpoints.

2016 was a tumultuous year. A spectacular summer Olympics, an unexpected UK referendum result and the election of the outsider candidate of Donald Trump to the presidency of the United States of America. These were the stories that dominated the newspaper column inches, or more accurately we read about on our smartphones and tablets. As such, many of us, including some in the information security community, may have missed the Royal Ascent of the Investigatory Powers Act onto the UK statute book. Certainly, the Bill did not register highly in the public consciousness, other than some headlines referring to the Act as the “Snoopers Charter”.

The Investigatory Powers Act 2016 is a mammoth piece of legislation covering a myriad of legislative areas. One of the standout parts of the Act is the introduction of Internet Connection Records (ICRs) which allow UK law enforcement, under certain circumstances, to request communications data from Internet Service Providers (ISPs). The aim of this article is to inform on:

- What actually are ICRs and what do they reveal?
- What is the purpose of ICRs?
- Are they likely to work in practice?
- Can they be avoided?

In doing so, we will highlight some surprising findings and discover if the criticism and the applause for ICRs are justified or not. The two conflicting views can be summarised as those who see ICRs as a necessary tool for law enforcement in a modern age, and those who view them as the incursion of the State on centuries old libertarian values.

What are ICRs?
To start with, the term initially caused much confusion. This is because ICR is not a term recognised by the computing industry. In fact, the term has been specifically created for the Act. ICRs are concerned with communications data. Communications data is the data that is acquired or derived from
transmitted data but not the actual content the message. This may contain data about the destination,
origin, time or type of underlying service used. For instance, communications data could tell that I had
looked at the NHS website but not the pages I had visited. Communications data might also be able to
tell that I had sent an email to a certain device and at a certain time but not what the contents of that
message was. ISPs have always kept records of some communications data for business purposes
(see Figure 1) but the Act obligates them, when served with a notice, to keep that data for 12 months.

Figure 1: An example of the type of information held by ISPs. (Written evidence to the Joint
Committee on the Draft Investigatory Powers Bill - National Police Chiefs Council, HM Revenue
committeeevidence.svc/evidencedocument/draft-investigatory-powers-bill-committee/
draft-investigatory-powers-bill/written/26427.html)

Why ICRs?

ICRs are concerned with using that communications data (see Figure 2) to enable law enforcement to
find out two key pieces of information:

- The identification of a device, and/or
- The identifications of services accessed.

Figure 2: How an ICR might look. (Written evidence to the Joint Committee on the Draft Investigatory
http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/
draft-investigatory-powers-bill-committee/draft-investigatory-powers-bill/written/
26427.html)
Why would law enforcement want to be able to access this communications data? The effect of the Internet on the way we communicate is so profound that many of us miss the generational change in the way we communicate. Many of us, in our lifetime, have seen the landline and fax (remember those?) give way to the desktop computer and the smartphone. Many of us now communicate almost exclusively by text, email, social media or messaging apps. In this space, law enforcement and the government talk about how criminals have embraced this technological change to communicate in different ways - the language used around the Act talks about retaining “lost capabilities” and in turn, the critics of the Act criticise the new “snooping powers of the state”.

However, both sides of the argument may contain exaggerations. In terms of Internet Connection Records, the Act produces no new capabilities. It should be noted that the ability to access this communications data is not new and has existed under legislation for some time. The much heralded 12-month retention period of data already existed under existing legislation, albeit emergency legislation that was due to expire in December 2016. The identification of a device, using IP resolutions already exists under counter terrorist legislation - in remarkably similar language to that used in the Act. Indeed, even the head of the National Crime Agency (NCA) giving evidence to Parliament at the scrutiny stage, accepted that ICRs brought no new capabilities to law enforcement.

Therefore the Act, in relation to ICRs, is merely consolidating powers that exist in other areas of legislation. This is important to consider, for both sides of the argument. It is equally incorrect to say that ICRs are more intrusive as it is to argue that ICRs retain lost capabilities.

What is clear is that under the previous mixture of legislation, law enforcement was requesting a lot of data. Certainly some 500,000 request for communications data are made on an annual basis. Proponents of ICRs would point to the fact that communications data is used in 95% of serious and organised crime cases. In conducting the research, it was certainly easy to construct examples where communications data would be invaluable, for example, in a missing person case or involving accessing child pornography - an illustrative example is included in Figure 3.

*Figure 3: An example of the use of ICRs.*
What was not easy to deduce from the actual statistics, is the utility of that communications data in the vast majority of cases. Simplistically, smartphone usage has become extremely widespread over the past decade for both criminals and law abiding citizens. However, crime detection and conviction rates have broadly remained stable despite access to communications data. So whilst communications data may be useful it may not necessarily be crucial. An important distinction. Indeed, the more difficult question we need to ask is “in what proportion of cases is communications data the sole determining factor in detection or conviction?” We don’t know the answer to this question. If the proportion of cases is low, then perhaps we need to question the utility of the collection of communications data.

The Danish experience

Interestingly this question has been asked and answered already, in a broadly similar scheme run by the Danish Government. The first point to make and to make clearly, is that the Danish use of session logging is not exactly comparable to ICRs. Certainly the aims were similar but the data collecting methods differ. The second key difference was the cost which in the Danish model, was borne by the ISPs, whereas in the UK model, the costs are reimbursed by the government. The cheaper solution in the Danish example was not the most effective solution. However, given that it is the nearest comparable scheme, it is worth taking note of the schemes failures and successes.

The Danish scheme ran from 2007 until 2014 before it was stopped. It was not stopped as the result of any civil liberties outcry but by the Ministry of Justice itself. The Ministry of Justice found that the retained records were of little use to law enforcement, a finding which was supported by the Danish Security and Intelligence Service. Specifically, it was found very difficult to identify devices, particularly for the growing use of mobile devices. Will the UK scheme fare much better?

The aims of ICRs

The main aims of accessing communications data is to identify a device and/or find out what services that device was accessing. These are two very simple aims but we should be wary of concluding that these aims are easily achieved. The main reason for this is that whilst our understanding of how the Internet operates is fairly simple, the actual workings of the Internet is highly complex.

Identification of devices

To illustrate this better, let’s take the issue of device identification, which relies on access to communication. For example, if I sent you a threatening message, we can use the communications data to work out from which device that message was sent from. This is certainly useful but is not as easy it seems. Most of us our aware that each device has a unique number or address - an IP address is assigned to identify the device and that can be associated with additional information such as the postcode, the mobile IMEI etc. However, the growth of the Internet meant that we were going to run out of IPv4 addresses, so a method called Network Address Translation (NAT), was created that enable a single public IP address to hide a large number of private IP addresses (these are a small range of IP addresses that are re-used around the world typically in home and business networks). There can be multiple levels of NAT - for example in your home network to your home gateway and then within your ISP you may also have NAT functionality deployed in gateways or middleboxes (networking devices that inspect, filter, manipulate or transform network packets as they flow through a network). This presents a major problem for law enforcement as they need to narrow down how that large number of private IP addresses are connected to the Internet through a variety of NAT gateways. This is possible by using identifying information such as port numbers and timestamps, and other information provided in plain text such as unencrypted web pages. It is this critical information that ISPs would have to retain when served a notice. Does this mean that all devices can be identified? No, not really. The more accurate answer is that it is possible to identify devices through accessing communications data but there are issues that make it extremely difficult, if not impossible, in some cases. Certainly,
the Home Office during the passage of the then Bill made clear that the issue is complex and draws back from the idea that this would be possible in all cases. The difficulties in IP resolution was one the main reason that the Danish abandoned their scheme.

**Identification of services accessed**

Other problems exist in the area of identifying services a device has accessed. Again on the face of it, knowing which site a criminal has accessed would seem to be of great use, for example, accessing airline booking sites could give evidence that the suspect is about to leave the country or provide the necessary intelligence to approach an airline directly.

One of the major problems is the near-permanent connections created for services like Twitter and Facebook. The very nature of these applications that alert their users to messages and updates, requires near-permanent connection to push content. However, that does not mean that you are accessing those services 24 hours a day. Is the knowledge that I connected to Twitter for most of the day actually useful?

We should also note that there is potential for abuse of this or at least confusion. Many websites contain advertising networks, or potentially more incriminating networks, that could show that you had visited a site which you hadn’t knowingly accessed. Thus it would be dangerous to rely on ICRs alone as evidence.

An interesting but very small scale experiment was conducted by a software developer and reported. He collected the address and port data for each Internet connection from his own firewall. This is the communications data that the government could request from an ISP. He was curious to see what this communications data would or would not reveal. He was able to identify an email sent to his account from a desktop computer but not from his smartphone. Similarly, he was able to identify an email account but not the webmail service he had used (such as Gmail). Where he had accessed services, he was surprised by the number of times a content delivery network was shown. In his case “Akamai”. This is a server that hosts other users content that is cached by Akamai for one of their customers at a location near to you to optimise the user experience. For example a request for the “New York Times” may well be hosted by a UK server and only this would be shown. It was also interesting to note that in his experiment none of the messenger apps he used were identified (such as Skype). Clearly this is a limited experiment based on one firewall device and the individual is simply collecting port data from the firewall, however it does illustrate the difficulties in making sense of communications data - a panacea it is not.

**Evading ICR**

It should be noted that not all ICRs notices served on the ISPs will be same. The government will take note of the “technical capabilities” of the ISP before issuing a notice. Therefore, not all requests for ICRs will result in the same level of data and confidence in that data. Indeed, in evidence to Parliamentary committees, the Home Office conceded that not all of the 200-300 ISPs in the UK are likely to be served a notice. Does this mean we will see individuals seeking out obscure and smaller ISPs in order to evade the threat of ICRs? The reality of ICRs, as the government concedes, is that they can be easily avoided.

The Achilles heel of ICRs and a real question mark of their utility, certainly for serious criminality, is the ease of their avoidance. This can be from the basic to the more complex but all within the grasp of the majority of technology savvy users.

The most obvious example is to shroud yourself in anonymity. As previously mentioned one of the aims of using communications data is to identify a device. This is commonly done by using the IP address and other communications data. By using anonymous communications software such as TOR, you can hide your location and identity from anyone conducting surveillance or traffic analysis on the internet, although there have been some experiments to show that TOR can be infiltrated and with sufficient compromised servers can be used to identify users. Anonymity networks such as TOR are well known
and used by a wide variety of users including terrorists and criminals alike but also journalists and individuals wishing to evade government censors.

TOR at present is a fairly niche application, but there is measurement evidence (see Figure 4) that its use increased significantly in 2013 to peak at 6 million which has been linked to the Edward Snowden release of information such as the NSA PRISM program. However TOR usage worldwide has since reduced down to circa 2 million (January 2017). Whether legitimate users of the Internet in the UK will start to use TOR and other similar services more often as a result of the introduced act, remains to be seen. Current data indicates that TOR users in the UK were approximately 50K in December 2016 and have risen to approximately 120K at the end of January 2017. There is no evidence to indicate that this is related to the introduction of the Investigatory Powers act.

Virtual Private Networks (VPNs) offer similar but different protection to communications data. A VPN allows a user to send his data via a third party before routing to an ISP. The data the ISP receives, including most of the communications data, is encrypted. So any deep packet inspection is useless unless you have the means to break the encryption. A crucial difference between TOR, which provides a decent level of anonymity, is that VPNs provide confidentiality end-to-end. With a VPN, the source and destination IP address is still visible but the remaining data is confidential, however there are techniques that can identify what sort of data is being sent based on their typical characteristics.

A similar use of confidentiality is provided by many of the social messaging apps such as “WhatsApp” or “iMessage”, which provide end-to-end encryption. This means that someone can only read the message if they can intercept messages and break the encryption; we have seen recently the claimed backdoor in “WhatsApp” that allows the encryption key to be changed, a feature that could allow message confidentiality to be compromised, although no evidence has been presented to say that this has been done.

It is important to remind ourselves that ICRs are only concerned with the communications data and not the content of the message, which would be covered by legal intercept. The ICRs would only reveal that you had sent a message and identify the destination. The practicality of this has been discussed in the evidence presented, as have the limitations to the gathering and usefulness of such data.

**Conclusion**

In this discussion we have identified that ICRs introduce no new capabilities to UK law enforcement that did not already exist. ICRs are unlikely to significantly contribute beyond the capabilities already available under the previous regime for accessing communications data. We have shown how ICRs are concerned with communications data which is used to identify a device and/or the services ac-
cessed, and that this communications data is increasingly requested. We have noted that there are potential problems in collecting communications data and that it is not possible in all cases to make meaningful conclusions from such data collection. The Danish experience was not a success and this must highlight the problems in this area, however the UK model put the financial burden on the government not the service provider.

What is unclear from the research is the criticality of that communications data for law enforcement, for example, what percentage of cases rely on communication evidence as a major line of enquiry or for conviction. This is worthy of much more detailed research if only to reassure an increasingly critical general public.

The potential evasion of ICRs by use of readily available techniques must also raise a concern. If only the technologically incompetent or negligent are likely to be convicted by utilising ICRs, then we must ask the question “are we using a sledgehammer to crack the nut?”

**Biographies**

*Daniel Coats* served as a Royal Marine officer after completing a law degree. He then moved into security consultancy. In this role he has worked in many countries but predominately North and West Africa. In 2015-2016 he completed his MSc in Information Security at Royal Holloway University with his dissertation focused on the role of ICRs in the then Investigatory Powers Bill. He remains working as a security consultant with a special interest in Information Security.

*Peter Komisarczuk* is a member of the Information Security Group at Royal Holloway, University of London, where he is Programme Director for the MSc Information Security (Distance Learning). He is a chartered engineer and has a PhD (Surrey), researches in networks and security and has worked in industry in various R&D roles at Ericsson, Fujitsu and Nortel Networks.