Do European data protection laws apply to the collection of WiFi network data for use in geolocation look-up services?

Mark Watts*, James Brunger**, and Kate Shires***

Abstract

- Digital mapping companies have built WiFi maps that provide a useful means for WiFi enabled devices to determine their current location. These WiFi maps were created by collecting WiFi network data publically broadcast from WiFi networks and associating such data with a calculated location.
- WiFi maps do not fall within the scope of existing European privacy legislation. They do not consist of 'location data' as defined by the E-Privacy Directive, nor do they consist of 'personal data' under the Data Protection Directive, except in highly unusual and very rare circumstances.
- Distinctions can be made between IP addresses and WiFi network data. As such, the Article 29 Working Party's reasons for considering IP addresses to be personal data are unlikely to apply to WiFi mapping.
- A growing industry would be hindered significantly if WiFi maps were to be designated as personal or location data, with knock-on consequences for the many geolocation services that rely on the WiFi maps. Digital mapping companies would find it extremely difficult to comply with European privacy legislation in practice, even though any risk to individual privacy posed by WiFi mapping is negligible.
- Open discussion on this complex issue should be initiated between all interested parties to ensure a proportionate, industry-wide response and to avoid individual countries' regulators adopting conflicting positions across Europe.

Introduction

Geolocation services and location-aware software applications have become increasingly popular over the last decade both online and on mobile phones. They cover a vast array of services that include mapping and navigation (such as Google Maps and Ovi by Nokia), social networking (such Facebook and Foursquare) and security for lost or stolen mobile phones (such as HTCSense.com). Such services can tell users (and if they wish, others) where they are and allow them to receive content relevant to their location without having to manually enter the address or postcode. Such are the benefits and convenience of geolocation services and location-aware applications, they are fast becoming an essential and expected aspect of being online for many users.

The first real growth in commercial location-aware services occurred in the early 2000s, with a variety of offerings being launched by cellular network operators around the world.[1] These services often relied on approximated location data provided by network operators calculated from information about a mobile handset's proximity to the cellular network's fixed base stations. Location information was also sourced from GPS satellites, and this is increasingly the case with many of today's smart phones making use of a built-in GPS receiver. However, while GPS can determine a location more precisely, it is only available where a signal can be received and typically cannot be used indoors. GPS receivers can also require a significant amount of power, quickly draining the limited batteries of mobile devices.[2]

The digital mapping industry has provided a third option for WiFi-enabled devices. There has been a

* Partner, Bristows, 100 Victoria Embankment, London EC4 0DH, <www.bristows.com>.

** Associate, Bristows.

*** Trainee Solicitor, Bristows.

1 Table 3, Location Based Services for Context Awareness—Moving from GSM to UMTS (Anthony S. Park, Steffen Lipperts, and Marc Wilhelm), <http://ssgrr2002w.atspace.com/papers/143.pdf>.

2 Jeongyeup Paek, Joongheon Kim, and Ramesh Govindan, 'Energy-Efficient Rate-Adaptive GPS-based Positioning for Smartphones' (14–18 June 2010) 10 MobiSys, <http://www.inf.ed.ac.uk/teaching/courses/cn/papers/raps.pdf> last accessed 12 May 2011.

huge growth in domestic and business WiFi networks[3] and commercial WiFi hotspots.[4] Significant parts of the world are now covered by at least one WiFi network. Many digital mapping companies have made extensive efforts to 'map' the locations at which each WiFi network can be received.[5] As a result, these digital mapping companies and their licensees are able to offer location look-up services which can tell a user where they are by noting the WiFi networks that are visible to that user's WiFi-enabled device and matching those networks to entries in a WiFi network map.

In order to match WiFi networks, a WiFi network map is typically populated with certain WiFi network data, for instance SSID (Service Set Identifier), BSSID (Basic Service Set Identifier), and MAC (Media Address Control) addresses. These data are found in the header of each communication transmitted over a WiFi network and are distinct from the content or 'payload' of such communications. Such data are publicly broadcast by every device that is WiFi-enabled and is accessible to anyone and any device within range, regardless of whether the WiFi network uses encryption. Such data are routinely collected and processed by WiFi-enabled devices every day, often automatically and without their users being aware. However, WiFi network data are only intended for use within the WiFi network and, unlike, for example, IP addresses, are not generally communicated across the Internet.

There is a lot of uncertainty over how this type of publicly broadcast data from WiFi networks should be treated and in particular if they should be considered 'personal data' for the purposes of international data protection and privacy laws. The collection of WiFi network data for the Google GLS database, for instance, has been subject to investigation by certain European data protection authorities, and the authors of this article are assisting Google in connection with these investigations. This article is meant as a contribution to this important debate and to raise awareness of the issues. It considers whether the collection of such publicly broadcast WiFi network data to build a geolocation database service presents a risk to individual privacy. It also considers whether this activity falls within current European data protection and e-Privacy regulations and the consequences for the digital mapping industry if it does.

WiFi mapping data

In order to build a map of WiFi networks, some digital mapping companies have literally travelled the world. Google's Street View cars are well-known, but there are also, for instance, Skyhook Wireless cars and scooters.[6] In general, these various services 'tag' information about a WiFi network with the GPS position of the vehicle recording the data. These data are then analysed and used to build up an approximate location 'map' of WiFi networks that were broadcasting at the time the vehicle went past.

Digital mapping companies can also populate their database each time their location services are used. For example, when a GPS-enabled smart phone communicates with a location service, the GPS coordinates (if available) of the phone may be sent to the digital mapping company together with details of any WiFi network and cellular base station that is in range. By linking the details of available WiFi networks with the phone's GPS reading, the company can update its WiFi network map.

In general, WiFi maps are comprised of two basic, linked components: (1) WiFi network data; and (2) approximated location data, the combination of which is referred to in this article as 'WiFi mapping data'. The following sections explain each component in turn.

WiFi network data

WiFi networks are based on the IEEE 802.11 standards for wireless local area networks published by the Institute of Electrical and Electronics Engineers ('IEEE 802.11'[7]). These standards ensure that WiFi devices from different manufacturers are compatible and can communicate with each other. As a result of the widespread adoption of IEEE 802.11, the messages publicly broadcast by every WiFi-compliant device generally follow the same structure and use the same address formats.

IEEE 802.11 uses technical terminology to refer to WiFi networks, devices and different address data. A group of one or more devices that have hosted and joined a common WiFi network is referred to as a 'basic service set', or 'BSS'. Alternatively, a WiFi device can act as a dedicated, central point for communication within the WiFi network and provides other devices

3 Wi-Fi chipsets shipped will pass one billion units per year by 2012, <http://www.instat.com/press.asp?Sku=IN1004768WS&ID=2925> last accessed 12 May 2011.

4 Wi-Fi Hotspot Sessions to Grow to Over 11 Billion by 2014, <http://www.instat.com/press.asp?ID=2966&sku=IN1004769WS> last accessed 12 May 2011.

5 See for example: <http://www.skyhookwireless.com/howitworks/coverage.php>, <http://wigle.net/gps/gps/main>, and <http://www.navizon.com/howitworks.php> last accessed 12 May 2011.

6 See Skyhook World Tour <http://www.flickr.com/photos/43105570@N03/sets/72157625605627090/> last accessed 12 May 2011.

7 See <http://standards.ieee.org/about/get/802/802.11.html> last accessed 12 May 2011.

with access to external networks, such as the Internet. IEEE 802.11 refers to this type of setup as 'infrastructure mode' and the central device as an 'access point'.

To facilitate communication within and between WiFi networks, IEEE 802.11 makes use of a number of identifiers in order to distinguish between different WiFi networks, access points and devices. These identifiers are primarily: Service Set Identifiers (SSIDs), Basic Service Set Identifiers (BSSIDs), and Media Access Control (MAC) addresses (which this article refers to as 'WiFi network data'). Each identifier is explored in more detail below.

Service Set Identifier (SSID)

An SSID is the name given to a WiFi network. It is typically made of human-readable characters to help people distinguish one WiFi network from another, though it is not necessarily unique.

In infrastructure mode, the SSID is assigned by the access point. Manufacturers of access points design their devices to be 'plug and play', that is, to work from the go as soon as the relevant power and network leads have been plugged in. Fresh from its box, an access point will have been configured with a 'default' SSID, often using the manufacturer's name and/or the product name or its capabilities, such as 'belkin54g' or similar. Some manufacturers set default SSIDs that incorporate an element which is unique to each access point it produces, for example, SpeedtouchBPCB8F.

However, an SSID can be any combination of up to 32 characters, and the owner or administrator of the WiFi network can choose to change the default to something more meaningful to its users. For example, most commercial WiFi hotspots will use the name of the organization that operates it, such as 'TMobile HotSpot 123'.

Access points periodically broadcast the presence of their WiFi networks, which includes the SSID. When a WiFi-enabled device tries to detect the WiFi networks in range, it will listen, in particular, for these broadcasts and display the associated SSID of any that it receives. This broadcast feature can be turned off in order to prevent an access point from advertising its presence. But even without this broadcast feature, an SSID will still be publicly broadcast in other network messages during normal operation (for instance when a WiFi device tries to connect or re-connect to a WiFi Network).

Media Access Control (MAC) address

Each WiFi-enabled device has a unique number that is assigned to it by the manufacturer, known as a MAC address. It is a 48-bit binary number that is usually represented by six groups of two hexadecimal digits, such as 00:01:02:0A:0B:0C. A MAC address is usually fixed for the life of a device and does not normally change. However, that is not always the case and it can be possible to alter the MAC address of a device (which is known as 'spoofing').

All messages sent across a WiFi network specify a number of MAC addresses in the 'header', or address section, of the message. Typically, the MAC addresses of the source device and the destination device, together with the BSSID (see below), will be specified. The header does not form part of the content, or 'payload' of a message. As a result, when WiFi encryption is enabled, the header will be broadcast, unencrypted, over the air for general reception, and it will be available to be picked up by any WiFi device in range.

Basic Service Set Identifier (BSSID)

The BSSID is the unique address given to each BSS (i.e. each WiFi network). A WiFi network operating in infrastructure mode will be assigned the MAC address of the access point as its BSSID.

As explained in the previous subsection, the BSSID is publically broadcast in the 'header' of every message sent by a device in a WiFi network. Any WiFi device in range can receive it, whether or not it is part of that network or the content of the message is encrypted.

The WiFi network data described in above (the SSID and MAC address) are generally used in a WiFi context for communicating within the WiFi network. They are not generally communicated outside the range of the WiFi network or across the Internet. Distinctions between IP addresses and WiFi network data are considered further below.

Approximated location data

In order to provide location look-up services, the publically broadcast WiFi network data that is captured by or provided to a digital mapping company for each WiFi network will be associated with approximated location data. The exact method of approximation will vary between digital mapping companies, but in general:

1. the position of the recording device (for example, the GPS coordinates of a digital mapping company's vehicle or the mobile phone of a user accessing the company's location-aware services) is stored on each occasion it collects publicly broadcast WiFi network data. In some cases, for example where the digital

mapping company's vehicle is unable to receive a GPS signal, the GPS coordinates may be approximated based on contextual information (such as the direction and speed of travel since the last GPS coordinate was received); and

2. the cluster of positions at which WiFi network data were collected is then used to calculate an approximate location for that network.

The degree of error in the approximation will depend on numerous factors; a typical value of error might be around 35 metres.[8]

Digital mapping companies and their licensees can use their WiFi network maps to offer location look-up services. A user of such services can request his or her location by sending the service provider the WiFi network data being publicly broadcast by the WiFi networks currently in range of their WiFi-enabled device. The service provider can compare this data with the WiFi network data of WiFi networks stored in its WiFi network map. The approximated locations of each successfully matched WiFi network can be triangulated to provide an approximate location of the user, which is sent back to the user's device. Clearly, the WiFi network maps are not 'live': they provide only an approximate position of WiFi-enabled devices at a point in time and are based on historic data.

WiFi mapping data is particularly useful for providing a user with his approximate location when conventional means of determining that user's location (such as via GPS or cellular network triangulation) are not available or are insufficient.

Relevant European regulation

There are two European Directives that could potentially restrict the activities of the digital mapping industry described above. First, Articles 6 and 9 of Directive 2002/58/EC on privacy and electronic communications[9] (the 'E-Privacy Directive') regulate the use of data that indicates the geographic position of user equipment. Secondly, Directive 95/46/EC on the processing of personal data[10] (the 'Data Protection Directive') regulates the use of data that relate to an identified or identifiable individual. The scope of each Directive and an analysis of whether they apply to WiFi mapping data are set out in the next four sections.

E-Privacy Directive

The E-Privacy Directive defines 'location data' as:

> any data processed in an electronic communications network, indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service (Article 2(c))

Recital 14 of the E-Privacy Directive provides some guidance as to the breadth of location data stating that it may refer to 'the latitude, longitude and altitude of the user's terminal equipment, to the direction of travel, to the level of accuracy of the location information, to the identification of the network cell in which the terminal equipment is located at a certain point in time and to the time the location information was recorded.'

Are WiFi Mapping Data 'location data'?

The definition of location data under the E-Privacy Directive can be separated into four distinct components:

(i) 'data processed in an electronic communications network'

(ii) 'indicating the geographic position'

(iii) 'of the terminal equipment of a user'

(iv) 'of a user of a publicly available electronic communications service'

Data processed in an electronic communications network

The term 'electronic communication network' is defined in Directive 2002/21/EC on a common regulatory framework for electronic communications networks and services (known as the 'Framework Directive').[11] It is a broad term covering any 'transmission system and, where applicable, switching or routing equipment and other resources which permit the conveyance of signals'. A WiFi network would fall within this definition. As SSIDs, BSSIDs, and MAC addresses are pro-

8 Paek, Kim, and Govindan (n 2).

9 Directive (EC) 2002/58 of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector [2002] OJ L201/37, amended by Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 [2009] OJ L337/11.

10 Directive (EC) 95/46 of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31.

11 Directive 2002/21/EC of the European Parliament and of the Council of 7 March 2002 on a common regulatory framework for electronic communications networks and services (Framework Directive) [2002] OJ L108/33.

cessed in WiFi networks, it seems likely such data would satisfy this first limb.

The location data that are combined with WiFi network data by the digital mapping industry, however, are not processed in WiFi networks. These data are approximated, rather than collected, based on the position of the vehicle or device collecting WiFi network data; they are not processed in an electronic communications network at all. Similarly, where GPS coordinates are used in the approximation, these are calculated by the GPS receiver of the device collecting the WiFi network data. Only the timing signals received from each GPS satellite are actually sent through an electronic communication network. As a result, it is not certain that the approximate location data would constitute 'data that is processed in an electronic communications network'.

Data that approximate the location of users of the location look-up service, and potentially WiFi mapping data uploaded to the service by a user's device, will be sent across the Internet when such services are accessed. However, the words 'processed in' seem to suggest that location data must be generated by or used by the electronic communications network, such as information about the location of mobile phones generated by the cellular network operators.[12] It is unlikely that the E-Privacy Directive was intended to cover location data generated outside of, and merely transmitted through, an electronic communications network.

Data indicating the geographic position

The location data associated with each WiFi access point or sent to users of the location look-up services by digital mapping companies are approximations, rather than exact positions. However, as these data are 'indicating' a geographic position and Recital 14 expressly references 'the level of accuracy of the location information', it seems likely that an approximation of a geographic position would be considered 'location data'. It is not clear if all levels of accuracy are covered (for example, if the entire United Kingdom would be considered to be a geographic position). But as the E-Privacy Directive was originally intended to cover network cell information, it seems likely that a geographic position with greater accuracy (as in the case of WiFi mapping data) would also be included.

Data of the terminal equipment of a user

The geographic position needs to refer to the terminal equipment of a user. The expression 'terminal equipment' is not defined in the E-Privacy Directive. On the face of it, terminal equipment would refer to the end-of-network devices which the user actually uses to initiate communications, for example a computer, laptop, or mobile phone, rather than an intermediate networking device such as an access point. It is WiFi networking data publicly broadcast by access points that digital mapping companies are interested in.

Recital 46 of the E-Privacy Directive references Directive 1999/5/EC on radio equipment and telecommunications terminal equipment and the mutual recognition of their conformity.[13] This Directive defines 'telecommunications terminal equipment' (in Article 2(b)) as 'a product enabling communication or a relevant component thereof which is intended to be connected directly or indirectly by any means whatsoever to interfaces of public telecommunications networks (that is to say, telecommunications networks used wholly or partly for the provision of publicly available telecommunications services)'. If this definition is used to interpret 'terminal equipment' under the E-Privacy Directive, it might be broad enough to include an access point that is connected to the Internet together with any WiFi-enabled devices connected to that access point (such as a laptop or mobile phone); they would be interfacing with a public telecommunications network. A wireless router not connected to the Internet, however, is unlikely to fall within this definition.

Data of a user of a publicly available electronic communications service

The terminal equipment must be of a user of a 'publicly available electronic communications service'. 'Electronic communications service' is defined in the Framework Directive as 'a service normally provided for remuneration which consists wholly or mainly in the conveyance of signals on electronic communications networks, including telecommunications services and transmission services in networks used for broadcasting, but excludes services providing, or exercising editorial control over, content transmitted using electronic communications networks and services . . .'.[14]

There are three services relevant to WiFi mapping data:

12 Recital 14 of the E-Privacy Directive specifically references such information.

13 Directive 1999/5/EC of the European Parliament and of the Council of 9 March 1999 on radio equipment and telecommunications terminal equipment and the mutual recognition of their conformity [1999] OJ L 91/10.

14 Article 2(c), Framework Directive

WiFi network access

Individuals accessing a WiFi network through an access point are unlikely to be making use of an electronic communications service, except where the access point is a commercial WiFi hotspot. Access to home or work WiFi access points is not normally provided for remuneration or available to the public.

Location services

The location look-up service offered by digital mapping companies provides content (including the approximated location) to its users; it is not wholly or mainly used for the conveyance of signals and as such it is not an electronic communications service.

Internet or mobile data access

In the case of users of an access point connected to the Internet, or of a location look-up service, they are all users of a publicly accessible electronic communications service, that is a data transmission service offered by an ISP or a mobile phone operator.

It is not clear from the definition of location data which service is relevant for determining whether the terminal equipment is of 'a user of a publicly available electronic communications service'. The services which are most immediately relevant to WiFi mapping data are WiFi network access and location services.

As such, unless digital mapping companies are collecting information concerning the location of user devices connected to commercial hotspots, WiFi mapping data does not indicate the location of the terminal equipment of a user of a publicly accessible electronic communications service. Even in the case of commercial hotspots, the approximate location data generated by digital mapping companies seems unlikely to be 'processed in' an electronic communications network. It is therefore highly unlikely that WiFi mapping data can be considered as 'location data'.

Data Protection Directive

The Data Protection Directive sets out the common principles for data privacy which must be adopted into the national law of each European country that is part of the European Economic Area. The cornerstone of European data protection law is the concept of 'personal data', which the Data Protection Directive defines as:

> any information *relating* to an *identified or identifiable* natural person (Article 2(a))

The key parts of this description, which are highlighted in italics above, are: (1) that the individual must be 'identified or identifiable' and (2) that the information must relate to that individual. If these two criteria are met, the information will be considered personal to that individual under the Data Protection Directive.

The Data Protection Directive provides some guidance on the meaning of 'identifiable' and explains that:

> an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity (Article 2(a));

and

> to determine whether a person is identifiable, *account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person*; whereas the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable (Recital 26).

This guidance can be summarized, using the wording highlighted in italics above, as a person will only be 'identifiable' if there are ways of identifying the individual that are reasonably likely to be used.

Are WiFi mapping data personal (and is there risk to personal privacy)?

Whether WiFi mapping data fall within the scope of the Data Protection Directive is a complex issue, which is usefully considered in stages. This section explores first whether an individual is 'identified or identifiable' from (i) the components of the WiFi network data (SSID and BSSID/MAC address) alone; or (ii) the combination of WiFi network data and an approximated location (ie WiFi mapping data) alone, and secondly whether there are ways that are reasonably likely to be used to identify an individual using the location look-up services provided by the digital mapping industry. The final part of this section considers in more detail the distinction between the identification of a device and the identification of an individual.

SSIDs

SSIDs are assigned to networks rather than devices or individuals. Unless changed by the network administrator, a wireless router will assign the default name set by the manufacturer. This name is not typically unique to the device and usually refers to the manufacturer or

the product name or capabilities of the access point, such as 'Belkin54g' or similar.

A person given only a list of default SSIDs would be faced with many repeating entries (ie there would be many WiFi networks listed that had been assigned the SSID 'Belkin54g' and so on). Without additional information, that person would not be able to identify a specific WiFi network, let alone a device or an individual. It would not even be possible for that person to distinguish between two networks using the same SSID (ie two networks assigned the name 'DIR-655' would be indistinguishable without further information). Clearly, no individual would be identified or identifiable from such default SSIDs.

In some cases, manufacturers set default SSIDs that incorporate a unique number, for example, 'SpeedtouchBPCB8F'. The unique part of the SSID, that is 'BPCB8F' in the previous example, would allow the WiFi network of this access point to be distinguished from that of another. But it does not 'identify' the WiFi network to which it has been assigned, in the sense of knowing where it is located or being able to affect it. Further, it does not identify a specific device operating within the WiFi network (the access point and each of its client devices are all represented by the same SSID). From a privacy perspective, a unique SSID is no different from a BSSID, which is also a unique number shared by devices within a WiFi network. There are serious doubts whether a unique SSID or a BSSID could be used to identify an individual, and these are considered in the next section.

Most wireless routers allow you to change the SSID assigned to the wireless network. It is theoretically possible that the person in control of an access point (the 'administrator') could set the SSID to be his full name and full address, but this situation is likely to be rare. A WiFi network administrator only seeks to distinguish his network from that of his neighbour and given the maximum 32 character limit, it is far more likely that a customized SSID would use short hand, such as 'James's router'. The name used might not even be the name of the person choosing the SSID. It could, for example, be the name of the administrator's favourite movie star or their pet. Even use of a full name—'James Jones's router'—or an address—'10 High Road's router'—is unlikely to identify a specific individual; there would be many people throughout the world who share that name or live at number 10 on another High Road.

Even if, in an unlikely event, an individual were identified by a customized SSID alone, there is no way for anyone (other than the users themselves) to know whether a specific user of the WiFi network could be attributed to the identified individual. This is because the identified individual might be one of a number of users of the network, all of whom would share the same SSID. There is also no guarantee that the individual referred to in the SSID actually uses the WiFi network at all; he or she may, for instance, have set up the network for a friend or family member. This considerable uncertainty is amplified in the case of an unsecured WiFi access point (ie an access point that allows any WiFi-enabled device in range to connect to its network). Anyone can use an unsecured WiFi network and they would share the same SSID as legitimate users. Such persons could be completely unrelated or unknown to the individual identified in the SSID.

In conclusion, it is hard to see how an individual could be identified from an SSID alone. Although an SSID could theoretically be customized to the extent that it would identify an individual, this is only likely to occur in a very small number of cases.

BSSIDs/wireless MAC addresses

As explained above, every WiFi device is assigned a MAC address by the manufacturer. Unlike dynamic IP addresses considered below, a MAC address is unique to that device. The MAC address of the wireless router is typically assigned as the network's BSSID. It is important to note that while both the MAC address and the BSSID are in most circumstances unique to a device, this fact has limited, if any, significance as to whether an individual can be identified. There are several reasons for this, which are explored below.

First, as above, suppose that a person was provided with a list of MAC addresses or BSSIDs as might be processed by a digital mapping company. Because of the unique nature of MAC addresses, it would normally be possible for that person to distinguish between different WiFi devices. However, without additional information, that person could not identify (in the sense of 'placing his hands on') the unique device that the MAC address or BSSID on the list corresponds to. A MAC address is set by the manufacturer rather than the end user, so there is nothing in the character string in itself that can identify that user.

Secondly, a MAC address is not assigned to a specific individual, only a device that in many cases will be shared by multiple individuals. This may not be immediately apparent but consider the following: many company buildings and households often contain more than one WiFi-enabled device each broadcasting its MAC address. These will change from time to time as mobile devices come and go and devices are purchased,

sold, and switched off. In its lifetime, one wireless device, and its MAC address, may have many owners each potentially in a different location. In the case of an unsecured wireless router, a network may be joined by other WiFi devices the users of which could be unknown and unrelated to the household or company administering the access point. Each wireless device may also be shared by a number of users. These will also change from time to time as, for instance, household guests stay, new employees join, or the device is sold. In addition, a user is also likely to use more than one of the WiFi devices within a household or company. Each MAC address is therefore highly likely to be associated with a number of users. Each user will also almost certainly be associated with more than one MAC address.

Finally, a MAC address of a specific device is not necessarily static and can be altered, for all effective purposes, in certain circumstances (for example, via MAC address 'spoofing'). This is the equivalent of driving around town with a number plate you have copied from another car. If this happens, the relevant MAC address is no longer unique to one device, and certainly not an individual.

The consequence of the above is that a MAC address alone cannot identify a specific individual. This is clearly the case for the MAC address of a wireless router, as this device (and its BSSID) will be shared by every WiFi device connected to the network, and by each of its users.

Combined with approximated location data

The previous sections consider WiFi network data in isolation. As explained above, this information is combined by the digital mapping industry with approximated location information (ie WiFi mapping data). This location data refers to the approximate area in which the digital mapping company has calculated that the relevant wireless device is likely to be. It is not an exact position. The GPS coordinates used to provide the approximation are those of the collecting devices (eg the digital mapping company's vehicles), not the relevant wireless device. In addition, the relatively low degree of accuracy of each approximation means that in most cases the location data will cover a number of street addresses and not individual buildings. This is particularly likely to be the case in areas of higher population density, such as towns and cities.

It would not be possible, therefore, for a person who is given an SSID, BSSID, or wireless MAC address combined with the location information approximated by a digital mapping company to identify the relevant wireless device (or devices in the case of the SSID), without further information. For example, the person would not be able to use the WiFi mapping data alone to walk up to the device and touch it. Neither could the person use the combined information alone to affect the device in any reasonable way, whether remotely or otherwise. Even a more accurate approximation, using for example signal strength, would not change this conclusion. The WiFi mapping data will still not provide an exact location for a WiFi device.

As a result, the approximated location data adds little to the likelihood that an individual will be identified from the WiFi network data processed by the digital mapping industry. Without additional information, the approximated location is unlikely to identify the house, flat, or building in which a WiFi device resides, let alone a specific individual.

Likelihood of use to identify

The definition of personal data in the Data Protection Directive includes information relating to individuals that can be identified from WiFi mapping data by means that are reasonably likely to be used (eg by combining the WiFi mapping data with other information that a person would be reasonably likely to obtain).

It has been suggested that a specific user of a WiFi device could be easily identified by visiting the approximate location area provided by a digital mapping company, checking the signal strength corresponding to that wireless MAC address and homing in on the specific user using the rule 'the stronger the signal, the closer the WiFi router'. However, if 'reasonably likely' is interpreted to mean that account must be taken of how difficult a given means of identifying a specific individual would be (ie practical likelihood) and the reasons why a person would undertake to identify an individual using such means (ie motivation), there are serious doubts as to whether anyone would reasonably, or could lawfully, use this 'trial and error' method to identify an individual. This is considerednext.

Practical likelihood

First, there are likely to be physical and legal restrictions that would make it difficult for someone to locate a house or apartment in which a WiFi device is located. For example, it may be necessary to cross physical boundaries such as a wall, fence, or hedge to get close enough to the buildings in question in order to determine the source of the wireless network. Significantly, it would not be legal to cross such a boundary without the property owner's consent (that would be trespassing or its equivalent in most countries). In the

case of apartment buildings, it may not even be possible to enter the building without a code or key in order to determine the relevant apartment.

Secondly, signal strength only provides an indication of how far away the relevant device is located. Nearby objects, such as buildings or terrain, can have an impact on signal strength and impede a person's ability to determine the particular building in which a wireless device is located. There are also numerous environmental factors which affect signal strength which can make determining a source building harder.

Thirdly, an average user armed with a laptop or mobile phone will typically be presented with limited information about the signal strength relating to a wireless access point (such as a '5 bar' strength symbol or a percentage figure) and in some cases only for 'broadcasting' wireless access points. At present, non-standard software would often be required in order to obtain a more precise reading and to locate 'hidden' access points. Trial and error would also still be necessary and a person trying to follow the strength of a signal might well arouse the suspicion of residents.

Finally, signal strength can only ever be used to determine a more precise location of a wireless device with a particular MAC address. As is explored below, it is hard to see how someone can use signal strength in order to identify a specific user of that device.

The identification of an individual using signal strength by 'trial and error' from the starting point given by WiFi mapping data therefore faces many practical and legal challenges. It does not seem 'reasonably likely' that this method would be used in practice.

Motivation

First, the location look-up services provided by digital mapping companies are primarily intended to provide the location of a user's WiFi-enabled device. Although it is possible for a person to use the location services offered to obtain the approximated location of a WiFi device that features in a WiFi network map, he can only do so by providing the digital mapping company with the MAC address of the WiFi device that he wants to locate. Given how the Internet currently works, WiFi MAC addresses are not usually relayed by an access point beyond the local WiFi network (unless specifically sent by user software). In order to obtain a MAC address, therefore, a person will likely need to either possess the relevant WiFi device or be in range when it publicly broadcasts messages across its WiFi network. Clearly, such a person will already know where the device is approximately located and would have no need to use a location look-up service based on WiFi mapping data.

Secondly, as they are not communicated outside a local WiFi network, SSIDs, BSSIDs, and wireless MAC addresses cannot be used to trace the source of a specific communication made across the Internet. Unlike IP addresses, ISPs will not log WiFi network data and these details are not stored with any substantive information such as internet traffic logs, website user logs or the content of electronic communications. There is little incentive, therefore, for someone, other than the owner, to obtain location information from a digital mapping company for a particular wireless MAC address.

One potential use of WiFi mapping data that has been raised is the possibility of tracking down an individual who has moved to a new address by looking up the MAC address of his or her access point. This example must, however, be placed in context. As mentioned above, WiFi network maps are not live but provide a historical snap-shot of the location of each access point at a given time. It would be unusual for digital mapping companies to immediately include new information about the approximate location of an access point before several queries have confirmed the new position. In addition, it is common for there to be some delay before an updated version of the WiFi network map is publicly accessible (for instance, Google currently refreshes its publicly accessible WiFi network map every two weeks). It is questionable, therefore, whether a location look-up service based on WiFi mapping data would or could be used to track the movements of a wireless device, and doing so would likely violate national laws in many jurisdictions.

Given the above, it would be hard to justify a conclusion that it is 'reasonably likely' that a person would use signal strength (or indeed any method) to identify an individual using the WiFi mapping data. It would be difficult and potentially unlawful to carry out in practice and such a person would have nothing to gain by doing so. It seems therefore that an individual would not be identifiable from WiFi mapping data except in very rare and highly unusual circumstances.

Identification of an individual

It is clear from the definition of personal data under the Data Protection Directive that data which only relate to an identified device will not be personal; an individual must be identified or identifiable. As touched upon above, there are serious doubts as to whether this is the case for WiFi mapping data and reasons for this are set out below.

First, even where it is possible determine that a device with a given BSSID/wireless MAC address is

located in a particular building, this does not mean an individual is identified. In most cases, more than one individual will use that building, and in some cases such as an apartment or office building, it could be several hundred.

Secondly, although a BSSID/wireless MAC address is unique to a particular device (subject, of course, to the practice of MAC address 'spoofing' mentioned above), it is not unique to a specific individual. Even if the exact location of a wireless access point is known, it is likely to be shared by numerous individuals—family, friends, coworkers, and other people who visit from time to time. In the case of an unsecured router, it may be unwittingly shared with many other users who are unknown and unrelated to its 'administrator'. Each individual may connect using more than one wireless device, and each such device may be shared by one or more other individuals. The fact that WiFi devices and access points are regularly used by multiple people, may be moved to new locations and passed on to new owners many times over means that the total number of individuals whose use might be associated with a particular MAC address is large. A person would, therefore, face considerable practical and technical difficulties when seeking to identify a specific individual using the WiFi mapping data.

Even if in a few extreme cases, and with extraordinary effort, WiFi mapping data could be traced to an individual, this should not affect the conclusion that such information is not personal. If under the Data Protection Directive 'identifiable' included any theoretical possibility that at some stage someone may be able to identify an individual, then almost all data could be regarded as personal data. This cannot be correct and such an extensive interpretation of personal data would have far-reaching and undesirable consequences (those relating to the digital mapping industry are outlined in the consequences section below).

WiFi network data compared to IP addresses

The influential group of representatives from national data protection regulators across Europe, known as the Article 29 Working Party, has at various times expressed its view that IP addresses are, or in some contexts might be, personal data (for example, in WP Paper 136 adopted 20 June 2007).[15] Although the Working Party has stated that a mere hypothetical or negligible possibility that an individual may be distinguished or identified is not sufficient, it has also expressed the view that a dynamic IP address allocated by an ISP should be regarded as personal data, even in the hands of a website owner who collected it from a user. This is on the basis that the website owner may in certain circumstances be able to obtain a court order compelling the ISP to provide details of the real world identity of the holder of a particular IP address.[16]

This 'one-size-fits-all' point of view in respect of IP addresses is not shared by all, and even the Working Party itself has not always been so unequivocal in its view that IP addresses are personal data. In its paper WP 37, the Working Party recognized that many third parties may receive a dynamic IP address without being able to link it to other data concerning the user that would make his/her identification possible.[17] Even in WP Paper 136, which appears to suggest that IP addresses are always personal data, the Working Party's analysis emphasizes the processing of IP addresses by ISPs and circumstances where IP addresses are being processed for the purpose of identification (giving the example of copyright holders) as key examples of when IP addresses should be considered personal data.[18] It also acknowledges circumstances, such as an internet café, where identification would not be possible using reasonable means.[19]

It should be noted that the Working Party's opinions are expressed in the context of static or dynamic IP addresses allocated by ISPs. Where Internet access is shared through a WiFi network, it is the access point that would use the ISP-allocated IP address but only when forwarding communications from within the WiFi network over the Internet. When the access point (and the other WiFi devices in its network) communicates within the WiFi network, it uses a different, local IP address. When used in this section (and in the article as a whole), the term 'IP address' refers to an IP address assigned by an ISP.

'Identification' should be considered on a sliding scale. At one end—'personal data'—are data from which an individual is identified (ie IP addresses linked to subscriber details in the hands of the relevant ISP). At the other end, 'anonymous data', ares data which are in the hands of a person who has no reasonable possibility of identifying the subject individual. The WP Papers appear to support a conclusion that an IP

15 Article 29 Working Party, 'Opinion 4/2007 on the concept of personal data' (WP 136, 20 June 2007).

16 First two paragraphs of Example 15, at 16 and 17, ibid.

17 See Article 29 Working Party, 'Privacy on the Internet—An integrated EU approach to online data protection' (WP 37, 21 November 2000), at 16.

18 WP 136 (n 14), at 17.

19 Ibid.

address could fall at different ends of the scale, depending on the context.

There are many reasons why SSIDs, BSSIDs, and wireless MAC addresses fall towards the anonymous end of this identification scale. This is largely on the basis of the points set out in the previous sections of this article, but is also based upon the following:

1. An individual does not need to subscribe for an SSID, BSSID, or wireless MAC address in order to obtain one. BSSIDs and MAC addresses are fixed by hardware manufacturers and SSIDs are set at default values by such manufacturers unless customized by individual network administrators. There will often be no single person, organization or body that holds a record of individuals' 'real world' identities alongside the SSIDs, BSSIDs, or MAC addresses of the WiFi-enabled devices which these individuals own or use.
2. As explained above, based on how the Internet currently works, SSIDs, BSSIDs, and wireless MAC addresses from user devices are not relayed by an access point beyond its WiFi network (unless specifically included by user software in the content rather than the header of a communication). ISP and website owners do not therefore retain a log of WiFi network data, as they generally do in respect of IP addresses. The fact that a particular user has visited an illegal site, for example, will not be associated by the relevant website owner or an ISP with a particular SSID, BSSID, or wireless MAC address.
3. Given (1) and (2) above, in order to obtain an SSID, BSSID, or wireless MAC address you need to be in the proximity of the relevant wireless network, that is you already know where it is. Therefore, the combination of SSID, BSSID, or wireless MAC address with an approximate location does not provide any additional information to a person which that person cannot already obtain for themselves.
4. It is not possible to use an SSID, BSSID, or wireless MAC address alone to communicate with a corresponding wireless device other than in the proximity of the relevant wireless network (assuming, also, the network is publicly accessible). It is not, therefore, possible to 'affect' someone in the same way that it is possible to do with an IP address.

Consequences for the digital mapping industry

There are several companies worldwide involved in the collection of WiFi network data or in associating it with GPS coordinates for digital mapping purposes. The digital map marketplace includes 'big players', such as Google, Navteq (a subsidiary of Nokia Corporation), and Tele Atlas, as well as several other smaller organizations, focused on particular countries, such as Zenrin (which focuses on Japan), or particular users of mapping data, such as intelligent transport systems and map enhancement products. The practice of collecting WiFi network data is also widespread, including for example, Skyhook, Inc. and Navizon, Inc.

According to the February 2010 industry report by Infiniti Research Ltd entitled 'TechNavio insights into the Global Digital Map Market 2009–2013', the global market for digital mapping is forecast to grow from $1.6 billion in 2009 to $2.8 billion in 2013. Almost 50 per cent of this market is in Europe, the Middle East, and Africa (EMEA), more than either the USA or the Asia Pacific. WiFi mapping data also form part of a framework that enables the creation of a wide range of location-aware applications. The enhanced functionality offered by location-aware services is rapidly becoming part of the fabric of what users expect when they go online.

The consequences of designating generally broadcast and commonly collected WiFi network data as personal data (whether in combination with GPS coordinates or not), would be significant for this large and fast-growing industry. Information that is routinely used by the digital mapping industry would become the subject of significant compliance obligations in circumstances where, on a practical level, compliance would be almost impossible (eg giving notice to every individual whose WiFi device broadcasts an SSID) and where there is minimal or no risk of identification or threat to individual privacy. Such a designation would impede the future growth of the digital mapping market in EMEA and hinder the development of new mapping-related products.

In the event that WiFi mapping data were to be considered location data, as these data are not 'traffic data' (ie they are not used in order to convey messages on an electronic communications network), they would be regulated under Article 9 of the E-Privacy Directive. This prohibits the processing of location data unless the data are anonymous or prior consent has been obtained. As is discussed above, it is not reasonably likely that WiFi mapping data could be used to identify a specific individual. The digital mapping industry is, therefore, processing WiFi mapping data anonymously and therefore should be operating in accordance with Article 9. Clearly, if the alternative conclusion is reached, it could cause insurmountable difficulties for WiFi mapping. It would not be feasible for digital

mapping companies to obtain prior consent from each 'user' of a WiFi access point.

In addition, there is also a considerable risk that if various EU countries' regulators adopt separate, contradictory positions regarding the status of WiFi network data as personal data and/or location data, they will create economic barriers to the free movements of mapping-related products and services.

Conclusion

It can be seen from the analysis set out in this article that, in all but very rare and highly unusual situations, WiFi network data alone cannot identify an individual. There is a risk that an SSID could be customized so as to contain identifying information, however, such cases would seem to be rare and without any certainty as to the relevance of such an identified individual to the WiFi network. The combination of WiFi network data with location information by digital mapping companies does not alter this position; such location information is approximated and is not reasonably likely to be used to identify a specific individual.

The inclusion of WiFi mapping data within the European privacy regime would affect a nascent but quickly growing industry. It would have significant consequences across the digital mapping industry, as many companies use publicly available SSID and MAC data to provide enhanced location services. It might well be impossible to meet all the theoretical obligations of European data protection and privacy law in the event WiFi mapping data is considered 'location data' and/or 'personal data'. There is clearly a need for more debate on these difficult and complex questions, to ensure that a balanced, proportionate, and fair decision is reached that protects both privacy and user choice.

Ultimately, it would seem disproportionate to effectively prevent the collection and use of WiFi mapping data on grounds of privacy and data protection. WiFi network data are publicly broadcast and anyone is free to record the locations at which each network can be received. Except in exceptional circumstances, there does not seem to be a significant privacy risk in allowing people to look up their current location by reference to these wireless 'landmarks'. But on the contrary, without access to this public data, thousands of online applications which rely on the WiFi mapping data of the digital mapping industry would be significantly hampered, preventing millions of people worldwide from using amazing and useful applications.

doi:10.1093/idpl/ipr013