Chapter 2
Anonymity Challenges in the Internet

Since information about people acting in the Internet (both, consciously or unconsciously provided by them) can be easily found, surfing on the World Wide Web is far from an anonymous activity of no consequences. With regard to the therewith associated risk of data abuses it is still a debatable point, whether the identification in the online world is essential, and if so to what extent, or whether there is a right to act anonymously within the World Wide Web.

In this sense, light will subsequently be shed on the motivations for the anonymous use of Internet services and the Internet participants' possibilities to make their activities on the Internet untraceable.

2.1 Risks for Anonymous Use of Internet Services

Manifold Internet activities cause risks for those persons being interested to remain anonymous when using the new communication channels and platforms. Some practices leading to data collection and consequently to the possibility of third persons to have access to personal data are discussed hereinafter.

2.1.1 Information Gathered by IP Addresses

Internet IP addresses[1] are used to route data from one host computer to another. Even though these numerical addresses do not directly identify particular Internet users, their identification can easily follow from the connected addresses by evaluating the gathered information (Schwartz and Solove 2011, pp. 1838/1839).

Initially, static IP addresses were used. A static IP address is a number (in the form of a dotted quad) that is assigned to a computer by an Internet Service

[1] See Sect. 1.3.2.2(1)(a).

R. H. Weber and U. I. Heinrich, *Anonymization*, SpringerBriefs in Cybersecurity,
DOI: 10.1007/978-1-4471-4066-5_2, © The Author(s) 2012

Provider (ISP) to be its permanent address. Accordingly, with each log on to Internet access the user is allocated the same IP address (Freund and Schnabel 2011, p. 496). In the end, this facilitates the tracing of the respective computer and therewith the identification of the Internet participant.

At the time of the Internet's inception, scarcity of IP address space seemed to be unlikely as information and communication technologies (ICT) were cost-intensive and therefore only few networks were interested in Internet connections (Edelmann 2009, pp. 1–13). In the course of the last 15 years the demand for IP addresses has enormously increased. Eventually, since IPv4 makes only available about four billion IP addresses, the exhaustion of the current Internet Protocol addressing system, Internet Protocol Version 4 (IPv4), occurred in February 2011.

Already more than ten years ago (in 1998) the substitute for IPv4, namely IPv6, was designed, aiming at providing quantitative and qualitative advantages compared to IPv4, the two Internet Protocols are currently not fully compatible (Weber and Heinrich 2011, p. 71). The problem of shortage could be mitigated by various techniques such as "Network Address Translation" (NAT) (Brunst 2009, p. 52), which hides multiple Internet hosts behind a single IP address by connecting private networks to the public Internet. However, such a procedure would have the disadvantage of breaking end-to-end connectivity. As a result, Internet activity would no longer be fully granted, making it difficult to establish Internet telephone calls directly between two hosts using standard Voice over Internet Protocols (VoIP) (Weber and Heinrich 2011, pp. 70/71). Furthermore, the method would increase complexity since there are two classes of computers (some with public and some with private addresses) as well as costs for design and maintenance of networks and for the development of applications (European Commission 2008).

Hence, with regard to the temporary scarcity of IP addresses and their associated sparing use, dynamic IP addresses were allocated by the Regional Internet Registries (RIRs) to the respective access providers which enable the access to the Internet and therewith serve as an interface between user and Internet; access provider administrate a small pool of IP addresses and allocate these addresses for the period of usage only (Brunst 2009, p. 51). Subsequently, a further allocation to a "new" user connecting to the Internet is possible. With regard to the impermanent allocation of IP addresses an exact tracing of the respective user is difficult and requires a recording at the material time; otherwise each of the access authorized computers could potentially have done the respective action (Brunst 2009, p. 51).

Since even Internet participation by using dynamic IP addresses is not qualified to preclude the respective Internet user's tracing with absolute certainty,[2] achieving the possibility of surfing on the Internet without revealing one's IP address and therewith the own identity must be seen as the most effective method to realize anonymity.

[2] Complete anonymity cannot be guaranteed.

2.1.2 Storage of Recorded Data

Although being partly (as far as scope and duration of storage is concerned) illegal according to most current national law (Freund and Schnabel 2011, p. 496)[3] many providers storage recorded data over a long period of time (Krause 2003, p. 161). In reality, data like the time of visit of a website, the used Internet IP address and the whole history of surfing are collected. The web page operators' prior intention to collect all these data usually is to conduct marketing analyses for streamlining their webpages and therewith increasing their business opportunities.

Furthermore, web page operators collect data for the protection of their own web page against misuse. Even if most of the individual data collected are insufficient to support a conclusion on the respective user the sum of data may have the ability to identify the user or his computer, respectively (Malin et al. 2003, p. 1); accordingly, the storage of data possesses a threat to anonymity. The period of time of data storage must (also) be put in relation with the right to be forgotten encompassing the right to have data deleted after a certain period of time.[4]

2.1.3 Insufficient Data Security Measures

With the development of new technologies, new attacking tools are also regularly developed. Therefore, security is and has to remain a topic of discussion. Since security and privacy of data are of particular importance for Internet participants both private and business, transactions and the interests of all parties involved have to be kept confidential in order to protect the Internet participants' privacy and ensure fair competition.

The online world is rich in possibilities; technical innovations and ingenuity allow the society to progress and prosper. However, the development of new forms of technical activity can also potentially be misused, among others by measures like denial of service attacks, dissemination of viruses, logical bombs or hacking (Weber 2009, p. 232):

- *Denial of service attacks* (DoS) consist of large streams of useless data directed towards particular network locations with the aim of overloading equipment and destroying its functionality. A denial-of-service attack does not steal passwords or manipulate data, but rather overloads the data traffic of certain systems (flood attack) or causes parts of the system's hardware or software to shut down.

[3] Example: According to German law, access providers are only allowed to use stored data for accounting purposes or for eliminating technical barriers.

[4] Extensively on the subject of Trojan horses, see Sect. 4.4.

In so-called distributed denial-of-service attacks (DDoS) multiple systems flood the bandwidth or resources of a targeted system.

- A *virus* is a program that can copy itself, and is therefore attached to or inserted in data documents or the boot sector of the hard disk. A virus is often capable of deleting data or of invalidating certain functions of computer software or the download of further Trojan horses.[5] Recently, attackers often bundle link virus programs with other malicious programs making viruses a major threat to private users and businesses (Graham et al. 2011, p. 92).
- Programs that are attached to any other program and lead to the shutdown of the system are called *logical bombs*.
- The most serious technical attack is arguably the actual *hacking* into a communication system; the term "hacking" is often used for a broad range of illegal objectives and technical activities.

During the past view years, experience has shown that hackers and attackers are breaking into vital portions of the global network infrastructure, causing problems and creating costs (Weber 2003, p. 105 ss). This was the scenario on December 24, 2011, when hackers using the pseudonym "Anonymous",[6] accessed to the database of Stratfor, a global security intelligence firm, and copied customer data like email addresses and credit card data. The goal of this action was to steal altogether one million dollar for gifting the money as Christmas donations to aid agencies.[7] Similarly, a group of people announcing to use the pseudonym "Anonymous" threatened to block certain servers or deviate some information flows if the US Congress would approve the pending proposal for a "Stop Online Piracy Act" (SOPA) in late January 2012

Within the last few years repeatedly individuals or groups of people using the alias "Anonymous" appeared on the Internet accomplishing hacker attacks. In so doing, among others in June 2011 "Anonymous" temporarily incapacitated the online presence of GEMA, a German collecting society, for protesting against the GEMA's claims to remuneration towards the video portal YouTube which result in the fact that most of the music videos cannot be accessed.

---

[5] In more detail see Sect. 4.4.1.1.

[6] Starting in 2008, a group of online activists acting under the synonym "Anonymous" appeared on the scene. In so doing, the name "Anonymous" itself was inspired by the (perceived) anonymity under which Internet participants post images and comments on the Internet. Representing the concept of any and all people as an unnamed collective the members of the group appear in public wearing the Guy Fawkes masks popularized by the comic book and film V for Vendetta. At the beginning, "Anonymous" provided warnings against the Church of Scientology and accomplished protest actions to support the right to freedom of speech and the Internet freedom. Initially acting only within the Internet, the activist meanwhile expanded their protest actions in sectors aside from the Internet. The activists sign their messages with "We are Anonymous. We are Legion. We do not forgive. We do not forget. Expect us."

[7] With regard to the fact that on behalf of "Anonymous" both a letter claiming responsibility and a denial was sent the perpetration of "Anonymous" is still unproved. According to the denial letter of December 25, 2011, "Anonymous" strongly condemned the action of being a violence of the freedom of press.

Regardless of the acting entity's intentions these and further incidents stress the relevance of data security in connection with the Internet. The actions have shown that a threat—for example, the shut down or attempt to shut down major sites used by an entire community to accomplish essential civil tasks—can go beyond a simple menace to economic safety and endanger national and international security. An umbrella term for such threats to infrastructure is "cyberterrorism", which is defined as an "extreme or intense force in an online setting, causing unexpected or unnatural results, and used for purposes of intimidating, coercing, or creating an atmosphere of anarchy, disorder, or chaos in a networked environment" (Biegel 2001, p. 232).[8]

In view of the wide difference between anonymous and untraceable acting (Solove 2007, p. 147; Schwartz and Solove 2011, p. 1837), as few as possible traces should be left in order to accomplish the aforementioned Internet users scope to achieve data protection and data security (Köhntopp 2000, p. 44). Hence, anonymizing services (also referred to as anonymizers) come into operation for masking the own IP address meanwhile surfing on the World Wide Web and therewith holding out the prospect of achieving data security and for realizing unobserved movements in the World Wide Web.

2.2 Technical Implementation of Anonymizing Services

Even though it is relatively easy to surf on the Internet without immediately revealing one's identity or to blog anonymously on the Internet, it is hard to be untraceable, too. With regard to the previously described informative content of IP addresses individuals can (more or less easily) be followed without them even knowing about it (Weber and Weber 2010, p. 45). Accordingly, in recent years the wish for anonymous communication on the Internet has motivated the development of a number of networking techniques.

2.2.1 Privacy Enhancing Technologies in General

Technological measures are available that increase privacy in the application layer. A number of technologies have been developed in order to achieve information privacy goals.[9] Privacy Enhancing Technologies (PET) can be oriented on the subject, the object, the transaction or the system. Subject-oriented PET aim at limiting the ability of other users to discern the identity of a particular business, object-oriented PET endeavour to protect identities through the use of particular

[8] In general to the problems of cyberterrorism see Council of Europe (2008).

[9] This subchapter is based on Weber and Weber 2010, pp. 47–50.

technology, transaction-oriented PET have the goal to protect transactional data through e.g. automated systems for destroying such data, and system-oriented PET want to create zones of interactions where users are hidden and objects bear no traces of businesses handling them nor records of interaction (Samuelson 2000, p. 1668; Froomkin 2000, pp. 1528–1553).

A further category is being developed by the World Wide Web Consortium (W3C) and is called a Platform for Privacy Preferences (P3P). P3P is supposed to enable individuals to program their browsers to identify which information they are willing and unwilling to disclose to the owners of the website (Samuelson 2000, p. 1668). This server-based filtering tool allows for identification and protection against deviations from the applicable codes of conduct in the privacy field (Weber 2009, p. 245).

2.2.2 Anonymizing Networking Techniques

In case encryption is not used almost all data retrieved by an Internet participant can be intercepted and seen by others. Insofar, as already said, the avoidance of collection of individual-related data best protects the informational and communicative self-determination of the persons concerned (Holznagel and Sonntag 2000, p. 72).

Applied by both Internet users (client anonymity) and service providers (server anonymity) anonymizers are among others used to hide the user's true physical location (Graham et al. 2011, p. 75) towards providers and other Internet participants for preventing conclusions on the respective identity by automatically anonymizing the Internet traffic (Brunst 2009, p. 131).[10] In that sense, light will be shed on some of the developed, partially cost-free services to facilitate anonymous Internet access hereinafter.

2.2.2.1 Client Anonymity

(1) Simple Proxy Service

The most utilized technical and easy to handle devices used for veiling the own activities are web-based proxy servers, also known as web-based proxies. Serving as intermediary between user and target page, a proxy server is a computer that forwards requests by other computers. By allowing actors to send network traffic through another computer the sender's IP address transmission is hampered by the proxy server (Graham et al. 2011, p. 75).

[10] However, anonymizing services do not automatically anonymize the communication's content.

Instead of connecting directly to the webserver, Internet participants make a circuit and connect to the proxy server first; afterwards, the proxy server connects to the requested page (Brunst 2009, pp. 52/53). As a result, the targeted server gets information solely about the proxy server's IP. Since the transmission of the user's IP is prevented, from the target page's point of view the Internet user makes no appearance (Krause 2003, p. 161).

(2) Mix Cascades

Although staying incognito to the target page operator when using a simple proxy server the Internet participant does not remain really anonymous; the proxy server's operator has the ability to ascertain the used computer. With regard to the ultimate aim of Internet anonymization to allow a host to communicate with an arbitrary server to an effect that nobody can determine the host's identity, newly anonymizing services connect proxy server in series, so called mix cascades or multiple proxies (Krause 2003, pp. 161, 173/74).

These independent devices mingle the incoming bitstreams and direct them through a large number of computers whereby an exact allocation of the requesting Internet participant is prevented or at least hampered since none of the servers involved has all information at his disposal. The final receiver can only discover the last proxy and is not directly communicating to any of the intermediary proxies or the sender of the information respectively his computer (Graham et al. 2011, p. 75).

(3) Onion Routing

The main idea of onion routing is to encrypt and mix Internet traffic from many different sources whereby onion routing protects the identity of the sender and the receiver of data both towards third parties and from each other (Berghel and Womack 2003, p. 18). With onion routing, data is wrapped into multiple encryption layers, using the public keys of the onion routers on the transmission path. This process would impede matching a particular IP packet to a particular source. A well-known anonymization service implementing this technique is the free software TOR ("The Onion Router").[11]

(4) Peer-to-Peer (P2P) Systems

Differently to the services explained above, within peer-to-peer (P2P) systems all computers enjoy equal rights to the effect that they utilize and allocate services. While P2P systems in the beginning still relied on a central root, the most

[11] See https://www.torproject.org/; Landau (2010), p. 139/40.

advanced forms of P2P systems operate without a centralized server. Data as well as inquiries for information are decentralized, and each peer only has access to his/her own communication data (Weber and Weber 2010, p. 50).

According to this system all peers are potential originators of the respective traffic and are also potential relays. Being part of the net each "peer" makes information available. Since none of the peers governs the net no participant knows the complete amount of forwarded data but just the peers he/she is collaborating with (Brunst 2009, p. 68). If communication is encrypted, the system enjoys a high degree of anonymity as communication cannot be intercepted and search of data is carried out indirectly through chains (Mayrhofer and Plöcklinger 2006, pp. 11–15).

The interest in utilizing P2P system has increased over the course of time, based on the wish to share files without revealing one's network identity and risking litigation, the distrust in governments and the increasing number of lawsuits against bloggers. The most common P2P type of use is the peer-to-peer filesharing application, in recent years frequently used for the illegal sharing of soundfiles and cinematic works protected by copyright. Besides, there are also legal grounds of justification for using peer-to-peer filesharing applications, as for instance the protection of free speech.

(5) Crowds

A further anonymizing technique is called "crowds". In contrast to the above described proxies that forward request by other computers, *crowds* work by hiding the actual source of data sent by an Internet user by "burying" it in the traffic of a "crowd" of users. Accordingly, each member of the crowd could be the sender of the received Internet traffic. Since this technique uses a just a single symmetric key there is less encryption necessary and data traffic can be forwarded faster (Brunst 2009, pp. 135–137).

2.2.2.2 Server Anonymity

As set out above, the most promising way to achieve data security and data protection is to mask or replace the own IP address. In certain cases not only the user of Internet services but also the service provider wishes to remain anonymous especially with regard to the fact that individuals often act as servers when participating in file sharing networks or hosting personal web pages (Bono et al. 2004, p. 1). The arguments given above regarding client anonymity are applicable on server anonymity, too. A service provider can also have an interest in staying incognito, as for instance in case a public interest group aims at publishing without taking on the risk of becoming subject to repressive measures (Demut and Rieke 2000, p. 40).

2.2.3 Virtue of Anonymizing Services

Anonymizing services are employed to accomplish the goal of achieving data security and therewith maintaining the power of control over the own data. Basically, anonymizers themselves and their use are not illegal (Graham et al. 2011, p. 78) even though the use to conduct an illegal activity is not allowed. Therefore, most anonymizing services provide rules within their general business terms, among others obliging the user to omit occurrences of illegal activity. As a consequence, infringements of the business terms may result in information exchanges between service providers and investigative authorities.[12]

In terms of efficiency of anonymizing services some critical annotations need to be made. Basically, proxy servers, mix cascades, onion routing, P2P systems and crowds have the ability to meet the envisaged goal.

With regard to possible technical failures or abuses the interposition of just one proxy server, however, involves the risk of missing the intended anonymity. Hence, the usage of mix cascades is preferable since these chains of proxy servers mingle the incoming bitstreams, direct them through a large number of computers and therewith to a great extent prevent the requesting Internet participant's IP address identification. However, since the encrypted bitstreams at the first and the last proxy remain without encryption, this kind of partial encryption cannot offer an adequate protection towards an observing attacker. Furthermore, the utilization of series-connected proxy servers noticeably decelerates the data stream.

Even though onion routing protects the identity of both the sender and receiver of data, this technique negatively affects the Object Naming Service (ONS)[13] and discovery services by increasing time of waiting and thereby resulting in performance issues. Furthermore, onion routing could only be used for the anonymization of traffic directed at EPCIS servers, thereby increasing anonymity, but not confidentiality or integrity of data.

Within regard to P2P systems, anonymity is not always given. Contrary to the general opinion of ordinary file-sharing applications being able to ensure anonymity, there is at most anonymity between the file-sharer and other users, but not necessarily vis-à-vis law enforcement agencies (Brunst 2009, p. 98). Only within anonymous P2P networks might it be possible to remain undetected by state control (Brunst 2009, pp. 98, 102).

Within a crowd the data traffic is routed through a great number of users thereby at first glance obliterating all traces (Berghel and Womack 2003, p. 18). Since there is no single server forwarding requests to receivers, every participant of the crowd could be the forwarder of traffic. However, this technical device's weak point consists of the fact that also the forwarder's IP address will be transmitted

[12] Compare for example Anonymizer, Terms of Use, http://www.anonymizer.com/legal/legal, Accessed 12 January 2012.

[13] The ONS is a service containing the network addresses of services; for further details see Weber and Weber 2010, p. 6.

which in case of investigative measures would lead to the computer having accepted the request at last before forwarding the requested data to the receiver (Brunst 2009, p. 137).

In a nutshell, it can be said that the use of anonymizing services is adapted for fulfilling the individuals' need to make an appearance on the Internet without revealing his/her identity even though complete anonymity seems to be a wishful thinking. However, it is debatable whether the advantages offered by such anonymizing services[14] do outweigh the disadvantages (Baeriswyl 2008, p. 4). Taking this assessment into account, subsequently the possible legal bases for the right to act anonymously on the Internet and therewith the use of (Internet) anonymization services are to be addressed.

References

Baeriswyl B (2008) Der Schatten über der Anonymität. Digma 1:4–5

Berghel H, Womack K (2003) Anonymizing the net: sanitizing packets for fun and profit. Communications of the ACM 46(4): 15–20. http://delivery.acm.org/10.1145/650000/641220/p15-berghel.pdf?ip=130.60.119.66&acc=ACTIVE%20SERVICE&CFID=62809855&CFTOKEN=52124798&__acm__=1327076953_549d640167b855013cd4ba7bd5ac87e2. Accessed 31 Jan 2012

Biegel S (2001) Beyond our control?: confronting the limits of our legal system in the age of cyberspace. MIT press, Cambridge

Bono SC, Soghoian CA, Monrose F (2004) Mantis: a lightweight, server-anonymity preserving, searchable P2P network. http://files.dubfire.net/jhu/publications/mantis-tr-b.pdf. Accessed 31 Jan 2012

Brunst PW (2009) Anonymität im Internet—rechtliche und tatsächliche Rahmenbedingungen. Duncker and Humblot, Berlin

Council of Europe (2008) Cyberterrorism: the use of the internet for terrorist purposes. Council of Europe Publishing, Strasbourg

Demut T, Rieke A (2000) Der Rewebber—Anonymität im World Wide Web. In: Sokol B (ed) Datenschutz und Anonymität. Toennes Satz + Druck GmbH, Düsseldorf

Edelmann, B (2009) Running out of numbers: scarcity of ip addresses and what to do about it. Working Paper Harvard Business School. http://www.hbs.edu/research/pdf/09-091.pdf. Accessed 31 Jan 2012

European Commission (2008) Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions: Advancing the Internet: Action plan for the deployment Internet Protocol version 6 (IPv6) in Europe. COM 2008(313). 27 May 2008. http://ec.europa.eu/information_society/policy/ipv6/docs/european_day/communication_final_27052008_en.pdf. Accessed 31 Jan 2012

Freund B, Schnabel C (2011) Bedeutet IPv6 das ende der anonymität im internet? MultiMedia und Recht 8:495–499

Froomkin AM (2000) The death of privacy? Stanford Law Rev 52:1461–1543

[14] During the so called "Jasmin Revolution" starting at the end of 2010 in Tunesia and continuing 2011 within the bordering Arab States governments (unsuccessfully) tried to silence the political opposition by shutting down important webpages. However, by using anonymizing services Internet users were able to bend this censorship of the Internet.

Graham J, Howard R, Olson R (eds) (2011) Cyber security essentials. Auerbach Publications, Boca Raton

Holznagel B, Sonntag M (2000) Rechtliche anforderungen an anonymisierungsdienste: das beispiel des janus-projektes der fernuniversität hagen. In: Sokol B (ed) Datenschutz und Anonymität. Toennes Satz + Druck GmbH, Düsseldorf

Köhntopp M (2000) Identitätsmanagement—anforderungen aus nutzersicht. In: Sokol B (ed) Datenschutz und anonymität. Toennes Satz + Druck GmbH, Düsseldorf

Krause C (2003) Tools für anonymität. In: Bäumler H, von Mutius A (eds) Anonymität im internet. Vieweg, Braunschweig

Landau S (2010) Surveillance or Security? The risks posed by new wiretapping technologies. The MIT Press, Cambridge and London

Malin B, Sweeney L, Newton E (2003) Trail re-identification: learning who you are from where you have been. LIDAP-WP12. Carnegie Mellon University. Laboratory for International Data Privacy. http://dataprivacylab.org/dataprivacy/projects/trails/paper3.pdf. Accessed 31 Jan 2012

Mayrhofer M, Plöcklinger O (2006) Aktuelles zum internetrecht: tagungsband zum symposium internet-recht vom 23. April 2005. Pro Libris, Engerwitzdorf

Samuelson P (2000) Privacy as intellectual property? Stanford Law Rev 52:1125–1173

Schwartz PM, Solove DJ (2011) The PII problem: privacy and a new concept of personally identifiable information. New York Univ Law Rev 86(6):1814–1894

Solove DJ (2007) The future of reputation: gossip, rumor, and privacy on the internet. Yale University Press, New Haven

Weber RH (2003) Towards a legal framework for the information society. Schulthess, Zurich

Weber RH (2009) Internet governance: regulatory challenges. Schulthess, Zurich

Weber RH, Heinrich UI (2011) IP Address allocation through the lenses of public goods and scarce resources theories. scripted 8(1): 69–92. http://www.law.ed.ac.uk/ahrc/script-ed/vol8-1/weber.pdf. Accessed 31 Jan 2012

Weber RH, Weber R (2010) Internet of things: legal perspectives. Schulthess, Zurich