Survey on Legal Data Protection Norms Relevant to Automated Network Infrastructure Analysis*

Ingo Ritter, Martin Kappes, Peter Wedde, Rüdiger Gad, Andreas Renner
University of Applied Sciences Frankfurt am Main, Germany
ritter@fb2.fh-frankfurt.de
kappes@fb2.fh-frankfurt.de
wedde@fb2.fh-frankfurt.de
rgad@fb2.fh-frankfurt.de
arenner@fb2.fh-frankfurt.de

Abstract:
In order to conduct an automated network infrastructure discovery yielding a complete description of the network and its components as possible, it is necessary to analyze data flows and services provided in the network. However, if such an analysis is done without restrictions, it may also capture, process and store personal data of network users. Therefore, such an analysis could be illegal according to data protection norms as personal rights of users might be affected.

In this paper, we identify the relevant German data protection norms in such scenarios and subsequently analyze the applicability of these norms to the layers of the network reference model. Furthermore, we study the relevance of application layer protocol header fields for some protocols and evaluate whether and in which cases the use of such fields is allowed with respect to data protection norms. Our main emphasis is the trade-off between data protection and legitimate IT security interests when conducting an automated network infrastructure analysis.

1 Scenario Characterisation

Computer networks are an essential, central, and crucial part of every business or governmental communication infrastructure. A successful attack on the infrastructure could lead to catastrophic consequences. Companies and authorities have taken up the challenge and are striving to secure their networks. Moreover, IT security has become an important element of risk management strategies in accordance with e.g. Basel-II guidelines for IT based business processes with “balance-sheet quality” (compare among others [Pin06]), the “Sarbanes-Oxley-Act” (SOX), and other legal requirements.

A common first step of an IT security audit is to analyze the security of the network infrastructure (see [fSidI08]). We focus on an automated scanning approach based on both active (“port scanning”) and passive (“sniffing”) methods (see [HK09]). Such a procedure results in the basic identification of services, components, and additional information like

*This work has been funded by the German Federal Ministry of Education and Research (grant 1783X08)

the operating systems used as well as traffic flows within the given network. This set of information is the starting point for our further consideration.

Additionally, actively generated data, such as from "port scanning" must be considered separately. Such active probing techniques that can be used to identify network infrastructures by creating traffic and connections not containing any personal data. Furthermore, no communication of users is analyzed in this process. The focus in this paper is on assessing the legal norms for personal data so we emphasis on "sniffing".

It is clear that the topic of our paper is interesting to very different audiences, namely legal as well as technical experts. Thus, the technical and legal background of potential readers is very heterogeneous. In the following, we will try to strike a balance between these fields keeping the paper readable for readers from both audiences.

Though we are primarily concerned with German laws, references to relevant European standards are also given.

2 Standards and Statutory Regulations

Judicature for IT always spans different areas of legislation. So, it is impossible to narrow down the relevant legal norms in general. Instead, they have to be identified on a case-by-case basis. Here, we have to consider network (tele-)communication, communication based teleservices and data protection laws. Therefore, we will first clarify the scope of valid legal norms for automated network infrastructure analysis.

2.1 Telecommunications Act (TKG)

The legislative purpose of the German Telecommunications Act[1] is according to §1 TKG "technology-neutral regulation, to promote competition and efficient infrastructures in telecommunications and to guarantee appropriate and adequate services". The regulation purpose is expanded by generic standards like the responsibility to "safeguard user, most notably consumer, interests in telecommunications and to safeguard telecommunications privacy" according to §2 (1) 1 TKG (see [G$^+$06, *Schuster* in Beck'scher TKG-Kommentar, released by Geppert et al, 3rd Edition Munich 2006, §§1 and 2 par. 2 , Rn. 3-37]). In European law context TKG is an implementation of several European guidelines[2].

As a basic rule electronic communication is subject to "Telecommunications Secrecy" according to Article 10 Basic Law for the Federal Republic of Germany (GG)[3]. Respective legal provisions can be found in §88 TKG. §88 par. 3 TKG limits the aforementioned Article for protection of "technical systems" as long as communication content knowledge

[1] Telecommunications Act (TKG) of 22 June 2004 (Federal Law Gazette I p 1190), as amended by Article 2 of the Law of 14. august 2009 (Federal Law Gazette I p 2821).

[2] see 2002/21/EC, 2002/20/EC, 2002/19/EC, 2002/22/EC all of 7.3.2002; 2002/58/EC of 12.7.2002;

[3] Promulgated by the Parliamentary Council on 23 May 1949 as amended up by Article 1 of the Law of 29. Juli 2009 (Federal Law Gazette I p 2248).

is restricted to the amount necessary to improve network security. §109 par. 1 TKG additionally demand "appropriate technical arrangements or take other measures in order to protect 1. the privacy of telecommunications and personal data; and 2. telecommunications and data processing systems against unauthorised access" to be taken by the service provider.

In our case, we deem the interception of data intervention as necessary, even if communication confidentiality is compromised (see [G^+06, *Bock* in Beck'scher TKG-Kommentar, released by Geppert et al, 3rd Edition Munich 2006, §88, Rn 26]). In the following, we compare the objectives of improving security against valid data protection laws according to §§91 TKG ff. In many cases, a comprehensive and accurate analysis of a given network infrastructure and valid assessment of the current state with respect to security cannot be achieved if the communication contents must remain confidential. In terms of §91 par. 1 TKG all information covered by the "Telecommunications Secrecy" is synonymous to personal data. From this basis we assume the analysis as legally possible and classify all "sniffed" network communication as personal data.

2.2 Teleservices Act (TMG)

The Teleservices Act[4] regulates so called "teleservices" and their usage. A limitation for priced services is contradicted according to §1 par. 1 TMG ("regardless of whether a fee is charged for"). Therefore we consider this law as applicable to our scenario of a business or authorities computer network. §7 par. 2 TMG also states that the "telecommunication secrecy according to §88 Telecommunication Act" is to be preserved. So in our case every communication regulated by the TMG is considered personal data. In European law context TMG is an implementation of several European guidelines[5].

§§12 f. TMG introduce a so called "ban with permit restriction". This ban can be lifted by the aggrieved party's consent. Therefore, especially important are §15 par. 1 TMG ("to enable utilisation and accounting of Teleservices (usage data)"), §15 par. 3 TMG ("for purposes of [...] (the) demand-oriented design of Teleservices the use of pseudonymised user profiles unless the user does not contradict" under the condition that these data is not "joint with the pseudonym carrier") and §15 par. 8 TMG (for detection and prosecution of theft of services by third parties).

To subsume we consider the possibility to gather pseudonymised usage data by sniffing of network communication with the objective to improve the network communication infrastructure or to identify theft of service by third parties. This restricts further processing of gathered information for other purposes and also the break-up of pseudonymised user profiles. Every subsequent action depends on user consent.

[4] Teleservices Act (TMG) of 26 February 2007 (Federal Law Gazette I p 179), as amended by Article 2 of the Law of 14. August 2009 (Federal Law Gazette I p 2814

[5] see 2000/31/EC of 8.6.2000; 2003/58/EC of 15.7.2003; 68/151/EEC of 22.6.1998; 98/48/EC of 20.7.1998.

2.3 Federal Data Protection Act (BDSG)

The Federal Data Protection Act[6] is a so called “backup act”. §1 par. 3 BDSG defines the subsidiarity principle of BDSG compared to other legal provisions with separate guidelines for handling of personal data. According to §1 par. 1 BDSG its purpose is “to protect individuals against infringement of their right to privacy as the result of the handling of their personal data”. This is primarily realised in “general estimations” that are concretised in separate legislation. In European law context BDSG is among others an implementation of 95/46/EC 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.

According to §3 par. 1 BDSG personal data is defined as “any information concerning the personal or material circumstances of an identified or identifiable natural person”. Information relating to people is data that reveals - in combination with other information - a direct personal reference (this is defined in more detail in [PtEC95] “as any information relating to an identified or identifiable natural person (’data subject’); 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”). It is without any doubt possible for administrative personal to combine personal related data from different sources to clearly identify a person. Due to this we consider both kinds of information synonymous and as equally worthy of protection.

In our scenario BDSG guidelines are applied when a fact is evaluated that is not treated in TMG or TKG (see et al [G$^+$06, *Robert* in Beck’scher TKG-Kommentar, released by Geppert et al, 3rd Edition Munich 2006, §91, Rn 4]) but also for general valid principles like “Data Avoidance” and “Data Economy”.

2.4 Juridical Determination of Validity

During data transfer (see §3 par. 24 TKG “conveyance of signals by means of telecommunications networks”) communication applies to the defaults of the “Telecommunications Secrecy” (see Article 10 (1) GG and §88 par. 1 TKG). Besides the content of communication these defaults also include the “closer circumstances of communication” (see Federal Constitutional Court (BVerfG) 107, 299, 312, constant jurisdiction) like “traffic data” (within the meaning of the directive 2002/58/EC in German law defined according to §3 par. 30 TKG as “data collected, processed or used in the provision of a telecommunications service”) or “location data” (defined according to §3 par. 19 TKG “any data collected or used in a telecommunications network, indicating the geographic position of the terminal equipment of an end-user of a publicly available telecommunications service). Article 10 (1) GG is directed against governmental interventions in telecommunications, whereas §88

[6] In the version promulgated on 14 January 2003 (Federal Law Gazette I, p. 66), last amended by Article 1 of the Act of 14 August 2009 (Federal Law Gazette I, p. 2814)

TKG is addressed towards private telecommunication service provider (see [G^+06, *Bock* in Beck'scher TKG-Kommentar, released by Geppert et al, 3rd Edition Munich 2006, §88, Rn 1-10]).

Before sending and right after the message is received by the recipient and the transmission process ends[7] data protection is subject to the "general personality right including the right of self-determination" (according to Article 2 par. 1 in connection with Article 1 par. 1 GG). This topic is not addressed in this paper.

The first question when assessing an automated topology discovery is whether the intercepted communication falls under the TKG which comprises legislation for "Telecommunications Secrecy". All data containing no personal information can be gathered for topology discovery without restrictions. This changes if collected information can be directly or indirectly matched to users of the infrastructure.

For intercepted user communication the TKG and BDSG are the applicable laws. TKG permits this operation for improvement of the technical infrastructure. Further guidelines of TKG and BDSG like the principle of "Data Avoidance" and "Data Economy" are to be taken into account.

Telecommunication services that "in addition to transmission service [offer] a special contentual service" (see the official statement of reasons for the German Teleservices Act that mentions as an example for telecommunication services internet access or mail services, whereas Voice-over-IP (VoIP) communication exhibits no externally visible difference to conventionally connected via wired telephony) are additionally subject to the regulations of TMG.

3 Layer Based Delimitation of Legal Specifications

Ensuing identification of German data protection norms we focus now on their applicability to the layers of a common layered network reference model (see [Tan02]). Our aim is to provide a clear layer dependent differentiation of data protection norms for automated network infrastructure analysis.

Physical layer information mostly defines electric, mechanic and functional interfaces besides physical addressing. According to the official statement of reasons of TMG, telecommunication services that merely provide "conveyance of signals on electronic communications networks" are to be evaluated only by the jurisdiction of TKG and BDSG. We consider this data as contextual non-critical.

Network, internet and transport layer are providing an addressing scheme and are responsible for the delivery of data to a device and an application. In contrast to information from the physical layer they contain information relatable to persons like MAC- or IP-addresses. This information opens the possibility to associate digital information to natural person so we consider them as sensitive. The TMG is only applicable to "special contentual service"

[7] see the judgement delivered by the Federal Constitutional Court 2 BvR 2099/04 (Second Senate) of 02.03.2006

that could not be identified in these layers so only TKG and BDSG are applicable.

Application layer data is classified as most critical due to its obviously strong personal reference for example information from sniffed SMTP mail traffic. Due to the vast differences in application layer protocols every protocol is to be treated separately to identify the relevant legal norms. TKG and BDSG are always applicable but in case of a "special contentual service" TMG must also be considered.

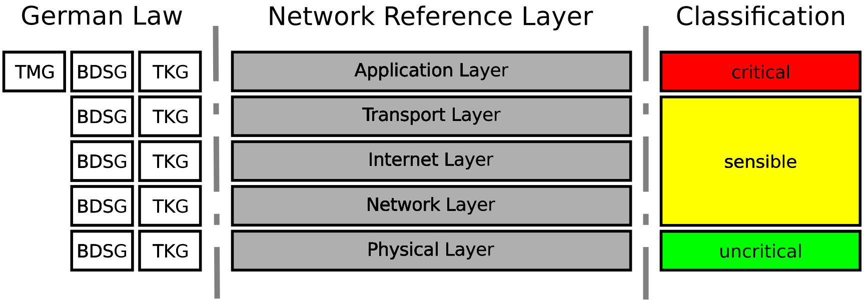

Figure 1: Applicable German legal norms and level of sensitivity considered per network layer

3.1 Information on Layers Two to Four

MAC- and IP-addresses are the most obvious person-based references acquired in layer two to four. Dealing with IP-addresses in the context of data protection in internet jurisdiction leads to a controversial debate[8]. In contrast the situation for corporate computer networks is non-ambiguous. It is rather easy for system administrators to connect IP-addresses to real persons. That is equally valid for MAC-addresses of network interfaces. This means by example that a complete pseudonymisation demands the change of both IP- and MAC-address. For our case we consider both MAC- and IP-address as personal-based references.

MAC- and IP-address should be at least pseudonymised with every occurrence on every layer in captured data.

[8] If you regard newer judgements of German courts for decisions about person-based reference of IP-addresses there is no clear result. District Court Darmstadt (see et al AZ 9 Qs 490/08 (721 Js 35200/08 - StA Darmstadt) of 09.10.2008) and District Court Berlin (see AZ 23 S 3/07 of 06.09.2007) support a personal reference, while Labour Court Munich (see AZ 133 C 5677/08 of 30.09.2008), Labour Court Bonn (see AZ 9 C 177/07 of 05.07.2007) and District Court Cologne (see AZ 28 O 339/07 of 12.09.2007) answer negative.

3.2 Information from the Application Layer

So far only header information has been examined. A technology called “Deep Packet Inspection” (DPI) can extract further information about application layer protocols by looking at the payload.

The next section evaluates data protection norms applicable to application layer protocol header fields. We recommend the removal of the application layer payload due to the irrelevance for topology discovery but its personal content.

4 Exemplary Examination of Application Layer Protocol Headers

Application layer protocol headers contain several data with relevance to both infrastructure discovery and data protection. That is why we provide guidelines for data protection officers and system administrators for dealing with exemplary header fields.

4.1 Hypertext Transfer Protocol (HTTP)

TMG, TKG and BDSG are all applicable to HTTP messages which provides a “transmission service with special contentual service”.

4.1.1 Authorization

The “Request Header” field could indicate the usage of “Basic Access Authentication”. For this purpose a base64 coded, composite string (user name and password) is used. Because of this credentials are easily reconstructable from intercepted data. This information is irrelevant for infrastructure discovery and should be removed. However the fact that sensible information is transmitted nearly in plain text over possibly unencrypted connections should prompt administrators and data protection officers to further improve network security.

One approach to improvement would be the use of “HTTP Digest Access Authentication” [FHBH$^+$99] when there is administrative access to the selected resource. A better but more complex alternative would be a “Public Key Authentication”.

4.1.2 Host and Referrer

Target and origin of an HTTP request are insignificant for infrastructure discovery. Furthermore they represent a strong personal information. We suggest removal of this information. By using this an illegal and comprehensive employee surveillance including behavioral observation (like reconstruction of a users communicative behaviour) would be possible.

4.1.3 Status Code

The “HTTP Status Code” contains valuable information for infrastructure analysis. For example a “305 Use Proxy” indicates that an authentication proxy has to be used when accessing an internal resource. This could point system administrators to a possibly misconfigured service.

4.2 Dynamic Host Configuration Protocol (DHCP)

The transfer of DHCP protocol messages is subject to TKG and BDSG as a “transmission service” without contentual service.

This further points to the trade-off between data protection and legitimate IT security interests. First, DHCP is special among application layer protocols as all the information is present only in header fields. A good example for this issue is the “Options” field that could contain a wide range of important information (see [AD97]) relevant to network infrastructure discovery.

On the other hand some fields also contain person-based information like for instance the “Vendor Class Identifier”. This could contain vendor specific information about used hardware or operating system. If this hardware information is rare or unique in a network this could be used as a “fingerprint” for real persons that could be used to deanonymize a record set. On one hand using this data a person could possibly be identified on the other hand a topology discovery without this information may loose quality. Due to this we classify this information as personal related. Any further usage must be weighted with security interests.

A combination with information like the time of a DHCP request is avoidable and could be misused for employee performance monitoring. For example a conjunction with MAC- or IP-address is not recommended to preserve data protection interests.

5 Conclusion

Considering that an (automated) infrastructure analysis is a common task for both IT administrators and people responsible for data protection, it is surprising that no detailed examinations of affected data protection norms were conducted before.[9] Moreover, we were unable to find any guidelines on this subject.

Since we were solely interested in personal data, only passive methods (“packet sniffing”) for analysis in the context of German data protection norms had to be studied in a case-based approach. Further, we analyzed the applicable German data protection norms for

[9] There are papers dealing with DPI and data protection norms in the context of net neutrality and Internet Service Providers (ISP) leveraging this technology (see, e.g., [Bed09]) but they mostly consider only the applicability of criminal code norms (Strafgesetzbuch STGB).

the different layers of the network protocol stack.

While “sniffing” collects information across all layers, we focused on application layer protocol header fields for common protocols. The actual payload of the protocol is considered not to be taken into account for the analysis as is of minimal relevance but might contain sensitive personal data. Furthermore, we demonstrated the trade-off between data protection and IT security interests by examining exemplary header fields.

The issue of this trade-off in terms of legislation is particularly clear when comparing the need for “appropriate technical arrangements” for protection in §109 par. 1 TKG against the basic principle of “Data Avoidance” in §3a BDSG. So, before conducting an automated network infrastructure analysis, it is necessary to balance the two.

This is also important for information gathered on layers below the application layer. In a simple analysis, e.g., the illustration of a network’s traffic flows, the pseudonymisation of person-based data such as MAC- or IP-address appears reasonable. However, this is impossible if detailed technical data is necessary for the task under consideration, e.g., the configuration of critical security interfaces like a company firewall as the option to subsequently resolve the pseudonymised data is required. Furthermore, depseudonymisation procedures could contain structured processes such as, e.g., the four-eye-principle.

In summary, capturing network communication is basically a violation of a users rights under data protection laws. Every possibly affected user must be informed about planned “sniffing” processes. Furthermore, the collected data has to be curtailed to the strictly minimal information necessary before analysis to keep the intervention as small as possible.

References

[AD97] S. Alexander and R. Droms. DHCP Options and BOOTP Vendor Extensions. RFC 2132 (Draft Standard), March 1997. Updated by RFCs 3442, 3942, 4361, 4833, 5494.

[Bed09] M. Bedner. Rechtmäßigkeit der Deep Packet Inspection, Nov 2009.

[FHBH$^+$99] J. Franks, P. Hallam-Baker, J. Hostetler, S. Lawrence, P. Leach, A. Luotonen, and L. Stewart. HTTP Authentication: Basic and Digest Access Authentication. RFC 2617 (Draft Standard), June 1999.

[fSidI08] Bundesamt für Sicherheit in der Informationstechnik. IT-Grundschutz-Vorgehensweise: BSI-Standard-100-2, 2008.

[G$^+$06] M. Geppert et al. Beck’scher TKG-Kommentar. 3. Auflage, 2006.

[HK09] F. Happel and M. Kappes. Tackling Network Security in Small to Medium Businesses. In *ITA09 – Third International Conference on Internet Technologies & Applications, Wrexham, North Wales, UK*, September 2009.

[Pin06] P. Pinder. Preparing Information Security for legal and regulatory compliance (Sarbanes–Oxley and Basel II), 2006.

[PtEC95] European Parliament and the European Council. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995, October 1995.

[Tan02] A. Tanenbaum. *Computer Networks 4th Edition*. Prentice Hall Professional Technical Reference, 2002.