Packet Inspection — Shifting the Paradigm of Fundamental Rights

Agata Królikowski

Humboldt-Universität zu Berlin, Department of Computer Science, Berlin, Germany
krolikow@informatik.hu-berlin.de

Abstract. In recent years deep packet inspection (DPI) has often been cited as a major factor in the debate concerning net neutrality. Packet inspection (PI) enables a profound analysis of the contents of IP-packets, especially with respect to the application layer and private data. To protect against this sort of privacy invading attack users are usually advised to encrypt as much of their data as possible in an online transaction. However, current PI-engines not only use plain text analysis but also employ a variety of statistical methods. This in turn allows the analysis and classification of packets even if encryption or obfuscation methods have been applied. It is possible to monitor and shape packet flows in real time and on a large scale. These PI-engines are deeply embedded in the current network infrastructure due to the requirements of lawful interception. This brings about a huge potential for misuse, because the engine's operation is not 'visible' to the end-user.

Keywords: packet inspection, fundamental rights

1 Introduction

The following paper reflects a recurring debate concerning how far technology may push the legal envelope. This aspect is particularly noticeable in technologies such as nuclear power plants, where the question arises: to what extent is the state under obligation to protect the rights of its citizens [1], [2]? In this example, the fundamental rights involved are the right to life or environmental protection — which for example in Germany is a constitutional principle, pursuant to Article 20 a Grundgesetz (GG) — and shows impressively the tension between a technological advance and legal boundaries. Technological developments may lead into a legal grey area where fundamental rights still apply but the lack of regulations makes it likely that these rights will not be enforced. This aspect also occurs in packet inspection (PI) since analyses in telecommunication systems also affect fundamental rights such as the protection of private data, the privacy of correspondence, posts and telecommunications[1], freedom

[1] Article 8 of the Charter of Fundamental Rights of the European Union, the German Bundesdatenschutzgesetz (BDSG) or § 88 and Section 2 of the Telekommunikationsgesetz (TKG).

of expression and information[2], freedom of the arts and science[3], freedom to conduct a business[4], freedom of thought, conscience and religion[5], or freedom of assembly and of association[6].

The question at hand is not the use of PI in the way it is used in non-democratic countries, e.g., to censor information [3], [4], but how PI may affect fundamental rights when it is used according to its purposes: to secure networks, to shape traffic, or to allow new pricing models.

Another concern is that in contrast to civil law, which regulates the matter between private persons, fundamental rights may only be enforceable against the state. Since networks are operated by private corporations this raises the question of how and with whom should protections of fundamental rights that may be affected by PI be enforced. This implies also both the questions of how to regulate the use of PI and how to implement regulations regarding the technology as a regulatory subject.

This paper addresses different aspects of packet inspection. Starting with general concepts such as the development and challenges of packet inspection it is shown that for Internet users it is virtually impossible to protect themselves against packet inspection in a technical manner. Section 3 describes the use cases of packet inspection. Section 4 describes why as a result from section 2 and 3 fundamental civil rights cannot be enforced by or guaranteed to the Internet users.

2 Technical Aspects of PI

Network packets are pieces of information that consist of a header and the payload. Packet inspection first emerged in the 1980s [5]. While early firewall systems performed analyses on the network layer, better hardware and new approaches lead to firewalls that were both able to examine packets on different OSI layers and to keep track of sessions and connections (thus, called stateful PI). These different concepts are unified in current PI systems that combine firewalls, stateful PI, and intrusion detection systems (IDS) [6].

An analysis consists of two phases. In the first phase, patterns (signatures) of packets or flows are examined. In the second phase, packets are classified based upon these signatures and – depending on their class – forwarded, delayed or dropped.

These signatures can be determined in different ways. Regarding the OSI model, PI performed on different layers it is called shallow, medium, or deep packet inspection (DPI). DPI therefore denotes the analysis of every bit of a packet including every header and personal data like emails, chat messages, pictures or passwords [6], [7].

However, generating signatures by analyzing data streams bit by bit involves many difficulties. String matching algorithms require a lot of storage and processing power.

---

[2] Article 11 of the Charter of Fundamental Rights of the European Union, Article 5 GG.
[3] Article 13 of the Charter of Fundamental Rights of the European Union, Article 5 GG.
[4] Article 16 of the Charter of Fundamental Rights of the European Union, Article 12 GG.
[5] Article 1 of the Charter of Fundamental Rights of the European Union, Article 4 GG.
[6] For example in Germany Article 12 of the Charter of Fundamental Rights of the European Union, Article 9 GG.

In addition to this, the analyses have to be performed at wire speed since packet delays may cause a network bottleneck. Common throughput rates in networks now are 10 Gbit per second, a newer standard called 100 GbE (with a throughput up to 100 Gbit per second) is on the horizon [6], [9].

Therefore, on the one hand, existing string matching algorithms have been evolved, while on the other hand special purpose hardware has been developed. String matching algorithms accomplish classification by either comparing predefined (exact) strings, regular expressions, or calculating hash values [6]. New hardware systems like the Advanced Telecommunications Computing Architecture or new processor designs like hybrid multi-core architectures are highly specialized and therefore able to meet the demands of real time and accurate (i.e., false positive rates are below 1%) analysis [10], [12].

But DPI fails as soon as the data are encrypted or obfuscated. Encryption not only replaces plain text with cipher text but also changes the distribution of occurring characters and thus removes links between the plain and the cipher text. Methods based on string matching algorithms or hash values are not able to reliably classify signatures anymore.

This is a significant drawback since encryption is used in standard protocols such as SSH, SSL/TLS, IPSec or in implementations like PGP, just to name a few. Therefore, numerous research projects have addressed this issue and developed new approaches to overcome DPI's deficiencies.

Instead of analyzing plain text, statistical packet inspection methods (SPI) exploit the fact that even if data are encrypted or obfuscated there is enough information left that can be classified without breaking the encryption. Signatures are then generated by analyzing patterns (in this context commonly referred to as fingerprints), that emerge within single packets, fragments, or packet flows. The patterns depend on the underlying protocol, the applied encryption algorithm, or the transmitted personal data.

Such patterns include different features and there are various ways to define and calculate them. Common features, for example, are the packet size or interarrival times between packets [13], [14]. As described in [15], the range of features is broad and all of them may be combined in a fingerprint if necessary. The more features are used to generate a pattern the more robust it is against network influences or user protection such as padding or traffic obfuscation.

Generally pattern recognition is implemented in machine learning algorithms such as clustering, Bayesian filters (as is the case with spam filters), and support vector machines, just to mention a few. These algorithms have been tried and tested for a long time in classification research and are, therefore, ideal as a basis for further developments [16].

By now these methods are able to calculate patterns and apply them as signatures to classify data in real-time as described in [17].

Usually traffic classification is performed on distinct application layer protocols such as BitTorrent or Skype, which also use strong encryption algorithms and obfuscation. But classification may also be applied to web sites, downloaded films or spoken language even if the packet flow is encrypted. The use of encryption or obfusca-

tion itself may even be utilized as information to pre-sort data into classes of encrypted and non-encrypted data and from there perform either DPI or SPI. It is also applicable in cases where any kind of data protection is prohibited or raises suspicion of malicious behavior to block this traffic. Thus, common protection such as tunnel software or encryption does not prevent traffic analysis. In combination with DPI PI-engines currently permit throughput rates of several hundreds of Gbit per second. [18]. In comparison, the greatest Internet exchange point of the world (Deutscher Commercial Internet Exchange DE-CIX) has a throughput of 1.85 Tbit per second [19].

In conclusion, PI inspection may be performed either in unencrypted or encrypted data in real time at any place and any time on any user [20].

3 Use Cases

As stated above, PI can be used to secure networks. But the development of elaborate algorithms and highly specialized high-end hardware has led to a further possibility of applications such as traffic shaping, new pricing models or personalized advertisements.

In all these cases the basic technology is the same allowing different system extensions [21].

Traffic shaping includes both quality of service (QoS) and specific content filtering. It is used to control the efficiency of a network that is to delay packets to avoid congestion and thus improve latency.

While QoS is a technical term and refers to packet handling depending on whether it is a real-time/non-real-time application (e. g. post office protocol (POP) vs. real-time transport protocol (RTP)), specific content filtering refers to blocking or delaying packets depending on the content of packets such as P2P-traffic, certain websites or VoIP [6], [22].

By analyzing the amount of data a user has downloaded or applications he or she may have used it is possible for ISPs to charge every service separately. For example, the company Plusnet uses different pricing depending on volume, amount of emails, or hours of game playing [23].

Another application of PI is the identification of users' interests to be able to customize advertisements. The perhaps most prominent company in this context is Phorm, which delivered personalizes ads to users depending on the data delivered by British Telecom (BT), TalkTalk and Virgin Media and thus triggered a popular outrage [24]. The company Kindsight on the other hand uses these kinds of ads in a no-cost version of their security services. While Kindsight analyzes a user's traffic by applying DPI to detect attacks on his or her network, at the same time this data is used to personalize ads. The user may prevent this by paying a monthly fee [25]. Furthermore, there is lawful interception that refers to surveillance of private communication pursuant to lawful authority by government agencies.

As a general rule, network operators are legally bound to provide a surveillance infrastructure in order to be granted permission to exploit their network commercially[7]. According to the European Council Resolution on the lawful interception of telecommunications, "Law enforcement agencies require a real-time, fulltime monitoring capability for the interception of telecommunications" [26]. Because the technology in lawful interception engines is basically the same as in traffic shaping machines, there is no possibility to distinguish if PI is performed for legal or commercial reasons. Thus the paradigm of fundamental rights may be shifted.

4 Impact on Fundamental Rights

4.1 Analyses and Storage of Data

To determine the legal implications of PI it is necessary to distinguish between the analysis of encrypted and unencrypted data.

While the analysis of unencrypted data may result in the examination of private data, in those cases where data is encrypted the examination of private data depends on the OSI-layer the encryption is applied to. For example, suppose that the entire payload of a packet is encrypted and only the IP-addresses are visible (as it is in the IPSec tunnel mode using Encapsulating Security Payload (ESP)) [27]. Then it might appear that a statistical analysis of the payload would not imply an examination of private data since the plain text is not visible. Thus the right of protection of private data would not be affected. But statistical analyses are specifically designed to overcome this obstacle and to derive the content from encrypted data.

As described above, it is possible to derive information from spoken words, downloaded movies or visited websites even if encryption or tunneling is applied. It is then possible to link the gathered information and to create a new context out of it. This picture can be so accurate that it profiles users even if this data is 'anonymized'. However, it is not anonymous, because it is possible to distinguish every single user [28].

With this in mind it can reasonably be assumed that the analysis of encrypted and unencrypted data may result in similar legal consequences regarding the invasion of privacy. The following discussion is based on this assumption.

4.2 Lawful Analyses

As stated above, law enforcement authorities are empowered to intercept communications based on different laws. Lawful interception is strictly regulated regarding the legal requirements involved in the application to monitor the communication of persons. In Germany only judges may order an interception and only in cases of serious

[7] Pursuant to § 110 TKG.

offenses such as high treason or murder[8]. Furthermore, the core area of the private conduct of life must be granted[9].

A critical point concerning which data may be monitored is so-called traffic data. This data includes all information regarding a telecommunication transaction, "indicating the communication's origin, destination, route, time, date, size, duration, or type of underlying service" [29]. Traffic data is very sensitive information, since it might provide a detailed knowledge about a user's habits, friends, lifestyle etc. For this reason, Germany's highest Constitutional Court (Bundesverfassungsgericht – BVerfG) decided to annul the implementation of the European directive regarding data retention [30].

On the other side, ISPs are allowed to collect traffic data for billing purposes. If this data is encrypted they have to distinguish different services by applying SPI methods. So traffic data may be collected in real time regarding every user at any point in time, though it may only be used for billing purposes. The resulting question therefore is whether it is possible to know which data is analyzed and for whatever reason. The answer is: We cannot know. Furthermore, in cases of traffic analyses for billing purposes this data has to be stored to issue the invoice. This information is then available and may be used for further analyses. Similar to data retention, this mirrors certain characteristics of preventive surveillance.

4.3 Commercial Analyses

Besides the right to privacy there are other fundamental rights that may be affected. Delaying or blocking packets may affect the right to freedom of expression and information.

One peril is that network operators discriminate against packets of competitors thus causing economic damage [31].

It is also worrying, however, when packets are discriminated against on the basis of their content or because of monetary reasons.

On the one hand there is political, religious and other information available on the Internet that determines freedom of information, when a person wants to form an opinion based on this information and there is the freedom of expression on the other hand when the same person wants to voice his or her opinion. Blocking or delaying packets within this process would affect fundamental rights regarding freedom of expression, thought, religion, art etc.

Besides, the Internet not only provides a platform for communication or downloading media, but is also a platform for concluding contracts or forming associations of any type. Interferences with this kind of use of the Internet may affect the right regarding the freedom to conduct a business or freedom of assembly and of association.

---

[8] For example, in Germany pursuant to § 3 Gesetz zur Beschränkung des Brief-, Post- und Fernmeldegeheimnisses (Artikel 10-Gesetz - G10), § 20l Bundeskriminalamtgesetz (BKAG) or § 100 a Strafprozessordnung (StPO).

[9] For example, example pursuant to European Council Resolution on the lawful interception of telecommunications (96/C 329/01) or § 3 a G10.

But it is a technical challenge to detect traffic discrimination. There are different projects concerning this issue, but they take standard protocols such as peer-to-peer, POP or video downloads into account, but not other benchmarks [32-34].

These projects may help us to understand if and how traffic shaping may be applied. Therefore it is necessary to raise awareness and to educate users to use software that may help detect packet discrimination.

5 Shifting the Paradigm of Fundamental Rights

Taking all this into consideration, the paradigm of the fundamental right regarding the protection of private data is shifting in many different ways:

Firstly, network operators are legally bound to provide surveillance infrastructure, which they in turn may use for their own interests.

Secondly, in order to meet their economic interests, private corporations such as ISPs are allowed to collect data that normally is restricted and may only be collected within the scope of lawful interception. In contrast to enforcement authorities, companies do not have to meet the high legal requirements regarding the cause of the collection and the extent of the collected data. Finally, the storage of such data corresponds to data retention, which may undermine the principle of presumption of innocence and right of defense.

Furthermore, due to the fact that corporations may collect data or shape traffic because of their economic interests, there is a shift in weighting particular rights because these interests appear to weigh more than the right to privacy, freedom of expression and information and the other rights that were mentioned above.

Finally, another shift should be noted. Fundamental rights apply in the relation between citizen and state. They serve the purpose to restrict the power of a state. However, companies are private institutions. This therefore raises the question of how to enforce fundamental rights against them, which is beyond the scope of this paper. But it should be noted that PI increases the number of fundamental rights affected by private companies.

6 Conclusion

PI-engines serve different purposes. On the one hand they have been designed to secure networks, help to avoid congestion and increase bandwidth efficiency. On the other hand PI may be used to deliver personalized ads, shape traffic or to monitor communication on behalf of law enforcement authorities. Due to the fact that network operators are obliged by law to provide a surveillance infrastructure, which in the case of PI may also be used to meet business interests, these operators are offered a loophole for misuse.

Thus, the essence of several fundamental rights is affected. It is necessary to draw the attention of politicians and legislators to these shifts and to close the gap between the technology and the law.

References

1. Boehme-Neßler, V.: Unscharfes Grundgesetz — Anmerkungen zum Verfassungsrecht in der digitalisierten Welt. In: Institut für Wirtschaftsrecht, 60 Jahre Grundgesetz, pp. 155 — 188, Kassel University Press (2010)
2. Roßnagel, A., Mayer-Tasch. P. C., Saladin, P. V.: Radioaktiver Zerfall der Grundrechte: Zur Verfassungsvertraglichkeit der Kernenergie. Beck, München (1984)
3. Champagne, A.: Watching over you. In: Le Monde diplomatique, http://mondediplo.com/2012/03/16internet, January (2012)
4. Brodkin, J: Iran reportedly blocking encrypted Internet traffic, http://arstechnica.com/tech-policy/news/2012/02/iran-reportedly-blocking-encrypted-internet-traffic.ars, February (2012)
5. Ingham, K., Forrest, S.: A History and Survey of Network Firewalls, A history and survey of network firewalls. Tech. Rep. TR-CS-2002-37, University of New Mexico Computer Science Department (2002)
6. Serpanos, D. N., Wolf, T.: Architecture of network systems. Morgan Kaufmann, Burlington, MA (2011)
7. Anderson, N.: Deep packet inspection meets 'Net neutrality, CALEA, vom 25 July (2007), http://arstechnica.com/hardware/news/2007/07/Deep-packet-inspection-meets-net-neutrality.ars/2
8. IEEE 802.3 Ethernet, http://standards.ieee.org/about/get/802/802.3.html
9. IEEE P802.3ba 40Gb/s and 100Gb/s Ethernet Task Force Public Area, http://www.ieee802.org/3/ba/public/index.html
10. Ipoque, Deep packet inspection solutions for network operators, http://ipoque.com/en/products/pace-network-analysis-with-deep-packet-inspection
11. Netronome, http://www.netronome.com/pages/heterogeneous-architecture
12. AdvancedTCA Specifications for Next Generation Telecommunications Equipment, http://www.picmg.org/v2internal/resourcepage2.cfm?id=2
13. Liberatore, M, Levine, B. N.: Inferring the Source of Encrypted HTTP Connections. In CCS '06: Proceedings of the 13th ACM Conference on Computer and Communications Security, pp. 255–263. ACM Press, New York (2006)
14. Alshammari, R., Nur Zincir-Heywood, A.: Can encrypted traffic be identified without port number, IP addresses and payload inspection? Computer Networks, 55, 1326–1350 (2010)
15. Hjelmvik, E., John, W.: Breaking and Improving Protocol Obfuscation, Dep. of Computer Science and Engineering, Chalmers University of Technology, Technical Report No. 2010-05, ISSN 1652- 926X (2010), http://publications.lib.chalmers.se/cpl/record/index.xsql?pubid=123751
16. Webb, A.: Statistical Pattern Recognition. Wiley, Chichester, UK (2003)
17. Ipoque, Net Reporter, http://ipoque.com/en/products/net-reporter
18. Procera,Products, http://www.proceranetworks.com/pdf/products/overview/Procera_Overview_Brochure_mech_2012-4-8.pdf
19. Packet Clearing House, Internet Exchange Directory, https://prefix.pch.net/applications/ixpdir/?show_active_only=0&sort=traffic&order=desc.
20. Nguyen, T. T. T., Armitage, G.: A survey of techniques for internet traffic classification using machine learning. In: Communications Surveys and Tutorials, IEEE, 10(4), 56 – 76 (2008)
21. Ipoque, Products, http://ipoque.com/en/products

22. Bendrath, R.: Global technology trends and national regulation: Explaining Variation in the Governance of Deep Packet Inspection (2009), http://userpage.fu-berlin.de/bendrath/ISA09_Paper_Ralf%20Bendrath_DPI.pdf
23. Plusnet, http://www.plus.net/broadband/?source=subBox
24. Pfanner, E.: 3 Internet Providers in Deal for Tailored Ads, http://www.nytimes.com/2008/02/18/technology/18target.html?_r=1&oref=slogin
25. Kindsight: Subscription options, http://www.kindsight.net/en/solution/subscription
26. European Council Resolution on the lawful interception of telecommunications (96/C 329/01), http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31996G1104:EN:HTML
27. Kent, S, Seo, K: Security Architecture for the Internet Protocol, RFC 4301 (2005), http://tools.ietf.org/html/rfc4301
28. Hoeren, T: Google Analytics — datenschutzrechtlich unbedenklich? In: Zeitschrift für Datenschutz (ZD), 1/2011, pp. 3–6 (2011)
29. Convention Committee on Cybercrime, Convention on Cybercrime, http://conventions.coe.int/Treaty/EN/Treaties/html/185.htm
30. Decision of the German Federal Constitutional Court, 2 March 2010, BVerfG, 1 BvR 256/08 vom 2.3.2010, http://www.bverfg.de/entscheidungen/rs20100302_1bvr025608.html
31. Gilroy, A. A.: Access to Broadband Networks: The Net Neutrality Debate, CRS Report for Congress (2011)
32. Dischinger, M., Marcon, M., Guha, S., Gummadi, K. P., Mahajan, R., Saroiu, S.: Glasnost: Enabling End Users to Detect Traffic Differentiation. In: Proceedings of the 7th USENIX conference on Networked systems design and implementation (2010)
33. Kanuparthy, P., Dovrolis, C.: Diffprobe: Detecting ISP service discrimination. In: INFOCOM, 2010 Proceedings IEEE, pp. 1–9 (2010)
34. S. Basso, A. Servetti, J. C. De Martin: The network neutrality bot architecture: a preliminary approach for self-monitoring of Internet access QoS. In: Proceedings of ISCC 2011, pp.1131–1136 (2011)