View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

ECONSTOR
WWW.ECONSTOR.EU

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationszentrum Wirtschaft
The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics

Knieps, Günter

Working Paper

Market driven network neutrality and the fallacies of Internet traffic quality regulation

Diskussionsbeiträge // Institut für Verkehrswissenschaft und Regionalpolitik, No. 136

Provided in cooperation with:
Albert-Ludwigs-Universität Freiburg im Breisgau

Suggested citation: Knieps, Günter (2011) : Market driven network neutrality and the fallacies of Internet traffic quality regulation, Diskussionsbeiträge // Institut für Verkehrswissenschaft und Regionalpolitik, No. 136, http://hdl.handle.net/10419/47440

Nutzungsbedingungen:
Die ZBW räumt Ihnen als Nutzerin/Nutzer das unentgeltliche, räumlich unbeschränkte und zeitlich auf die Dauer des Schutzrechts beschränkte einfache Recht ein, das ausgewählte Werk im Rahmen der unter
→ http://www.econstor.eu/dspace/Nutzungsbedingungen
nachzulesenden vollständigen Nutzungsbedingungen zu vervielfältigen, mit denen die Nutzerin/der Nutzer sich durch die erste Nutzung einverstanden erklärt.

Terms of use:
The ZBW grants you, the user, the non-exclusive right to use the selected work free of charge, territorially unrestricted and within the time limit of the term of the property rights according to the terms specified at
→ http://www.econstor.eu/dspace/Nutzungsbedingungen
By the first use of the selected work the user agrees and declares to comply with these terms of use.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der Leibniz-Gemeinschaft

Market driven network neutrality and the fallacies of Internet traffic quality regulation

by Günter Knieps

Discussion Paper
Institut für Verkehrswissenschaft und Regionalpolitik
No. 136 – Revised Version: March 2011

Abstract:

In the U.S. paying for priority arrangements between Internet access service providers and Internet application providers to favor some traffic over other traffic is considered unreasonable discrimination. In Europe the focus is on minimum traffic quality requirements. It can be shown that neither market power nor universal service arguments can justify traffic quality regulation. In particular, heterogeneous demand for traffic quality for delay sensitive versus delay insensitive applications requires traffic quality differentiation, priority pricing and evolutionary development of minimal traffic qualities.

Prof. Dr. Günter Knieps
Institut für Verkehrswissenschaft und Regionalpolitik
Universität Freiburg
Platz der Alten Synagoge, 79085 Freiburg i. Br.
Phone: (+49) - (0)761 - 203 - 2370
Fax: (+49) - (0)761 - 203 - 2372
E-mail: guenter.knieps@vwl.uni-freiburg.de

The electronic version of this paper is available under:
https://portal.uni-freiburg.de/vw/files/files-publikationen/disk/disk136

1. Introduction: Alternative forms of traffic quality regulation in the U.S. and in Europe

In order to transport data packets from one Internet access services network to another, transmission via Internet backbone services networks is required. Both are Internet traffic services based on telecommunications capacities, combined with Internet logistics (transmission protocols, Internet protocol, routers etc.). Telecommunications capacities are produced by local telecommunications infrastructure as well as long-distance telecommunications infrastructure. Internet access services require not only local telecommunications infrastructure but also long distance telecommunications infrastructure capacities. Complementary to Internet traffic services, Internet application services are provided, including e-mails, file sharing, Voice over IP, Video games, search machines and other content delivery services.

The network neutrality debate is gaining increasing momentum worldwide. Its focus is on the regulation of Internet traffic management. The most extreme version of mandatory network neutrality would be to transform the traditional Transmission Control Protocol (TCP) into a technical regulatory concept. Since TCP assigns all data packets equal priority, best effort average traffic quality results endogenously without any traffic quality guarantee. Due to the transition from narrowband access to broadband access congestion management by TCP is challenged. As a consequence, a growing debate on traffic quality regulation has developed. Whereas in Schwartz, Weiser (2009, 1) the term network neutrality has been used as a regulatory concept addressing which deviations from the TCP should be permitted, the focus of the current debate increasingly moves towards specific forms of traffic management regulation.

Two different opposing approaches to traffic management regulation became relevant in the U.S. and in Europe. The net neutrality debate in the U.S. has already a rather complex history. The Network Neutrality Act of 2006, focusing on net neutrality regulations, in particular the obligation of broadband network providers not to impose a surcharge for prioritization or quality of service, never became law.[1] Since then the FCC and related Supreme Court Cases on its statutory competence to deal with net neutrality have played an increasingly larger role (FCC 2010). In this context the FCC differentiates between reasonable network management and unreasonable discrimination (FCC, 2010b, 40 ff.). Paying for priority arrangements between Internet access service providers and Internet application providers to favor some traffic over other traffic is considered unreasonable discrimination, because an incentive would occur to erode the quality of service provided to non-prioritized traffic (FCC, 2010b, 18, 43). In contrast, differential treatment of traffic enabling users to choose quality of service enhancements could be reasonable (FCC, 2010b, 41).

In Europe the starting point of the network neutrality debate was the Commission declaration on net neutrality (European Commission, 2009b, C 308/2): "…the creation of safeguard powers for national regulatory authorities to prevent the degradation of services and the hindering or slowing down of traffic over public networks". Furthermore, Directive 2009/136/EC, Article 22 states: "…, Member States shall ensure that national regulatory authorities are able to set minimum quality of service requirements on an undertaking or undertakings providing public communications networks".[2] Based on this EU framework, there is currently a review of the German telecommunications law under consideration focusing explicitly on requirements of minimal service quality.[3] In particular § 45o deals with requirements of minimal service quality, which the Fed-

---

[1] Text of H.R. 5273 [109th]: Network Neutrality Act of 2006, http://www.govtrack.us/congress/bill.xpd?bill=h109-5273

[2] Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws, Official Journal of the European Union, L 337/11. See also Annex 3, Quality of service parameters (L 337/35).

[3] Entwurf eines Gesetzes zur Änderung telekommunikationsrechtlicher Regelungen, Stand 15. 09. 2010, 3. http://www.bmwi.de/BMWi/Redaktion/PDF/Gesetz/referentenentwurf-tkg,property=pdf,bereich=bmwi,sprache=de,rwb=true.pdf

eral Ministry of Economics and Technology may issue and transmit to the Federal Network Agency.

Although minimal quality regulation and prohibition of priority pricing are different regulatory interventions, they are both forms of traffic management regulation. Their focus is on the best effort average traffic quality of the TCP and on regulations aimed at avoiding the deterioration of best effort traffic quality of the non-prioritized traffic. In the U.S., the regulatory focus of the FCC is in particular on broadband Internet access service providers with market power (FCC, 2010b, 19 ff.). In contrast, the focus of the net neutrality debate in Europe is on minimal traffic quality requirements for universal service obligations for providers of public communications networks rather than market power regulation of Internet access service providers. However, the aim of both approaches of traffic quality regulation is to avoid the deterioration of best effort quality as a consequence of traffic quality differentiation (FCC, 2010b, 43; Holznagel, 2010, 2).

It is important to differentiate between mandatory network neutrality and market driven network neutrality. Mandatory network neutrality consists of ex ante regulation of traffic management based on the traditional TCP. In contrast, market driven network neutrality means an entrepreneurial search for traffic allocation in such a way that there are no incentives for the Internet traffic service provider to discriminate against possible Internet application services on the basis of network capacity requirements. This is the case if any application is charged according to the opportunity costs of the traffic capacities it requires (Knieps, 2010).

The paper is organized as follows. In the subsequent section 2 the transition from edge based technical congestion management to an economically based traffic management in broadband Internet is demonstrated, taking into account the opportunity costs of traffic capacity. In section 3 a critical appraisal of traffic quality regulation is provided. It can be shown that neither market power nor universal service arguments can justify the prohibition of priority pricing or the regulatory requirement of minimum traffic qualities. In section 4 it is pointed out that market driven network neutrality requires the entrepreneurial search for traffic quality differentiation and priority pricing as well as the evolutionary development of minimal traffic qualities.

2. From TCP to traffic quality management

2.1 TCP based best effort average traffic quality

The reference point of traffic quality regulation is the best effort traffic quality resulting from TCP, assigning all data packets equal priority. Best effort average traffic quality results endogenously, depending on traffic level for each given capacity. The congestion level results from the amount of data packets sent according to the technical protocol TCP at the end user's computer and the available bandwidth capacity. There is no traffic management and consequently no traffic quality guarantee.

In narrowband Internet TCP-based best effort average traffic quality was sufficient for two reasons. Firstly, all users required traffic for (at least roughly) homogenous narrowband applications, in particular e-mails, thus there was no need for heterogeneous traffic qualities. Secondly, the underlying technical algorithm to solve congestion problems was sufficient, because the possibility of the sender computers causing congestions was strongly limited due to the low narrowband Internet access capacities.

Due to the transition from narrowband to broadband Internet there has been a change in the meaning of best effort traffic quality based on Jacobson's (1988) congestion and control mechanism. Whereas in narrowband Internet the resulting best effort average traffic quality remains sufficient for all applications feasible, in broadband Internet there is an increasing relevance of capacity intense peer-to-peer (P2P) applications. As more users are provided with very high speed Internet access, their capability to cause congestion in Internet traffic becomes larger: "Many P2P applications consume as much bandwidth as they can find" (FCC, 2008, 66). As a consequence, average best effort traffic quality may deteriorate strongly. Moreover, capacity intense, time insensitive applications (e.g. P2P file sharing) are discriminating against non-capacity intense, time sensitive applications (e.g. voice over IP), so that average best effort quality is no longer sufficient to allow the provision of time sensitive applications.

User restrictions complementary to TCP, e.g. a cap on the average user’s capacity, a throttling back of the connection speeds of high capacity users or other forms of contractual restrictions between end-user and traffic service provider, can limit the amount of data packets of the sender and thus influence congestion levels (FCC, 2008, 30; Wu, 2003, 158 ff.). However, they are rather inefficient measures for dealing with these increasing congestion problems and cannot prevent deterioration of best effort traffic quality. Moreover, the requirements of sufficient traffic quality for delay sensitive applications cannot be guaranteed. As a consequence, best effort average traffic quality may be too low for delay sensitive applications and too high for some delay insensitive applications. Moreover, applications for which earlier best effort quality had been sufficient may require higher traffic quality levels than is provided by the deteriorated best effort quality as a consequence of capacity intense peer-to-peer applications.

Incentives arise to shift from best effort TCP-based congestion management at the edge to traffic management by Internet traffic service providers. The increasing importance of traffic management in broadband Internet has also been recognized by the FCC, differentiating between end-user controlled and broadband-provider controlled practices (FCC, 2010b, 41).

2.2 Innovations in traffic quality management architecture

The Comcast Case provides important insights into the future role of traffic management. Due to immense potentials and incentives for using capacity intense and delay insensitive applications by means of P2P swarming systems, P2P programs, such as BitTorrent, enable all peer users in the swarm to share large files. As a consequence there results an increasing scarcity of Internet traffic capacities. The Comcast network management was not based on economic congestion pricing. Instead, Comcast applied an intransparent traffic management practice blocking the BitTorrent P2P applications (FCC, 2008, 2-6).

The question arises how transparent and non-discriminatory traffic management practices can solve the overusage problem of traffic capacities setting the proper incentives to take into account congestion externalities and to provide the required traffic qualities for delay sensitive as well as delay insensitive applications.

In broadband Internet TCP-based best effort average traffic quality faces two serious problems. Firstly, there is a heterogeneous demand for traffic qualities, depending on the delay sensitivity of the applications (voice over IP vs. P2P exchange). Best effort average traffic quality is not sufficient for delay sensitive applications during periods of network congestion. Secondly, capacity intense, delay insensitive applications can deteriorate best effort quality in such a way that more and more applications are hampered. Thus TCP-based best effort average traffic quality can be strongly impaired by a relatively small number of capacity intense applications. In the worst case not even e-mails could arrive in time.

Due to the problems of TCP in broadband Internet there is an increasing need for traffic management by traffic service providers. This requires the transition from TCP to flexible traffic management architectures, in order to reflect the heterogeneous demand for traffic quality and to implement economically based allocation mechanisms for traffic capacity.

There is a variety of traffic architectures capable of implementing quality of service differentiation. Two basic types of traffic quality differentiation architectures have been developed: Integrated Services (IntServ) and Differentiated Services (DiffServ). IntServ architecture guarantees quality of service by reserving resources along the path for each quality of service sensitive flow. Each router has to maintain per-flow state information which causes large scaling problems (Li et al., 2004, 90). IntServ supports end-to-end quality of service for a wide variety of IP applications, applying the Resource Reservation Protocol (RSVP) (similar to circuit-switched telephone networks) based on the priority information in the packet header. Thus, the specification of quality of service of packet transmission is directly related to the specific application service provided. Each application service has the specific guaranteed transmission quality it requires.

In contrast, DiffServ architecture uses a much less subtle differentiation approach, classifying packets into an exogenously determined number of classes at the network edge. Only the edge routers (ingress or egress edge routers) perform packet classification based on the priority information in the packet header, whereas core routers inside each DiffServ domain only deal with aggregated traffic for given service classes (Chen, Zhang, 2004, 370 ff.). Within each DiffServ-enabled domain, forwarding of packets by routers is performed according to traffic classes, not according to the flow they belong to. Thus, within a DiffServ domain, all packets belonging to a given quality of service class may receive the same treatment, such that within one service class no priority rule is applied. Due to its scalability compared to IntServ, the DiffServ framework is considered particularly suitable for larger scale networks (Bouras, Sevasti, 2004, 167).

The DiffServ scheduler router offers a predefined number of traffic classes using a strict priority schedule. A packet is inserted into the transmission buffer behind previous packets of the same traffic class but ahead of packets of a lower traffic class. The scheduler transmits the packets which are at the head of the buffer; packets at the tail of the buffer are dropped as soon as the buffer is full. Traffic quality can be measured by mean packet delay and packet loss. Applying a strict priority schedule, traffic classes are monotone with respect to traffic quality. Packets within a higher traffic class will be transported with lower delay and lower loss than packets within lower traffic classes (Jin, Jordan, 2005, 842).

Depending on the demand for high quality, medium quality and low quality traffic, quality of service in different classes results endogenously. The carrier may provide a quality of service guarantee for the data packet transport within a quality of service class – irrespective of the forwarding rate of lower classes (Chen, Zhang, 2004, 374 ff.) – defining a maximum allowable delay and packet loss. Alternatively, no quality of service guarantees may be offered.

2.3 Economically based traffic management

Quality of service based price differentiation is well-known from road pricing experiments. Experiments in the U.S. have attempted to make road pricing more appealing by giving motorists the option of travelling without toll on regular lanes or paying for congestion free high-quality lanes (Verhoef, Small, 2004, 133 ff.). In an example with two parallel roads A and B and two heterogeneous user groups Q_A and Q_B, where users in group Q_A have a strong preference for uncongested trips and users in group Q_B have lower time preferences, price differentiation can be applied, i.e., there will be a high price guaranteeing absence of congestion on road A, and toll free travel on road B. These partial road pricing rules already result in high social welfare improvements, provided there are substantial heterogeneities in the preferences of the users (Small, Winston, Yan, 2005).

A simple price differentiation rule to allow travelers with heterogeneous preferences for avoiding congestion to self-select was provided years ago by Paris Metro. Two classes of (homogenous) carriages were provided (first class and second class). The more expensive first class carriages resulted in lower congestion, although all other characteristics of travel were identical. Odlyzko's proposal is to apply the same rule to Internet traffic (Odlyzko, 1999). Thus, a network would be partitioned into separate logical channels, with different charges applied on each channel. Although no guarantee of traffic service quality is provided, on average higher priced channels will be less congested. Users self-select into the different channels according to their preferences for avoiding congestion and the prices charged on the channels. "Paris Metro Pricing" has not only been considered for separate channels but also for priority scheduling. Odlyzko (1999, 141) also considers the case of priorities, where all packets with higher priorities would always be forwarded before packets with lower priorities. According to "Paris Metro Pricing" packet prices should be monotonically decreasing. Premium class packets with the lowest congestion level should pay the highest packet price, medium class packets should pay a medium price, and third class packets with the highest congestion level should pay the lowest price (which may be zero).

In the “Paris Metro Pricing” approach there are no qualities of service guarantees for each class. In fact, traffic quality in each class results endogenously. But in order to provide sufficient levels of traffic quality for delay sensitive applications, guarantees of traffic qualities for prioritized traffic classes are required. Since traffic quality is directly related to congestion the goal is to combine congestion management with traffic quality differentiation. The corresponding economic principle is reflected in the concept of market driven network neutrality. Any application is charged according to the opportunity costs of the traffic capacities it requires, taking into account the increasing delay of lower class packets due to the transmission of higher class packets (see section 4). Applied to a three class differentiation a necessary condition for incentive compatibility of traffic quality differentiation is that priority service has to pay a higher price than medium service and medium class a higher price than the lowest class. Thus, there should be no regulatory prohibition of priority pricing.

3. A critical appraisal of traffic quality regulation

In the following a critical evaluation of regulatory traffic management is provided. It can be shown that neither market power nor universal service considerations can justify the regulation of traffic management. In particular, heterogeneous demand for traffic quality for delay sensitive versus delay insensitive applications requires quality of service differentiation and complementary priority pricing. In section 4 the concept of market driven network neutrality is applied to point out the importance of entrepreneurial unregulated traffic quality management. In particular, the irrelevance of best effort quality and the endogenous evolution of minimal traffic qualities are pointed out.

3.1 The argument of market power

The focus of the debate is on the question whether Internet application providers should be protected from the abuse of market power of Internet access providers (Economides, 2008, 210; FCC, 2010b, 19 ff.). Criteria such as relative market share, financial strength and access to input and service markets can only serve as a starting point for evaluating the existence of market power. Conjecturing a dominant position on the basis of market shares, for example, can lead to economically unjustified criteria for government intervention in network industries. From a competition economics point of view, the use of ex ante, sector-specific regulatory intervention constitutes massive interference in the market process and thus requires a particularly well-founded justification based on modern network economics. Obviously, the development of an ex ante regulatory criterion creates a need for a more clear-cut definition of market power.

It is necessary to differentiate between those areas in which active and potential competition can work and other areas, characterized by the combination of a natural monopoly and irreversible costs (monopolistic bottlenecks). A natural monopoly situation exists, if a single supplier can serve the relevant market at lower costs than several suppliers. Irreversible costs are no longer relevant for the decision making by established firms, in contrast irreversible costs are a crucial factor for potential competitors, insofar as they must decide whether to invest such costs in the market. A monopolistic bottleneck is characterized by network specific market power, because, due to the lower decision relevant costs, the incumbent has a credible threat that may discourage a second network operator from entering the market (Knieps, 2011, 18).

The transmission of data packets belongs to the markets for Internet traffic services. Due to the absence of irreversible costs in the provision of Internet logistics, Internet traffic services do not possess the characteristics of monopolistic bottlenecks. They are, however, based on telecommunications infrastructure. Insofar as the markets for telecommunications infrastructure capacities have monopolistic bottleneck components, regulation may be necessary to guarantee competitive markets for traffic services.

Since the markets for long-distance telecommunications infrastructure capacities are competitive, the markets for Internet backbone services are also competitive. However, local telecommunications infrastructure capacities may possess the characteristics of monopolistic bottlenecks for which neither active nor potential substitutes are available. In the meantime increasing competition within local telecommunications infrastructures (local loops) can be observed worldwide. Competition has led to a considerable variety in technological platforms, e.g. optical fiber (FTTx), wireless networks (e.g. WiMax, UMTS or LTE), Community Antenna Television (CATV) networks and satellite technology, as well as to an increase in product variety. In addition, because of these rapid developments, the local loop facilities in larger cities and agglomerations are increasingly losing their status as monopolistic bottlenecks. Although it is not possible at this point to predict exactly how long it will take for the monopolistic bottlenecks in the local loop to disappear completely, the development of alternative access networks indicates that the potential for phasing out sector-specific regulation in telecommunications is strongly increasing (Blankart, Knieps, Zenhäusern, 2007).

To the extent that local networks constitute monopolistic bottlenecks, ex ante regulation appears justified. Since unregulated tariffs would enable owners of monopolistic bottlenecks to generate excessive profits, disaggregated price-cap regulation should be introduced to regulate the price level. It is important to restrict such price-cap regulation to those areas of telecommunications networks where market power due to monopolistic bottlenecks is a regulatory problem. In all other subparts of telecommunications networks, price setting should be left to competitive market forces.

In this context it is interesting to note the different regulatory frameworks in the U.S. and in Europe. Whereas in the U.S. broadband access has been strongly deregulated (FCC, 2004), in Europe access to the unbundled local copper loop as well as wholesale access to ducts of the last mile are still heavily regulated (Knieps, Zenhäusern, 2010, 1003 ff.). In any case, remaining monopoly problems with respect to local telecommunications infrastructures should be regulated at the roots and not used as justification for forbidding Internet access providers price and quality differentiation of Internet traffic.

By means of access regulation of local loop bottleneck components the transfer of market power from the telecommunications network bottleneck components into the complementary Internet access service markets can be avoided. Thus, any regulation of contracts between access service providers and Internet application service providers is not justified. As a result, the avoidance of network neutrality regulation is of great importance, because only then can the adequate market signals (congestion tariffs, quality differentiation etc.) be supplied to the content provider, leading to an endogenous and more efficient exploitation of the Internet traffic resources.

3.2 The argument of Universal Service

In liberalized telecommunications markets, universal services are an important political objective worldwide, focusing on the provision of nation-wide adequate and sufficient services. The number of questions initially arising from the demand for universal services exceeds the number of clear political guidelines that it provides. These questions are, in particular: Which services should be provided universally as services of general economic interest? Which quality of universal services should be provided? Is a lowering of quality levels at the margins of an area of universal service provision acceptable or not? At what rates should universal services be offered? How should the suppliers of universal services be chosen? How should the provision of universal services be financed?

Answering these questions requires political decisions. A society’s view as to which services should be subsidized is realized through the political process and may vary considerably over time. The question is whether technological change will in the future lead, at least in the long run, to a phasing-out of universal services due to shrinking costs for the provision of traditional universal services. In this context, an increasing variety of the standards for universal service (scope, minimum quality, prices, etc.) in different countries and regions is also conceivable. When determining the scope of non-profitable universal services, the division of labour, for instance between the federal authority, individual federal states, counties and municipalities must also be considered (Blankart, 2003).

Considering the disaggregated representation of the Internet (Knieps, Zenhäusern, 2008, fig. 1, 122) the question arises at which submarkets universal service issues occur. Again it is important to differentiate between the markets for communications infrastructure capacities and the markets for Internet traffic services.

On the infrastructure level, the problem of a universal service and the related subsidy arises only for the case that in some areas local broadband access is not yet available. On March 16, 2010 the FCC released the National Broadband Plan (FCC, 2010a). Although the number of Americans who have broadband access at home increased from 8 million in 2000 to nearly 200 million in 2009, there are still 100 million Americans who do not have broadband access at home. The goal of the FCC is to reform universal service mechanisms to support the deployment of broadband access in high cost areas and to ensure that low income inhabitants can afford broadband (FCC, 2010a, XI). The goal of national broadband availability is that everybody in the U.S. should have access to affordable broadband infrastructure, which allows the provision of acceptable quality of service for the most common Internet applications (FCC, 2010a, 135).

In September 2009 the European Commission issued guidelines for the application of state aid rules to support the deployment of broadband infrastructure in areas where such infrastructures do not yet exist, in particular rural areas. The state aid guideline for broadband networks does not only focus on the funding of broadband networks as ADSL, but also of very high speed Next Generation Networks (European Commission, 2009a, C 235/8, C 235/12 ff.). In September 2010 the European Commission initiated its Broadband Package. Among others the package consists of a communication entitled “European Broadband: investing in digitally driven growth” (European Commission, 2010). In particular, the Commission asked all member states to adopt plans for high and ultra high speed broadband networks with implementation strategies including the provisions for the necessary funding. In January 2011 the Commission approved the amount of € 1.8 billion in public funds (state aid) for the deployment of broadband networks aiming to ensure that everybody in the European Union has access to broadband infrastructures: “The approach has ensured that broadband networks are built in areas where nothing was available before and are made accessible to competing Internet service providers on non discriminatory terms”.[4]

An increase of the penetration rate of end-users with high speed broadband access strongly increases the capacity usage of access service networks as well as backbone service networks. The Comcast case has shown that even a rather small group of users can strongly increase capacity usage of the traffic service provider’s network by using a peer-to-peer application (BitTorrent). Thus, a large number of users of high speed distribution services will potentially lead to strongly increasing capacity usage. As more users possess the capability to send or receive large amounts of data packets via broadband access, traffic capacity management by Internet access service and backbone service providers becomes increasingly more important. In conclusion, the increasing access to high-speed broadband infrastructure does not reduce the problem of scarcity of Internet traffic capacities, but may strongly aggravate it. It creates an urgent need for congestion based quality differentiation for broadband Internet traffic.

On the level of Internet traffic services universal service problems may also arise. To the extent that Internet traffic for specific socially desired applications (tele-medicine, interactive video for schools etc.) should be used, best effort TCP average quality is not sufficient. Such applications require priority traffic with guaranteed traffic qualities. Instead of prohibiting priority pricing, a subsidy for the required premium traffic services may be necessary.

---

[4] State aid: Commission approves record amount of state aid for the deployment of broadband networks in 2010, Press releases, http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/54&format=HTML IP/11/54, Brussels, 20.01.2011.

4. Market driven network neutrality, priority pricing, and endogenous minimal traffic qualities

Internet traffic markets are characterized by active and potential competition. This includes competition between Internet Access Service Providers as well as between Internet backbone service providers. Internet traffic providers have free access to the market, since high profits achieved by one firm would have the immediate effect of attracting others. Often a newcomer enters the market with no intention of duplicating the established firm. The relevant point is active competition, achieved by means of technological and product differentiation and the introduction of new products and processes (Faratin et al. 2007; Knieps, Zenhäusern, 2008, 127 ff.).

Since traffic quality is directly related to congestion externalities, the starting point for an economically founded approach is the pricing model of MacKie-Mason and Varian (1995) for only one traffic class. The authors show the interrelation between optimal congestion charges and optimal investment in traffic capacities, based on the well-known congestion pricing model from transportation economics. An extension to multi-channel congestion pricing does not lead to traffic quality differentiation with quality guarantees. Instead, the introduction of traffic quality classes with traffic quality guarantees in the upper traffic classes is unavoidable. The interrelationship between congestion pricing and incentive compatible traffic quality differentiation can be achieved under application of the DiffServ architecture (Knieps, 2010).

In the following the congestion pricing model for different traffic quality classes is applied, in order to show the importance of priority pricing as well as the evolutionary development of minimal traffic qualities.

4.1 Priority pricing and endogenous traffic capacity choice

The potentials of quality and price differentiation are to be exploited by each traffic service provider requiring entrepreneurial decisions on the number of traffic qualities. The aim is to derive a pricing rule applying a quality of service-based price differentiation of traffic classes. In order to achieve an incentive compatible traffic quality differentiation, prices in higher service classes should be higher than prices in lower service classes. The starting point for the development of such price differentiation strategies are the opportunity costs of capacity usage.

It is important to differentiate between intraclass congestion externalities within a traffic class and interclass congestion externalities between traffic classes. Intraclass externalities reflect the delays which an additional data packet causes for all other data packets of the same class. Interclass externalities reflect the delays which an additional data packet imposes on the data packets in the other quality classes. Due to the strict priority rule only upper traffic classes cause interclass externalities to lower classes, but not vice versa.

Assume that a traffic service provider has decided that data packets are classified and grouped into *n* different traffic classes. Q_{it} denotes the number of data packets belonging to the same traffic class *i*, *i*=1,…, *n* in period *t*. $P_{it}(Q_{it})$ denotes the inverse demand for aggregated traffic in traffic class *i*. $\rho(w)$ denotes the capacity costs of the channel with the bandwidth *w*.

Let $k_{it}(Q_{1t},...,Q_{nt},w) \qquad i=1,...,n$ be the private (average) variable costs of a data packet transmission within traffic class *i*, which may also depend on the flow of data packets in the other traffic classes.

$\frac{\partial k_{it}}{\partial Q_{it}} > 0$ if capacity w remains constant, additional traffic within traffic class *i* will decelerate data packets, thereby raising costs.

$\frac{\partial k_{it}}{\partial w} < 0$ if traffic remains constant, additional bandwidth capacity will allow to speed up every data packet.

Optimal prices for the different quality classes and optimal traffic capacity are determined simultaneously. Competitive behavior of the traffic service provider results in an optimal packet price for each traffic class and optimal traffic capacities. We assume that there are zero income effects associated with the demand function P_{it}, so that the social net benefit of all packet transmission on the different traffic classes i=1,…, n over periods T is defined by:

$$(1) \qquad \max_{(Q_{it},\dots,Q_{nt},w)} S = \sum_{t=1}^{T}\left[\sum_{i=1}^{n}\int_{0}^{Q_{it}} P_{it}(\tilde{Q}_{it})d\tilde{Q}_{it} - \sum_{i=1}^{n} k_{it}(Q_{1t},\dots Q_{nt},w)Q_{it}\right] - \rho(w)$$

Necessary conditions for the maximum may be found by differentiating (1) with respect to $Q_{1t},\dots,Q_{nt}$ for each t=1,…,T and with respect to w and setting each derivative to zero.

The optimal pricing rule requires for a packet transmission within traffic class i to include intraclass as well as interclass externalities:

$$(2) \qquad \tau_{it} = P_{it} - k_{it} = \frac{\partial k_{it}(\cdot,w)}{\partial Q_{it}}\cdot Q_{it} + \sum_{j=i+1}^{n}\frac{\partial k_{j}(\cdot,w)}{\partial Q_{it}}\cdot Q_{jt} \quad t=1,\dots,T; \quad i=1,\dots,n$$

The optimal rule for bandwidth capacity w is given by:

$$(3) \qquad \rho'(w) = -\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{\partial k_{it}(Q_{1t},\dots,Q_{nt},w)}{\partial w}\cdot Q_{it}$$

Capacity is extended to the point where the marginal cost of an extra unit of capacity is equal to its marginal benefits of reduced congestion within the different traffic classes.

Simultaneous solutions of equation (2) and (3) provide first-best allocation of traffic flows $Q_{1t}^{*},\dots,Q_{nt}^{*}$ t=1,…,T as well as first-best capacity dimension w^{*}.

Even if traffic in the premium class is low, the delay imposed by high priority traffic to the traffic of subsequent classes may be substantial. The opportunity costs of the transmission of data packets under strict priority scheduling are strongly determined by interclass externalities, the increasing delay of lower class packets being due to the transmission of premium class packets. In contrast, intraclass externalities in the upper classes may be neglected, if the quality standard is defined high enough, such that transmission quality is sufficient for all relevant applications independent of the traffic load in this class. If only interclass externalities are taken into account, congestion-fee based traffic class prices are monotone

$$(4) \qquad \tau_{1t} = \sum_{j=2}^{n} \frac{\partial k_{jt}(\cdot, w)}{\partial Q_{1t}} \cdot Q_{jt} > \\ \tau_{it} = \sum_{j=i+1}^{n} \frac{\partial k_{jt}(\cdot, w)}{\partial Q_{it}} \cdot Q_{jt} > \tau_{n-1t} = \frac{\partial k_{nt}(\cdot, w)}{\partial Q_{n-1}} \cdot Q_{nt} > \tau_{nt} = 0$$

and the lowest traffic class has a data packet transmission price of zero.

Quality of service-based price differentiation allows users with heterogeneous demands for traffic quality to self select. Whereas users with high preference for priority traffic services have the possibility to pay a high user charge for high-quality less congested traffic services, users with preferences for low quality more congested traffic services have to pay lower user charges.

4.2 Endogenous minimal traffic qualities

In the lowest traffic class the user charges according to interclass externalities are zero. Congestion results endogenously without traffic quality guarantee depending on the demand for traffic in the different traffic class and the level of capacity which is used for all traffic classes. This minimal traffic quality of the lowest traffic class should not be confused with the best effort average traffic quality of TCP.

Since in the lowest traffic class *n* no quality guarantee is provided, intraclass externalities are of particular relevance. If intraclass externality pricing is applied, socially inefficient delay in traffic class *n* can be avoided. Within DiffServ architecture all data packets within the same class are treated equally, thus only average delay within a quality class is considered but not the individual delay of a packet depending on the position of the data packet within the queue at the router. The intraclass externality price in class *n* is always lower than the interclass externality price of class *n*-1. Due to top priority scheduling an additional data packet in class *n*-1 has priority before all packets in class *n* and therefore causes a larger delay on the packets in class *n* than an additional data packet within class *n*. Thus, the monotony requirements of traffic class prices are still fulfilled.

Capacity is allocated endogenously between the different quality classes according to the degree of heterogeneity between the different consumers. Since capacity is chosen endogenously, an increase in the demand for high quality transmission with the resultant high opportunity costs of additional high quality traffic will lead to an incentive compatible capacity extension. When there is a population of users who need more high priority transmission, a larger share of capacity is used for the high priority transmission. When there is a population of users who have strong preferences for low priority transmission, more capacity will be used for low priority transmission. As a consequence, minimal traffic quality in the lowest traffic class results endogenously.

5. Summary and discussion of policy implications

The network neutrality debate is gaining increasing momentum worldwide. Its focus is on the regulation of Internet traffic management. Two different opposing approaches to traffic management regulation became relevant in the U.S. and in Europe. In the U.S. the FCC differentiates between reasonable network management and unreasonable discrimination. Paying for priority arrangements between Internet access service providers and Internet application providers to favor some traffic over other traffic is considered unreasonable discrimination. In Europe the focus is on minimum quality of service requirements.

In this paper a critical appraisal of traffic quality regulation has been provided. It is shown that neither market power nor universal service arguments can justify the prohibition of priority pricing or the regulatory requirement of minimum traffic qualities. In particular, heterogeneous demand for traffic quality for delay sensitive versus delay insensitive applications requires traffic quality differentiation and complementary priority pricing.

The transmission of data packets belongs to the markets for Internet traffic services. Due to the absence of irreversible costs in the provision of Internet logistics, Internet traffic services do not possess the characteristics of monopolistic bottlenecks. They are, however, based on telecommunications infrastructure. By means of access regulation of local loop bottleneck components the transfer of market power from the telecommunications network bottleneck components into the complementary Internet access service markets can be avoided. Thus, any regulation of contracts between access service providers and Internet application service providers is not justified.

Considering the disaggregated representation of the Internet the question arises at which submarkets universal service issues occur. Again it is important to differentiate between the markets for communications infrastructure capacities and the markets for Internet traffic services. On the infrastructure level, the problem of a universal service and the related subsidy arises only for the case that in some areas local broadband access is not yet available. As more users possess the capability to send or receive large amounts of data packets via broadband access, traffic capacity management by Internet access service and backbone service providers becomes increasingly more important. In conclusion, the increasing access to high-speed broadband infrastructure does not reduce the problem of scarcity of Internet traffic capacities, but may strongly aggravate it. It creates an urgent need for congestion based quality differentiation for broadband Internet traffic.

On the level of Internet traffic services universal service problems may also arise. To the extent that Internet traffic for specific socially desired applications (tele-medicine, interactive video for schools etc.) should be used, best effort

TCP average quality is not sufficient. Such applications require priority traffic with guaranteed traffic qualities. Instead of prohibiting priority pricing, a subsidy for the required premium traffic services may be necessary.

Based on innovations in Internet traffic quality management architecture the basic characteristics of an economic pricing approach for traffic quality differentiation are demonstrated. Since traffic quality is directly related to congestion externalities, the well-known congestion pricing models from transportation economics are extended for different traffic quality classes in order to show the importance of priority pricing as well as the evolutionary development of minimal traffic qualities.

References

Blankart, CH.B. (2003), Universaldienst und Liberalisierung: Die föderale Dimension - Konsequenzen für das neue TKG - , Telekommunikations- & Medienrecht, TKMR-Tagungsband, 13-17

Blankart, Ch.B., Knieps, G., Zenhäusern, P. (2007), Regulation of New Markets in Telecommunications: Market Dynamics and Shrinking Monopolistic Bottlenecks, in: European Business Organization Law Review, Vol. 8, 413-428

Bouras, C., Sevasti, A. (2004), SLA-based QoS pricing in DiffServ networks, Computer Communications 27, 1868-1880

Chen, J.-C., Zhang, T. (2004), IP-Based Next-Generation Wireless Networks: Systems, Architectures, and Protocols, John Wiley & Sons

Economides, N. (2008), “Net Neutrality,” Non-Discrimination and Digital Distribution of Content Through the Internet, I/S: A Journal of Law and Policy for the Information Society 4/2, 209-233

Entwurf eines Gesetzes zur Änderung telekommunikationsrechtlicher Regelungen, Stand 15.09. 2010

European Commission (2009a), Communication from the Commission, Community Guidelines for the application of State aid rules in relation to rapid deployment of broadband networks, (2009/C 235/4), Official Journal of the European Union, C 235/7, 30.09.2009

European Commission (2009b), Information from European Institutions and Bodies, Commission declaration on net neutrality, (2009/C 308/02), Official Journal of the European Union, C 308/2, 18.12.2009

European Commission (2010), Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions, COM (2010) 472, European Broadband: investing in digitally driven growth, 21.09.2010

Faratin, P., Clark, D., Gilmore, P., Bauer, S., Berger, A., Lehr, W. (2007), Complexity of Internet Interconnections: Technology, Incentives and Implications for Policy. Paper prepared for 35th Annual Telecommunications Policy Research Conference, George Mason University, September, 2007

Federal Communications Commission (FCC) (2004), In the Matter of Unbundled Access to Network Elements, Review of the Section 251 Unbundling Obligations of Incumbent Local Exchange Carriers, Order on Remand, WC Docket No. 04-313, CC Docket No. 01-338, FCC 04-290, Adopted December 15, 2004

Federal Communications Commission (FCC) (2007), Formal Complaint of Free Press and Public Knowledge against Comcast Cooperation for secretly degrading peer-to-peer applications, November 1, 2007

Federal Communications Commission (FCC) (2008), In the Matter of Formal Complaint of Free Press and Public Knowledge against Comcast Corporation for Secretly Degrading Peer-to-Peer Applications, File No. EB-08-IH-1518 (FCC 08-183), Memorandum Opinion and Order, Adopted: August 1, 2008

Federal Communications Commission (FCC) (2010a), Connecting America: The National Broadband Plan, http://www.broadband.gov

Federal Communications Commission (FCC) (2010b), In the matter of preserving the open Internet; Broadband industry practices, Report and Order, GN Docket No. 09-191, WC Docket No.07-52, FCC 10-201, Adopted: December 21, 2010

Holznagel, B. (2010), Netzneutralität im TKG-Entwurf, http://www.telemedicus.info/article/1866-Netzneutralitaet-im-TKG-Entwurf.html

Jacobson, V. (1988), Congestion avoidance and control, in: Proceedings of SIGCOMM ’88 (Stanford, CA, Aug. 1988) ACM.

Jin, N., Jordan, S. (2005), Information exchange in diffServ pricing, IEEE Globecom 2005 proceeding, 841-846

Knieps, G. (2010), Network Neutrality and the Evolution of the Internet, Discussion Paper No. 135 of the Institute of Transport Economics and Regional Policy, University of Freiburg; forthcoming in: International Journal of Management and Network Economics

Knieps, G. (2011), The Three Criteria Test, the Essential Facilities Doctrine and the Theory of Monopolistic Bottlenecks, in: Intereconomics, Vol. 46/1, 2011, S. 17-21

Knieps, G., Zenhäusern, P. (2008), The fallacies of network neutrality regulation, Competition and Regulation in Network Industries, 9/2, 119-134

Knieps, G., Zenhäusern, P. (2010), Phasing out sector-specific regulation in European telecommunications, in: Journal of Competition Law & Economics, Vol. 6/4, 995-1006

Li, T., Iraqi, Y., Boutaba, R. (2004), Pricing and admission control for QoS-enabled Internet, Computer Networks, 46/1, 87-110

MacKie-Mason, J.K., Varian, H.R. (1995), Pricing the Internet, in W. Sichel and D.L. Alexander, editors, Public Access to the Internet. MIT Press, Cambridge, 1995, 269-314

Odlyzko, A. (1999), Paris Metro Pricing for the Internet, Proceedings of the 1st Association for Computing Machinery (ACM) conference on Electronic Commerce, 140-147

Small, K.A., Winston, C., Yan, J. (2005), Uncovering the Distribution of Motorists’ Preferences for Travel Time and Reliability, Econometrica, 73/4, 1367-1382

Verhoef, E.T., Small, K.A. (2004), Product Differentiation on Roads – Constraint Congestion Pricing with Heterogeneous Users, Journal of Transport Economics and Policy, 38/1, 127-156

Wu, T. (2003), Network Neutrality, Broadband Discrimination, Journal on Telecommunication and High Technology Law, 2, 141-179

Als Diskussionsbeiträge des
Instituts für Verkehrswissenschaft und Regionalpolitik
Albert-Ludwigs-Universität Freiburg i. Br.
sind zuletzt erschienen:

109. G. Knieps: Telecommunications markets in the stranglehold of EU regulation: On the need for a disaggregated regulatory contract, erschienen in: Journal of Network Industries, Vol. 6, 2005, S. 75-93

110. H.-J. Weiß: Die Probleme des ÖPNV aus netzökonomischer Sicht, erschienen in: Lasch, Rainer/Lemke, Arne (Hrsg.), Wege zu einem zukunftsträchtigen ÖPNV: Rahmenbedingungen und Strategien im Spannungsfeld von Markt und Politik, Erich Schmidt Verlag, Berlin, 2006, S. 119-147

111. G. Knieps: Die LKW-Maut und die drei Grundprobleme der Verkehrsinfrastrukturpolitik, erschienen in: Schriftenreihe der Deutschen Verkehrswissenschaftlichen Gesellschaft: Die LKW-Maut als erster Schritt in eine neue Verkehrsinfrastrukturpolitik, Reihe B, B 292, 2006, S. 56-72

112. C.B. Blankart, G. Knieps, P. Zenhäusern: Regulation of New Markets in Telecommunications? Market dynamics and shrinking monopolistic bottlenecks, erschienen in: European Business Organization Law Review (EBOR), Vol. 8, 2007, S. 413-428

113. G. Knieps: Wettbewerbspotenziale im Nahverkehr: Perspektiven und institutionelle Barrieren, erschienen in: Schriftenreihe der Deutschen Verkehrswissenschaftlichen Gesellschaft: Warten auf Wettbewerb: ÖPNV in Deutschland, Reihe B, 2007, S. 11-23

114. F. Birke: Universaldienstregulierung in der Telekommunikation heute: Herausforderungen, Chancen und Risiken – Ein historischer Ansatz, Mai 2007

115. G. Knieps, P. Zenhäusern: The fallacies of network neutrality regulation, erschienen in: Competition and Regulation in Network Industries, Vol 9/2, 2008, S. 119-134

116. G. Knieps: Disaggregierte Regulierung in Netzsektoren: Normative und positive Theorie, erschienen in: Zeitschrift für Energiewirtschaft, 31/3, 2007, S. 229-236

117. G. Knieps, H.-J. Weiß: Reduction of Regulatory Risk: A Network Economic Approach, erschienen in: in: U. Blum (Hrsg.), Regulatorische Risiken – Das Ergebnis staatlicher Anmaßung oder ökonomisch notwendiger Interventionen?, Schriften des Instituts für Wirtschaftsforschung Halle, Bd. 29, Nomos Verlag, Baden-Baden, 2009, S. 99-109

118. G. Knieps, H.-J. Weiß: Regulatory Agencies and Regulatory Risk, Revised Version: January 2008

119. G. Knieps: Regulatorische Entbündelung in Netzindustrien, erschienen in: Tagungsbände der Gesellschaft für Verkehrswissenschaft und Regionalpolitik an der Universität Freiburg e.V., 40. Freiburger Verkehrsseminar 2007, Entbündelung in Netzindustrien: Chancen und Risiken, Freiburg, S. 11-25

120. G. Knieps: The Net Neutrality Debate and the German Communications and Competition Law, erscheint in: Proceedings of the 3rd Annual International Conference on „Net Neutrality“, Center for Law and Public Utilities, Seoul National University, 17 October 2008

121. **G. Knieps:** Verkehrsinfrastrukturen zwischen Wettbewerb und Regulierung, erschienen in: Wirtschaftspolitische Blätter, 56/1, 2009, S. 11-21

122. **G. Knieps:** Wettbewerb und Netzevolutorik, erschienen in: Vanberg, V.J. (Hrsg.), Evolution und freiheitlicher Wettbewerb – Erich Hoppmann und die aktuelle Diskussion, Walter Eucken Institut, Untersuchungen zur Ordnungstheorie und Ordnungspolitik, Band 58, Mohr Siebeck, Tübingen, 2009, S. 193-210

123. **G. Knieps, P. Zenhäusern:** ‘Stepping stones’ and ‘access holidays’: The fallacies of regulatory micro-management, erschienen in: P. Baake, R. Borck (eds.), Public Economics and Public Choice: Contributions in Honour of Charles B. Blankart, Springer Verlag, Berlin u.a., 2007, S. 257-277

124. **G. Knieps:** Qualitäts- und Preisdifferenzierung im Internet, erschienen in: Jörn Kruse, Ralf Dewenter, Hrsg., Wettbewerbsprobleme im Internet, Nomos Verlag, Baden-Baden, 2009, S. 103-113

125. **G. Knieps:** Wettbewerb im transeuropäischen Eisenbahnverkehr, erschienen in: Zeitschrift für Verkehrswissenschaft, 81. Jg., Heft 1, 2010, S. 1-12

126. **G. Knieps:** Sektorsymmetrische Regulierung in Netzsektoren: Ein Vergleich zwischen Gas und Elektrizität, erschienen in: Netzwirtschaften & Recht (N&R), 6. Jg., Nr. 3, 2009, S. 138-143

127. **G. Knieps:** Theorie und Praxis der Price-Cap-Regulierung, in: Netzwirtschaften & Recht (N&R), 7. Jg., Nr. 2, 2010, S. 66-71

128. **G. Knieps, P. Zenhäusern:** The reform of the European regulatory framework for electronic communications: The unexploited phasing-out potentials, erschienen als: Phasing out sector-specific regulation in European telecommunications, erschienen in: Journal of Competition Law & Economics, Vol. 6/4, 2010, S. 995-1006

129. **G. Knieps:** Regulatory Reforms of European Network Industries and the Courts, February 2010 – Revised Version: May 2010

130. **G. Knieps:** Preis- und Qualitätsdifferenzierung in Verkehrsnetzen, erschienen in: DIW Vierteljahreshefte zur Wirtschaftsforschung, 79/2, 2010, S. 211-220

131. **G. Knieps:** Zur Evolutorik von Marktmechanismen für Flughafenslots, Juli 2010

132. **G. Knieps:** The Three Criteria Test, the Essential Facilities Doctrine and the Theory of Monopolistic Bottlenecks, erschienen in: Intereconomics, 11/1, 2011, S. 17-21

133. **G. Knieps:** Der Einfluss von Gerichtsentscheidungen auf Regulierungsreformen, erschienen in: Tagungsbände der Gesellschaft für Verkehrswissenschaft und Regionalpolitik an der Universität Freiburg e.V., 43. Freiburger Verkehrsseminar 2010, Regulierung und Wettbewerb in Netzen: Zur Rolle der Gerichte, Freiburg , 2010, S. 5-21

134. **H.-J. Weiß:** Markt und Staat in der Verkehrswirtschaft, Oktober 2010

135. **G. Knieps:** Network neutrality and the Evolution of the Internet, paper presented at the 2010 European ITS Conference in Copenhagen, 13-15 September 2010

136. **G. Knieps:** Market driven network neutrality and the fallacies of Internet traffic quality regulation, Revised Version: March 2011