

Statistische Analyse von Delay-Messungen zur Performance-Evaluation in Netzwerken

Thomas Holleczeck

holleczeck@ife.ee.ethz.ch

Abstract: Die Messung von Paketlaufzeit und Paketverlust durch dedizierte Testpakete in Computernetzwerken ermöglicht die Beurteilung der Dienstgüte auf dem Pfad, den diese durchlaufen. Aus diesem Grund werden im X-WiN, dem Deutschen Forschungsnetz, und GÉANT2, seinem europäischen Ebenbild, seit vielen Jahren vom WiN-Labor der Universität Erlangen-Nürnberg auf sämtlichen Strecken IP Performance-Messungen durchgeführt. Dieses Paper gibt einen Überblick darüber, wie die Ergebnisse dieser Messungen durch statistische Methoden analysiert werden können, um mehr über den Zustand und die Performance der überwachten Netzwerk-Strecken zu erfahren.

1 Einleitung

Für viele Multimediaanwendungen wie VoIP oder Videokonferenzen ist ein hoher Dienstgüte-Standard unabdingbar. Deshalb entwickelt das WiN-Labor an der Universität Erlangen-Nürnberg seit 1997 das aktive Messsystem Hades, mit Hilfe dessen sich IP Performance Metrics (IPPM) wie One-Way Delay (OWD), IP Delay Variation (IPDV) und One-Way Packet Loss (OWPL) in Computernetzwerken bestimmen lassen.

Statistische Analysen der durch Hades erzeugten Messwerte bilden nun die Grundlage eines in der Entwicklung stehenden Alarmsystems, das in der Lage ist, Verschlechterungen der Netzwerk-Performance automatisch zu erkennen und entsprechende Warnungen zu generieren.

2 Hades-Messsystem

2.1 Hardware

Das Messsystem besteht aus Intel[®] Pentium[®] 4 PCs mit Fedora Core als Betriebssystem, die an alle wesentlichen Router im X-WiN und GÉANT2 angeschlossen sind. Jede dieser Messboxen ist mit einer GPS-Karte ausgestattet, um für eine Messgenauigkeit von mindestens 7 Mikrosekunden zu sorgen.

2.2 Software

Die Hades-Software besteht aus einem Sende- und einem Empfangsprozess, von denen jeweils eine Instanz auf jeder Messbox läuft.

Der Sendeprozess auf dem Quellrechner einer Messung generiert UDP-Messpakete und versieht diese mit einem Zeitstempel. Zusätzlich beinhaltet der Paket-Header eine eindeutige Sequenznummer, um den Verlust oder Vertauschung von Testpaketen erkennbar zu machen.

Auf dem Empfangsrechner einer Messung nimmt der entsprechende Empfangsprozess eintreffende Pakete entgegen, versieht sie ebenfalls mit einem Zeitstempel und speichert diesen mit der Sequenznummer und dem Startzeitstempel in einer lokalen Datei. Danach stehen die Messdaten zur Auswertung bereit.

Die Messungen finden vollvermascht statt, so dass Statusinformationen für jede beliebige Strecke vorhanden sind. Derzeit generiert ein Sendeprozess für jeden bekannten Empfangsrechner alle 60 Sekunden einen Burst von 9 Testpaketen der Größe 41 Bytes. Der Abstand zwischen zwei aufeinander folgenden Paketen eines Bursts beträgt 30 Millisekunden, um Kollisionen an der Netzwerkschicht zu vermeiden.

3 Definitionen

Jede Strecke, d.h. jedes Paar von Quell- und Zielmessbox, zeigt im Bezug auf den gemessenen OWD ihr eigenes Verhalten. Die statistische Analyse betrachtet die in einem Zeitintervall zwischen zwei Messboxen gemessenen OWD-Werte $\mathbf{Y} = (y_1, \dots, y_n)$. Wie in Abbildung 1 illustriert, wird der OWD offensichtlich durch zwei Komponenten beeinflusst:

1. **Intrinsic Delay.** Der *intrinsic delay* c zwischen zwei Messboxen wird dominiert durch die Ausbreitungsverzögerung, d.h. diejenige Zeit, die das elektrische bzw. optische Signal benötigt, um die Links des IP-Pfades zu durchlaufen, sowie die Übertragungszeit des Testpakets. Normalerweise ist c zwischen zwei Hosts konstant, lediglich im Falle einer Veränderung des durchlaufenen Pfades kommt es zu Veränderungen. Der *intrinsic delay* darf also als konstante untere Grenze betrachtet werden, die kein OWD-Wert unterschreiten kann.
2. **Routing Delay.** Zusätzlich besteht der OWD aus einem variablen Anteil, der auf das Verhalten der Router zurückzuführen ist, die die beiden Messboxen verbinden. Er wird als *routing delay* bezeichnet, da es sich bei diesem um diejenige Zeitspanne handelt, die alle involvierten Router benötigen, um das Testpaket zu verarbeiten. Der *routing delay* stellt also den interessanten Bestandteil des OWD dar, da dessen Analyse Rückschlüsse auf den dynamischen Zustand der Router erlaubt. Die entsprechende Folge von Werten des *routing delay* wird mit \mathbf{X} gekennzeichnet und

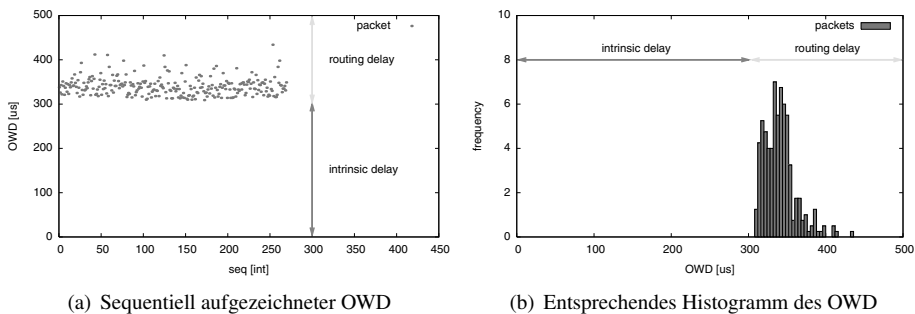


Abbildung 1: Innerhalb einer halben Stunde zwischen zwei beliebigen Messboxen im X-WiN aufgezeichneter OWD

lässt sich in Abhängigkeit des OWD und des *intrinsic delay* angeben:

$$\mathbf{X} = (x_1, \dots, x_n) = (y_1 - c, \dots, y_n - c) \quad (1)$$

Der *routing delay* \mathbf{X} kann also durch die Subtraktion des *intrinsic delay* c vom gemessenen OWD berechnet werden. Da c prinzipiell unbekannt ist, richtet die Vorverarbeitung der Messwerte das Hauptaugenmerk auf dessen Bestimmung.

4 Vorverarbeitung

Bevor der in einem beliebigen Intervall gemessene OWD \mathbf{Y} statistisch analysiert werden kann, muss dieser vorverarbeitet werden. Dies umfasst die Säuberung der Messwerte: aufgrund von Uhrenfehlern ungültige Werte müssen verworfen werden. Zusätzlich werden alle aufgetretenen Pfadveränderungen—die immer dann auftreten, wenn Router oder Links auf dem verbindenden IP-Pfad ausfallen—durch die Anwendung eines Clustering-Algorithmus erkannt. Das Wissen um ein solches aufgetretenes Umrouting erlaubt es schließlich Netzwerkadministratoren, entsprechende Gegenmaßnahmen einzuleiten.

Vorausgesetzt, dass keine Sprungstelle aufgetreten ist, wird im Anschluss der *intrinsic delay* als die untere Schranke des OWD bestimmt. Da die spätere statistische Analyse den *routing delay* modelliert, kann dieser nun gemäß Gleichung (1) berechnet werden.

5 Datenanalyse

Nach der Vorverarbeitung der Messwerte ist es die Aufgabe der Datenanalyse, diese durch ein mathematisches Modell zu beschreiben, aus welchem sich Rückschlüsse auf den Zustand der entsprechenden Strecke ziehen lassen. Hier geht der Analyseprozess nun davon aus, dass der *routing delay* innerhalb des zu analysierenden Zeitfensters von einer Wahrscheinlichkeitsdichte $f(x|\theta)$ mit dem Parametervektor θ erzeugt wurde. Aufbauend auf

dieser Hypothese wird eine Parameterschätzung durchgeführt, z.B. gemäß der *maximum-likelihood*-Methode oder dem Prinzip *expectation-maximization*. Im Zuge dieser wird auf der Basis der Werte \mathbf{X} eine Schätzung $\hat{\theta}$ für θ berechnet, so dass $f(x|\hat{\theta})$ mit großer Wahrscheinlichkeit für die Erzeugung von \mathbf{X} verantwortlich war. Durch eine Interpretation der Schätzung $\hat{\theta}$ kann schließlich mehr über den Status der entsprechenden Strecke erfahren werden.

5.1 Einfache Wahrscheinlichkeitsverteilungen

Gut geeignet zur Modellierung des *routing delay* ist die Gamma-Verteilung \mathcal{G} , charakterisiert durch die zwei Parameter α und β und die folgende Dichtefunktion:

$$\mathcal{G}(x|\alpha, \beta) = x^{\alpha-1} \cdot \frac{\beta^\alpha \cdot e^{-\beta x}}{\Gamma(\alpha)} \quad (2)$$

Eine einfachere Alternative mit nur einem Parameter σ^2 und ähnlicher Form ist die Rayleigh-Verteilung \mathcal{R} mit der Wahrscheinlichkeitsdichte:

$$\mathcal{R}(x|\sigma^2) = \frac{x \cdot \exp\left\{-\frac{x^2}{2\sigma^2}\right\}}{\sigma^2} \quad (3)$$

5.2 Mischverteilungen

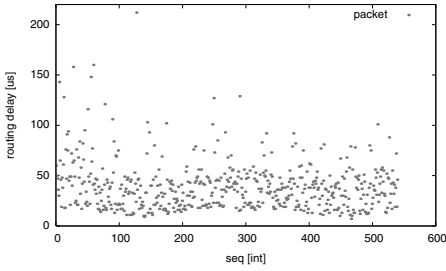
In einigen Situationen sind einfache Verteilungen zur Modellierung des *routing delay* jedoch nur bedingt geeignet—z.B. wenn das zu approximierende Histogramm wie in Abbildung 2 mehrere Peaks enthält. Hier ist es empfehlenswert, die Annahme zu machen, dass \mathbf{X} von einer linearen Überlagerung von K gewichteten Gamma-Verteilungen, einer sog. Gamma-Mischverteilung (engl. *Gamma mixture model*), erzeugt wurde. Dies führt zu folgender Dichtefunktion:

$$f(x|\theta) = \sum_{k=1}^K \pi_k \cdot \mathcal{G}(x|\alpha_k, \beta_k) \quad (4)$$

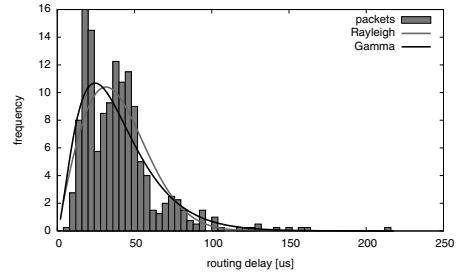
wobei π_k das Gewicht der k -ten Komponente ist.

5.3 Ergebnisse

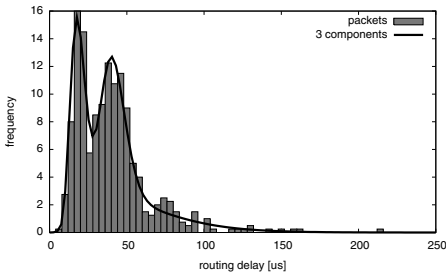
Die Adäquatheit der vorgestellten Modelle wurde im X-WiN für Zeitfenster der Breite 30 Minuten auf zahlreichen Strecken überprüft. Dabei konnten folgende Beobachtungen gemacht werden:



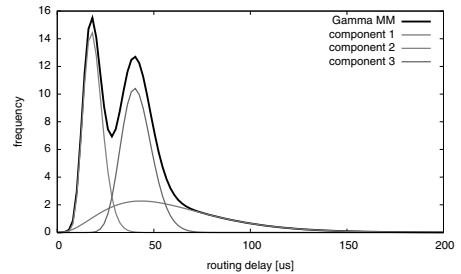
(a) Routing delay, der am 21. Oktober 2007 von 09:00 bis 10:00 Uhr zwischen GÉANT2-Messboxen in Thessaloniki und Wien aufgezeichnet wurde



(b) Aus dem Histogramm bestimmte einfache-Gamma und Rayleigh-Verteilung



(c) Gamma-Mischverteilung mit drei Komponenten, die aus dem Histogramm geschätzt wurde. Offensichtlich approximiert diese das Verhalten besser, da die Peaks des Histogramms klar erkannt werden.



(d) Verteilungen der drei Komponenten der Gamma-Mischverteilung. Die dritte Komponente (rot) modelliert den langsamen Abstieg des Histogramms nach rechts, was als Indikation für langsam entstehende Überlast interpretiert werden kann.

Abbildung 2: Verschiedene geschätzte Wahrscheinlichkeitsverteilungen

1. Die Gamma-Verteilung schneidet bei einer Evaluation aufgrund ihrer höheren Flexibilität leicht besser ab als die Rayleigh-Verteilung. Allerdings konnte nur in maximal der Hälfte aller untersuchten Fälle davon ausgegangen werden, dass diese einfachen Verteilungen als statistischer Generator des *routing delay* fungierten.
2. Bedingt durch die noch höhere Flexibilität können Gamma-Mischverteilungen die Form der beobachteten Histogramme wesentlich besser approximieren. Dies spiegelt sich in den Erfolgsquoten der Parameterschätzungen wieder: schon bei der Verwendung von drei Komponenten konnten in mehr als 80 Prozent aller Fälle die wiedergewonnen Gamma-Mischverteilungen als statistischer Generator des beobachteten *routing delay* angenommen werden. Darüber hinaus eignet sich dieses Modell auch zur frühzeitigen Erkennung entstehender Überlast einer Strecke: Diese äußert sich durch eine Komponente, die extrem langsam abfällt, wie in [Abb. 2d] illustriert.

Ein Beispieldiagramm mit dem *routing delay* einer Stunde sowie die aus diesem bestimmten Wahrscheinlichkeitsverteilungen findet sich in Abbildung 2.

6 Performance-Klassifizierung

In ausführlichen Analysen wurde das Verhalten sämtlicher Strecken im X-WiN untersucht. Dabei stellte sich heraus, dass sich die Performance einer Strecke im Hinblick auf den *routing delay* o.B.d.A. grob in vier Kategorien, die sog. Performance-Klassen C_k , einteilen lässt:

1. **Excellent.** Der bestmögliche Zustand einer Strecke ist charakterisiert durch einen stabilen *routing delay* und trägt den Namen *excellent*.
2. **Fair.** Eine leichte Verschlechterung der Performance einer Strecke deutet sich durch eine wachsende Varianz des *routing delay* und einzelne statistische Ausreißer an. Dieser Zustand wird als *fair* bezeichnet.
3. **Poor.** Der Zustand *poor* repräsentiert leichte Überlast auf einer Strecke und ist charakterisiert durch eine wiederum größere Streuung der Messwerte.
4. **Bad.** Der verbleibende, schlechteste Zustand einer Strecke steht für starke Überlast des Netzwerks und wird als *bad* bezeichnet.

Das durchschnittliche Verhalten dieser Klassen lässt sich am besten durch Rayleigh-Verteilungen beschreiben, da Gamma-Verteilungen und Gamma-Mischverteilungen zu individuell für Performance-Klassen sind. Dadurch ist sichergestellt, dass diese die für jede beliebige Strecke im X-WiN anwendbar sind. Die dabei für die einzelnen Performance-Klassen bestimmten Rayleigh-Parameter wurden in Tabelle 1 zusammengefasst.

Die Aufgabe der Performance-Klassifizierung ist es nun, herauszufinden, welche der obigen Klassen für die Erzeugung der Messwerte des *routing delay* eines bestimmten Zeitintervalls am ehesten in Frage kommt. Die Wahrscheinlichkeit, dass die beobachtete Abfolge \mathbf{X} von der Performance-Klasse C_k generiert wurde, wird als $P(C_k|\mathbf{X})$ bezeichnet und lässt sich durch den Satz von Bayes bestimmen als:

$$P(C_k|\mathbf{X}) = \frac{\prod_{i=1}^n \mathcal{R}(x_i|\sigma_k^2)}{\sum_{m=1}^4 \prod_{i=1}^n \mathcal{R}(x_i|\sigma_m^2)} \quad (5)$$

wobei es sich bei $\mathcal{R}(x|\sigma_k^2)$ die für Klasse C_k geschätzte Rayleigh-Verteilung handelt.

Nach der Berechnung der Wahrscheinlichkeiten $P(C_k|\mathbf{X})$ für alle Klassen wird das beobachtete Muster des *routing delay* schließlich derjenigen Performance-Klasse zugeordnet, für die $P(C_k|\mathbf{X})$ maximal ist, die also am wahrscheinlichsten für die Erzeugung von \mathbf{X} verantwortlich ist. Dadurch lassen sich kritische Situationen wie Überlast auf einer Strecke leicht identifizieren.

Eine beispielhafte Anwendung der Performance-Klassifizierung für die Messwerte eines Tages ist in Abbildung 3 illustriert.

Excellent	Fair	Poor	Bad
932.4	1 492.6	210 511.7	20 492 767.0

Tabelle 1: Geschätzte Rayleigh-Parameter σ_k der Performance-Klassen

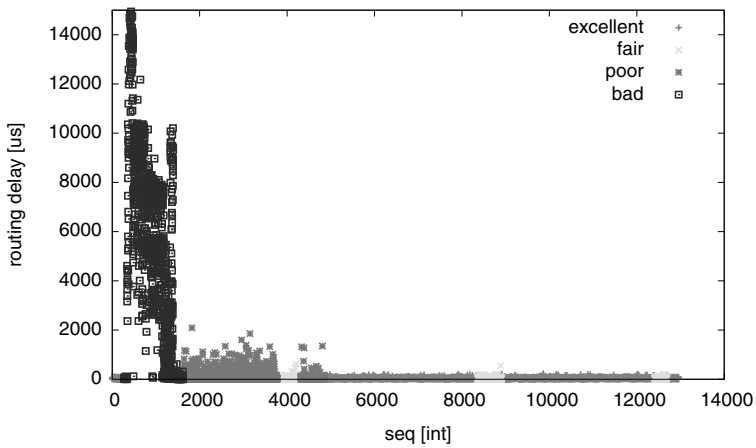


Abbildung 3: Ergebnisse der Performance-Klassifizierung für den am 21. Oktober 2007 gemessenen *routing delay* zwischen GÉANT2-Messboxen in Thessaloniki und Wien. Kurz nach Mitternacht wird die Performance dominiert durch starke Überlast. Erst gegen Morgen normalisiert sich der Zustand der Strecke wieder.

SRC	DST	#Excellent	#Fair	#Poor	#Bad	Score
A	B	3	4	20	21	37
A	C	7	0	29	12	58
...
K	P	35	13	0	0	131
C	F	41	6	1	0	136

Tabelle 2: Illustration eines Rankings abstrakter Strecken. Die Spalte #Excellent gibt die Anzahl der Zeitintervalle an, in denen die Performance-Klassifizierung für die entsprechende Strecke *excellent* ergab, #Fair die Anzahl für *fair* usw. Der Score wird schließlich als gewichtete Summe der Anzahlen berechnet, wobei der Faktor für *excellent* 3 ist, für *fair* 2, für *poor* 1 und für *bad* 0. Je geringer am Ende der Score einer Strecke ist, desto schlechter schneidet sie im Vergleich ab.

7 Ausblick

Auf Basis dieser Forschungsergebnisse entwickelt das WiN-Labor aktuell ein System, das in der Lage ist, einerseits im Falle kritischer Situationen Alarme zu generieren und auf der anderen Seite Langzeit-Auswertungen der Performance aller Strecken durchzuführen:

1. **Echtzeit-Alarmsystem.** Das Alarmsystem analysiert die neu generierten Messwerte in Form eines Datenstromsystems in Echtzeit und generiert in erkannten kritischen Netzwerksituationen wie z.B. Umrouting oder mittlerer bis starker Überlast Warnungen und Alarme.
2. **Langzeit-Analysesystem.** Darüber hinaus wird ein Analysetool entwickelt, das darauf ausgelegt ist, die durchschnittliche Performance aller verfügbaren Strecken im X-WiN über einen längeren Zeitraum zu bestimmen. In einem Ranking können dann schließlich die Strecken, wie in Tabelle 2 illustriert, miteinander verglichen werden. Dies ermöglicht u.a. die Identifikation der *low performers* unter ihnen, und liefert somit wichtige Informationen für Netzwerkspezialisten im Hinblick auf Schwachstellen im Netzwerk, die es zu beheben gilt.

Literatur

- [AKZ99a] G. Almes, S. Kalidindi, and M. Zekauskas. A One-way Delay Metric for IPPM. <http://www.rfc-editor.org/rfc/rfc2679.txt>, September 1999. Online resource, accessed 2008-05-08.
- [AKZ99b] G. Almes, S. Kalidindi, and M. Zekauskas. A One-way Packet Loss Metric for IPPM. <http://www.rfc-editor.org/rfc/rfc2680.txt>, September 1999. Online resource, accessed 2008-05-08.
- [Bis06] C. Bishop. *Pattern Recognition and Machine Learning*. Springer, Berlin, 2006.

- [CCM⁺04] B.-K. Choi, B.-K. Choi, S. Moon, Zhi-Li Zhang, K. Papagiannaki, and C. Diot. Analysis of point-to-point packet delay in an operational network. In S. Moon, editor, *Proceedings INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 3, pages 1797–1807, 2004.
- [DC02] C. Demichelis and P. Chimento. IP Packet Delay Variation Metric for IP Performance Metrics (IPPM). <http://www.rfc-editor.org/rfc/rfc3393.txt>, November 2002. Online resource, accessed 2008-05-08.
- [DLR77] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.
- [EM99] T. Elteto and S. Molnar. On the distribution of round-trip times in TCP/IP networks. In *Proceedings of the 24th Conference on Local Computer Networks*, pages 172–181, 1999.
- [FPP98] D. Freedman, R. Pisani, and R. Purves. *Statistics*. W.W. Norton & Company, New York, 3rd edition, 1998.
- [GEA] The GÉANT2 Network. <http://www.geant2.net/>. Online resource, accessed 2008-05-08.
- [HAD] The Hades measurement tool. Performance measurement in Internet Backbones. <http://www-win.rrze.uni-erlangen.de/ippm/messprogramm.html>. Online resource, accessed 2008-05-08.
- [HC78] R. Hogg and A. Craig. *Introduction to Mathematical Statistics*. Macmillan, New York, 4th edition, 1978.
- [HKK⁺06] P. Holleczeck, R. Karch, R. Kleineisel, S. Kraft, J. Reinwand, and V. Venus. Statistical Characteristics of Active IP One Way Delay Measurements. In R. Karch, editor, *Proc. International Conference on Networking and Services ICNS '06*, pages 1–1, 2006.
- [Hof01] G. Hofmann. Implementation eines Programms zur Bestimmung von Dienstgüte in IP-Netzen. Master's thesis, Friedrich-Alexander University of Erlangen-Nuremberg, 2001.
- [Hol08] T. Holleczeck. Statistical Analysis of IP Performance Metrics in International Research and Educational Networks. Master's thesis, Friedrich-Alexander University of Erlangen-Nuremberg, May 2008.
- [HP06] J.A. Hernández and I.W. Phillips. Weibull mixture model to characterise end-to-end Internet delay at coarse time-scales. *IEE Proceedings-Communications*, 153(2):295–304, 2006.
- [KR07] J.F. Kurose and K.W. Ross. *Computer Networking: A Top-Down Approach*. Addison-Wesley, 4th edition, 2007.
- [Mac67] J.B. Macqueen. Some methods of classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–297, 1967.
- [MC06] A. Morton and L. Ciavattone. Packet Reordering Metrics. <http://www.rfc-editor.org/rfc/rfc4737.txt>, November 2006. Online resource, accessed 2008-05-08.

- [MP99] J. Mahdavi and V. Paxson. IPPM Metrics for Measuring Connectivity. <http://www.rfc-editor.org/rfc/rfc2678.txt>, September 1999. Online resource, accessed 2008-05-08.
- [Muk94] A. Mukherjee. On the Dynamics and Significance of Low Frequency Components of Internet Load. *Internetworking: Research and Experience*, 5:163–205, 1994.
- [NJ06] S. Naegele-Jackson. *Network-QoS and Quality Perception of Compressed and Uncompressed High-Resolution Video Transmissions*. PhD thesis, Friedrich-Alexander University of Erlangen-Nuremberg, 2006.
- [NJKH04] S. Naegele-Jackson, R. Kleineisel, and P. Holleccek. IPPM Measurements and Network Load Behavior of the German Research Network G-WiN. In *Proceedings of the International Conference on Computing, Communications and Control Technologies (CCCT 04)*, pages 390–395, Austin, 2004.
- [PAMM98] V. Paxson, G. Almes, J. Mahdavi, and M. Mathis. Framework for IP Performance Metrics. <http://www.rfc-editor.org/rfc/rfc2330.txt>, May 1998. Online resource, accessed 2008-05-08.
- [Pap91] A. Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York, 3rd edition, 1991.
- [PMF⁺03] K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, and C. Diot. Measurement and analysis of single-hop delay on an IP backbone networks. *IEEE Journal on Selected Areas in Communications*, 21(6):908–921, 2003.
- [Ric07] J.A. Rice. *Mathematical Statistics and Data Analysis*. Thomson, Belmont, 3rd edition, 2007.
- [RN02] S. Russel and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, Upper Saddle River, 2nd edition, 2002.
- [STK⁺06] S. Shalunov, B. Teitelbaum, A. Karp, J. Boote, and M. Zekauskas. A One-way Active Measurement Protocol (OWAMP). <http://www.rfc-editor.org/rfc/rfc4656.txt>, September 2006. Online resource, accessed 2008-05-08.
- [Tan03] A.S. Tanenbaum. *Computer Networks*. Pearson Education International, 3rd edition, 2003.
- [XWI] X-WiN – Germany’s National Research and Educational Network. <http://www.dfn.de/content/xwin/>. Online resource, accessed 2008-05-08.