

# Modellierung verlustinduzierender Lasttransformationen für markovsche Ankunftsprozesse

Stephan Heckmüller, Bernd E. Wolfinger

Dept. Informatik, TKRN, Universität Hamburg  
Vogt-Kölln-Str. 30, D-22527 Hamburg  
E-mail: {heckmueller|wolfinger}@informatik.uni-hamburg.de

## Kurzfassung

Realistische Lastmodellierung ist für eine exakte Leistungsbewertung von Kommunikationsnetzen unverzichtbar. Innerhalb dieser Netze erfolgt die Dienstleistung typischerweise durch eine Reihe von Zwischenschritten, welche die Lasteigenschaften, wie sie von nachfolgenden Stationen gesehen werden, beeinflussen. In diesem Beitrag werden diese Zwischenschritte unter Verwendung des Konzepts der Lasttransformation explizit modelliert. Neben der Präsentation der allgemeinen Methodik erfolgt die Anwendung auf markovsche Ankunftsprozesse als Lastmodelle. Hierbei gilt die Konzentration einer Lasttransformation, welche insbesondere in drahtlosen Netzen von hervorragender Bedeutung ist, dem Auftragsverlust. Diese Auftragsverluste werden hierbei direkt in die Lastbeschreibung integriert, ohne die analytische Handhabbarkeit zu verlieren. Dies gilt sowohl für unabhängig verteilte Verluste als auch für markovsche Verlustprozesse. Darüber hinaus werden die Kombinationen verschiedener Transformationen und die hieraus resultierende Ankunftsrate untersucht.

## 1 Einleitung und Problemstellung

Netzwerklasten, wie sie von heutigen (verteilten) Anwendungen induziert werden, weisen im Allgemeinen komplexe statistische Eigenschaften auf, welche wiederum einen starken Einfluss auf die Leistungskennwerte realer Systeme haben. Neben der Entstehung neuer Leistungsbewertungsverfahren und Messtechniken sind daher auch im Bereich der Lastmodellierung eine Vielzahl von Neu- und Weiterentwicklungen zu verzeichnen.

Unglücklicherweise werden die Lasteigenschaften, wie sie an höheren Schnittstellen einer Netzarchitektur festgestellt werden können, durch die Verarbeitungsschritte innerhalb des übertragenden Rechnernetzes in signifikanter Weise verändert. Zu diesen Schritten zählen Fragmentierung, Header-Generierung und Maßnahmen zur Überlastvermeidung. Die Methode der Lasttransformation bietet die Möglichkeit diese Schritte explizit zu modellieren, um so zu einer Beschreibung der Lasteigenschaften an tieferen Schnittstellen des Netzes zu gelangen. Neben den genannten Verarbeitungsmechanismen kann eine Beeinflussung der Lasteigenschaften durch fehlerhafte Verarbeitung erfolgen, wobei im Kontext der Rechnernetze insbesondere Paketverluste zu nennen sind. Diese dienen im Gegensatz zu den vorgenannten Verfahren nicht der Dienstleistung, auch wenn das Verwerfen von Paketen zur Überlastvermeidung ein Mittel zur Aufrechterhaltung der Dienstbereitschaft eines Rechnernetzes ist.

Im vorliegenden Beitrag soll gezeigt werden, dass sich eine Reihe wichtiger Verlustprozesse – unabhängig verteilte Verluste und markovsche Verlustprozesse – direkt als Lasttransformationen auf markovschen Ankunftsprozessen angeben lassen. Somit kann die analytische Handhabbarkeit der transformierten Lastmodelle sichergestellt werden, bei gleichzeitiger Beibehaltung guter Realitätsnähe.

Das Konzept der Lasttransformation ist bereits in einer Reihe von Arbeiten untersucht worden [WZHB02, Bai99, Zad01, HW07]. Die in den vorgenannten Arbeiten verwendete Lastbeschreibungstechnik legt besondere Betonung auf die Schnittstelle, welche das Bediensystem (z.B. ein paketübertragendes Netz) von seiner Umgebung (z.B. die Netzdienste beanspruchenden Benutzer) separiert. Hierdurch wird die Lastbeschreibung in Abhängigkeit vom betrachteten Bediensystem möglich. Hierauf aufbauend führen die Autoren die Unterscheidung von Primär- und Sekundärlast ein, welche die formale Definition der Lasttransformation ermöglicht. In [WZHB02] werden analytische Lasttransformationen für Rechnernetze untersucht und auf ihre Realitätsnähe überprüft. Die Autoren konzentrieren sich auf die Transformation

von Paketlängen, wie sie durch Fragmentierung und Header-Generierung hervorgerufen wird. Weiterhin wird in [Bai99] die simulative Lasttransformation untersucht. In [HW07] werden Transformationen auf markovschen Ankunftsprozessen (BMAPs) für Fragmentierung und Schiebefenstermechanismen vorgeschlagen, welche sowohl die Transformation von Paketattributen als auch des zeitlichen Verhaltens modellieren.

*Batch Markovian Arrival Processes* (BMAP) wurden in abweichender Notation in [Neu79] eingeführt. Der Autor verwendet die Bezeichnung *versatile Markovian Point Process*. BMAPs in der in vorliegender Arbeit verwendeten Notation wurden in [Luc91] vorgeschlagen. Aufbauend auf diesen Arbeiten wurden BMAPs in einer Vielzahl weiterer Arbeiten untersucht. Betrachtet wurden beispielsweise die Blockierungswahrscheinlichkeit einer *BMAP/G/1*-Warteschlange [CW06], der Abgangsprozess einer *BMAP/G/1*-Warteschlange [FC01] oder die Wahrscheinlichkeitsverteilung der Arbeitslast von Systemen mit Schwellwert [LB06]. In letztgenannter Arbeit wird angenommen, dass der Bediener erst aktiv wird, wenn die Anzahl wartender Aufträge einen festzulegenden Schwellwert überschreitet. Eine Reihe anderer Arbeiten verallgemeinern die *BMAP/G/1*-Warteschlange: In [Hof01] wird die *BMAP/G/1*-Warteschlange erweitert, um insbesondere eine Abhängigkeit des Ankunftsstromes vom Systemzustand zu gestatten.

In [KLL03] werden BMAPs genutzt, um den Verkehr in Netzen auf IP-Ebene zu modellieren. Die Autoren schlagen hierzu Techniken zur Anpassung der Modellparameter an vorliegende Messdaten vor. Einige Arbeiten nutzen BMAPs zur Modellierung von Videolasten. In [BC92] werden Lasten mit variabler Bitrate (VBR) mit Hilfe von BMAPs modelliert. In [HSB04] werden Videolasten durch diskrete *BMAPs* modelliert, um ausgehend hiervon warteschlangen-basierte Untersuchungen vorzunehmen. Die so erzielten Ergebnisse dienen weiterhin der Dimensionierung von Puffern.

Die im Folgenden verwendeten markovschen Fehlermodelle finden insbesondere in nachrichtentechnischen Untersuchungen drahtloser Übertragungskanäle Verwendung. In [ZRM95] werden diese auf ihre Tauglichkeit zur Modellierung des 'Fadings' von Kanälen untersucht. Die Autoren sind der Ansicht, dass auch einfache Markov-Modelle als gute Approximationen des Kanalverhaltens angesehen werden können. Dementsprechend werden diese Modelle in [ZK99, KL00] zur Modellierung dieses Phänomens eingesetzt. Darüber hinaus finden insbesondere die verhältnismäßig einfachen Gilbert-Elliot-Modelle mit zwei Zuständen auch als Modelle korrelierter Verluste in Rechnernetzen Anwendung (siehe z.B. [LZ06]). Der Rest des Beitrags ist wie folgt gegliedert: Im Abschnitt 2 wird zunächst Lasttransformation definiert. Weiterhin rekapitulieren wir die für die folgenden Untersuchungen notwendigen Resultate zur Fragmentierungs-Transformation von BMAPs und schlagen Header-Transformationen für BMAPs vor. Im hierauf folgenden Abschnitt 3 wird eine Transformation auf *Batch Markovian Arrival Processes* zur Berücksichtigung von Auftragsverlusten vorgestellt. Darüber hinaus werden Formeln zur Berechnung des mittleren Durchsatzes von Ankunftsprozessen angegeben, die ihrerseits bereits (Paket-)Fragmentierung und Verluste modellieren. Die gemäß dieser Transformationen erzeugten Ankunftsprozesse werden im Abschnitt 4 mit simulativ transformierten Lasten verglichen, um die Genauigkeit der Transformationen zu überprüfen. Abschließend wird in Abschnitt 5 ein Fazit gezogen und ein kurzer Ausblick gegeben.

## 2 Lasttransformation

Innerhalb von Netzen von Bedienstationen ist die Last, wie sie an einer einzelnen Station anliegt, im Allgemeinen von der Diensterbringung anderer Stationen abhängig. Dies gilt im besonderen für gegenwärtige Rechnernetze, welche in Schichten und verbundenen Einzelsystemen organisiert sind. Jede Station beeinflusst hier die Last der jeweilig nachfolgenden Station, wie in Abbildung 1 schematisch dargestellt. Die so hervorgerufene Veränderung der Lastcharakteristiken wird im Folgenden als Lasttransformation von Primär- zu Sekundärlast bezeichnet. In der vorgenannten Abbildung können diese Transformationen beispielsweise Fragmentierung und Headergenerierung in einer *UDP*-Protokollinstanz oder Paketverlust bei der drahtlosen Übertragung modellieren.

Um diese Transformationen formal beschreiben zu können, wird Last, wie in Definition 1 [Wol99] dargestellt, definiert.

**Definition 1** Die Last  $L = L(E, S, IF, T)$  wird definiert als eine Sequenz von Aufträgen, die während des Beobachtungsintervalls  $T$  an das Bediensystem  $S$  durch seine Umgebung  $E$  übergeben werden. Die Aufträge werden über die Schnittstelle  $IF$  übergeben, welche das Bediensystem von seiner Umgebung trennt.  $\diamond$

Die Last kann somit durch eine Sequenz von Aufträgen  $a_i$ , die während des betrachteten Zeitintervalls  $T$  eintreffen, beschrieben werden. Für wohldefinierte Lasten sei der Ankunftsprozess definiert als Tupel aus

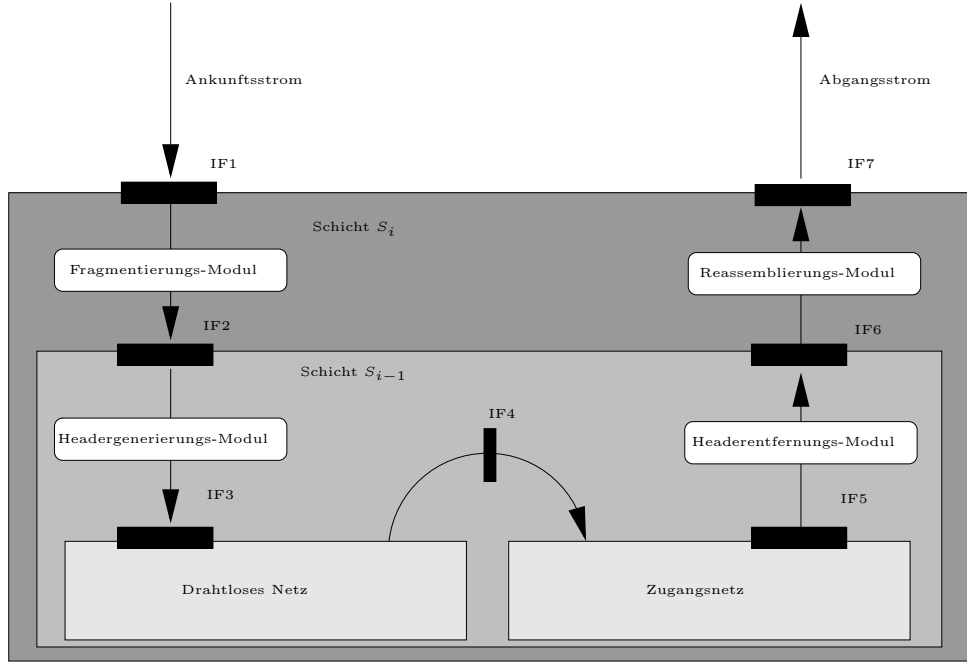


Abbildung 1: Übertragung eines Paketes als Ablauf von Transformationen

Ankunftszeitpunkten  $t_i$  und den Aufträgen  $a_i$

$$\{(a_i, t_i) | a_i \in \mathcal{A}_i, t_1 \leq t_2 \leq \dots \leq t_N, t_1, \dots, t_N \in T\} \quad (1)$$

Einzelne Aufträge  $a_i$  können hierbei beispielsweise Datenübertragungs- oder Verbindungsaufbauwünsche repräsentieren. Das vorgestellte Konzept ist weiterhin nicht auf Modellierung von Rechnernetzen beschränkt. Zur Modellierung von Datenbanksystemen könnten Transaktionen als Aufträge modelliert werden, um so zu einer Lastbeschreibung gemäß Definition 1 zu gelangen. Um eine detaillierte Spezifikation der Aufträge zu ermöglichen, verwenden wir darüber hinaus typisierte Aufträge sowie eine typabhängige Menge von Auftragsattributen. Aufbauend auf Definition 1 seien die folgenden 4 Klassen von Transformationen definiert:

1. Wir definieren eine Auftragstransformation als Abbildung  $T_A$ , welche eine Sequenz von Primärlastaufträgen  $A^p = (a_1^p, \dots, a_N^p)$  auf eine Sequenz von Sekundärlastaufträgen  $A^s = (a_1^s, \dots, a_K^s)$  für eine gegebene Lasttransformation abbildet.

$$T_A : A^p \rightarrow A^s$$

2. Die Transformation des zeitlichen Verhaltens sei als Abbildung der Ankunftszeitpunkte der Primärlast  $T^p = (t_1^p, \dots, t_N^p)$  auf die Ankunftszeitpunkte der Sekundärlast  $T^s = (t_1^s, \dots, t_K^s)$  definiert.

$$T_T : T^p \rightarrow T^s$$

3. Falls die Transformation von Attributen nicht unabhängig vom zeitlichen Verhalten erfolgen kann, so wird die Transformation der Aufträge beschrieben durch die Abbildung:

$$T'_A : A^p \times T^p \rightarrow A^s$$

4. Analog hierzu:

$$T'_T : A^p \times T^p \rightarrow T^s$$

Hiervon ausgehend ist es prinzipiell möglich, simulative Transformationen durchzuführen, was jedoch aus zwei Gründen nachteilig ist: Einerseits können bei einem simulativen Ansatz auch aus einem stochastischen Primärlastmodell nur deterministische Sequenzen konkreter Ankünfte als Sekundärlast gewonnen werden, indem Auftragssequenzen in Übereinstimmung mit dem Primärlastmodell erzeugt werden und

diese dann transformiert werden. Dies stellt eine Einschränkung der Allgemeingültigkeit des Modells dar. Andererseits ist Last, welche durch simulative Transformation erzeugt wurde, warteschlangentheoretischen Methoden unzugänglich, was insbesondere die Leistungsbewertung erschwert.

Um eine weitgehende analytische Handhabbarkeit der transformierten Modelle zu sichern, ohne gleichzeitig die Beschreibungsmächtigkeit der Modelle allzu sehr einzuschränken, werden im vorliegenden Beitrag Transformationen auf *Batch Markovian Arrival Processes* (BMAPs) untersucht [Luc93], wie sie in Definition 2 dargestellt sind.

**Definition 2** *Ein Batch Markovian Arrival Process (BMAP) wird definiert als markovscher Ankunftsprozess mit infinitesimaler Generatormatrix:*

$$\mathbf{Q} = \begin{pmatrix} D_0 & D_1 & D_2 & D_3 & \dots \\ & D_0 & D_1 & D_2 & \dots \\ & & D_0 & D_1 & \dots \\ & & & \vdots & \ddots \end{pmatrix}$$

Weiterhin seien die  $(m \times m)$  Matrizen  $D_0, \dots, \infty$  definiert durch:

$$(D_0)_{ii} = -\lambda_i, \quad 1 \leq i \leq m$$

$$(D_0)_{ij} = \lambda_i p_i(0, j), \quad 1 \leq i, j \leq m \wedge i \neq j$$

$$(D_k)_{ij} = \lambda_i p_i(k, j), \quad 1 \leq i, j \leq m \wedge k > 0$$

$$\sum_{\substack{j=1 \\ j \neq i}}^m p_i(0, j) + \sum_{k=1}^{\infty} \sum_{j=1}^m p_i(k, j) = 1, \quad 1 \leq i \leq m$$

◇

Hierbei bezeichnet  $p_i(k, j)$  die Wahrscheinlichkeit, dass beim Übergang von Zustand  $i$  in Zustand  $j$  eine Batch-Ankunft der Größe  $k$  erfolgt. Es können somit beliebige diskrete Verteilungen mit möglicherweise unendlichem Träger modelliert werden. Die Gesamtrate, mit der Zustand  $i$  verlassen wird, ist weiterhin gegeben durch  $\lambda_i$ . BMAP-Prozesse sind abgeschlossen gegenüber Aggregation, so dass auch Transformationen aggregierter Lasten oder Aggregationen transformierter Lasten angegeben werden können.

In [HW07] wurde bereits eine BMAP-Transformation für Fragmentierungsvorgänge vorgeschlagen. Da im Folgenden auch die Kombination mehrerer Transformationen betrachtet werden soll, ist die dort vorgeschlagene Transformationsvorschrift in Definition 3 dargestellt. Hierbei wird angenommen, dass die Zwischenankunftszeit von Fragmenten  $\tau_f$  konstant ist mit  $\tau_f > 0$ . Diese konstante Zwischenankunftszeit kann im Kontext kontinuierlicher Markov-Prozesse nur approximativ modelliert werden; durch die Einführung zusätzlicher Zustände kann jedoch – falls nötig – ein höherer Übereinstimmungsgrad erreicht werden.

Da durch die konstante Zwischenankunftszeit zwischen den Fragmenten  $\tau_f$  auch die Zwischenankunftszeit zwischen dem letzten Fragment eines Paketes und dem ersten Fragment des nachfolgenden Paketes verändert wird, erfolgt eine Reduktion der mittleren Verweilzeit des Prozesses in den Zuständen  $w_{i,j}$ . Diese wird um die mittlere Zeit  $E[T_i^F]$ , die für die Fragmentierung benötigt wird, reduziert. Die Übergangsraten wurden so gewählt, dass die durch eine Paketankunft im Zustand  $i$  induzierte Anzahl Fragmente im Mittel mit der durch den transformierten Prozess generierten Anzahl übereinstimmen. Die durch den transformierten Prozess generierte Fragmentanzahl ist hierbei geometrisch verteilt. In [HW07] wurden weiterhin Methoden angegeben mit Hilfe derer auch andere Verteilungen modelliert werden können.

Jeder Zustand des untransformierten BMAPs resultiert in mindestens zwei Zuständen der transformierten BMAPs, wobei die Zustände  $w_{i,j}$  die eigentlichen Transitionsraten des Primärlastmodells repräsentieren und die Fragmenterzeugung durch die Zustände  $g_i$  modelliert wird.

Hierbei steht  $p_i^{TF}(k, j)$  für die Wahrscheinlichkeit, dass aus Zustand  $i$  ein Übergang in Zustand  $j$  erfolgt, der ein abschließendes Teilfragment der Länge  $k$  induziert. Weiterhin entspricht  $m_p$  der Anzahl der Zustände von  $B^P$ .<sup>1</sup> Die Anzahl der im Zustand  $i$  generierten Fragmente sei des Weiteren durch die Zufallsvariable  $L_i^F$  gegeben.

<sup>1</sup>Zur Verdeutlichung benutzen wir für die Anzahl der Zustände des Primär- bzw. Sekundärlastmodells die Bezeichnung  $m_p$  bzw.  $m_s$ .

**Definition 3** Sei die Primärlast definiert durch BMAP  $B^P$ . Die Sekundärlast nach der Fragmentierung mit maximaler Fragmentlänge  $M$  wird dann modelliert durch BMAP  $B_F^S$  mit der Menge der erzeugenden Zustände <sup>2</sup>

$$G = \{g_i, 0 < i \leq m_p\},$$

der Menge der Wartezustände

$$W = \{w_{i,j}, \forall i, j, \exists k. p_i(j, k) > 0\}$$

und den Matrizen  $D_0, \dots, \infty$

$$(D_k)_{w_{i,j}, g_j} = \begin{cases} \frac{\lambda_j}{1 - \lambda_j \cdot E[T_i^{TF}]}, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

$$(D_k)_{w_{i,j}, w_{i,j}} = \begin{cases} -\frac{\lambda_j}{1 - \lambda_j \cdot E[T_i^{TF}]}, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

$$(D_k)_{g_i, g_i} = \begin{cases} \lambda_f \cdot \left(1 - \frac{1}{E[L_i^F] + 1}\right), & k = M \\ 0, & k \neq M \wedge k \neq 0 \\ -\lambda_f, & k = 0 \end{cases}$$

$$(D_k)_{g_i, w_{i,j}} = \begin{cases} \lambda_f \cdot p_i^{TF}(k, j) \cdot \frac{1}{E[L_i^F] + 1}, & k \leq M \\ 0, & k > M \end{cases},$$

wobei  $\lambda_f = \tau_f^{-1}$ .

◇

Neben der Fragmentierung stellt – insbesondere für Effizienzbetrachtungen – das Hinzufügen von zusätzlichen Informationen (*Header*) eine wichtige Transformation dar. Dies gilt insbesondere für verlustbehaftete Übertragungswege, weil dort die Fragmentierung großer Informationseinheiten einen Weg zur Reduktion der Fehlerhäufigkeit darstellt und sich diese nur unter Berücksichtigung des zusätzlichen Overheads realistisch betrachten lässt. Daher sei die Headertransformation für BMAPs definiert, wie in Definition 4 dargestellt.

**Definition 4** Sei die Primärlast definiert durch BMAP  $B^P$  mit den Matrizen  $(D_i^P)_{j,k}$ . Das Hinzufügen von  $h$  Aufträgen zu einer Batchankunft der Länge  $i$  ergibt BMAP  $B_H^S$ , welcher definiert ist durch die Matrizen

$$D_0^S = D_0^P$$

$$D_{i+h}^S = D_i^P, \quad i > 0$$

◇

Der vorgenannten Transformation liegt dabei die Annahme zugrunde, dass Zustandsübergänge mit Ankünften der Größe 0 keinen Paketversand zur Folge haben – also auch kein Headerversand nötig wird. Weiterhin zieht jede Ankunft der Größe  $i$  den Versand eines Paketes der Länge  $i + h$  nach sich, so dass die korrespondierenden Matrizen  $D_{i+h}^S$  und  $D_i^P$  identisch sind.

### 3 Verlusttransformation

Neben den beiden oben erwähnten Transformationen spielen insbesondere bei drahtlosen Übertragungskanälen Auftrags- bzw- Paketverluste eine große Rolle. Nachfolgend soll daher eine Transformation vorgeschlagen werden, welche die Berechnung des transformierten BMAPs sowohl für Fehlermodelle mit konstanter Fehlerrate als auch für markovsche Fehlermodelle erlaubt. Diese ist in Definition 5 dargestellt.

<sup>2</sup>Wir bezeichnen im Kontext der Fragmentierungstransformation diejenigen Zustände, bei deren Verlassen Batch-Ankünfte größer 0 mit Wahrscheinlichkeit größer 0 auftreten, als *erzeugende Zustände*.

**Definition 5** Die augenblickliche Verlustrate zum Zeitpunkt  $t$  werde durch den Markovprozess  $X = (X(t) : t \geq 0)$  mit infinitesimaler Generatormatrix  $Q_L$  gesteuert; die Fehlerwahrscheinlichkeit im Zustand  $i$  für ein Paket der Länge  $j$  sei  $r_{i,j}$ . Dann ergibt sich aus BMAP  $B^P$  der verlustbehaftete BMAP  $B_L^S$  mit der Zustandsmenge  $Z = \{Z_B \times Z_X\}$  und den Matrizen

$$(D_0^L) = D_0^P \oplus Q_L + \sum_{i=1}^{\infty} E_i \otimes D_i^P$$

$$(D_i^L) = (I - E_i) \otimes D_i^P, \quad i \geq 1$$

Hierbei sei  $E_i$  definiert als

$$E_i = \text{diag}(r_{1,i}, \dots, r_{m_L,i}) = \begin{pmatrix} r_{1,i} & 0 & \cdots & 0 \\ 0 & r_{2,i} & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{m_L,i} \end{pmatrix}.$$

◇

Hierbei steht  $\otimes$  für das Kroneckerprodukt zweier Matrizen,  $\oplus$  für die Kroneckersumme (siehe z.B. [Lau05]),  $m_L$  für die Anzahl der Zustände von  $X(t)$  und  $I$  für eine Einheitsmatrix entsprechender Größe. Im folgenden soll nun gezeigt werden, dass durch die Transformation die zeitlichen Eigenschaften der beiden Prozesse – also Ankunfts- und Fehlerprozess – erhalten bleiben. Hierzu seien auf den verlustbehafteten BMAPs zwei Äquivalenzrelationen  $\sim_B$  bzw.  $\sim_L$  definiert, die jeweils diejenigen Zustände umfassen, bei denen sich der Ankunfts- bzw. der Fehlerprozess im selben Zustand befinden. (Hierbei wird der endliche zweidimensionale Zustandsraum aus Def. 5 eindimensional betrachtet in der Ordnung wie sie durch die Kroneckeroperationen induziert wird.) Es wird sich zeigen, dass die Raten zwischen Zustandsmengen, die in verschiedenen Äquivalenzklassen enthalten sind, den Raten des korrespondierenden Ausgangsprozesses gleichen.

$$i \sim_B j \iff i \equiv j \pmod{|Z_B|} \quad (2)$$

$$i \sim_L j \iff \left\lfloor \frac{i}{|Z_B|} \right\rfloor = \left\lfloor \frac{j}{|Z_B|} \right\rfloor \quad (3)$$

Hierauf aufbauend steht  $\approx_L$  bzw.  $\approx_B$  für die Nicht-Äquivalenz zweier Zustände bezüglich der in den Formeln (2) und (3) eingeführten Relationen. Weiterhin sei die Rate zwischen zwei Zuständen  $i$  und  $j$  des Fehlerprozesses mit  $\lambda_{i,j}^L$  bzw. die Rate zwischen zwei Zuständen  $i$  und  $j$  des transformierten Prozesses mit  $\lambda_{i,j}$  bezeichnet.

**Satz 1** Für zwei Zustände  $i, j$  mit  $i \approx_L j$  bleiben die Übergangsraten des Fehlerprozesses zwischen den Äquivalenzklassen  $[i], [j]$  durch die Verlusttransformation erhalten, d.h. es gilt:

$$\lambda_{[i],[j]} = \lambda_{k,l}^L, \quad k = \left\lfloor \frac{i}{|Z_B|} \right\rfloor, \quad l = \left\lfloor \frac{j}{|Z_B|} \right\rfloor$$

**Beweis** Per Konstruktion gilt für zwei Zustände  $m, n$ ,  $m \sim_L i$ ,  $n \sim_L j$ ,  $m \sim_B n$ :

$$\begin{aligned} \lambda_{m,n} &= (D_0^P \oplus Q_L)_{m,n} + \left( \sum_{i=1}^{\infty} E_i \otimes D_i^P \right)_{m,n} + \left( \sum_{i=1}^{\infty} (I - E_i) \otimes D_i^P \right)_{m,n} \\ &= (D_0^P \oplus Q_L)_{m,n} \\ &= (I \otimes D_0^P)_{m,n} + (Q_L \otimes I)_{m,n} \\ &= (Q_L \otimes I)_{m,n} \\ &= (Q_L)_{k,l} = \lambda_{k,l}^L \end{aligned}$$

Hieraus folgt:

$$\begin{aligned} \sum_{\forall m \sim_L i} \sum_{\forall n \sim_L j} \pi_m \cdot \lambda_{m,n} &= \lambda_{k,l}^L \sum_{\forall m \sim_L i} \pi_m \\ &= \lambda_{k,l}^L \end{aligned}$$

□

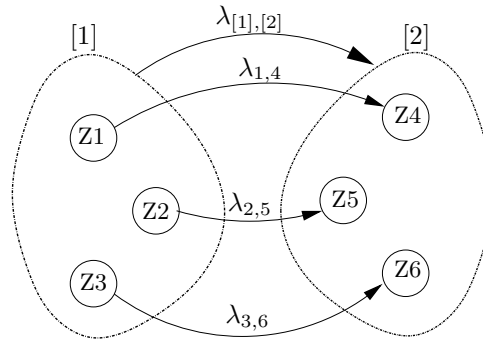


Abbildung 2: Raten zwischen Einzelzuständen und den korrespondierenden Äquivalenzklassen

Hierbei gibt  $\pi_m$  die Zustandswahrscheinlichkeit des Zustands  $m$  für den verlustbehafteten Ankunftsprozess an, unter der Bedingung, dass sich dieser in einem zu  $m$  äquivalenten Zustand (bezüglich  $\sim_L$ ) befindet. Zur Verdeutlichung ist der Zusammenhang zwischen den Raten der Einzelzustände des transformierten Modells und den Raten zwischen den Äquivalenzklassen in Abbildung 2 exemplarisch dargestellt: Da die Transformation so vorgenommen wurde, dass die Raten zwischen den Einzelzuständen der Äquivalenzklassen [1] und [2] identisch sind und der Rate des Fehlermodells entsprechen, ist auch  $\lambda_{[1],[2]}$  gleich dieser Rate. Dies entspricht dem Konzept der *Strongly Lumpable Markov Chains* (siehe z.B. [JB03]), wobei die Partitionierung der Zustandsmenge jeweils entsprechend der Äquivalenzrelation durchzuführen ist.

**Satz 2** Für zwei Zustände  $i, j$  mit  $i \approx_B j$  bleiben die Übergangsraten des Ankunftsprozesses durch die Verlusttransformation erhalten, d.h. es gilt:

$$\lambda_{[i],[j]} = \lambda_{k,l}^B, \quad k = i \bmod |Z_B|, \quad l = j \bmod |Z_B|$$

**Beweis** Folgt analog. □

Die Sätze 1 und 2 besagen insbesondere, dass durch die Kombination der beiden Markovprozesse die zeitlichen Eigenschaften sowohl von Ankunfts- als auch von Fehlerprozess erhalten bleiben. Da weiterhin die Übergangsraten zwischen den Äquivalenzklassen erhalten bleiben, entsprechen auch die jeweiligen stationären Wahrscheinlichkeiten denen der Ausgangsprozesse. Darüber hinaus wird im vorgeschlagenen Modell – aufgrund der Zeitlosigkeit der Batch-Ankünfte – davon ausgegangen, dass sich der Zustand des Fehlerprozesses während der Übertragung eines Fragmentes nicht ändert. Um den hierdurch induzierten Fehler nicht zu groß werden zu lassen, sollten die Änderungsraten nicht erheblich größer als die Übertragungsdauer eines Fragmentes sein (z.B. *Slow Fading*). Weiterhin muss bedacht werden, dass die vorgeschlagene Transformation in erster Linie auf Szenarien mit Vorwärtsfehlerkontrolle (*FEC*) anwendbar ist. Hierbei würden sich all diejenigen Bitfehler als Verlust im Sinne der Transformation niederschlagen, welche durch die Vorwärtsfehlerkontrolle nicht korrigiert werden können. Im Falle von Übertragungswiederholung (*ARQ*) kann diese jedoch ebenfalls verwendet werden, falls durch die eventuellen Wiederholungen nur mit geringer Wahrscheinlichkeit Rückstaus induziert werden. Ist dies der Fall, so müssten darüber hinaus die zeitlichen Verschiebungen, wie sie durch Übertragungswiederholungen hervorgerufen werden, in geeigneter Weise berücksichtigt werden.

Mit Hilfe der vorgeschlagenen Transformation lässt sich nun die Veränderung der Lasteigenschaften bis zur Ankunft im Zugangsnetz nachbilden (siehe Abbildung 1), wobei in Abhängigkeit vom vorliegenden Protokollstapel u.U. mehrfache Fragmentierungs- und Headertransformationen anzuwenden sind.

Der Effekt einer solchen Komposition von Transformationen ist in Abbildung 3 illustriert: Während in der linken Hälfte der Abbildung die Auswirkungen von Fragmentierungs-, Header- und Verlusttransformation auf Paketebene dargestellt ist, zeigt die rechte Hälfte der Abbildung die Kombination auf Prozessebene in schematischer Art und Weise. Hierbei sind die Zustände, welche die Prozesse zum jeweiligen Zeitpunkt

einnehmen, gemäß ihres Indexes ( $Z$ ) dargestellt.  $BMAP$  und Verlustprozess haben jeweils 2 Zustände und resultieren dementsprechend in einem transformierten Prozess mit 4 Zuständen. Bezüglich der Illustration auf Paketebene ist besonders auf die Tatsache hinzuweisen, dass nicht zwangsläufig sämtliche Fragmente eines Paketes verloren gehen müssen.

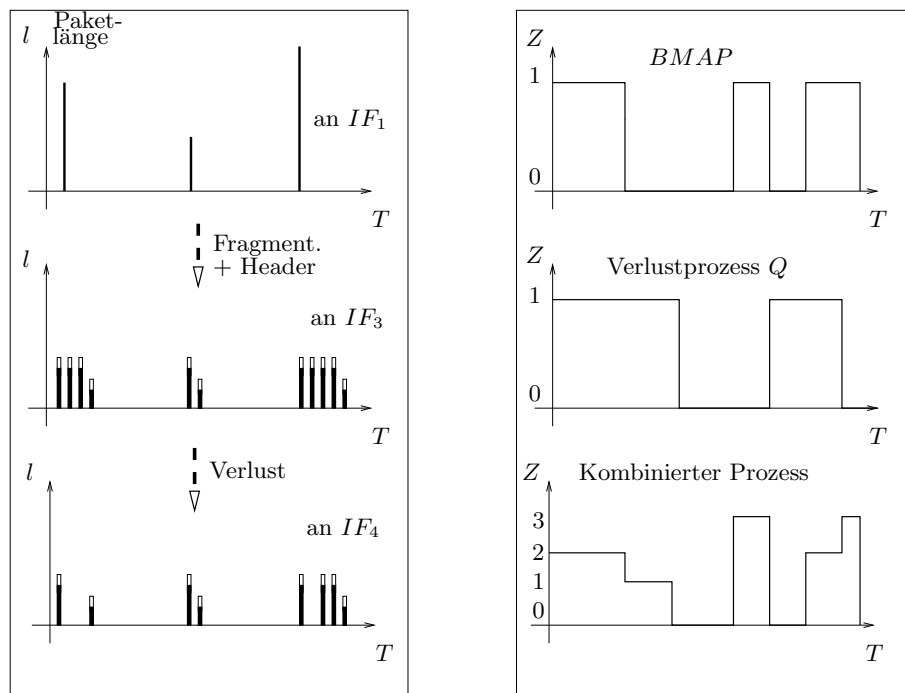


Abbildung 3: Illustration des Transformationsvorgangs auf Prozessebene und am Beispiel konkreter Auftragsankünfte (Bem.: Die genannten Schnittstellen  $IF_i$  (linke Graphik) beziehen sich auf Abb. 1.)

### 3.1 Berechnung der mittleren Rate

Da im vorliegenden Modell Paketankünfte als zeitlose Ereignisse angenommen werden, hängt die Wahrscheinlichkeit der Ankunft eines fehlerhaften Paketes ausschließlich von den stationären Zustandswahrscheinlichkeiten des Fehlerprozesses  $\pi^L$  ab. Hieraus folgt, dass sich die mittlere Rate des verlustbehafteten  $BMAP$ s, wie in Formel (4) dargestellt, angeben lässt, wobei  $e$  dem Spaltenvektor  $[1 \ 1 \ \dots \ 1]^T$  entsprechender Länge entspricht. Weiterhin gibt  $\pi$  die stationären Zustandswahrscheinlichkeiten des  $BMAP$ s an, welcher die Primärlast beschreibt.

$$R = \sum_{i=1}^{\infty} i \left( (\pi \cdot D_i \cdot e) \cdot (\pi^L \cdot E^i \cdot e) \right) \quad (4)$$

Erfolgt zusätzlich eine Fragmentierung der Ankünfte, so lässt sich die mittlere Rate des fragmentierten Verlustprozesses aufbauend auf der folgenden Überlegung angeben: Der fragmentierte Ankunftsprozess, welcher Paketverlusten ausgesetzt ist, besteht aus einer Abfolge von Ankünften der maximalen Batch-Größe  $M$  unterbrochen von Ankünften einer Batch-Größe kleiner  $M$ . Dies entspricht der Unterteilung eines Paketes auf Primärlastebene in  $n - 1$  Fragmente voller Länge, sowie einem Paket der Länge 0 bis  $M - 1$ , welches die restlichen Daten des Paketes enthält.<sup>3</sup> Die Fragmente voller Länge erfahren eine Verlustwahrscheinlichkeit  $E^M$ , wobei pro Paket im Mittel  $E[L^F]$  solcher Fragmente versendet werden. Entsprechend folgt:

$$R = \sum_{i=1}^{M-1} i \left( (\pi \cdot D_i^{TF} \cdot e) \cdot (\pi^L \cdot E^i \cdot e) \right) + M \cdot \left( (\pi \cdot \lambda) \cdot (\pi^L \cdot E^M \cdot e) \right) \cdot E[L^F] \quad (5)$$

<sup>3</sup>Die Fragmentlänge 0 des  $n - ten$  Paketes entspricht genau dem Fall, in dem die maximale Fragmentlänge die Paketlänge ohne Rest teilt und modelliert das "Nicht-Versenden" des Teilfragmentes.

Hierbei steht  $\underline{\lambda}$  für den Zeilenvektor bestehend aus den Abgangsraten aus den Zuständen  $1, \dots, n$  und  $D_i^{TF}$  ist die Ratenmatrix der Teilfragmente der Länge  $i$ . Zur realistischen Betrachtung des Durchsatzes gilt es allerdings zu bedenken, dass einerseits jedes Fragment zusätzlichen Overhead induziert und andererseits eine Unterteilung in kleinere Fragmente grundsätzlich eine geringere Fehlerwahrscheinlichkeit zur Folge hat. Um diesen Randbedingungen gerecht zu werden, erfolgt die Erweiterung von Formel (5) um den Faktor  $\frac{i}{i+h}$ , so dass folgt:

$$R' = \sum_{i=1}^{M-1} i \left( (\pi \cdot D_i^{TF} \cdot e) \cdot (\pi^L \cdot E^i \cdot e) \right) \cdot \frac{i}{i+h} + M \cdot \left( (\pi \cdot \underline{\lambda}) \cdot (\pi^L \cdot E^M \cdot e) \right) \cdot E[L^F] \cdot \frac{M}{M+h} \quad (6)$$

Mit Hilfe von Formel (6) ist es jetzt insbesondere möglich, für gegebene Kombinationen von Ankunfts- und Fehlerprozess eine hinsichtlich des Gesamtdurchsatzes optimale Fragmentgröße zu ermitteln. Hierbei ist zu beachten, dass nicht mit jedem Zustandswechsel eine Neuberechnung dieses Wertes nötig wird, sondern nur die Prozesse als solche eine Rolle spielen. Weiterhin lässt sich erkennen, dass bezüglich des Ankunftsprozesses in erster Linie die mittlere Größe der Primärlastankünfte von Bedeutung ist. Die eigentliche Verteilung wirkt sich nur im ersten Summanden von Formel (6) aus. Dessen Gewicht nimmt aber mit zunehmender Gesamtgröße ab. Weiterhin lässt sich für die Verteilungen der Teilfragmentlänge in vielen Fällen mit guter Näherung eine Gleichverteilung annehmen [HW07], so dass die eigentliche Paketlängenverteilung einen geringen Einfluss auf das Ergebnis der Summe hat.

Dies wird auch durch den in Abbildung 4 dargestellten Durchsatz in Abhängigkeit von der Fragmentgröße bestätigt. Wir betrachten das folgende Szenario: Als Fehlermodell fand dort ein *Gilbert-Elliot-Modell* Verwendung.<sup>4</sup> Ein solches Modell besteht aus zwei Zuständen, wobei im Zustand 1 eine geringe Fehlerrate  $r_1$  angenommen wird, im Zustand 2 jedoch eine höhere Fehlerrate  $r_2$ . Als Fehlerraten pro Auftrag wurden  $r_1 = 0$  und  $r_2 = 10^{-3}$  gewählt. Die Generatormatrix  $Q_1$  für das verwendete Fehlermodell  $G_1$  ist in Formel (7) dargestellt.

$$Q_1 = \begin{pmatrix} -8 & 8 \\ 2 & -2 \end{pmatrix} \quad (7)$$

Weiterhin erzeugen alle betrachteten Primärlastmodelle 138000 Aufträge pro Sekunde, wobei der konkrete Auftragsprozess jeweils variiert wurde. BMAP  $B_1$  erzeugt Pakete gleichverteilter Länge mit einer Rate von 100 Aufträgen pro Sekunde, während BMAP  $B_2$  ein einfaches Modell eines Videostromes darstellt, bei welchem mit einer Rate von 25 Aufträgen pro Sekunde in 3 Zuständen Pakete mit normalverteilter Länge erzeugt werden. BMAP  $B_3$  erzeugt weiterhin Aufträge konstanter Länge mit einer Rate von 100 Aufträgen pro Sekunde. Die mittlere Ankunftsrate der vorgenannten BMAPs in Abhängigkeit von der maximalen Fragmentlänge wurde nach Formel (8) berechnet. Hierbei erfolgte in Analogie zu Formel (6) die Erweiterung der Formel zur Ermittlung des mittleren Durchsatzes eines BMAPs (siehe z.B. [Luc93]) um den Faktor  $\frac{i}{i+h}$  zur Gewichtung des Overheads.

$$\lambda^* = \pi \left( \sum_{i=1}^{\infty} \frac{i}{i+h} i D_i \right) e \quad (8)$$

In der linken Hälfte der Abbildung lässt sich erkennen, dass für alle 3 Ankunftsprozesse der optimale Durchsatz bei  $M \approx 290$  erreicht wird und dass trotz der unterschiedlichen Charakteristiken alle Kurven einen ähnlichen Verlauf nehmen.

In der rechten Hälfte der Abbildung erfolgt der Vergleich zwischen dem Durchsatz der transformierten BMAPs  $\lambda^*$  und dem direkt berechneten Durchsatz  $R'$ . Es lässt sich (erfreulicherweise) kein Unterschied zwischen beiden Kurven erkennen.

## 4 Vergleich mit Simulationsresultaten

Neben dem analytischen Vergleich der mittleren Rate soll im Folgenden eine detailliertere Validierung der mittels Transformation erzeugten Ankunftsprozesse vorgenommen werden. Diese erfolgt mit Hilfe des Vergleichs von Auftragssequenzen, wie sie in Übereinstimmung mit transformierten BMAPs erzeugt wurden und simulativ transformierten Auftragssequenzen. Es soll somit gezeigt werden, dass unter Nutzung der oben dargestellten Transformationen realitätsnahe Simulationsszenarien mit hoher Genauigkeit nachgebildet werden können. Die Simulation erfolgte unter Zuhilfenahme des Simulationspaketes *ns2*, wobei das

<sup>4</sup>Die im Folgenden genutzten Parametrisierungen haben ausschließlich exemplarischen Charakter. Konkrete Parametrisierungen für reale Kanäle können beispielsweise den in Abschnitt 1 erwähnten Arbeiten entnommen werden.

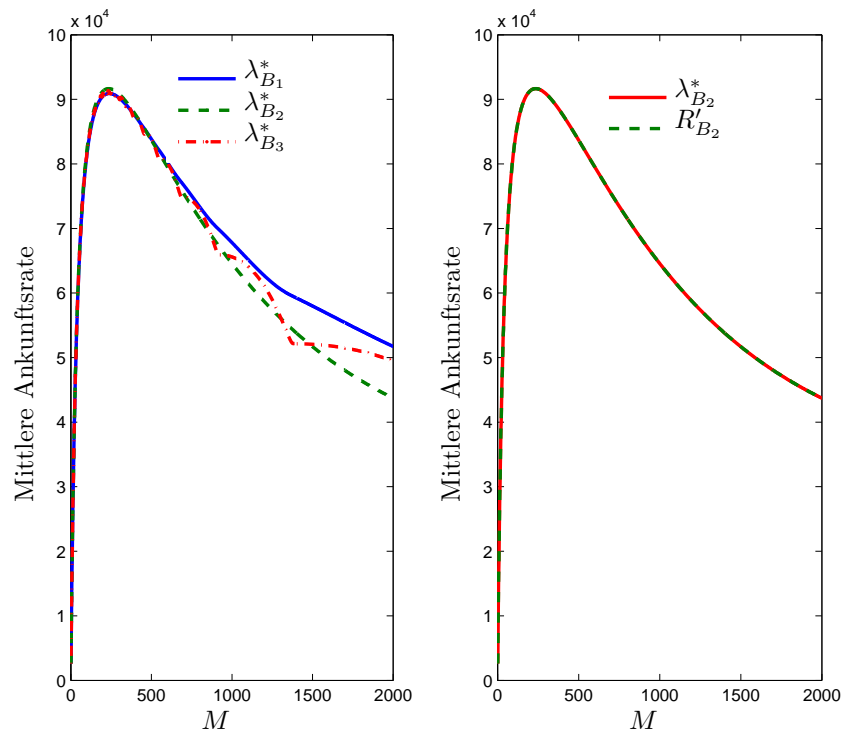


Abbildung 4: Mittlere Ankunftsraten  $\lambda^*$  der transformierten BMAPs  $B_{1,2,3}$  und die direkt berechnete Rate  $R'_{B_2}$

untersuchte Simulationsszenario in Abbildung 5 dargestellt ist: Die jeweilig betrachteten Ankunftsprozesse an der Schnittstelle zur Transportschicht (als vorausgesetzter Primärlastschnittstelle  $IF_P$ ) wurden per UDP übertragen; die Fragmentierung der Pakete entsprechend der Maximalgröße  $M$  wurde innerhalb des *UDP-Moduls* vorgenommen. Die so fragmentierten Daten wurden simulativ über eine  $100\text{Mbit/s}$ -Leitung übertragen, auf der die Paketverluste induziert wurden. Die Messung der Ankunftszeiten und der Paketgrößen erfolgte, wie in Abbildung 5 dargestellt, an der Schnittstelle  $IF_S$ . Diese Schnittstelle repräsentiert die Schnittstelle oberhalb der Datensicherungsschicht (beobachtet in sämtlichen als Empfänger fungierenden Stationen/Rechnern).

In der Folge sollen dabei zwei Ankunftsprozesse mit abweichenden Charakteristiken, aber identischer

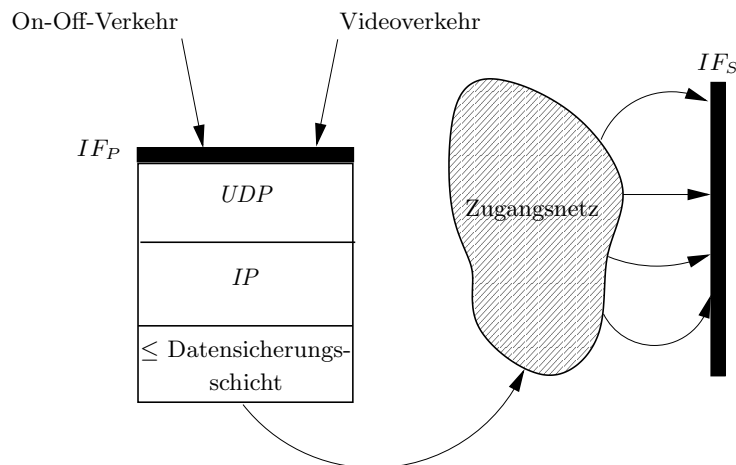


Abbildung 5: Übertragung eines Paketes in der Simulation

mittlerer Ankunftsrate, Gegenstand der Untersuchung sein. Diese sind:

- *Videoverkehr*  $B_2$  Hierbei findet ein Modell Verwendung, in dem die 3 Rahmentypen eines MPEG-

Videos (siehe z.B. [Sym01]) durch jeweils eigene Zustände modelliert werden. Die Wahrscheinlichkeitsverteilung der Batch-Größen ist jeweils diskretisiert normalverteilt mit von einander abweichenden rahmenspezifischen Parametern.

- *On-Off-Verkehr*  $B_4$  Dieses Modell verfügt über zwei Zustände mit einer mittleren Verweilzeit von jeweils 0,5s. Im *On*-Zustand werden Pakete der Größe 2766 mit der Rate 100 erzeugt. Im *Off*-Zustand erfolgen keine Ankünfte.

Als Fehlermodell fand neben dem bereits in Abschnitt 3.1 verwendeten Modell  $G_1$  das Modell  $G_2$  mit abweichender Generatormatrix  $Q_2$ , siehe Formel (9), Verwendung.

$$Q_2 = \begin{pmatrix} -2 & 2 \\ 8 & -8 \end{pmatrix} \quad (9)$$

Die Fehlerprozess  $G_2$  wurde hierbei wegen seiner vom Fehlerprozess  $G_1$  stark abweichenden stationären Zustandswahrscheinlichkeiten gewählt. Weiterhin wurden neben den im Folgenden dargestellten Experimenten eine Vielzahl weiterer Experimente mit abweichenden Modellen und Startwerten des Zufallszahlengenerators durchgeführt, die indes bezüglich der Übereinstimmung der simulativ bzw. analytischen transformierten Prozesse vergleichbare Resultate lieferten. Anhand des On-Off-Modells  $B_4$  wurde zunächst der Vergleich der erzeugten Gesamtlast vorgenommen. Da aufgrund der konstanten Paketlänge die Länge der Teilfragmente a priori bekannt ist, ist in Tabelle 1 jeweils die Zahl der Fragmente maximaler Länge und der Teilfragmente in Abhängigkeit der maximalen Fragmentgröße  $M$  und der Zustandswahrscheinlichkeit des fehlerbehafteten Zustands 2 angegeben. Hierbei entspricht  $\pi_2 = 0,8$  dem Fehlerprozess  $G_1$  bzw.  $\pi_2 = 0,2$  dem Prozess  $G_2$ . Die Abweichungen sind durchgängig gering, wobei für die Modelle mit höherer Fehlerrate diesbezüglich eine leichte Zunahme zu beobachten ist. Geringere Abweichungen sind indes a priori zu erwarten, da im Verlustfall die Zwischenankunftszeit zwischen den angrenzenden Paketen nicht mehr exakt nachgebildet werden kann: Geht in der Simulation ein Fragment eines Paketes verloren, so beeinflusst dies nur die Zwischenankunftszeit der beiden angrenzenden Fragmente, die dann nicht mehr negativ exponentiell verteilt ist. Von dieser Tatsache muss jedoch in der Transformation aus Komplexitätsgründen abstrahiert werden, so dass besagte Zwischenankunftszeiten im Verlustfall weiterhin exponentiell verteilt sind.

Auch für den zweiten zu betrachteten Ankunftsprozess  $B_2$  lassen sich sowohl Paketlängen als auch Zwi-

$M$	$\pi_2$	Simulation	Transformation
1000	0,8	177571/102529	178424/103478
1000	0,6	223380/226147	226147/124235
1000	0,2	313638/160041	312331/160004
1500	0,8	67854/76136	67747/76152
1500	0,6	94948/101194	95782/101367
1500	0,2	150640/152633	151454/153578
$\infty$	0,8	45266	43841
$\infty$	0,6	79264	78289
$\infty$	0,2	146549	145374

Tabelle 1: Anzahl der ankommenden Batch-Aufträge maximaler/nicht maximaler Größe

schenankunftszeiten mit hoher Genauigkeit nachbilden.<sup>5</sup> Anhand der in Abbildung 6 dargestellten empirischen Verteilung der Fragmentlängen lässt sich erkennen, dass auch diese Verteilung von der Transformation mit hoher Genauigkeit nachgebildet wird – zwischen den beiden Kurven ist jeweils kein Unterschied erkennbar. Hierbei ist im linken Teil der Abbildung die Verteilung für den Fehlerprozess  $Q_1$  dargestellt, während der rechte Teil die Verteilung für  $Q_2$  zeigt. Weiterhin lässt sich erkennen, dass mit zunehmender Bitfehlerwahrscheinlichkeit die Wahrscheinlichkeit von Fragmenten mit maximaler Größe abnimmt, was auf die steigende Fehlerrate mit zunehmender Paketlänge zurückzuführen ist. Aus dem selben Grund kann mit zunehmender Bitfehlerwahrscheinlichkeit auch keine Gleichverteilung für die Länge der Teilfragmente mehr angenommen werden: Während im rechten Teil von Abbildung 6 der Verlauf für  $X < 1500$  noch mit guter Näherung linear ist – was für Gleichverteilung spricht –, ist dies im linken Teil nicht mehr der Fall und kürzere Pakete sind deutlich wahrscheinlicher. Neben der Übereinstimmung der Längenverteilung

<sup>5</sup>Den dargestellten Resultaten liegen eine Vielzahl von Simulationsläufen mit unterschiedlichen Zufallszahlengenerator-Startwerten und einer Simulationsdauer von 6000s zugrunde, welche alle nur sehr geringfügig voneinander abweichende Ergebnisse lieferten.

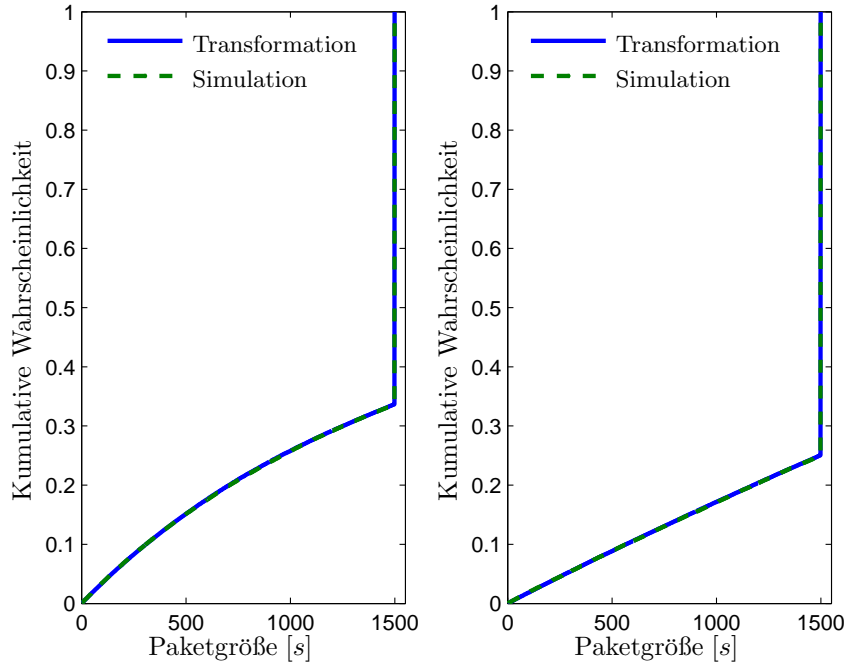


Abbildung 6: Kumulative Fragmentlängen-Wahrscheinlichkeit des verlustbehafteten Ankunftsprozesses bei maximaler Fragmentlänge  $M = 1500$  für die Fehlerprozesse  $Q_1$  (links) und  $Q_2$  (rechts)

der erzeugten Fragmente ist die zeitliche Übereinstimmung von großer Wichtigkeit. Nach der analytischen Untersuchung dieses Aspektes in Abschnitt 3 soll hier nun am Beispiel von  $B_2$  gezeigt werden, dass die vorgeschlagenen Transformationsverfahren darüber hinaus mit den simulativ gewonnenen Resultaten mit hoher Genauigkeit übereinstimmen, so dass auf eine sehr gute Realitätsnähe des analytischen Modellierungsansatzes geschlossen werden kann. Hierzu ist in Abbildung 7 die empirische Verteilung der Zwischenankunftszeiten für den Fehlerprozesse  $Q_1$ , wie sie jeweils simulativ und durch Transformation gewonnen wurden, dargestellt. Während in der linken Hälfte der Abbildung die Verteilung für  $M = 1500$  zu sehen ist, zeigt die rechte Hälfte die Verteilung bei ausbleibender Fragmentierung. Für beide Szenarien bildet die Transformation die simulativ erzielten Ergebnisse mit hoher Genauigkeit nach. Geringe Abweichungen sind insbesondere für  $M = 1500$  erkennbar. Dies steht im Einklang mit dem weiter oben erwähnten Problem der nicht-exponentiellen Verteilung der Zwischenankunftszeit im Verlustfall. Weiterhin ist zu beobachten, dass mit zunehmender Fehlerwahrscheinlichkeit die Wahrscheinlichkeit höherer Zwischenankunftszeiten erwartungsgemäß zunimmt.

## 5 Resümee und Ausblick

Dieser Beitrag beschäftigte sich mit dem Problem einer realitätsnahen Prädiktion von Lasten (im Sinne von Auftragssequenzen betrachtet an wohldefinierten Schnittstellen), ausgehend von unterschiedlichem (End-)Benutzerverhalten.

Dabei wurden die folgenden Ergebnisse erzielt: Es wurde aufgezeigt, wie durch das Konzept der Lasttransformation die Lasten an tieferen Schnittstellen einer Rechnernetzarchitektur (sog. Sekundärlastschnittstellen) ausgehend von bekannten Lasten an endbenutzernahen bzw. anwendungsorientierten Schnittstellen (sog. Primärlastschnittstellen) prognostiziert werden können. Eine Klassifikation von Lasttransformationen wurde vorgeschlagen, die u.a. berücksichtigt, ob der Auftragsankunftsprozess, die Auftragstypen und /oder die Auftragsattribute bei der entsprechenden Lasttransformation verändert werden. Im Zentrum des Beitrags stand die analytische Modellierung der Lasttransformationen “Headergenerierung”, “(Nachrichten-/Paket-) Fragmentierung” sowie “(Nachrichten-/Paket-) Verlust”, die allesamt auf Lastmodelle vom Typ BMAP (d.h. auf Markov’sche Ankunftsprozesse mit Gruppenankünften) angewendet wurden. Dabei wurden die Implikationen der entsprechenden Lasttransformationen untersucht sowohl im Hinblick auf den Auftragsabgangsprozess aus der betrachteten lasttransformierenden Komponente

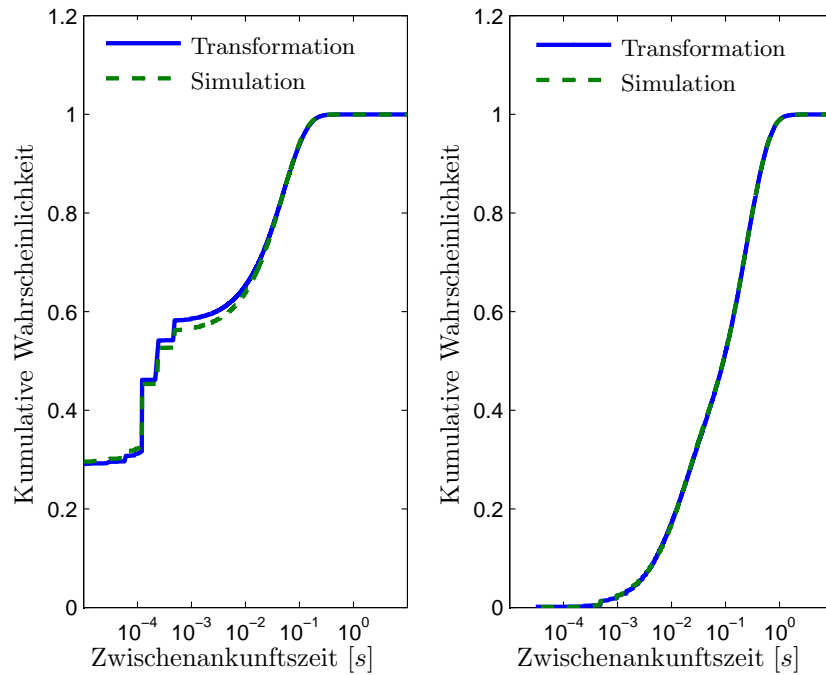


Abbildung 7: Kumulative Zwischenankunftszeit-Wahrscheinlichkeit des verlustbehafteten Ankunftsprozesses bei Verwendung des Fehlerprozesses  $Q_1$  für  $M = 1500$  (links) und für ausbleibende Fragmentierung (rechts)

als auch in Bezug auf die abgehenden Aufträge selbst, z.B. Veränderung von deren Auftragsattributwerten (insbesondere die Auswirkungen auf die Längen – als zumeist zentrales Auftragsattribut – der zu übertragenden Daten bei Datenübertragungsaufträgen). Es gelang uns, die Auswirkungen der Lasttransformationen “Headergenerierung”, “Fragmentierung” und “Auftragsverlust” auf die resultierenden Sekundärlasten weiterhin in Form von BMAP-Modellen zu beschreiben und diese, als Funktion der betrachteten Lasttransformation, zugleich geeignet zu parametrisieren. Unsere Simulationsstudien (auf *ns2*-Basis) zeigten auf, dass selbst hochkomplexe Lasttransformationen in realen Netzen – unter Annahme überlagerter Video- und Sprachübertragungen – noch auf realistische Weise durch die von uns vorgeschlagenen analytischen Lasttransformationen auf BMAP-Modellen berücksichtigt werden können. Hierbei bewährte sich u.a. auch unser Vorschlag, komplexe Lasttransformationen als Hintereinanderausführung elementarer Lasttransformationen zu modellieren. Überdies demonstrierte die in den Simulationsexperimenten betrachtete Fallstudie auch, wie Prognosen von Sekundärlasten auf Basis analytisch modellierter Lasttransformationen herangezogen werden können, um ein ‘Tuning’ freier Parameter des Netzes (wie z.B. eine im Netzbetrieb zu wählende maximale Fragmentgröße) in optimaler Weise vorzunehmen.

Unsere weiteren geplanten Arbeiten zur analytischen Modellierung von Lasttransformationen werden sich auf die Berücksichtigung weiterer Klassen von Transformationen beziehen, insbesondere auch auf Lasttransformationen, deren Verhalten noch stärker vom Netzzustand geprägt ist. Überdies streben wir eine vertiefte Modellvalidierung an, bei der direkte Messungen resultierender Sekundärlasten in exemplarisch gewählten Netzen (WLANs, Intranets, etc.) Verwendung finden sollen.

## Literatur

- [Bai99] G. Bai, Load measurements and modeling for distributed multimedia applications in high-speed networks, Ph.D. thesis, University of Hamburg, 1999.
- [BC92] C. Blondia and O. Casals, Performance analysis of statistical multiplexing of VBR sources, IEEE INFOCOM '92 (Vol. 2) (Los Alamitos, CA, USA), IEEE Computer Society Press, 1992, 828–838.

- [CW06] A. Chydzinski and R. Winiarczyk, Blocking Probability in a BMAP Queue, ISCC '06: Proc. 11th IEEE Symp. on Computers and Communications (Washington, DC, USA), IEEE Computer Society, 2006, 547–553.
- [FC01] H.-W. Ferng and J.-F. Chang, Departure Processes of BMAP/G/1 Queues, Queueing Syst. Theory Appl. **39** (2001), no. 2-3, 109–135.
- [Hof01] J. Hofmann, The BMAP/G/1 Queue with Level-Dependent Arrivals - An Overview, Telecommunication Systems **16** (2001), no. 3-4, 347–359.
- [HSB04] T. Hofkens, K. Spaey, and C. Blondia, Transient Analysis of the D-BMAP/G/1 Queue with an Application to the Dimensioning of a Playout Buffer for VBR Video, NETWORKING, 2004, 1338–1343.
- [HW07] S. Heckmüller and B.E. Wolfinger, Load Transformations for Markovian Arrival Processes, ASMTA 2007, June 2007, 35–43.
- [JB03] R.W. Jernigan and R.H. Baran, Testing lumpability in Markov chains, Statistics and Probability Letters **64** (2003), 17–23.
- [KL00] Y. Y. Kim and S. Li, Modeling multipath fading channel dynamics for packet data performance analysis, Wireless Networks **6** (2000), no. 6, 481–492.
- [KLL03] A. Klemm, C. Lindemann, and M. Lohmann, Modeling IP traffic using the batch Markovian arrival process, Perform. Eval. **54** (2003), no. 2, 149–173.
- [Lau05] A. J. Laub, Matrix Analysis for Scientists and Engineers, SIAM, 2005.
- [LB06] H. W. Lee and J. W. Baek, Threshold Workload Control in the BMAP/G/1 Queue, International Conference on the Quantitative Evaluation of Systems (QEST), IEEE, 2006, 353–364.
- [Luc91] D. M. Lucantoni, New results for the single server queue with a batch Markovian arrival process, Stoch. Mod. **7** (1991), 1–46.
- [Luc93] D. M. Lucantoni, The BMAP/G/1 queue: a tutorial, Models and Techniques for Performance Evaluation of Computer and Communication Systems (L. Donatiello and R. Nelson, eds.), Springer-Verlag, New York, 1993, 330–358.
- [LZ06] H. Levy and H. Zlatokrilov, The effect of packet dispersion on voice applications in IP networks, IEEE/ACM Trans. Netw. **14** (2006), no. 2, 277–288.
- [Neu79] M. F. Neuts, A versatile Markovian point process, J. Appl. Prob. **16** (1979), 764–779.
- [Sym01] P. Symes, Digital Video Compression, McGraw-Hill, 2001.
- [Wol99] B. E. Wolfinger, Characterization of Mixed Traffic Load in Service-Integrated Networks, Systems Science Journal **25** (1999), no. 2, 65–86.
- [WZHB02] B. E. Wolfinger, M. Zaddach, K. D. Heidtmann, and G. Bai, Analytical modeling of primary and secondary load as induced by video applications using UDP/IP, Computer Communications **25** (2002), no. 11-12, 1094–1102.
- [Zad01] M. Zaddach, Charakterisierung, Modellierung und Transformation von Videoverkehrslasten, Ph.D. thesis, University of Hamburg, 2001.
- [ZK99] Q. Zhang and S. A. Kassam, Finite-State Markov Model for Rayleigh Fading Channels, IEEE/ACM Trans. Communic. **47** (1999), no. 11, 1688–1692.
- [ZRM95] M. Zorzi, R. Rao, and L. Milstein, On the accuracy of a first-order Markov Model for data transmission on fading channels, Proc. IEEE ICUPC'95, November 1995, 211–215.