

HOLGER KARL

## STRUKTUR AUS ZUFALL: ENTSTEHUNG VON ABHÄNGIGKEITEN IN TELEKOMMUNIKATIONSSYSTEMEN

### Zusammenfassung

Der Verkehr in Telekommunikationssystemen ist in wesentlichen Aspekten durch die zufälligen und unabhängigen Handlungsweisen einzelner Benutzer geprägt. Dennoch ergibt sich aus solchem unabhängigem Zufall in oft überraschender Weise Struktur, die auch über größere Skalen (insbesondere Zeitskalen) erhalten bleiben kann. Dieser Artikel beschreibt als Beispiel hierfür ein bekanntes Phänomen aus dem World Wide Web: Die Art und Weise, wie Seitengrößen auf einem Webserver augenscheinlich zufällig gemischt sind, reicht bereits aus, um selbst unter zufälligen Zugriffen eine komplexe, über mehrere Zeitskalen hinweg stabile Struktur der Netzauslastung zu erzeugen. Dem Benutzer drängt sich dadurch der Anschein eines komplexen Automatismus auf.

### Einleitung

Der Begriff des Automatismus ist eng mit der Idee der Strukturbildung verknüpft. Werden Strukturen – im Raum, in der Zeit– beobachtet, die sich augenfällig von Zufall unterscheiden, so ist man geneigt, nach einer Ursache zu suchen. Ist eine solche Ursache nicht in einem expliziten, zielgerichteten Handeln zu finden, so steht die Frage nach einer automatischen Bildung dieser Struktur im Raum; der strukturbildende Automatismus mag dabei explizit oder implizit, intentional oder als Nebeneffekt entstanden sein.

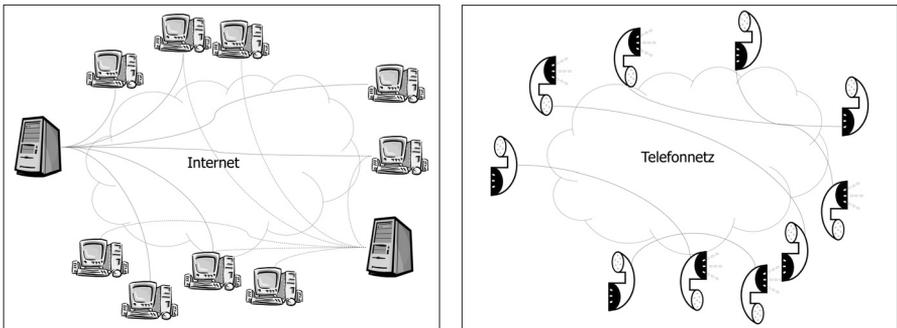
Dieser Artikel stellt ein Beispiel für einen nicht beabsichtigen, sich nur implizit bildenden Automatismus vor, der zu einer augenfälligen und in ihrer technischen Konsequenz relevanten Strukturbildung führt: Die Charakteristik von World-Wide-Web-Verkehr im Internet.

### Zufall in Telekommunikationssystemen – World Wide Web als Beispiel

Die grundlegende Struktur des World Wide Web ist inzwischen auch dem Laien in Grundzügen vertraut. Nutzer dieses Systems bedienen sich eines Rechners, eines mobilen Gerätes (z. B. Mobiltelefon) oder einer ähnlichen Vorrichtung, um auf sogenannte „Webseiten“ zuzugreifen. Dabei wird die Anfrage

nach einer Webseite von einem Webbrowser an einen Webserver übermittelt und von diesem mit der gewünschten Webseite beantwortet. Die eigentliche Übermittlung erfolgt durch das Internet, an das Webserver und Webbrowser angeschlossen sind.

Diese Grundstruktur ist, auf den ersten Blick, der des konventionellen Telefonnetzes sehr ähnlich – Abbildung 1 vermittelt einen Überblick. Auch im Telefonnetz stehen Kommunikationspartner durch Vermittlung eines Netzes miteinander in Kontakt, und Daten werden zwischen diesen Partnern übertragen. Demnach könnte man erwarten, dass die wesentlichen Eigenschaften dieser beiden Systeme zumindest in Grundzügen vergleichbar sind.



1 – Struktur der Kommunikationsbeziehungen in Internet und Telefonnetz

Ein augenfälliger Unterschied ist, dass bei einem Telefonsystem (in aller Regel) zwei Partner in direkter Kommunikationsbeziehung stehen, während im World Wide Web ein Webserver häufig viele unterschiedliche Browser, d. h. Nutzer, bedient, viele davon auch zeitgleich. Tatsächlich ist dies jedoch für die Charakteristik des Verkehrs, der durch diese beiden unterschiedlichen Systeme in ein Kommunikationsnetz injiziert wird, nur von geringem Belang – es ist leicht, sich den Webserver in diesem Kontext als eine Sammlung von einzelnen Geräten vorzustellen, die jeweils separat mit dem zugehörigen Webbrowser kommunizieren. Sicherlich wird dadurch der eingespeiste Verkehr stärker lokalisiert sein; über größere Distanzen und über eine größere Anzahl von Webservern sollte sich dieser Effekt jedoch ausmitteln und keinen allzu großen Beitrag zu einer möglichen Struktur mehr leisten. Dies ist tatsächlich der Fall.

Dennoch ist in diesen Strukturen noch reichlich Platz für Zufall oder, genauer gesagt, für eine zufällige Modellierung von Effekten und Aspekten, die möglicherweise zwar deterministisch sind, für die aber lediglich eine stochastische Beschreibung<sup>1</sup> zur Verfügung steht oder angemessen ist. Wesentliche Aspekte sind:

<sup>1</sup> Stochastik: Teilgebiet der Mathematik, das sich mit der Beschreibung und der Analyse von Zufall befasst. Wichtige Teilgebiete sind die Wahrscheinlichkeitsrechnung (mengentheoretisch)

1. Welcher Teilnehmer möchte mit welchem anderen Teilnehmer kommunizieren?

Hier mag, wie im vorigen Abschnitt diskutiert, ein Unterschied zwischen Internet und Telefon liegen, der jedoch nicht von allzu großem Belang ist.

2. Wie groß ist die Menge der Teilnehmer, bzw. der Webserver und Webbrowser?

3. Zu welchem Zeitpunkt erfolgt ein Kommunikationsversuch? Alternativ: Wie viel Zeit vergeht zwischen dem Ende einer Kommunikation und dem Beginn der nächsten?

In der praktischen Modellierung ist die Zeit zwischen Kommunikationsvorgängen meist einfacher zu behandeln als die absoluten Zeitpunkte. Letztere spielen tatsächlich auch nur eine untergeordnete Rolle, wenn man quantitative Aspekte wie etwa „Auslastung“ untersuchen möchte.

In der Tat kann man für die Größe bei Telefonnetz und Internet in erster Näherung zunächst ähnliche Annahmen treffen. Eine typische Beschreibung würde die *Zwischenkommunikationszeit* in beiden Fällen als eine stochastische Größe auffassen, hierzu eine Verteilungsfunktion bestimmen, und ebenfalls typischerweise eine Unabhängigkeitsannahme treffen – d. h., es gibt keine Struktur und keine Abhängigkeit unter aufeinander folgenden Zwischenkommunikationszeiten eines initiiierenden Teilnehmers noch gibt es solche Abhängigkeiten unter den Teilnehmern. Diese Unabhängigkeitsannahmen sind aus der Beobachtung näherungsweise zu rechtfertigen, wenn auch nicht in allen Szenarien korrekt (beispielsweise führen bestimmte Ereignisse – Naturkatastrophen, Sportveranstaltungen u. Ä. – zu einem koordinierten Verhalten der Nutzer und zu einem korrelierten Anschwellen der Kommunikationshäufigkeit; man spricht hier von einem *flash crowd effect*). Verteilungsannahmen sollten aus empirischen Beobachtungen gewonnen werden; eine typische Vereinfachung ist es, für diese Zwischenankunftszeiten eine Exponentialverteilung anzunehmen. Dies ist grob in Einklang mit empirischen Befunden und vereinfacht die analytische Behandlung erheblich.

4. Wie lange dauert ein Kommunikationsvorgang? Bestehen Abhängigkeiten zwischen diesen Längen?

Abhängigkeiten zwischen Längen wird man auch hier aus der Intuition heraus zunächst verneinen (tatsächlich ist dies insbesondere für präzisere Webmodelle nicht ganz korrekt) und damit auch ein einfaches Modell erhalten. Entscheidend sind in der Tat die Verteilungsfunktionen der Länge eines Kommunikationsvorgangs. In beiden Fällen ist man dabei auf die Empirie angewiesen.

Für ein Telefongespräch ist die Dauer, ähnlich wie die Zeit zwischen Telefonaten, in sehr guter Näherung durch eine Exponentialverteilung beschrieben (Abbildung 2 veranschaulicht diese Verteilungsfunktion). Damit liegt beim Telefonsystem insgesamt eine einfache stochastische Struktur vor, wenn man

---

scher und axiomatischer Aufbau des Wahrscheinlichkeitsbegriffes) und die Statistik (die Beschreibung von Daten und das Ziehen von Schlüssen aus Daten). [Anm. d. Hg.]

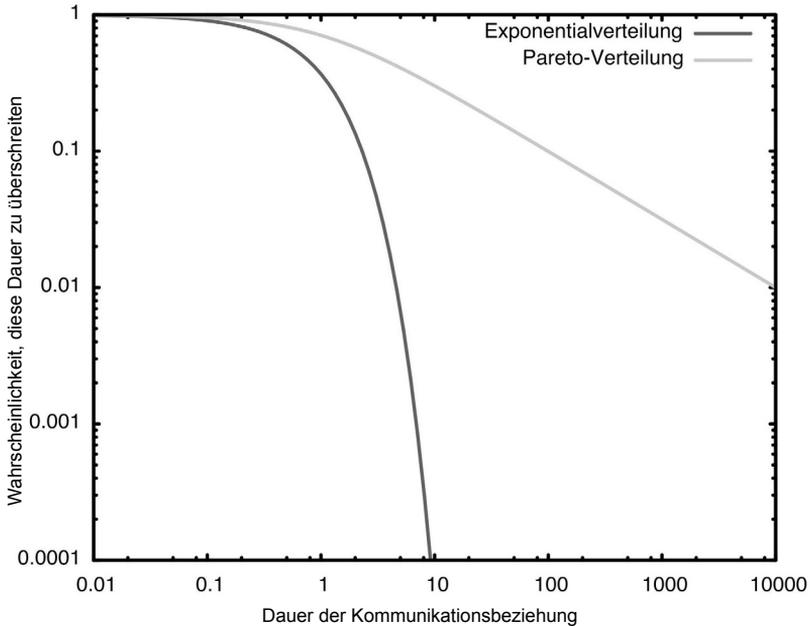
insbesondere eine Verteilerstelle in der Mitte des Netzes betrachtet. Von einer großen, näherungsweise als unendlich groß auffassbaren Menge von Teilnehmern treffen „Dienstanfragen“ (= die Anfrage nach Herstellung einer Verbindung) ein, die jeweils für sich exponential verteilt lange dauern und bei denen zwischen den einzelnen Anfragen eine ebenfalls exponential verteilte lange Zeit verstreicht. Alle relevanten Zufallsgrößen sind in diesem Modell unabhängig. Es handelt sich damit um das typische Beispiel eines sogenannten *Poisson-Prozesses*<sup>2</sup> als dem wohl einfachsten Fall eines nichttrivialen stochastischen Prozesses.

Bei Webbrowsering ist die Lage anders. Im Wesentlichen hängt die Dauer einer Übertragung von der Größe der angefragten Webseite ab; zusätzlich bestimmt die momentane Auslastung des Webservers diese Dauer. Die Statistik der Seitengröße ist dabei vergleichsweise einfach zu bestimmen: Man nehme einen Webserver und zähle alle darauf gespeicherten Seiten samt ihrer Größen auf. Unterstellt man gleich verteilten Zugriff auf die einzelnen Seiten (auch dies ist nicht ganz korrekt, typischerweise sind Popularitätsverteilung z. B. nach *Zipf's Law*<sup>3</sup> anzutreffen), so erhält man unmittelbar die empirische Verteilungsfunktion der einzelnen Seiten. Es stellt sich nun bei typischen Webservern heraus, dass diese Verteilungsfunktion der Größe – und damit im Wesentlichen auch die Länge der Kommunikation – bei Webverkehr eben nicht der Exponentialverteilung oder einer ähnlichen Verteilung gehorcht, sondern im Gegenteil einer sogenannten *Heavy-tail-Verteilung*<sup>4</sup>, oft einer *Pareto-Verteilung* (Abbildung 2). Diese Verteilungen zeichnen sich durch ein ganz anderes Verhalten der Restverteilung aus: Während bei der Exponentialverteilung die Wahrscheinlichkeit, einen Wert „x“ zu überschreiten, exponentiell schnell in „x“ gegen 0 geht (daher der Name), geschieht dies bei *Heavy-tail-Verteilungen* nur polynomiell schnell – sehr große Werte kommen also mit einer deutlich größeren Wahrscheinlichkeit vor als dies bei einer Exponentialverteilung bei gleichem Mittelwert zu erwarten wäre.

<sup>2</sup> Der *Poisson-Prozess* ist ein stochastischer Prozess, der die Wahrscheinlichkeit beschreibt, in einem bestimmten Zeitintervall eine bestimmte Anzahl von Zufallsereignissen zu beobachten. Dieser Prozess formalisiert die Anschauung, dass Ereignisse durch eine sehr große Anzahl unabhängiger Quellen entstehen. Der *Poisson-Prozess* ist ein diskreter Prozess in stetiger (d. h. kontinuierlicher) Zeit. [Anm. d. Hg.]

<sup>3</sup> Das Zipfsche Gesetz ist ein Modell, mit dessen Hilfe man bei bestimmten Größen, die in eine Rangfolge gebracht werden, deren Wert aus ihrem Rang abschätzen kann. Häufige Verwendung findet das Gesetz in der Linguistik, speziell in der Korpuslinguistik und Quantitativen Linguistik, wo es zum Beispiel die Häufigkeit von Wörtern in einem Text zur Rangfolge in Beziehung setzt. [Anm. d. Hg.]

<sup>4</sup> In der Wahrscheinlichkeitstheorie ist eine *Heavy-tail-Verteilung* bzw. endlastige Verteilung eine Wahrscheinlichkeitsverteilung mit einer unendlichen Varianz. Anschaulich besagt der Begriff, dass auch sehr große Werte mit nicht vernachlässigbarer Wahrscheinlichkeit auftreten. [Anm. d. Hg.]



2 – Exponential- und Pareto-Verteilung als komplementäre Verteilungsfunktion in doppelt-logarithmischer Darstellung

Für den Vergleich von Telefon- und Webverkehr ist damit der entscheidende Unterschied gefunden: Es ist sehr viel wahrscheinlicher, dass eine Webübertragung eine (im Vergleich gesehen) sehr viel längere Zeit braucht als dies bei einem Telefongespräch der Fall wäre. Dieser Effekt liegt letztlich im Verhalten von Menschen beim Telefonieren oder beim Erstellen von Webseiten begründet; es ist dabei keine wesentliche Intention oder bewusstes, zielgesteuertes Verhalten erkennbar.

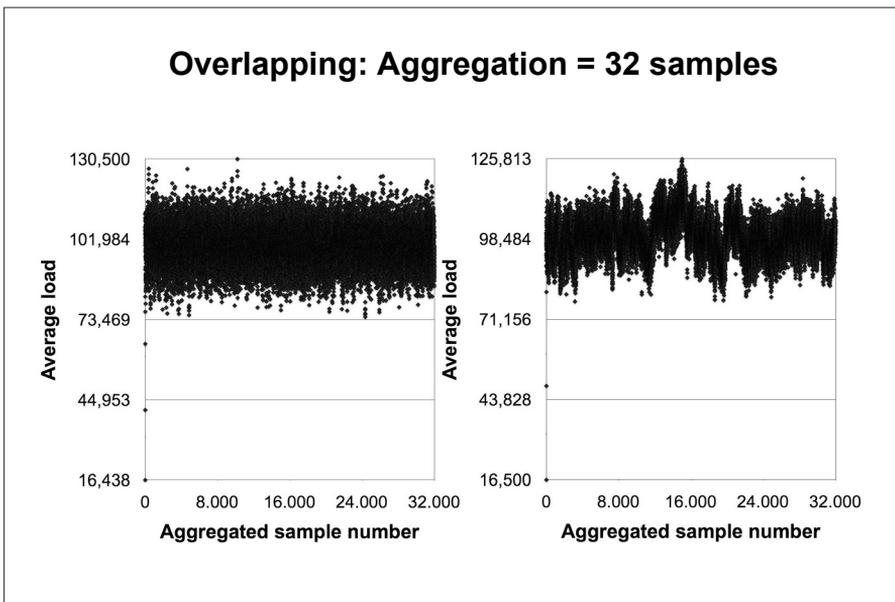
Aber führt dieses unterschiedliche Verhalten auch zu nennenswert unterschiedlichen Effekten, ergibt sich daraus eine wie auch immer geartete Struktur? Der folgende Abschnitt gibt darauf eine Antwort.

### Struktur durch Überlagerung

Der vorherige Abschnitt hat sich auf das Verhalten eines einzelnen Teilnehmers beschränkt – wie lange wartet dieser zwischen Kommunikationsvorgängen, wie lange dauert ein einzelner solcher Vorgang? Wechselt man nun die Perspektive und nimmt den Standpunkt eines Webservers oder eines Vermittlungsknotens in einem Telefonnetz oder Internet ein, so ist individuelles Verhalten mehr oder minder ohne Belang. Relevant ist das Verhalten vieler Teilnehmer gleichzeitig, und damit die Frage, wie viele Teilnehmer zeitgleich ak-

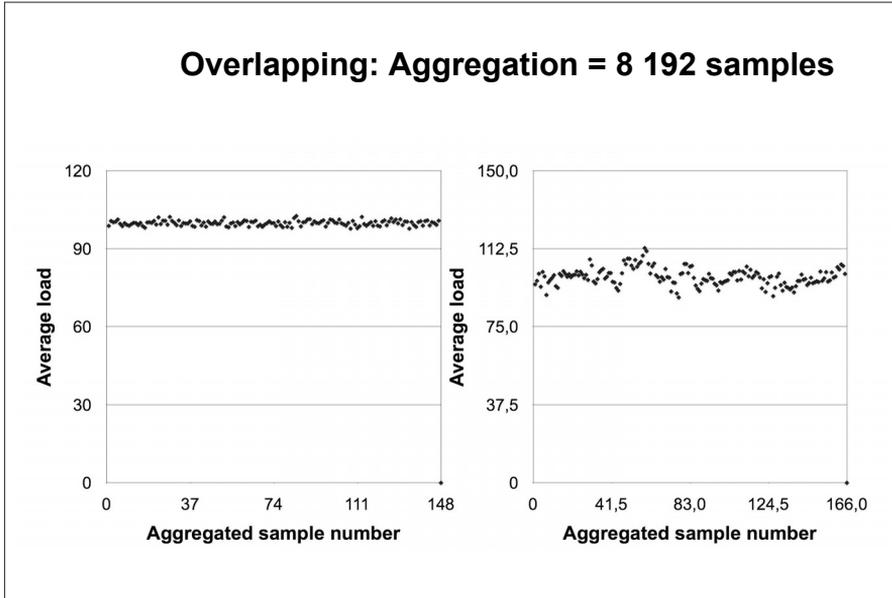
tiv sind. Dies entspricht der Anzahl gleichzeitig zu behandelnder Telefongespräche (und ist damit für Telefongesellschaften zur Auslegung der Netze relevant) oder entsprechend der Anzahl gleichzeitig stattfindender Webseiten-Übertragungen.

Zur Illustration zeigt Abbildung 3 einen solchen Auslastungsverlauf [*average load*] sowohl für das Telefonnetz als auch für das Internet. Die entsprechende Anzahl von Telefonaten bzw. Webseiten-Übertragungen entstammen dem oben beschriebenen Modell. Dabei wurde für diese Abbildung leicht über die Zeit aggregiert (32 Zeitpunkte wurden durch Mittelwertbildung zusammengefasst). Offenbar sind beide Systeme ähnlich parametrisiert, die mittlere Last liegt bei ca. 100; auf dieser Zeitskala ist bei beiden Systemen noch deutliche zufällige Streuung zu sehen.



3 – Aggregierter Verkehr (Telefonnetz links, Internet rechts);  
über 32 Zeiteinheiten geglättet

Welches Verhalten sollte sich nun bei einer stärkeren Mittelung über einen größeren Zeithorizont einstellen? Mitteln entspricht einer Glättung; Schwankungen werden entfernt, indem über größere Zeithorizonte beobachtet wird. In der Tat geschieht genau dies auch im Fall des Telefons; der linke Teil von Abbildung 4 zeigt dies als Beispiel – hier wurden über 256-mal längere Zeitintervalle geglättet. Der Internet-Fall verhält sich dagegen ganz anders (Abbildung 4 rechts): Selbst bei starker Glättung bleibt hier deutlich Struktur sichtbar, die über lange Zeiträume hinweg anhält und durch Glättung nicht (bzw. nur sehr langsam) verschwindet.



4 – Aggregierter Verkehr (Telefonnetz links, Internet rechts);  
über 8192 Zeiteinheiten geglättet

### Interpretation und Zusammenfassung

Damit ist das beabsichtigte Beispiel vollständig. Durch eine augenscheinlich minimale Veränderung eines Verhaltens bzw. einer Modellannahme ändern sich die Charakteristiken eines Gesamtsystems fundamental. Die Konsequenzen hieraus sind vielfältig: Sowohl für die Benutzer des World Wide Web, die sich mit einem unerwartetem Verhalten und mit stark schwankenden Wartezeiten konfrontiert sehen, die es in dieser Form beim Telefon natürlich nicht gibt, als auch für die Anbieter von Webservern und Telekommunikationssystemen, die Schwierigkeiten haben ihre Systeme korrekt auszulegen, weil das nichtexponentielle Modell deutlich schwerer zu analysieren und zudem kostenaufwendiger ist, weil die starke Schwankung trotz Glättung größere Puffer und eine höher ausgelegte Maximalkapazität verlangt.

Während diese Aspekte in den letzten Jahren fast durchweg geklärt werden konnten, ist der eigentliche Grund bisher noch weitgehend unklar: Warum unterscheiden sich die Statistiken von Webseiten in so erheblichem Maße von denen von Telefongesprächen? Diese Frage entzieht sich letztlich der technischen Beantwortung und ist nur durch Überlegungen zu Nutzungsmustern und Kommunikationsformen zu klären. Der genannte Effekt ist insofern Beispiel für einen *Automatismus*, der in technischen Parametern beschreibbar ist, des-

sen Zentrum aber außerhalb der Technik, in den Bottom-up-Aktivitäten der Nutzer, liegt. Hier könnte eine interdisziplinäre Kooperation zwischen der Informatik und den Kulturwissenschaften einen wertvollen Beitrag leisten.