University of Zurich
Department of Informatics

Burkhard Stiller
Thomas Bocek
Fabio Hecht
Cristian Morariu
Peter Racz
Andrei Vancea
Martin Waldburger
(Eds.)

Communication Systems III

TECHNICAL REPORT – No. ifi-2009.03

June 2009

University of Zurich
Department of Informatics (IFI)
Binzmühlestrasse 14, CH-8050 Zürich, Switzerland

ifi

B. Stiller et al. (Eds.): Communication Systems III
Technical Report No. ifi-2009.03, June 2009
Communication Systems Group (CSG)
Department of Informatics (IFI)
University of Zurich
Binzmühlestrasse 14, CH-8050 Zurich, Switzerland
URL: http://www.csg.uzh.ch/

Introduction

The Department of Informatics (IFI) of the University of Zürich, Switzerland works on research and teaching in the area of computer networks and communication systems. Communication systems include a wide range of topics and drive many research and development activities. Therefore, during the spring term FS 2009 a new instance of the Communication Systems seminar has been prepared and students as well as supervisors worked on this topic.

The areas of communication systems include among others wired and wireless network technologies, various network protocols, network management, Quality-of-Service (QoS) provisioning, mobility, security aspects, peer-to-peer systems, multimedia communication, and manifold applications, determining important parts of future networks. Therefore, this year's seminar addressed such areas in more depth. The understanding and clear identification of problems in technical and organizational terms have been prepared and challenges as well as weaknesses of existing approaches have been addressed. All talks in this seminar provide a systematic approach to judge dedicated pieces of systems or proposals and their suitability.

Content

This new edition of the seminar entitled "Communication Systems III" discusses a number of selected topics in the area of computer networks and communication systems. The first talk on "P2P Anreizmechanismen" addresses cooperation strategies and incentive mechanisms in Peer-to-Peer networks. Talk two on "Overview and Trends of Video Streaming over IP" gives an overview of different techniques for video streaming over IP, presenting client-server, multicast, and Peer-to-Peer approaches. Talk three on "P2P Collaboration in Support for Spam Detection" discusses different Peer-to-Peer-based collaborative anti-spam systems. Talk four on "Honeypots" presents various types of honeypots in security architectures and discusses their strengths and weaknesses. "Inter-domain Routing" as talk five addresses routing between autonomous systems in the Internet and presents the BGP routing protocol. Finally, talk six on "Information and Communication Technology for the Needy" gives an overview and comparison on the deployment and usage of information and communication technology in different countries.

Seminar Operation

Based on well-developed experiences of former seminars, held in different academic environments, all interested students worked on an initially offered set of papers and book chapters. Those relate to the topic titles as presented in the Table of Content below. They prepared a written essay as a clearly focused presentation, an evaluation, and a summary of those topics. Each of these essays is included in this technical report as a separate section and allows for an overview on important areas of concern, technology architectures and functionality, sometimes business models in operation, and problems encountered.

In addition, every group of students prepared a slide presentation of approximately 45 minutes to present its findings and summaries to the audience of students attending the seminar and other interested students, research assistants, and professors. Following a general question and answer phase, a student-lead discussion debated open issues and critical statements with the audience.

Local IFI support for preparing talks, reports, and their preparation by students had been granted by Thomas Bocek, Fabio Hecht, Cristian Morariu, Peter Racz, Andrei Vancea, Martin Waldburger, and Burkhard Stiller. In particular, many thanks are addressed to Peter Racz for his strong commitment on getting this technical report ready and quickly published. A larger number of pre-presentation discussions have provided valuable insights in the emerging and moving field of communication systems, both for all groups of students and supervisors. Many thanks to all people contributing to the success of this event, which has happened in a lively group of highly motivated and technically qualified students and people.

Zürich, June 2009

Contents

Kapitel 1

P2P Anreizmechanismen

Konstantin Benz

Diese Arbeit befasst mit Kooperation in P2P-Netzwerken und Anreizmechanismen. Es handelt sich dabei um Verfahren, welche kooperatives Verhalten zwischen Peers fördern. Anhand des in BitTorrent verwendeten Mechanismen wird aufgezeigt, weshalb Anreizmechanismen nicht nur die Kooperation fördern, sondern auch entscheidend dazu beitragen, dass P2P-Netze effizienter genutzt werden. Zu diesem Zweck werden die verschiedenen Anreizmechanismen klassifiziert sowie deren Vor- und Nachteile erläutert. Anschliessend werden vier Ansätze vorgestellt, die aufzeigen sollen, wie man einen Anreizmechanismus implementieren kann. Dies soll letztlich als Basis dafür dienen, Verbesserungsmöglichkeiten für den in BitTorrent verwendeten Anreizmechanismus aufzuzeigen.

Inhaltsverzeichnis

1.1 Einleitung

In den folgenden Abschnitten wird beschrieben, weshalb sich Entwickler von P2P-Netzwerken mit Anreizmechanismen auseinandersetzen. Zudem wird eine kurze Übersicht über Aufbau und Struktur der Arbeit gegeben.

1.1.1 Motivation

In dieser Arbeit geht es um ein Problem, welches in P2P-Netzwerken auftritt: In einem P2P-Netzwerk beanspruchen Peers Ressourcen von jeweils anderen Peers. Die Anzahl jener Peers, die Ressourcen zur Verfügung stellen, ist jedoch häufig deutlich kleiner als die Anzahl jener Peers, die Ressourcen für sich beanspruchen [1]. Dadurch kommt es zu einer einseitigen Belastung jener Peers, die Ressourcen zur Verfügung stellen [2]. Dies wiederum kann zu Problemen wie bspw. Übernutzung des P2P-Netzwerks und mangelnde Bereitschaft zur Bereitstellung von Ressourcen führen. Anreizmechanismen wirken diesen Problemen entgegen. Es geht dabei um Mechanismen, die die kooperative Nutzung von P2P Netzwerken fördern. In dieser Arbeit werden einige der gebräuchlichsten Anreizmechanismen vorgestellt und klassifiziert. Anhand eines Praxisbeispiels (BitTorrent) soll gezeigt werden, dass Anreizmechanismen für P2P-Netzwerke relevant sind. Anschliessend soll gezeigt werden, wie diese Anreizmechanismen in verschiedenen Ansätzen zur Anwendung kommen. Aufgrund der Vor- und Nachteile dieser Ansätze sollen Möglichkeiten aufgezeigt werden, wie diese Ansätze eine reale P2P-Applikation wie bspw. BitTorrent verbessern könnten.

1.1.2 Umfang der Arbeit

In dieser Arbeit wird anhand des in BitTorrent verwendeten Mechanismus aufgezeigt, warum Anreizmechanismen in P2P-Netzwerken benötigt werden und weshalb sich die Forschung mit diesem Thema beschäftigt. Danach werden Anreizmechanismen sowie Ansätze, welche diese Anreizmechanismen anwenden, kategorisiert. Als Grundlage dieser Kategorisierung dient eine Forschungsarbeit von Philipp Obreitner und Jens Nimis, in welcher die Anreizmechanismen als “Incentive Patterns” bezeichnet werden [4]. Die Kategorisierung gemäss Obreitner et al. soll dazu dienen, Verbesserungsmöglichkeiten hinsichtlich der Förderung kooperativen Verhaltens in P2P-Netzwerken aufzuzeigen.

1.2 Kooperatives und unkooperatives Verhalten

Ein bekanntes Beispiel für P2P-Netzwerke ist das Filesharing-Netzwerk von BitTorrent. BitTorrent kennt tatsächlich einen Mechanismus, der kooperatives Verhalten unter den Peers fördert. Dieser Mechanismus wird nun im folgenden Abschnitt erklärt, um zu zeigen, wie sich Anreizmechanismen auf P2P-Netzwerke auswirken. Anschliessend werden verschiedene Kategorien von Anreizmechanismen aufgezeigt, um Verbesserungsvorschläge für BitTorrent aufzuzeigen.

1.2.1 BitTorrent: ein Praxisbeispiel

Um den Grund zu verstehen, weshalb in P2P-Netzwerken so genannte "Anreizmechanismen" zur Anwendung kommen, sei im folgenden ein Beispiel aus der Praxis genannt. Ein typisches Praxis-Beispiel für ein Incentive Mechanismus ist der in BitTorrent verwendete Upload-/Download-Algorithmus [3]. In einem Client-/Server-Netzwerk stellt derjenige Nutzer, der eine Datei zum Download anbietet, seine gesamte Bandbreite für den Download zur Verfügung. Dies kann für die Stabilität des Netzwerks sehr problematisch werden, wenn zahlreiche Benutzer ein- und dasselbe File gleichzeitig herunterladen wollen. Die gesamte Last des Uploads liegt bei demjenigen Nutzer, der in diesem Moment die Server-Rolle übernimmt. Dabei kann es zu Überlastung dieses Nutzers kommen. Um diesem Problem entgegenzutreten haben die Entwickler von BitTorrent einen Mechanismus eingeführt, der sich von dem in Client-/Server-Architekturen verwendeten Ansatz wesentlich unterscheidet. Im P2P-Netzwerk BitTorrent wird die Upload-Last auf die downloadenden Peers verteilt. Dies wird ermöglicht, indem das angebotene File während des Downloads in kleine Pakete aufgeteilt wird. Wenn mehrere Peers gleichzeitig dasselbe File herunterladen, erhalten die verschiedene Downloader jeweils verschiedene Teile der Datei, wie in Abbildung 1.1 dargestellt. Nach dem ersten Downloadschritt stellt jeder Downloader seine bereits heruntergeladenen Pakete den anderen Downloadern zur Verfügung. Abbildung 1.2 zeigt, wie die Downloader den Download fehlender Pakete bei jeweils anderen Downloadern abfragen. Die Downloader laden in der Folge weitere Pakete nicht mehr vom ursprünglichen Anbieter des Files herunter. Stattdessen greifen sie auf die bereits heruntergeladenen Pakete der jeweils anderen Downloader zu, wie in Abbildung 1.3 gezeigt. Dadurch entsteht unter den Downloadern ein Austausch der heruntergeladenen Pakete. Der ursprüngliche Anbieter des Files wird somit entlastet. Dies geht aus Abbildung 1.4 hervor. Der BitTorrent-Algorithmus verhindert punktweise Überlastungen des P2P-Netzwerks und sorgt für eine effizientere Nutzung der vorhandenen Bandbreite innerhalb des P2P-Netzwerks. Der BitTorrent-Algorithmus inspirierte viele Forscher dazu, sich näher mit Ansätzen zu befassen, die die Effizienz von P2P-Netzwerken steigern.

1.2.2 Unkooperatives Verhalten in P2P-Netzwerken

Eine Vielzahl von Nutzern kann für die Verwaltung von P2P-Netzwerken eine grosse Herausforderung darstellen. P2P-Nutzer unterscheiden sich in der Art des unkooperativen Verhaltens. Einige Peers sind eher "selfish": sie versuchen, die vorhandenen Ressourcen anderer Peers übermässig zu nutzen; andere Peers verhalten sich "lavish": sie weigern sich, ihre Ressourcen mit anderen Peers zu teilen [4]. Zusätzlich gibt es auch in P2P-Netzwerken Peers, welche bewusst anderen Peers schaden wollen, indem sie Schadprogramme, Viren etc. verbreiten. Der Umgang mit den verschiedenen Arten des unkooperativen Verhaltens ist eine der Hauptaufgaben von Anreizmechanismen. Es geht darum, von den Peers Verhaltensweisen zu erzwingen, die zu einer effizienteren und besseren Nutzung des Gesamtnetzwerks führen. Einige Charakteristika von Anreizmechanismen sind daher [4]:

- Zuweisung von Rollen: Anreizmechanismen weisen den Peers verschiedene Rollen zu, die diese während einer Transaktion einnehmen. Eine typische Zuweisung ist die

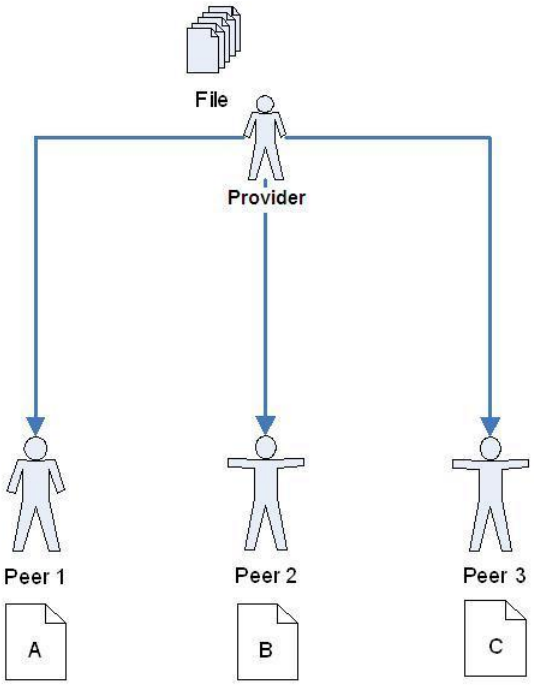

Abbildung 1.1: Der BitTorrent-Algorithmus startet wie in einer Client/Server-Architektur: das File wird in die Teile A, B und C aufgeteilt. Der Anbieter des Files (Provider) versendet zunächst je ein Teil der Datei an alle downloadenden Peers.

Producer/Consumer-Beziehung, bei der das eine Peer als Anbieter einer Dienstleistung auftritt, während das andere Peer diese Dienstleistung konsumiert. Anreizmechanismen definieren dabei, welche Rolle welche Aktionen während einer Transaktion ausführt.

- Belohnungsmechanismen: Anreizmechanismen nutzen Belohnungsmechanismen, um unter bei Transaktionen zwischen den Peers das von den P2P-Netzwerkbetreibern gewünschte Verhalten zu fördern. Die Belohnung kann auf verschiedene Arten erfolgen, bspw. kann ein Peer bei kooperativem Verhalten eine höhere Reputation erlangen oder zusätzliche Ressourcen beanspruchen.

- Sanktionierung von unkooperativem Verhalten: typischerweise bietet ein Incentive Mechanismus Massnahmen an, die unerwünschtes Verhalten bestrafen. Bspw. können bestimmte Dienstleistungen oder Ressourcen für bestimmte Peers blockiert werden.

- Vertrauensbasis: Anreizmechanismen sorgen dafür, dass Peers einander vertrauen können. Die Schaffung einer Vertrauensbasis ist ein zentrales Element für die Kooperation von Peers.

- Anonymität: P2P-Nutzer sind können daran interessiert sein, aus rechtlichen Gründen anonym zu bleiben. Deshalb ist Anonymität ebenfalls ein wichtiger Punkt, der in Anreizmechanismen berücksichtigt werden muss.

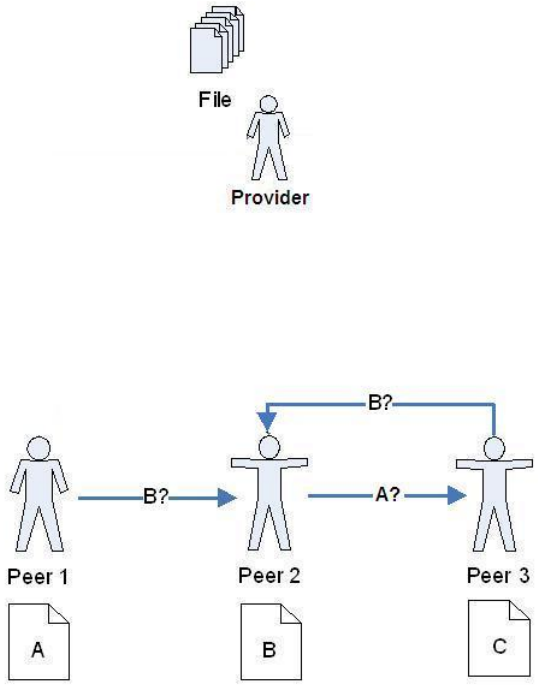

Abbildung 1.2: Die weiteren Teile des Files erhalten die Downloader nicht mehr vom Provider. Stattdessen fragen sie bei den anderen Downloadern nach, wer welche Partition des Files besitzt.

- Skalierbarkeit: Da Anreizmechanismen dafür konzipiert werden, auch in grossen P2P-Netzwerken angewendet zu werden, muss ein Incentive Mechanismus-Algorithmus auch für grosse Netzwerke skalierbar sein. Hohe Komplexität wird deshalb bei Incentive Mechanismus-Algorithmen in der Regel vermieden.
- Implementierungsprobleme: Anreizmechanismen können in der Implementierung verschiedene Probleme hervorrufen, auf die im folgenden ebenfalls eingegangen wird.

1.3 Anreizmechanismen

Bei Anreizmechanismen handelt es sich um Algorithmen, die aufgrund von Erkenntnissen aus den Wirtschafts- und Sozialwissenschaften konzipiert wurden. Ein wesentliches Merkmal von Anreizmechanismen ist die Steuerung oder Förderung eines bestimmten individuellen Verhaltens der P2P-Nutzer, so dass sich für das Gesamtnetzwerk ein absehbarer Nutzen ergibt. Die verschiedenen Anreizmechanismen lassen sich in zwei grössere Gruppen einteilen: einerseits gibt es Anreizmechanismen, welche versuchen, durch Authentifizierungs- oder Reputationsmechanismen das gegenseitige Vertrauen (“Trust”) zwischen Peers herzustellen. Das Ziel ist es, missbräuchliches Verhalten durch eine kollektive Vertrauensbasis zwischen den Peers zu verhindern oder zu sanktionieren. Diese Gruppe der Anreizmechanismen nennt man “trust-based”. Die zweite Gruppe der Anreizmechanismen orientiert sich an Modellen aus den Wirtschaftswissenschaften. Die Idee ist es, Interaktionen zwischen Peers wie ein Handels- oder Tauschgeschäft abzuwickeln.

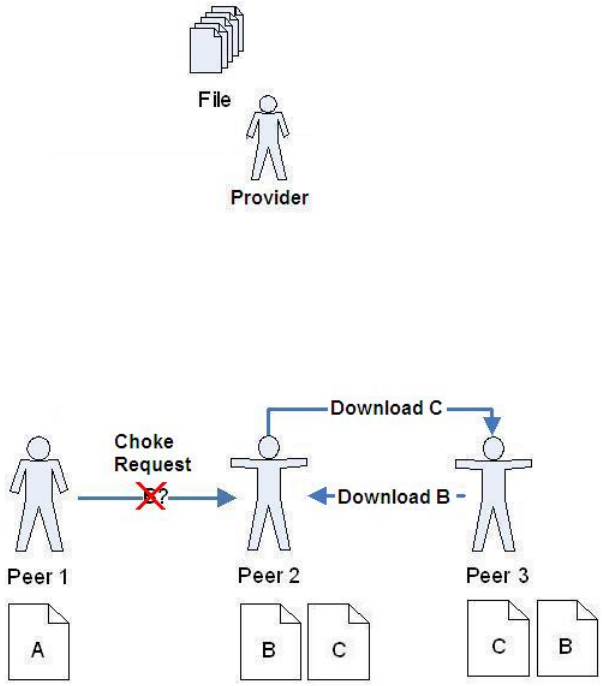

Abbildung 1.3: Jedes Peer kann nun auf Anfrage die fehlenden Teile des Files herunterladen, wenn sich diese beim entsprechenden Peer befinden. Um einer möglichen Überlastung einzelner Peers durch zahlreiche gleichzeitige Anfragen zu entgehen, bietet BitTorrent die Möglichkeit, dass einzelne Peers Anfragen blockieren können.

Mittels eines Preismodells soll kooperatives Verhalten belohnt und unkooperatives Verhalten bestraft werden. Die zweite Gruppe der Anreizmechanismen nennt man "trade-based". Trust-based und trade-based Mechanismen lassen sich in weitere Klassen (siehe Tabelle 1.1) unterteilen. In den folgenden Abschnitten werden die einzelnen Klassen der Anreizmechanismen im Detail erläutert sowie deren Vor- und Nachteile aufgezählt.

Tabelle 1.1: Klassifizierung der Anreizmechanismen

Trust-based Mechanismen	Trade-based Mechanismen
Collective Mechanismen Community Mechanismen Bearer Notes Mechanismen Bearer Bills Mechanismen Banking Mechanismen Banknotes Mechanismen	Barter Trade Mechanismen

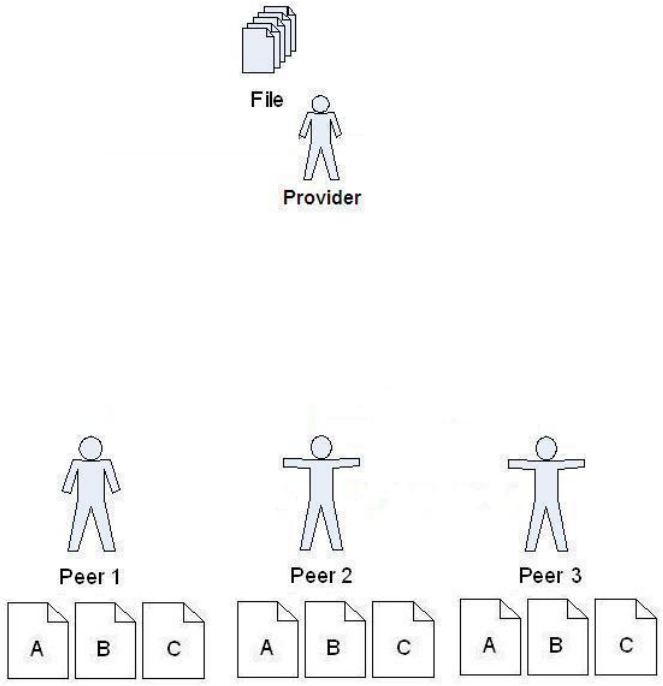

Abbildung 1.4: Die Anfragen (und anschliessende Downloads) werden solange fortgesetzt, bis jedes Peer alle Teile des Files besitzt. Der Provider wird somit entlastet. Nach dem initialen Download muss er weitere Teile des Files nur dann bereitstellen, wenn ein bestimmtes Teil des Files verloren gehen sollte.

1.3.1 Trust-based Anreizmechanismen

Ein wichtiger Grund für kooperatives Verhalten zwischen den Peers ist die Bildung von gegenseitigem Vertrauen. Die Trust-based Anreizmechanismen versuchen, mittels Reputation oder mittels Zugehörigkeit zu einer bestimmten Gruppe das gegenseitige Vertrauen in P2P-Netzwerken sicherzustellen.

1.3.1.1 Collective Mechanismen

Collective Mechanismen basieren auf der Idee, dass sich P2P-Netzwerke als eine Ansammlung von Gruppen (so genannte "Kollektive") betrachten lassen, wobei die Zugehörigkeit zu einem Kollektiv als gemeinsame Vertrauensbasis für die Mitglieder des Kollektivs dient. Peers, die Mitglied in ein- und demselben Kollektiv sind, können Ressourcen uneingeschränkt untereinander austauschen. Ein Austausch von Ressourcen zwischen Peers, welche nicht demselben Kollektiv angehören, ist vollständig untersagt. Dadurch wird verhindert, dass Peers mit böswilligen Absichten sich die Ressourcen anderer Peers aneignen können. Will ein Peer eine Transaktion mit einem anderen Peer starten, muss es sich zunächst gegenüber dem anderen Peer authentifizieren und dadurch nachweisen, dass es Mitglied in demselben Kollektiv ist. Jegliche Weigerung, sich zu authentifizieren, wird in einem Collective Mechanismus als unkooperatives Verhalten gedeutet. Der Vorteil dieses

Verfahrens besteht darin, dass es in einem Collective Mechanismus unkooperatives Verhalten immer aufgedeckt wird: ein Peer, dass die Ressourcen eines anderen Peers nutzen will, muss sich auf jeden Fall authentifizieren. Unkooperatives Verhalten kann jederzeit zurückverfolgt werden. Für böswillige P2P-Nutzer stellt dies eine Hemmschwelle dar, welche sie bspw. daran hindert, schadhaften Programmcode zu verbreiten. Ein Nachteil dieses Verfahrens besteht darin, dass die Anonymität im P2P-Netzwerk aufgegeben wird. Dies wirkt sich insofern nachteilig aus, weil P2P-Nutzer ihre Daten aus juristischen Gründen lieber anonym austauschen wollen. Die Authentifzierung und Preisgabe der eigenen Identität kann auch für legitime P2P-Nutzer eine Hemmschwelle darstellen.

1.3.1.2 Community Mechanismen

In Community Mechanismen werden die Peers in "Communities" aufgeteilt. Die Grundidee ist es, den Peers innerhalb ihrer Community eine bestimmte lokale Reputation zukommen zu lassen, die zugleich als Belohnung für Peers dient, welche sich kooperativ verhalten. Peers, die Ressourcen zur Verfügung stellen, werden durch eine Erhöhung des Reputationswerts belohnt. Optional können Peers, welche Ressourcen beanspruchen, durch Peers, welche die Ressourcen anbieten, durch eine Verringerung ihrer Reputation bestraft werden. Im Gegensatz zu Collective Mechanismen ist in Community Mechanismen gruppenübergreifende Kooperation möglich. Die lokale Reputation gilt nur für Peers, die bereits zuvor kooperiert haben. Dies soll davor schützen, dass gewisse Peers diffamiert oder von ihrer Community ungerechtfertigt belohnt werden. Gute Reputation zahlt sich nur innerhalb einer stabilen lokalen Community aus. Der Vorteil von Community Mechanismen liegt darin, dass die Kooperationsbeziehungen zwischen den Peers flexibler sind als bei Collective Mechanismen. Dieser Vorteil wird allerdings durch die Tatsache relativiert, dass Peers auf ungerechtfertigte Weise von ihrer Community benachteiligt oder belohnt werden können. Ein gewichtiger Nachteil von Community Mechanismen ist auch die mangelnde Skalierbarkeit des Mechanismen [5]. Da Reputation nur dann eine Rolle spielt, wenn Peers zuvor zusammengearbeitet haben, müssen innerhalb der Community zahlreiche Kooperationen stattfinden. Dies führt dazu, dass Communities im Vergleich zu den daran teilnehmenden Entitäten sehr gross sein müssen.

1.3.2 Trade-based Mechanismen

Die Trade-based Mechanismen bauen auf der Logik von Angebot und Nachfrage auf. Jede Nutzung von Ressourcen wird dabei als "Handel" aufgefasst. Trade-based Mechanismen fördern die Kooperation zwischen Peers, indem Peers, welche Ressourcen freigeben, durch entsprechende Gegenangebote anderer Peers belohnt werden.

1.3.2.1 Barter Trade Mechanismen

Im Gegensatz zu Trust-based Mechanismen basiert das Barter Trade Mechanismus auf der Idee, dass jede Transaktion zwischen Peers auf Basis eines Tauschhandels erfolgen soll. Jedes Peer, das eine Ressource konsumieren will, ist im Barter Trade Mechanismus zugleich

Anbieter einer Ressource. Beansprucht ein Peer eine Ressource eines anderen Peers, so muss sie gleichzeitig dem anderen Peer eine eigene Ressource zur Verfügung stellen. Die Gleichzeitigkeit dieses Tauschhandels verhindert, dass ein Peer ein anderes Peer durch Blockieren der eigenen Ressourcen benachteiligen kann. Jedes Angebot einer Ressource benötigt ein Gegenangebot. Der Vorteil dieses Mechanismen liegt darin, dass einseitige Nutzung von Ressourcen verunmöglicht wird. Die Gleichzeitigkeit des Tauschhandels ermöglicht es, dass die Peers anonym bleiben können. Es werden keine vertrauensbildende Massnahmen wie bspw. Authentifizierungsmechanismen benötigt. Der Barter Trade Algorithmus ist skalierbar, weil nur zwei Entitäten an diesem Algorithmus beteiligt sind. Der Nachteil des Barter Trade Mechanismus liegt darin, dass die Nutzung einer Ressource immer zeitgleich mit dem Produzieren einer Ressource zusammenfallen muss. Manche P2P-Nutzer können oder wollen aber nicht gleichzeitig Ressourcen anbieten und Ressourcen beanspruchen.

1.3.2.2 Bearer Notes Mechanismen

Wie bei allen Trade-based Mechanismen sind auch Bearer Notes Mechanismen ein Ansatz, um Transaktionen zwischen Peers als Tauschhandel von Ressourcen ablaufen zu lassen. Im Gegensatz zu Barter Trade Mechanismen muss bei einem Bearer Notes Mechanismus der Tauschhandel nicht gleichzeitig ablaufen. Dies geschieht mit Hilfe eines "Schuldscheins", der seinem Träger erlaubt, im Gegenzug für einmal zur Verfügung gestellte Ressourcen, Ressourcen bei einem Schuldner einzufordern. Wenn ein Peer eine Ressource von einem anderen Peer beansprucht, übergibt es diesem eine Bearer Note und wird dadurch zum Schuldner gegenüber dem produzierenden Peer. Das produzierende Peer kann dann zu einem späteren Zeitpunkt mit dieser Bearer Note eine Ressource vom konsumierenden Peer einfordern. Die Bearer Notes dienen somit als Belohnung für kooperatives Verhalten. Der Vorteil des Bearer Notes Mechanismen liegt darin, dass der Tauschhandel keine einseitige Nutzung von Ressourcen zulässt. Kooperative Peers werden belohnt. Im Gegensatz zu einem Community Mechanismus sind diese Peers im Besitz ihrer Belohnung. Der Nachteil der Bearer Notes Mechanismen ist, dass Bearer Notes Mechanismen darauf angewiesen sind, dass ein Produzent nach Erhalt einer Bearer Note auch die vereinbarte Ressource freigibt und kein Betrug stattfindet. Dies ist allerdings nur möglich, wenn die an einer Transaktion beteiligten Peers durch eine Drittinstanz authentifiziert werden. Die Anonymität ist in diesem Fall aber beeinträchtigt, wenn keine geeigneten kryptographischen Verschlüsselungsverfahren verwendet werden.

1.3.2.3 Bearer Bills Mechanismen

Bearer Bills Mechanismen sind eng verwandt mit den Bearer Notes Mechanismen. In Bearer Bills Mechanismen werden ebenfalls Schuldscheine für erbrachte Dienstleistungen ausgestellt. Der Unterschied zwischen einer Bearer Bill und einer Bearer Note besteht allerdings darin, dass der Schuldner einer Bearer Bill nicht der Konsument einer Ressource ist, sondern eine von den beteiligten Peers ausgewählte Drittinstanz. Die Vor- und Nachteile des Bearer Bills Mechanismen sind die gleichen wie beim Bearer Notes Mechanismus. Der einzige Unterschied zwischen Bearer Bills und Bearer Notes Mechanismen besteht darin,

dass für die Bearer Bills Mechanismen nicht nur das eine Peer sich dem anderen gegenüber authentifzieren muss: zusätzlich dazu muss sich noch ein drittes Peer authentifizieren, bei dem die Bearer Bill eingelöst werden kann. Bearer Bills Mechanismen und Bearer Notes Mechanismen verlangen daher komplexe Verschlüsselungsmechanismen, welche ihrerseits die Skalierbarkeit dieser Mechanismen beeinträchtigen können. Ein weiteres Risiko besteht darin, dass die Gläubiger im Bearer Bills Mechanismus wenige Möglichkeiten haben, die Vertrauenswürdigkeit des Schuldners zu prüfen (da es sich um eine Drittinstanz handelt). Umgekehrt haben die Schuldner nur eine geringe Motivation Ressourcen im Austausch für eine Bearer Bill abzuliefern, da ihnen die Identität des Bearer Bill-Besitzers unbekannt ist. Dies kann im Extremfall dazu führen, dass eine Bearer Bill faktisch nicht handelbar ist.

1.3.2.4 Banking Mechanismen

Da die gegenseitige Authentifizierung von Peers (und Drittinstanzen) sowohl beim Bearer Notes als auch beim Bearer Bills Mechanismus problematisch sein kann, gibt es ein Bedürfnis, erhaltene Schuldscheine bei einer einzigen im gesamten P2P-Netzwerk anerkannten Entität einzulösen. Dadurch entfällt die Suche nach vertrauenswürdigen Drittinstanzen, welche beim Bearer Bills Mechanismus immer wieder anfällt. Das Banking Mechanismus basiert auf der Idee, dass Kooperationen zwischen Peers mit einem "Scheck" belohnt werden, der bei einer "Bank" (also einer im P2P-Netz allgemein als Schuldner anerkannten Entität) eingelöst werden können. Im Bearer Bills Mechanismus trägt der Gläubiger das Risiko, dass sein Schuldner nicht vertrauenswürdig ist. Dieses Risiko entfällt im Banking Mechanismus, da nur Banken als Schuldner auftreten dürfen. Diese Banken kennen die Identität der Schuldner. Das Risiko, eine eingelöste Bearer Bill nicht loszuwerden, verringert sich. Der Vorteil des Banking Mechanismen liegt darin, dass die Anonymität, welche in Bearer Bills und Bearer Notes Mechanismen verloren geht, durch den Einsatz von Banken wiederhergestellt wird. Ein Peer kann die Ressourcen eines anderen Peers konsumieren und das produzierende Peer mit einer Banking Bill belohnen, ohne dass sich die Peers gegenseitig authentifizieren müssen. Der Tausch einer Banking Bill gegen eine Ressource erfolgt gleichzeitig. Ein Vorenthalten von Ressourcen ist nicht möglich. Soll nun die Banking Bill bei einer Bank im Austausch für eine Ressource eingelöst werden, kann auch dieser Tausch anonym erfolgen. Banken können gegenüber ihren Schuldnern anonym bleiben, da sie grundsätzlich als vertrauenswürdig anerkannt sind. Der Nachteil des Banking Mechanismen liegt darin, dass die Banken generell für alle Peers zugänglich sein müssen. Dies bedingt dass Peers, welche als Banken agieren, mit speziellen Mechanismen ausgestattet sein müssen, welche diese Peers als vertrauenswürdige Banken im gesamten Netzwerk auszeichnen. Die Implementierung eines Banking Mechanismen wird dadurch komplexer. Banking Mechanismen sind deshalb weniger gut skalierbar als bspw. Barter Trade Mechanismen.

1.3.2.5 Banknotes Mechanismen

Im Gegensatz zu Bearer Notes Mechanismen basieren Bank Notes Mechanismen auf der Idee, dass die Schuldscheine a priori bei einer zertifizierten Entität, einer Bank, eingelöst

werden können. Der Bearer Notes Mechanismus birgt den Nachteil, dass ein Peer nach dem Erhalt einer Bearer Note diese von seinem Schuldner zu einem späteren Zeitpunkt einfordern muss. Dabei ist es wichtig, dass er seinem Schuldner vertrauen kann. Dies wird im Banknotes Mechanismus dadurch gewährleistet, dass ein Peer mit einem allgemeingültigen Zahlungsmittel (also "Banknoten") den Konsum einer Ressource bezahlt. Die Bank Notes sind (im Gegensatz zum Banking Mechanismus) handelbar. Dies bedingt, dass ein Peer, das eine Dienstleistung bereitstellt, im Gegenzug nicht einen einzelnen Schuldschein erhält, sondern je nach angebotener Ressource mehrere Banknoten. Das Banknotes Mechanismus basiert also auf einem Preissystem. Der Vorteil des Banknotes Mechanismen liegt darin, dass Anonymität vollständig gewährleistet ist. Hinzu kommt, dass Banken nicht jederzeit im gesamten Netzwerk erreicht werden müssen wie beim Banking Mechanismus. Da die Banknoten handelbar sind, kann jedes Peer mittels Banknoten Ressourcen von einem anderen Peer einfordern, ohne dass die Bank als Hauptschuldner auftreten muss.

1.4 Ansätze

Die verschiedenen Anreizmechanismen sind nicht als Implementierungen, sondern viel eher als Implementierungsvorschriften zu verstehen. Um die verschiedenen Anreizmechanismen in realen P2P-Netzwerken zu implementieren, wurden verschiedene Ansätze gewählt, die einige Aspekte der zuvor genannten Anreizmechanismen umsetzen. Im Folgenden werden einige der wichtigsten Ansätze und deren Struktur vorgestellt.

1.4.1 Tit-for-Tat

Tit-for-Tat ist ein Ansatz, um den Barter Trade Mechanismus zu implementieren. Im Tit-for-Tat-Ansatz geht es darum, dass ein Peer eine Ressource immer als direktes Tauschmittel für eine andere Ressource erhält. Es wird dabei kein Zahlungsmittel gewählt. Die Ressourcen werden direkt gegeneinander ausgetauscht. Dieser Austausch von Ressourcen erfolgt gleichzeitig. In einem Filesharing-P2P-Netz könnte Tit-for-Tat z.B. so implementiert werden, dass Files, welche zum Download angeboten werden, zunächst in kleine Pakete aufgeteilt werden. Da typischerweise immer mehrere Peers ein File von einem anderen Peer herunterladen, erfolgt zunächst nur ein initialer Download verschiedener Pakete zu allen Downloadern. Nun beginnt der eigentliche Tauschhandel: die weiteren Pakete werden nun nicht mehr direkt von jenem Peer heruntergeladen, das das File anbietet. Stattdessen tauschen die Downloader ihre bereits heruntergeladene Pakete untereinander aus. Wenn bspw. ein Peer bereits Paket A erhalten hat und ein Peer Paket B besitzt, so können die beiden Peers nun ihre jeweiligen Pakete austauschen. Dieser Tausch erfolgt in Tit-for-Tat gleichzeitig, d. h. das erste Peer erhält vom zweiten das Paket B, während das zweite Peer vom ersten das Paket A erhält. Anschliessend können die Peers weitere Pakete direkt untereinander austauschen oder sich andere Peers suchen, welche über das jeweils benötigte Paket verfügen. Ein Vorteil besteht darin, dass Tit-for-Tat leicht implementierbar und auch skalierbar ist. Der Nachteil besteht darin, dass die Peers jeweils wissen müssen, über welche Pakete die jeweils anderen Peers verfügen. Ausserdem kann

es ein grosser Nachteil sein, wenn als Tauschpartner ausgerechnet ein Peer ausgewählt wurde, das nur über eine geringe Bandbreite verfügt. Ebenfalls problematisch ist, dass die Tauschpartner physisch weit voneinander getrennt sein können, da P2P-Applikationen mittels Overlay-Netzwerken arbeiten, die die Lokalität von Peers nicht berücksichtigen müssen. Folglich können bei Tit-for-Tat Verklemmungen vorkommen. Um Verklemmungen entgegenzuwirken muss der Tit-for-Tat-Ansatz zusätzlich mit einem Verfahren ausgestattet werden, welches unzuverlässige Peers blockieren kann [6]. In BitTorrent ist dies der Blockierungs-/Deblockierungs-Algorithmus, welcher es erlaubt Peers welche eine Verklemmung verursachen zu blockieren.

1.4.2 Give-to-Get

Give-to-Get ist ein Ansatz der auf der Reputation von Peers beruht. Im Gegensatz zu Tit-for-Tat erfolgt kein gleichzeitiger Tausch von Ressourcen. Stattdessen wird ein sogenannter Chocking/Unchocking-Mechanismus eingesetzt, welcher bestimmt, wer die Ressource eines Peers benutzen darf und wer nicht. Give-to-Get implementiert hauptsächlich Aspekte des Community Mechanismen. Typischerweise werden bei Give-to-Get zunächst alle Peers abgeblockt ("chocked"). Ein Peer kann nur Ressourcen von einem anderen Peer einfordern, wenn es unchocked ist. Will ein Peer eine Ressource einfordern, so muss es eine Anfrage für einen Download stellen. Dann wird das Peer deblockiert und kann nun Ressourcen anfragen. Um Überflutung von Anfragen zu vermeiden, gibt es beim Deblockieren eine Obergrenze. Die deblockierten Peers erhalten nun gemeinsam einzelne Filepakete, die sie einander gegenseitig forwarden. Anschliessend werden die Peers nach ihrer Reputation sortiert. Diejenigen Peers, die viele bereits erhaltene Ressourcen forwarden, erhalten eine höhere Reputation. Give-to-Get nützt dabei den Aspekt aus, dass Peers, welche viele Pakete weiterleiten, entweder über eine höhere Bandbreite verfügen oder als Tauschpartner zuverlässiger sind als andere. Peers mit einer hohen Reputation werden vom File-Anbieter bevorzugt: sie erhalten die Pakete des Files direkt vom File-Anbieter. Die Downloadzeit verkürzt sich damit für jene Peers, die viele Datenpakete schnell weiterleiten. Hat ein Peer die gesamte Datei heruntergeladen, wird es wieder blockiert, so dass die übriggebliebenen Downloader schneller bedient werden können. Der Vorteil von Give-to-Get ist, dass die Selektion der Tauschpartner weniger umständlich ist als bei Tit-for-Tat. Durch die Verteilung der Downloadlast auf besonders kooperative Peers wird die Effizienz von Downloads erhöht. Der Nachteil besteht darin, dass Peers mit geringer Bandbreite grundsätzlich benachteiligt werden, da geringes Forwarding als unkooperatives Verhalten interpretiert wird. Die Peers mit geringer Bandbreite können weniger Pakete direkt vom File-Anbieter beziehen und müssen deshalb länger auf den Abschluss eines Downloads warten [7].

1.4.3 Private Shared History

Der Private Shared History (PSH) liegt die Idee zugrunde, dass Peers, welche sich in der Vergangenheit kooperativ gezeigt haben, bei Downloads bevorzugt behandelt werden müssen, damit diese eine Belohnung für ihre Kooperation erhalten. PSH basiert auf zwei Komponenten: einer privaten History, welche einem Peer das Kooperationsverhalten von

anderen Peers zeigt, mit denen es früher zusammengearbeitet hat; und einer Shared History, welche allen Peers das allgemeine Kooperationsverhalten eines Peers anzeigt [8]. Die private History dient einem Peer dazu, abzuschätzen, welche Peers mit ihm in der Vergangenheit kooperiert haben und welche Peers sich unkooperativ verhielten. Wenn Peer B einem Peer A ein File zum Download zur Verfügung gestellt hat, so hat Peer B in Peer As privater History einen höheren Wert. Für Peer B wird es dadurch wahrscheinlicher, in der Zukunft von Peer A ein File als Gegenleistung zu beziehen. Die private History ist folglich mit einem Bearer Notes Mechanismus vergleichbar, bei dem A als Schuldner von B auftritt. Verhält sich A bei einer Ressourcenanfrage von B unkooperativ, so wird dies in der privaten History von B sichtbar. Es wird dadurch unwahrscheinlicher, dass A erneut von B eine Ressource beanspruchen kann. Der Vorteil der Private History liegt darin, dass Kooperation gefördert wird, während unkooperatives Verhalten jeglicher Art bestraft wird. Der Nachteil liegt darin, dass private Historys nur direkte Transaktionen zwischen zwei Peers A und B aufzeichnen. Peer A kann in der Private History nur feststellen, ob B vertrauenswürdig ist, wenn A zuvor mit B zusammengearbeitet hat. Möchte ein in der Private History von Peer A unbekanntes Peer C mit Peer A kooperieren, so kann Peer A nicht feststellen, ob C vertrauenswürdig ist, da A und C bislang nicht zusammengearbeitet haben. Die Shared History nimmt dieses Problem auf: sie zeichnet das kooperative Verhalten aller Peers im Netz auf. Kooperiert ein Peer, so wird sein Wert in der Shared History erhöht. Peer A kann somit Peer C in der Shared History aufsuchen und den Reputationswert von C in der Shared History in As Private History übertragen. Die Shared History hat somit Ähnlichkeit zu einer Bank in einem Banking Mechanismus, wobei ausstehende Schulden nicht bei der Bank, sondern bei anderen Peers eingefordert werden. Im PSH-Ansatz werden nun Private und Shared History kombiniert: wenn Peer B mit Peer A kooperiert, steigt Bs Wert in As Private History. Gleichzeitig steigt auch Bs Wert in der Shared History. Wenn A nun bei einer Ressourcenanfrage von B nicht kooperiert, so sinkt As Wert in Bs Private History sowie in der Shared History. Peer B wurde für sein kooperatives Verhalten belohnt, während Peer A für sein unkooperatives Verhalten nicht belohnt wird. Der Vorteil von PSH liegt darin, dass es sehr schwierig ist, unkooperatives Verhalten anzuwenden, ohne dafür sanktioniert zu werden. Die Shared History hält unkooperatives Verhalten jederzeit fest. Das Problem bei PSH ist, dass die Shared History für alle Peers jederzeit zugänglich sein muss. Die Shared History muss über das gesamte Netzwerk propagiert werden. Dies erfordert einerseits zusätzliche Rechenleistung und andererseits auch kryptographische Verfahren, um die Shared History vor Verfälschung abzusichern. Zusätzlich beinhaltet die Shared History die Problematik, dass sie als zentrales Element in einem P2P-Netz auftreten kann. Die Propagierung der Shared History ist folglich ein systemkritisches Element.

1.4.4 BarterCast

BarterCast verfolgt denselben Grundgedanken wie PSH: die Peers sollen mittels einer Erhöhung globaler und lokaler Reputationsraten für kooperatives Verhalten belohnt werden. BarterCast benutzt Statistiken, die ähnlich wie die Bank in einem Banking Mechanismus funktionieren: sie zeichnen die Reputation von Peers anhand ihres bisherigen Kooperationsverhaltens auf. Im Gegensatz zu PSH wird die Bank jedoch nicht dauerhaft erstellt. BarterCast lädt zuerst eine Statistik über das Kooperationsverhalten von

Peers und propagiert diese in einer bestimmten Community. In BitTorrent könnte dies bspw. alle Downloader eines bestimmten Files sein. Die durch diese Statistik informierten Peers versenden nun bei jeder Nutzung von Ressourcen eines Peers durch ein anderes sog. "BarterCast-Nachrichten", die innerhalb der Community über das Kooperationsverhalten der Peers informieren. Um zu verhindern, dass diese BarterCast-Nachrichten verfälscht werden (bspw. indem ein Peer fälschlicherweise bekannt gibt, dass es mehr Files verteilt hat, als das tatsächlich der Fall war), wird eine Höchstgrenze für Kooperation innerhalb eines bestimmten Zeitraums festgelegt. Die via BarterCast-Nachrichten informierten Peers können jederzeit Peers, welche sich innerhalb der Community unkooperativ verhalten haben, mittels Blockierung von Ressourcen bestrafen. Umgekehrt können kooperative Peers mit der Freigabe von Ressourcen belohnt werden. Der Vorteil des BarterCast-Algorithmus ist die gute Skalierbarkeit. Der Nachteil liegt allerdings darin, dass BarterCast-Nachrichten verfälscht werden können. BarterCast-Nachrichten benötigen deswegen Algorithmen, welche Übertreibungen aufdecken [9].

1.5 Zusammenfassung und Diskussion

Für P2P-Netzwerke existieren zahlreiche Anreizmechanismen, welche auf verschiedene Art und Weise implementiert werden und kooperatives Verhalten unter den Peers fördern. Diese Anreizmechanismen können auch ein grosses P2P-Netzwerk entscheidend verbessern.

1.5.1 Zusammenfassung

In den vorangehenden Abschnitten wurde gezeigt, dass Anreizmechanismen Verfahren sind, welche kooperatives Verhalten in P2P-Netzwerken fördern. Es wurde gezeigt, dass Anreizmechanismen eine praktische Bedeutung für P2P-Applikationen wie BitTorrent haben. Die Anreizmechanismen wurden klassifiziert und ihre Vor- und Nachteile aufgezeigt. Es wurde gezeigt, dass Anreizmechanismen einerseits of Vertrauen und andererseits auf der Belohnung von kooperativem Verhalten basieren. Es wurden vier verschiedene Ansätze vorgestellt, welche auf ihre Weise einige Aspekte dieser Anreizmechanismen aufgreifen. Dabei wurde festgestellt, dass Tit-for-Tat Ähnlichkeit mit dem Barter Trade Mechanismus hat, während Give-to-Get eher einem Community Mechanismus und PSH sowie BarterCast eher einem Banking Mechanismus gleichen. In den einzelnen Abschnitten wurden die Vor- und Nachteile dieser Ansätze gezeigt.

1.5.2 Diskussion

BitTorrent verwendet Tit-for-Tat als Approach, um kooperatives Verhalten zu fördern. Da Tit-for-Tat aber den Nachteil hat, dass die Wahl der Kooperationspartner problematisch ist, kann BitTorrent verbessert werden. Give-to-Get bietet die Möglichkeit, die Effizienz von BitTorrent-Downloads zu steigern, indem es schnellere Downloads für kooperative Peers anbietet und diese zudem dazu bringt, bereits heruntergeladene Pakete mit anderen

Peers zu teilen. PSH und BarterCast sorgen für ein Belohnungssystem, welches unkooperative Peers dazu bringt, mehr File-Pakete anderen Peers zu überlassen. Der gegenwärtige BitTorrent-Algorithmus könnte also durch Ansätze, welche verschiedene Aspekte der Anreizmechanismen berücksichtigen, effizienter und fairer gestaltet werden.

Literaturverzeichnis

[1] Michael Feldman, John Chuang. Overcoming Free-Riding Behavior in Peer-to-Peer Systems. ACM SIGecom Exchanges, volume 5, number 4, 2005.

[2] Kevin Lai, Michal Feldman, Ion Stoica, John Chuang. Incentives for Cooperation in Peer-to-Peer Networks. In *Workshop on Economics of Peer-to-Peer Systems*, Cambridge, MA, USA, June, 2003.

[3] Bram Cohen. Incentives Build Robustness in BitTorrent. In *Proceedings of the 1st Workshop on Economics of Peer-to-Peer Systems (P2PECON)*, Berkeley, CA, USA, June 2003.

[4] Philipp Obreiter and Jens Nimis. A Taxonomy of Incentive Patterns - The Design Space of Incentives for Cooperation. In *Proceedings of the Second International Workshop on Agents and Peer-to-Peer Computing (AP2PC'03)*, Melbourne, Australia, July 2003. Springer LNCS 2872.

[5] Philipp Obreitner, Birgitta König, Michael Klein. Stimulating Cooperative Behavior of Autonomous Devices: An Analysis of Requirements and Existing Approaches. In *Second International Workshop on Wireless Information Systems (WIS2003)*, Angers, France, April 2003.

[6] L.G. Alex Sung, Herman H.Y. Li. Neighbour selection strategies for P2P systems using titfortat exchange algorithm. University of Waterloo, 2004. `http://www.flamenet.ca/uw/courses/cs856_w05/project/web-project-tft.pdf`

[7] J. J. D. Mol, J. A. Pouwelse, M. Meulpolder, D. H. J. Epema, and H. J. Sips. Give-to-Get: free-riding resilient video-on-demand in P2P systems. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 6818, January 2008.

[8] Thomas Bocek, Wang Kun, Fabio Victora Hecht, David Hausheer, and Burkhard Stiller. PSH: A Private and Shared History-based Incentive Mechanism. In *Proceedings of the 2nd International Conference on Autonomous Infrastructure, Management and Security Resilient Networks and Services (AIMS 2008)*, Bremen, Germany, July 2008.

[9] Michel Meulpolder, Johan Pouwelse, Dick Epema, , and Henk Sips. BarterCast: Fully Distributed Sharing-Ratio Enforcement in BitTorrent. Technical report, Delft University of Technology, 2008.

Chapter 2

Overview and Trends of Video Streaming over IP

Frank Neugebauer, Aleksandar Markovic

Video streaming over IP has become a wide-spread phenomenon in the last few years. With increasing number of users, the classic client-server based streaming models were pushed to their limits. New multicast architectures were introduced in order to deal with the growing amount of users. This paper provides an overview over different techniques for video streaming over IP and shows the advantages and drawbacks of different approaches like Client-Server, IP Multicast and Peer to Peer transmission (also known as Application Layer Mutlicast). It also provides information on video streaming over the air and an outlook on technologies and problems to come.

Contents

2.1 Introduction

One of the main purposes of the internet is the exchange of data. It was mainly designed for transport of web documents, files and electronic mails. With its increasing popularity and availability of high speed access, this network is more and more used for video streaming, which means a technique that makes it possible to play a video file without the need to download the entire file first, i.e. a client receives a part of the file, begins with playback and downloads the missing parts at the same time. There are two major classes of video streaming: live and on-demand. Live streaming means that a user gets the video content in realtime; in video-on-demand, a user can choose what video she wants to watch at any time. Video streaming systems need to satisfy not only the requirement for high bandwidth, but also for scalability, reliability, latency et al. In Section 2.2, a basic solution for such system is described, namely the client-server model, where a client connects to the server and gets the stream. This works fine for a small number of clients, but if there are too many, the server will be overloaded. In order to achieve higher scalability, multicast systems are proposed, which allow a source to efficiently send data to many recipients. In Section 2.3, multicast on network layer is discussed, followed by the description of multicast on application layer (peer-to-peer) in Section 2.4. Support for streaming in wireless-based networks is discussed in Section 2.5.

2.2 Client-Server Model

The client-server paradigm is utilized by many distributed applications, also by those used in video streaming. The process of starting a video stream playout is illustrated in Figure 2.1 on page 28. A web browser sends an HTTP request to a web server. The web server responds with a meta file, which contains the information about the requested video file. This meta file is forwarded to a media player like Windows Media Player or Flash Player. According to the information in the meta file, the media player establishes the connection to the web server and starts with downloading of the video file. The traffic is transported either by TCP or by UDP protocol. The media player also performs other important functions, including the following:

- **Decompression.** As video file is almost always compressed in order to save disk storage and network bandwidth, the media player needs to decompress the video during playout.
- **Jitter removal.** Packet jitter means the variability of delays of packets within the transfer. Video must be played as the same sequence of pictures and sounds it was recorded, so the media player will buffer received packets for a short period of time to remove this jitter.

As mentioned above, the media player first downloads a certain part of the video file and stores it in the buffer. After a certain amount of data arrived, the media player starts with playout. During this process, the data from the buffer is being drained. A continuous

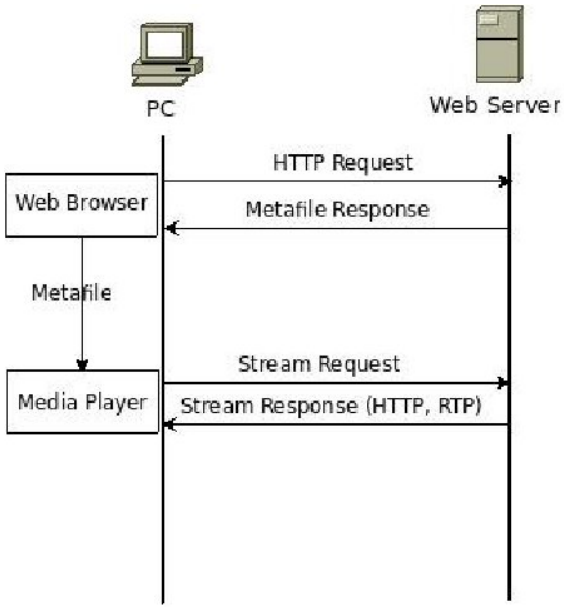

Figure 2.1: Client-Server Model for Video Streaming

playout of the video requires that the buffer is being filled with new data as fast or faster then data is being drained from there. If bandwidth is too low, a congestion or any other network problem occurs, the buffer may be emptied, and the media player will be forced to pause the playout. This process is known as starvation.

The described model could be further optimized by splitting the functionality of requesting a video and streaming it. Web servers can very well send video data over HTTP, but dedicated, for streaming optimized servers can do it better. The previously mentioned meta file shall contain information that would redirect the media player to the streaming server. A protocol, specifically developed for media streaming like RTP, can be used between the server and the media player and allow rich interaction between the user and the video stream [22].

2.2.1 Content Distributed Networks

If a video content is stored only on one server, problems may result in following situations:

- When the client and the server are too far away from each other, packets may need to pass through many ISPs. With each passing, the chance for delay and loss increases.
- When many clients request a video at the same time, the video will likely be sent many times through the same ISPs, consuming a lot of bandwidth.

If for example many people from Europe are interested in watching a video that is saved on a server in California, every packet has to pass many routers across the entire USA until east coast and then travel along the submarine cable to London, from where it could be further distributed. In order to avoid such a scenario, a replicate can be installed in

Europe (and other places of the world). Each time the content of the main server changes, the new content is sent to all replicates. Of course, a mechanism is provided by which the most optimal replicate for a position can be determined [22].

2.3 Multicast Systems

Multicast is defined as a one-to-many routing topology, with the purpose of saving network bandwidth by transmitting the desired data as sole consistent stream over the backbones and only multiplying this stream at the end of the route [14]. An example would be the transmission of a live media stream to multiple receivers.

2.3.1 History of Multicast

In the 1980s scientists developed the idea of optimizing the transport of data over the early internet infrastructure. These considerations culminated in 1985 in the RFC 966 [14] which discussed the possibility of adding multicast to the network structure, hence avoiding to send unnecessary copies of the same data. A few years later, in 1989, the RFC 1112 [13] outlined the basis for multicast over the Internet Protocol (IP). Additionally the Internet Group Management Protocol (IGMP), which will be presented later, was defined in this RFC.

Founding on these RFCs the MBone multicast network was build, which tunneled multicast traffic through the unicast internet. The first audio multicast over the MBone took place in 1992, followed by the first video multicast the same year [16]. In the following years the MBone network rapidly developed into a commonly used tool in the academic sphere, but failed to inspire the mass of users. [3] Reasons for this underdevelopment will be presented in the next sections.

2.3.2 Technical Details of Multicast in IPv4

Multicast has to be distinguished from unicast and broadcast, because unlike unicast it does not just address one receiver or all of them (broadcast) but a variable subset of receivers called the multicast group. Any such group is identified by a multicast address which lies in the class D address space of the internet protocol (with a few addresses reserved by the IANA [2]). As multicast usually utilizes UDP for data transmission it is considered unreliable and any host can join or leave a multicast group dynamically, as there is no access control [3].

Another vital point is the Internet Group Management Protocol (IGMP) [13] which is used by multicast routers to survey the nodes in a multicast group. The IGMP 1 defines joining a group, membership query and the sending of membership reports. There is no process for explicitly leaving a group, other than to wait for the router to realize the host is no longer responding; leaving a group was introduced in IGMP 2 (RFC 2236) [18].

Table 2.1: Chronological overview of common multicast algorithms

Protocol	Dense mode	Sparse mode	Implicit join	Explicit Join	Information
DVMRP	x		x		RFC 1075 [35]
MOSPF	x			x	RFC 1584 [27]
CBT		x		x	RFC 2189 [6]
PIM-DM	x		x		RFC 3973 [1]
PIM-SM		x		x	RFC 4601 [17]

A lot of research has been invested into the field of routing algorithms for multicast, as it turned out to be very complex to efficiently route information from a variable number of senders to a changing multicast group. The best known algorithms to fulfill this task are the Distance Vector Multicast Routing Protocol, Multicast Open Shortest Path First, Core Based Tree and Protocol Independent Multicast in dense and sparse mode [3]. Dense mode describes an operational mode, which assumes a receiver at most locations, like a companywide presentation of a new CIO, whereas sparse mode assumes only few data sinks to be present, like an introductory video of a small workgroup [19]. Their main difference lies in the initial behavior of these algorithms. In dense mode the multicast stream is send to every router in the network, a characteristic known as flooding. This data flow is then stopped by the prune messages, which report if the multicast date is really consumed by any sink. If there are no receiving sinks behind a router, the data stream is no longer sent to this router [1]. Typically flooding occurs every few minutes, making it only suitable for relatively low amounts of transferred data.

In contrast sparse mode requires the source to wait for a specific request of the sink. These are called join messages guarantee that the data packets are only passing the routers that have requesting sinks behind them. If a receiver does not require the data stream anymore, a prune message is send back, cutting the data flow off [17]. Due to this bandwidth preserving approach sparse mode multicast scales better than dense mode, but requires a more sophisticated architecture [11]. Their main functionality and the corresponding RFCs are specified briefly in Table 2.1.

2.3.3 Multicast's Problems

Multicast suffers from a variety of incompletely resolved issues which pose a constant drawback for multicast's proliferation in the world. One of these problems is the unreliability of data transmission, due to the UDP foundation of multicast. Additionally there is no congestion control, no error correction and no way to guarantee the chronologically correct arrival of datagrams. The solution to this shortcoming could be the Reliable Multicast Transfer Protocol which implements these features [24].

Another rather banal reason for the slow development of multicast lies in the satisfactory performance of unicast serving small groups of users with information. A complex multicast installation is seldom needed, when just small amounts of data are to be transferred over the well tested unicast architecture [8]. Furthermore the majority of internet routers and internet service providers are still unable to handle multicast traffic. Most IP multicast routing algorithms claim a lot of the computing power of today's routers and tend

to clog their lookup tables thus limiting the scalability. Until more effort is put into the development of multicast routers and connecting the multicast islands it is unlikely that multicast will fulfill its' potential [4].

2.3.4 Future of Multicast

While multicast is often considered to be too complex to be implemented in a reasonable amount of time compared to IPv4 one has to wonder for how long this will stay true. The address range of IPv4 is almost exhausted and the deployment of the 6st incarnation of the internet protocol is imminent. In this situation it is likely that there will be another push in the research and use of multicast. Not only will there be fewer difficulties concerning the scalability, but users will be forced to use multicast, as it will replace broadcast [34].

Additionally IPv6 was developed for the era of ubiquitous computing, where a large amount of systems will be mobile [29]. Multicast can profit from this growth, although researchers will have to contrive new algorithms for ad-hoc multicasting. Possible uses of multicast in the near future would find itself in the computer game sector. As Massively Multiplayer Online Games attract huge amounts of players, providers have to resort to service a lot of independent servers. With multicast providers would be able to consolidate and decrease the workload on these servers [33].

For achieving the commercial breakthrough multicast had been missing a killer application up till now. It is most likely that Internet Protocol TeleVision (IPTV) will provide multicast the mass market it needs to reach a critical mass. While video streaming over IP is nothing new, the circumstances we live in are: Internet access is practically unlimited, Bandwidth is of no concern to the users and an increasing number of people is turning its back on the inflexible TV schedules. Even today a remarkable number of people are using devices to automatically capture television in order to watch these multimedia streams in due time. Multicast would push today's boundaries of television a lot further and finally enable people to live a more flexible life [38].

2.4 Peer-To-Peer Systems

In previous section, IP multicast systems were discussed. Another approach of multicasting is the implementation of the functionality on the application layer, instead on the network layer. The main idea of P2P is that users act both as clients and servers at same time, namely as peers. IP Multicast is able to avoid multiple copies of a packet on the same link and can construct optimal trees, but peer-to-peer systems can be deployed easier and immediately on the internet, they can easier be maintened and adapt better to a specific application. Participating end hosts fully control constructing of the multicast delivery structure and data forwarding, no support of intermediate nodes such as routers is required. Therefore, creation and maintenance of good topologies is the main challange of P2P systems. Good topology means that the connections of peers allow exchange of data in an efficient, cost-aware manner, the delay between peers and the number of connections between them are minimized. Other requirements are robustness (dealing with

members who leave), scalability (system can efficiently deal with increasing number of nodes) and low control overhead (the exchange of control messages is minimized) [21].

There are two categories of P2P streaming systems: tree-based and mesh based. The classification is based on the overlay network structure. The overlay structure of tree-based systems is well organized and video data is typically distributed by pushing data from a peer to its children peers. But such systems suffer from vulnerability to peer churn: if a peers leaves, the data flow to all peers in the subtree rooted at the leaving peer will temporarily stop. In a mesh-based system, connections between peers are established/terminated based on the availability of content and bandwidth on peers. According to definied time intervals, peers exchange information about data they have and pull the content they need from their neighbors who already obtained that data. Those systems are very robust to peer churns, but there is no guarantee that a peer will receive all data it needs at the right time. As a consequence, the video playback may suffer from startup delay, playback freezes ect. [25]

2.4.1 Single-Tree Streaming

In single-tree streaming, the construction of the tree starts with source server as root. Each peer joins the tree at certain level, receives the data from its parent peer and forwards already received data to children peers. With an arbitrary amount of peers, there are many possible ways to construct a streaming tree. An example of a single tree construction is illustrated in Figure 2.2.

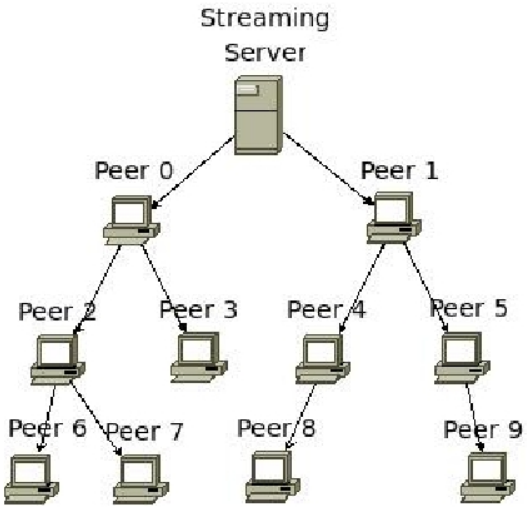

Figure 2.2: Single-tree structure

The major considerations are the depth of the tree and the fan-out of the nodes. Peers at the lower levels receive video after peers at upper levels. In order to reduce the delay for peers at bottom level, a tree with fewest levels possible is required, that is to say at each level, the tree topology should fan out as wide as possible. The number of children a peer can have (and upload data to them) depends on its uploading bandwidth. Apart from tree construction, tree maintanance is another important point. Each user might

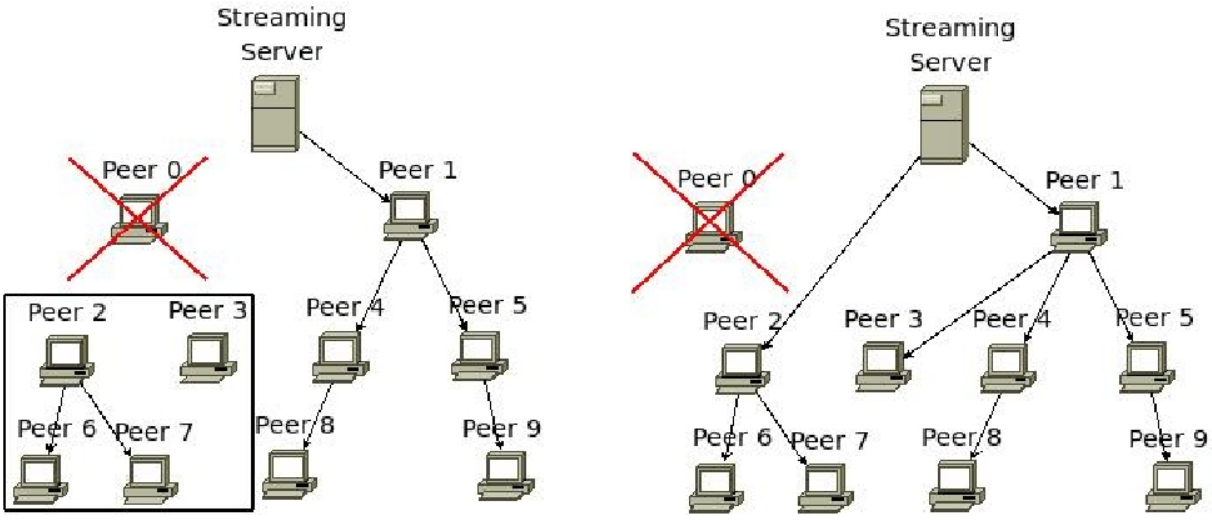

Figure 2.3: Peer churn and reconstruction of the streaming tree

leave the streaming session unexpectedly. The reason could be a machine crash, network crash or other. This has very negative consequences for the descendants in the streaming tree: they are cut off receiving video data. In that case, the tree shall be recovered by reassigning the affected peers to another peer, which may take to long.

The construction and maintenance of a streaming tree can be done in a centralized or distributed way. In a centralized solution, the construction and recovery is controlled by a central server. In order to join the system, a peer needs to contact the central server, that would compute the appropriate position and inform the peer what parent peers it shall connect to. The central server can detect changes in topologies, e.g. peer departures, recalculate the tree topology and instruct the remaining peers how they have to form the new topology. Therefore, it is crucial for the system that the central server works fine [21].

2.4.2 Multi-Tree Streaming

Two major drawbacks of the single-tree systems are the fact that the bandwidth of all the leaf nodes (which are numerous) is not utilized and interruption of the stream in case of peer churn. In multi-tree systems, the stream is divided into multiple sub-streams. For each substream, a streaming tree is constructed. Each peer is part of all sub-trees (it needs to receive some of sub-streams), but at different positions. In some sub-trees, a peer acts as an internal node, in others as a leaf node. For an efficient bandwidth utilization, the number of sub-trees in which a peer is placed on an internal node can be set according to its uploading bandwidth [21].

Multi-tree streaming approach can be further optimized by combining it with SVC (scalable video coding). The idea of SVC is the following: a video stream is encoded into a base layer and one or more enhancement layers. Each layer is pushed to the clients as separate stream. A client must receive the stream of the base layer in order to be able

to play the video in low quality. With each additional stream it gets, the quality will improve.

2.4.3 Mesh-based Systems

In a single tree-based system, each peer has only one possible source for getting data, namely its parent. If the parent peers leaves, all its descendant cannot get data until they are connected to another source, so peer churns pose a large challange for the management of streaming trees. Mesh-based streaming systems encounter this problem by providing a very dynamic topology. At any time, a peer maintains peering connections with multiple neighboring peers, e.g. a peer may download/upload video from/to multiple other peers at the same time, as illustrated in Figure 2.4 on page 34. In case of a peer churn, a peer can get data from the remaining, available peers.

In mesh-based systems, a tracker keeps the information about the active peers in the video session. In order to join a session, the first step for a peer is to contact the tracker and report its own information like IP and port number. Afterwards, it receives the list with information about a random subset of available peers. The joining peer will then try to connect to peers from that list and start the exchange of video content. In order to deal with peer churn and arrivals, each peer constantly updates the list about its neighbors during the session. If a peer is able to leave the system gracefully, it will inform the tracker and its neighbors before, so that they can update their lists. In order to discover unexpected leavings, peers exchange keep-alive messages.

In mesh-based systems, the basic data unit is a video chunk. A video file is devided into small media chunks at the server. Each chunk contains data for a small time interval and is labeled by a unique sequence number. A lower sequence number means that this chunk contains earlier video parts. Because chunks are distributed along different paths, they may arrive to late to some peers. Each peer informs periodically its neighbors about buffered chunks, pulls the chunks that it needs from its neighbors and buffers received chunks for a certain time [21].

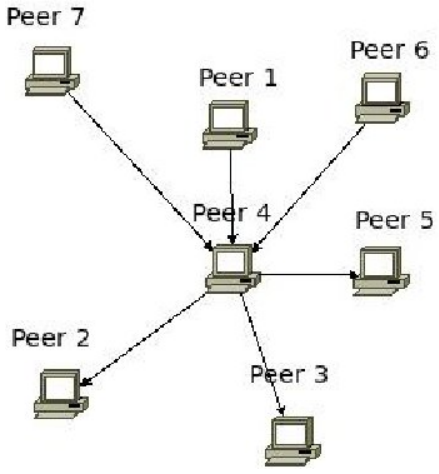

Figure 2.4: Mesh structure

Table 2.2: Comparison of tree-based vs. mesh-based systems

	tree-based systems	mesh-based systems
approach	push	pull
form of data	stream	chunk
topology	static (if no churn)	dynamic
maintenance	central server / distributed	tracker, exchange of info list
main problem	peer churn	receiving chunks not guaranteed

2.4.4 Examples

The company Joost with its homonymous software is a good example for the pitfalls of IPTV. Joost started in 2006 as a standalone application that enabled the user to watch Video on Demand (VOD) on his computer monitor. It received a lot of premature praise from the media, but Joosts popularity and user base dwindled away in the next two years. But where did Joost fail?

To understand the downfall of Joost one has to understand the supporting architecture. First Joost had to cope with the legal geographic limitations the Rights Owners imposed, and thus the provided range of media was disappointing for customers outside the USA. Secondly Joost created artificial barriers for the user, as the only way to view a Joost video stream was to install the closed source software Joost provided [26]. This step excluded mainly corporate users, who didn't have the privileges to install software.

Additionally Joost tried to safe maintenance costs by building their application on top of a hybrid network architecture. It combined peer to peer transfer with massive client server architecture. This promising move was intended to save bandwidth, but because Joost offered a wide but mediocre selection of titles it was very unlikely for a viable peer to peer network to form. The Joost software buffered only approximately one hour of recent played movies on the client side, which simply wasn't sufficient to create enough seeding peers. The result was that the majority of data was transferred from the Joost servers, leaving the peer to peer architecture deserted [20]. But all this would not have bothered the users, if it was not for Joosts low quality. Even though Joost used about 90 KBps for its streaming and the H.263 Codec for video encoding [26] the picture quality was inferior compared to other same-sized data streams. This raised the assumption that Joosts hybrid architecture created a large overhead, while not generating any benefit. But it has to be remarked, that this is unproven, because Joost utilized an asymmetric encryption for all its transferred data, which made the analysis very complex [28].

The problems mentioned above led to Joost's decision to drop the standalone application and the hybrid network architecture in favor of a flash based browser player. But where lies the distinction between Joost and more successful competitors like Zattoo? Zattoo employs an easier procurement scheme than Joost did and Zattoo offers a much more condensed library of content. This simple strategy is the key point, as it enables Zattoo to efficiently use a peer to peer network for its multimedia streaming. While Joost entered the competition with a big media library, Zattoo concentrates on near-live streaming of a variety of channels. This constraint leads to bigger peer clouds that can efficiently distribute the data between them.

2.4.5 Lessons Learned

These examples show that while Joost started with more backup from the media industry the underlying architecture played an important part in its downfall, and in the rise of Zattoo as a European competitor. Zattoo could easily pocket large parts of the European market, as it offers the major TV stations in better quality than Joost does with only a short delay compared to live TV. Due to this restriction Zattoo has no need to sustain several data centers like Joost does to fill the bandwidth gap and to ensure that there are enough seeding peers available for every stream. It has become obvious that peer to peer technology is no magical cure in itself, but needs even more thought than simple client-server philosophy. In this environment it is also very important to be aware of the bandwidth gap, which emerges from the imbalance of user upload and download capabilities [20].

Furthermore it also becomes clear, that while users like a big selection of media, they favor image quality over quantity. A reason for this is that the user expects that IPTV should be at least perceived as good as his well-known single definition television (SDTV), which would need a higher resolution than the common PAL/NTSC format offer. [30] This contradiction is explained by the fact that SDTV was developed to match the physical criteria of cathode ray tubes (CRT), and don't really match today's Liquid Crystal Displays (LCD). The same image, with the same resolution, is rated to be of lower quality on a LCD, because errors and inadequacies are more easily perceived by the human eye.

Another lesson learned is that users want to become more independent from the static television program and that they use video on demand services if they are easily available [25]. All these reasons show that multicast, as described in the previous sections, will most likely become a key factor in the operation of video streaming when the network and the algorithms are ready.

2.5 Streaming over the Air

With streaming over wired connections slowly reaching a mature state, new research is directed into hitherto viewy fields. One of these research areas is streaming over the air (OTA) which focuses on techniques to enable devices to seamlessly communicate with each other in a mobile environment. In this chapter different approaches and their characteristics are provided, ranging from soon to be deployed technologies like IPTV 2.0 to visionary ones like the IPTV after next, scalable video codecs and ubiquitous streaming.

2.5.1 IPTV 2.0

IPTV 2.0 is the next generation network (NGN) based incarnation of IPTV. Its key feature are multimedia services being delivered over wired and wireless converged networks, not only providing a high bandwidth but also qualitative features like QoS, new security

concepts, interactivity and reliability of the transmission [38] [36]. Additionally it should be possible for anyone to create or consume content wherever, whenever he desires.

The next version of IPTV is planned to outperform simple IPTV in various aspects. Not only should the image resolution rise, but IPTV 2.0 should also be viewable from mobile devices. Users should be able to take their personal media library with them or at least to access it from any point. This leads to the next point: Users will no longer be chained to a single telecommunications provider and its walled garden architecture, but be able to select from a variety of interoperable providers [23]. Another addition to IPTV are location based and context sensitive media streams (ie. in a museum there is instant availability of a videostream for each object in proximity to the user) which aim to enrich our daily life with media. It goes without saying that seamless mobility is also part of the IPTV 2.0 idea [32].

2.5.2 Wireless Streaming

One near term solution for video streaming OTA is wireless transmission according to IEEE-802.11 [12]. The benefit from this streaming lies in the grown mobility for the user. But there is not only this transmission method: WiMAX (IEEE-802.16d/e [15]) is also a hot candidate for wireless streaming in metropolitan areas [5]. But as promising as these ideas sound, one has to remember, that wireless transmission of data is error prone and that all participants share the same transmission frequencies, thus limiting the possible maximum usage of anyone.

It is yet unclear if wireless transmission technologies will rise to the challenge of providing users with bandwith-consuming services as IPTV. It is most likely that in the beginning wireless streaming will be used for devices with small screens as mobile phones, music players and portable gaming consoles; in addition to a stationary cabled large screen television. A point in favor of streaming OTA is the maximization of possible subscribers in cities. The setup of a city wide wireless network is cheap compared to putting up lines to every customer. This means that wireless streaming services are likely to be offered at low cost [31].

There are a few technical difficulties to master though: What if one broadcasted stream is being received by many participants who have different reception [10]? Is a high forward error correction (FEC) the best way to ensure that all users receive a continuous data stream; even if it would reduce the resolution due to the redundant data in the stream? Or is it a better idea to split the available bandwidth into two data streams, one of which is based upon the other and carrying additional special information [37].

Another candidate for wireless streaming is Bluetooth, which was developed for mobile devices with low power consumption [7]. Bluetooth is using time and frequency multiplexing, hopping almost randomly through 72 possible channels and is less prone to interference. Bluetooth has a much shorter perimeter and smaller bandwidth than WiMAX, but for near area streaming Bluetooth could be a worthy contestant[7]. There is research pointing out that Bluetooth personal area networks (PANs) can be combined into scatter nets which cover a bigger area. These scatter nets could be used to send low resolution video

(with approx. 50 kBps) to different users [39]. This restriction may seem severe, but this data rate is adequate for most user generated content or small video clips.

2.5.3 Future of Streaming

The future of streaming is mobile and ubiquitous. With the help of scalable video codecs and high bandwidth mobile carrier networks media will be everywhere. For example, it will be possible to start watching your favorite TV series on your mobile internet device (MID) in the tube on your way home from work and seamlessly continue to watch it at home in a higher resolution on a big LCD. Personal media servers will be established in every household and gather all information the user wants, starting with the weather forecast and not ending with recommendations on what to cook or watch on TV [9].

Thanks to scalable video codecs the content will not be needed to be encoded in different formats over and over again, they just have to be encoded once and only the necessary parts will get transmitted redundancy free, depending on the capabilities and the connection of the receiving device.

2.6 Conclusion

In this report, an overview was provided over the different approaches to streaming of IPTV. In the first chapters the different known and used distribution methods were elaborated. This started with the Client-Server Model and the widely used Content Distribution Networks and continued with more complex multicast systems. These multicast systems are yet far from maturity but the advantages are undisputed. The next chapter dealt with peer to peer systems which realize some kind of application layer multicast and are commonly used today, but the example presented here shows that there are dure considerations to be made before deploying a peer to peer system for streaming. The last part dealt with the future of streaming, which will presumably be dominated by wireless transmission.

We find it most likely, that peer to peer networks will continue to dominate IPTV in the next few years, but start to lose support with the introduction of IPv6, which is anticipated to relieve multicast from its infancy. The expanding proliferation of wireless networks will profit from cooperation with multicast, as this offers wider user range and faster data transmission, which will be needed to satisfy the users' demand for increasingly higher image resolutions and mobility. Presumably the trend towards mobile media usage will continue, and originate not only new codecs like SVC but also new devices (Mobile Internet Devices) in the niche between today's netbooks and smartphones. The trend continues towards a new kind of user, a mobile prosumer, who will be able to create, share and consume wherever he is staying with the same ease as we watch television today.

Bibliography

[1] J Nicholas A Adams and W. Siadak. Protocol independent multicast-dense mode (pim-dm): Protocol specification (revised). *Internet Engineering Task Force (IETF) Request For Comments*, RFC 3973, 2005.

[2] Z Albanna, K Almeroth, D Meyer, and M Schipper. Iana guidelines for ipv4 multicast address assignments. *Internet Engineering Task Force (IETF) Request For Comments*, RFC 3171, 2001.

[3] KC Almeroth. The evolution of multicast: from the mbone to interdomain multicast to internet2 deployment. *IEEE Network*, 2000.

[4] KC Almeroth. A long-term analysis of growth and usage patterns in the multicast-backbone (mbone). *IEEE INFOCOM*, 2000.

[5] JG Andrews, A Ghosh, and R Muhamed. *Fundamentals of WiMAX*. ulb.tu-darmstadt.de, 2007.

[6] A Ballardie, B Cain, and Z Zhang. Core based trees (cbt version 2) multicast routing-protocol specification. *Internet Engineering Task Force (IETF) Request For Comments*, RFC 2189, 1997.

[7] C Bisdikian. An overview of the bluetooth wireless technology. *IEEE COMMUN MAG*, 2001.

[8] MS Blumenthal and DD Clark. Rethinking the design of the internet: the end-to-end arguments vs. the brave new world. *ACM Transactions on Internet Technology (TOIT)*, 2001.

[9] T Cagenius, A Fasbender, J Hjelm, U Horn, and IM Ivars. Evolving the tv experience: Anytime, anywhere, any device. *Ericsson Review*, 2006.

[10] DH Cho, JH Song, MS Kim, and KJ Han. Performance analysis of the ieee 802.16 wireless metropolitan area network. *Distributed Frameworks for Multimedia Applications*, 2005.

[11] JC Chuang and MA Sirbu. Pricing multicast communication: A cost-based approach. *Telecommunication Systems*, 2001.

[12] BP Crow, I Widjaja, LG Kim, and PT Sakai. Ieee 802.11 wireless local area networks. *IEEE Communications magazine*, 1997.

[13] S Deering. Host extensions for ip multicast. *Internet Engineering Task Force (IETF) Request For Comments*, RFC 1112, 1989.

[14] SE Deering and DR Cheriton. Host groups: A multicast extension to the internet protocol. *Internet Engineering Task Force (IETF) Request For Comments*, RFC 966, 1985.

[15] C Eklund, RB Marks, KL Stanwood, and S Wang. Ieee standard 802.16: a technical overview of the wirelessman tm air interface for broadband. *IEEE communications magazine*, 2002.

[16] H Eriksson. Mbone: The multicast backbone. *Communications of the ACM*, 1994.

[17] B Fenner, M Handley, H Holbrook, and I Kouvelas. Protocol independent multicast-sparse mode (pim-sm): Protocol specification (revised). *Internet Engineering Task Force (IETF) Request For Comments*, RFC 4601, 2006.

[18] W Fenner. Internet group management protocol, version 2. *Internet Engineering Task Force (IETF) Request For Comments*, RFC 2236, 1997.

[19] S Floyd, CG Liu, S McCanne, and L Zhang. A reliable multicast framework for light-weight sessions and application level framing. *IEEE/ACM Transactions on Networking (TON)*, 1997.

[20] YJ Hall, P Piemonte, and M Weyant. *Joost: A measurement study*. pembrokeballet.com, 2007.

[21] M Hosseini, D T Anvir, and S Shirmohammadi. A survey of application-layer multicast protocols. *IEEE COMMUNICATIONS SURVEYS*, 2007.

[22] J F Kurose and K W Ross. *Computer Networking, A Top-Down Approach*. Addison Wesley, 2007.

[23] CS Lee and G ETRI. Iptv over next generation networks in itu-t. *2nd IEEE/IFIP International Workshop on Broadband*, 2007.

[24] JC Lin and S Paul. Rmtp: A reliable multicast transport protocol. *Proceedings IEEE INFOCOM'96. Fifteenth Annual Joint*, 1996.

[25] J Liu, SG Rao, B Li, and H Zhang. Opportunities and challenges of peer-to-peer internet video broadcast. *PROCEEDINGS-IEEE*, 2008.

[26] C MacCarthaigh. Joost network architecture. *7th UK Network Operators Forum*, 2007.

[27] J Moy. Multicast extensions to ospf, internet engineering task force. *Internet Engineering Task Force (IETF) Request For Comments*, RFC 1584, 1994.

[28] F Neugebauer. Next generation media: Analysen zu joost. *Seminarbeit Informationsmanagement*, 2007.

[29] CE Perkins and DB Johnson. Mobility support in ipv6. *Proceedings of the second annual ACM/IEEE International Conference on Mobile Computing and Networking, MobiCom '96*, pp. 27–37, White Plains, New York, USA, 1996.

[30] UH Reimers. Dvb-the family of international standards for digital video broadcasting. *Proceedings of the IEEE*, 2006.

[31] J She, F Hou, PH Ho, and LL Xie. Iptv over wimax: Key success factors, challenges, and solutions [advances in mobile multimedia]. *IEEE Communications Magazine*, 2007.

[32] DH Shin. Potential user factors driving adoption of iptv. what are customers expecting from iptv? *Technological Forecasting & Social Change*, 2007.

[33] J Smed, T Kaukoranta, and H Hakonen. Aspects of networking in multiplayer computer games. *The Electronic Library*, 2002.

[34] W Stallings. Ipv6: The new internet protocol. *IEEE Communications Magazine*, 1996.

[35] D Waitzman, C Partridge, and S Deering. Distance vector multicast routing protocol (dvmrp). *Internet Engineering Task Force (IETF) Request For Comments*, RFC 1075, 1988.

[36] J Walko. I love my iptv. *Communications Engineer*, 2005.

[37] HY Wei, S Ganguly, R Izmailov, and ZJ Haas. Interference-aware ieee 802.16 wimax mesh networks. *2005 IEEE 61st Vehicular Technology Conference (VTC)*, 2005.

[38] Y Xiao, X Du, J Zhang, and F Hu. Internet protocol television (iptv): the killer application for the next-generation internet. *IEEE Communications Magazine*, 2007.

[39] GV Zaruba, S Basagni, and I Chlamtac. Bluetrees-scatternet formation to enable bluetooth-based ad hocnetworks. *IEEE International Conference on Communications, 2001. ICC*, 2001.

Chapter 3

P2P Collaboration in Support for Spam Detection

Alexander Schäfer, Rilind Balazi

A lot of spam is entering the mail inbox every day, filling space wastefully and taking away precious time, dispite the current employed high tech anti-spam systems, which can identify and eliminate most of the spam emails. There are currently collaborative tools, like Razor or DCC, which harness the spam knowledge of a collective to hinder the delivery of spam emails very successfully. In this small paper, several peer-to-peer based collaborative anti-spam system proposals are analyzed and necessary comparisons were made of key technologies employed by each of the systems. Technologies similar in each paper are discussed together, differences are mentioned inside the proposal introductions.

Contents

3.1 Introduction

Today the e-mail is commonly used for communicating between people, because of its easiest and cheapest form of communication. With the grow of the volume of e-mail, spam has become a major problem. Spam spans all the e-mail traffic where a e-mail user receives unwanted mail, usually not personalized to that user, with the intention to sell some arbitrary product the user has probably never heard about. More then the half of the e-mails we receive can be spam, depending on the mail service, or lack thereof, this can easily rise up to 90 per cent spam volume. In the following we will introduce some definitions, necessary to understand the whole work.

3.1.1 E-Mail System

Electronic mail – often abbreviated as e-mail or email – is any method of creating, transmitting, or storing primarily text-based human communications with digital communications systems [10]. E-mail messages consists of two major sections: header and body. The header contains meta data, like sender and receiver e-mail, subject etc.. The body contains the whole message (unstructured text, images, movies, audio etc).

3.1.2 Voice over IP

Voice over Internet Protocol (VoIP) is a general term for a family of transmission technologies for delivery of voice communications over IP networks such as the Internet or other packet-switched networks [11]. VoIP is different from the e-mail because it is a real time service. It uses the signaling Protocol, SIP (Session Initiation Protocol) for setting up and tearing down multimedia communication sessions [11].

3.1.3 SPAM and HAM

E-mail spam, also known as junk e-mail, is a subset of spam that involves nearly identical messages sent to numerous recipients by e-mail. The costs caused by spam are increased each year. According to the spam statistics made by ITS (Information Technology Services) of Yale University [13], on August 2008 94.54% of all incoming emails are detected as spam. Daily average of incoming spam is 4 Million. Nowadays there are some anti-spam methods. For explaining the existing methods, we take two different aspects into account: the filter's software architecture and their internal filtering engines [2]. Ham is called the desired e-mail for the user.

3.1.3.1 Architectural differences

The existing spam solutions can be split in three major categories:

Email client plug-in. On every computer that needs spam filtering are installed software plug-ins on client-side. It is commonly used by organizations with few computers whose users are in charge of managing the filters. Web-based or mobile access to email is not protected by plug-ins [2].

Centralized filtering server. This architecture consists just of one anti-spam filter running on a centralized mail server. Drawback of this architecture is the ignoring of user opinions and preferences [2].

Gateway filtering. All emails by this approach are filtered before being delivered to the email server. Disadvantage of this architecture is the fact, that it is a single point of failure and may be difficult to manage in presence of multiple mail servers within an organization [2].

3.1.3.2 Filtering methods

There are several techniques to be used for actual spam filtering, some of the most important methods used are listed below.

List-based Filtering. This filtering-method uses white-/blacklisting for all the email-sender. Emails-sender of a spam is inserted into a blacklist and other are inserted into a white one. This approach is considered to be ineffective, although server-based solutions adopt them as an auxiliary technique often to be integrated with challenge-response [15] in [2]. This method seems to be very slow and often costly and difficult for an organization to be removed from these lists, even if the insertion was not correct or due to a transitory condition [2].

Rule-based Filtering. This kind of filter assign a score to each email based on the information if the email contains features typical of spam messages, such as fake SMTP components, keywords, HTML. But spammers learns quickly these rules and find ways to overcome them [2].

Source authentication/challenge-response. The sender of a suspect email which uses this form of filtering should reply to a recipients challenge by simply clicking on reply and then send [2].

Monitoring by a human operator. This form of filtering seems to be the best solution, but the rate of human error should not be underestimated. This filtering method is very expensive and violates the final recipient's privacy [2].

Bayesian word distribution filters. Bayesian filters use the frequency-based probability to words as spam indicators. A drawback of this method is that they have the hardest time blocking messages that do not lexically look like spam, e.g. messages sent by many recipients for representing them a serious URL [2].

Collaborative spam filtering. In this approach the filtering is based on the information taken from a server-side automatic monitoring systems that compare incoming message to known spam as classified by other peers. The main advantage of this method is the overcome of the single point of failure typical of centralized architecture [2].

The main problem is the existing of false positives. On the following we will introduce some methods that want to overcome this problem.

3.1.4 Peer-to-Peer

Peer-to-Peer network is a network in which all participants share their resources. Overlay is the virtual signaling network established via TCP [and/or UDP] connections between the peers [14]. A peer is a node actively participating in the overlay. At the same time is it a content provider, content requester and router in the overlay network [14].

The main features of the peer-to-peer networks are:

1. Resources are shared between the peers.
2. Resources can be accessed directly from other peers.
3. Peer is provider and requester (Servant concept) [14].

There are some advantages using P2P networks in comparing with client-server architectures. As the number of the participants on the network and the demand on resources increase, the total capacity of the system increases. By client-server architecture adding more clients means slower data transfer for all users [12]. It is to note that Peer-to-peer networks are stable, because there is no single point of failure.

3.2 Key Technologies

The technologies discussed here are important for the P2P anti spam systems. Unique identifiers are there to allow correct identification of emails, and p2p protocols give more information about that stuff.

3.2.1 Unique Identifiers

In order to refer messages in a peer-to-peer network we use a unique identifier (digest, hash value). It is important to use unique identifiers in spam filtering, because the unique identifier never disclose the message content. There are some methods for creating unique identifiers (digests, hash values), such as *nilsimsa*[5] and *Crossed Tri-gram*[4]. According to [6] an unique identifier should satisfy following requirements:

1. The digest identifying each message should not vary significantly for changes that can be produced automatically.
2. The encoding must be robust against intentional attacks.

3. The encoding should support an extremely low risk of false positives.

Now will be briefly explained how the nilsimsa algorithm works for creating unique identifiers. Nilsimsa takes content of the message as input and then generate a digest consisting of 32-Byte code representing the distribution of trigrams in the text [6]. After generating the digest, it will be compared with other mail-digests in order to see if they are similar. By comparing all the bits that are not equal with each other at the same position are counted and if the number of bits is greater then a predefined threshold the mail will be identified as a spam.

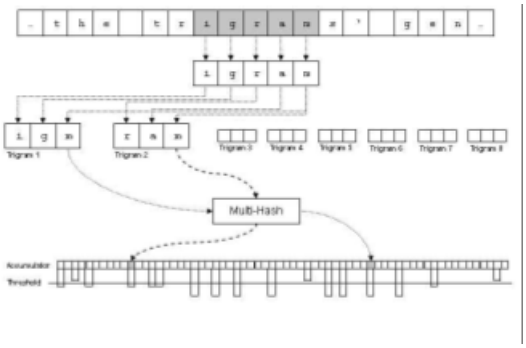

Figure 3.1: Nilsimsa working [5]

Unique identifiers are used in all the system proposals. The Three-Tier System uses unique identifiers in following two cases: each mailer identifies the user via a unique identifier and for each message a digest is computed. Multi-Agent Interaction System uses the digests (unique identifiers) too.

3.2.2 Peer-to-Peer DHT Protocols

The different P2P protocols suggested by the proposals are there to build a network with the following functionality: discovery, connect and disconnect nodes, insert and retrieve values of the network. Furthermore all of the protocols are DHT-based, allowing them to scale to the Internet, as they partition the network for efficiency.

All of the protocols are based on the old style of searching a key and retrieving a value. This is very useful for a Gnutella like network, where a person can search for a filename and retrieve the data behind the filename. However if someone searches for an evaluation of a certain email, there must be a use of a locality sensitive hash (see Unique Identifier) as key to counter random character insertion in spam messages, which could have several values associated with it. A search range for the hash will even broaden or shrink the amount of values to be found. Therefore proposal [4] has adapted their P2P protocol to accommodate this, and it can be assumed the other proposals adjusted theirs to the same extent for their testing.

All the protocols also include connect and disconnect. Each node typically holds a set of values within a key range. For caching nodes additionally hold the data of near neighbors. Upon connect and disconnect data has to be shifted around to accommodate newly created or changed segmentation within the network.

3.2.2.1 Chord

Chord [7] is one of the first DHT protocols, a P2P-technique which brings scalability to P2P. The protocol only supports one operation, it assigns a key to a node. This can be used for finding data and retrieving it or just sending processing instruction to complete complex algorithm in parallel in a grid. There must be consistent hashing implemented in order to correctly associate and distribute the keys on all nodes participating in the network.

Furthermore, each node will only need to know O(logN) other nodes in order for the network to work at maximum design efficiency, which differs from other earlier P2P-approaches, which require almost all nodes to be known. Searches also run only at O(logN) complexity. The network is organized in a ring structure, each node has a predecessor and a sucessor, and knows exactly which nodes it should ask if a given key is searched. That information is saved in a finger table, which is filled with node information which are certain through a algorithm defined distance away from a current node. This makes chord so fast and stable, however it also hinders more dynamic routing and requires that the network fixes itself all the time when nodes join and leave.

Chord has been implemented as a official reference implementation in C++, but is also available in several other languages like C (portable) and Java.

3.2.2.2 Kademlia

This P2P-System has come out after chord, and could therefore learn alot of it. One of the things Kademlia [8] has over Chord is the ability to dynamically store more routing information, and therefore use information gathered while searching for keys.

Routing information is combined with the data that is searched for, meaning that the node id is the same as the key id. With that information, newly found nodes can directly be linked to keys that are searched for.

It also has a single routing algorithm, which can be used to find near-matches and exact-matches. In order to do that it has to be able to calculate the distance of two nodes/keys. That's possible with the help of the XOR algorithm, which is called that way since the XOR of two key ids will give back the distance between the two. Note that the distance does therefore not include information like latency or geolocation which could help network performance.

Kademlia is probably the most used DHT-Protocol, as the algorithm is used in Bittorrent programs like uTorrent or Vuze/Azureus, and other filesharing applications. In order to find the additional routing information, potential nodes can be pinged and then prioritize, though this is not a feature of kademlia itself.

3.2.2.3 Percolation

Percolation [9] search is based on observations of the topology of social networks. It's heavily dependent on random parallel and random length searches. Another feature are high degree nodes, nodes can act as high degree nodes to be used more heavily as others. This protocol also uses other notations, e.g. for query it says "query implantation".

3.3 Peer-to-Peer Anti-Spam Collaboration Introduction

In the following sections we will introduce the four proposed peer-to-peer systems for Anti-Spam. First the system proposed by Ernesto Damiani et al. [2] will be explained. Section 3.2 will illustrate another Anti-Spam-System proposed by Guoging Mo et al. [4], which is based on Multi-Agent-Interaction. After that the system proposed by Fei Wang et al [1] will be explained,an anti-spam based on the VOIP technology. The last proposals will address interest diversity, by Fang Weidong et al [3].

3.3.1 Three Tier System

This proposal uses the P2P-collaboration integrated with existing filtering methods for detection of spam messages. Through the P2P-based collaborative network important spam-information is shared between the peers (mail-servers).

This system is based on a network topology with a three-tier architecture. The first layer of this architecture is the **user tier**. The P2P network between the mail servers is placed on the second tier (**peer-tier**) and some mail servers can play the role of **super-peers** (on the third layer) [2].

The communication between users and peers is named as intra-cluster data communication, while inter-cluster data communication is the communication between the peers in the peer-to-peer network. The users in this approach do not participate as peers of P2P-network, because of their performance and privacy [2].

In the following we will explain each tier separately.

3.3.1.1 User tier

At this tier, users receive emails. There are two ways to report the fact that the received message is a spam: by clicking a spam or junk button offered by mail client or by addressing the message to a spam trap. If the user does not agree with the mail server judgment, whether a message is spam or not spam, she can send a *contrary report* to her mailer [2].

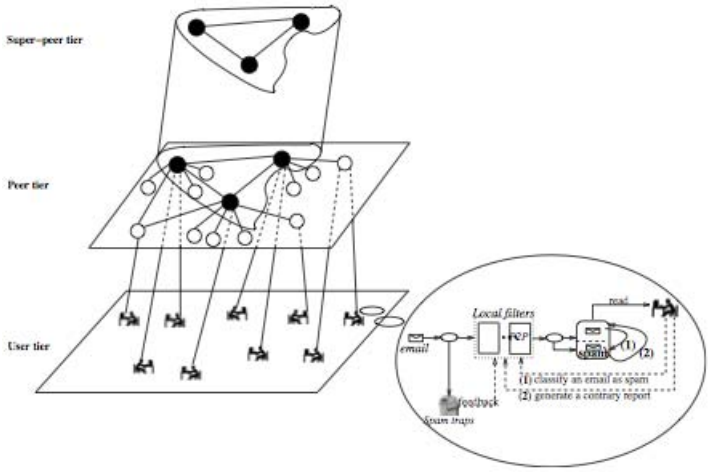

Figure 3.2: System Architecture [2]

3.3.1.2 Peer tier

At this layer the mail server receives emails directed to its users as well as contrary reports or spam notifications reports from its users [2]. This layer consists of its three main steps, that are: *Email processing, user spam report evaluation* and *user contrary report evaluation.*

3.3.1.3 Super-Peer tier

At the super-peer tier the mail servers serve as collectors and pollers of spam reports. They manage the spam reports and the inquires coming from the mail servers that refer to it or from other super-peers [2]. The two steps of this tier are the *spam report processing and spam inquiry processing.*

In this approach P2P is used in the peer-tier and super-peer tier. If a message in the peer-tier is identified as spam, then this information is shared with other peers. If a peer from the peer-tier needs information about a message, it sends a query to the third level (super-tier). If there are no information available for this message, the super-peer requires information from other super-peers.

3.3.2 Multi-Agent-Interaction

The main idea of this approach is to utilize the interaction between multi agents in a peer-to-peer system for spam filtering [4]. There are two kinds of agents: *social agent* and *local agent.* The social agent is located by mail server, while the local agent in mail user's host computer.

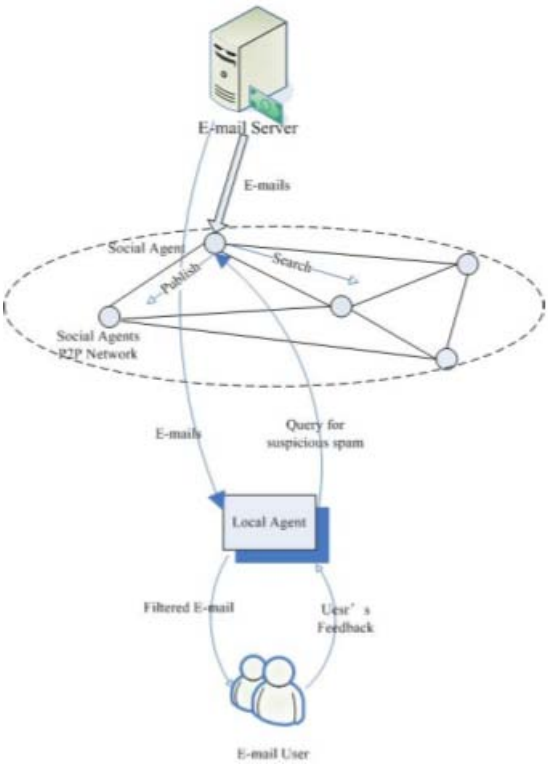

Figure 3.3: Multi-Agent System Architecture [4]

3.3.2.1 Social Agent

Social agent plays the role of the node (peer) in the peer-to-peer network. He is activated by two kinds of events: when an email arrives and by receiving a query from the local agent. This kind of agent takes the email digest, when an email comes, and publishes it to the P2P network [4]. This is done with help of digest agent and P2P agent, as part of the social agent.

Digest Agent – Digest agent calculate the digest of an incoming mail. For generating the mail digest Crossed Tri-gram [4] method is used. This method is like Nilsimsa code[5], but is more long than it since Chinese character set is more large than English. When a document is changed partly, only little bits of its digest would change [4].

P2P Agent – The job of this kind of agent is publishing the mail-digests to the P2P network. After this the digest should be stored at one of the existing peers in network. The P2P-network used here is an extend Chord[7], because the normal Chord support only exact query [4].

3.3.2.2 Local Agent

This kind of agent is activated when locale MUA (mail user agent) connects to email server. When the emails is downloaded to local agent, local agent uses Bayesian filter to classify the incoming mail. The filtered mail is then classified into *spam, not spam*

or *suspicious spam*. After that the emails are put into user's spam folder, inbox and suspicious spam folder respectively. Only the suspicious spam is sent to social agent for collaborative judgment [4].

3.3.2.3 Multi-Agent Interaction

By accepting a query coming from the local agent, the social agent computes its digest and compares it with other mail digests. The total number of approximate emails of the queried mail is found with the extended Chord API *LookupN (digest, range)*, the value of range is determined by local agent. If the LookupN returns a number that exceeds a threshold, then the email is classified as spam [4].

3.3.3 Interest Diversity

The following content is based on paper [3]. Everyone with an anti-spam filtering system for their email have had an experience where email was wrongly flagged as spam. A lot of times this results in email getting into the wrong folder an sometimes those emails don't even arrive at all. When email is wrongly identified as spam, it is called a "false positive" match (if spam gets in without them being detected it's a "false negative').

Usually this happens if a spam system detects email as spam even though it shouldn't according to certain mindset. However, most anti-spam filtering systems use the knowledge of some individuals to create some sort of spam detection lists, which intelligently can then be used to identify spam messages. Two basic anti-spam filtering system types can be identified which rely on such information, "content-based" filters like bayesian and "list-based" filters. The bayesian filters can learn what a personal actually thinks to be unwanted email, but in order to have some capabilities in detecting spam from day one, it's trained to detect the most obvious spam messages already. That training was made by some other individual who might have included spam messages which again others woudln't. The list based one is just a list of emails that probably are spam and is being put together again by some selected individuals.

The paper of Fang Weidong and Dong Shoubin tries to address the discrepancy between a persons interests and those of others. The main idea they are pursuing is by creating several interest groups which have certain interests in message categories. Then for new messages it will try to figure out what a persons (P2P-)neighbors with similar interests have flagged the message as, and use that information to flag the new message someone just received.

3.3.3.1 Message Categories

Message categories are used to group messages together. It has to be differentiated between categories which are defined by a person and categories probably used in the P2P

network itself, to which the person's own categories are only linked. That way it's possible to use different views on categories. Next categories are either viewed as full of spam or full of ham. The main algorithm used is measuring the similarity of spam/ham messages in a certain category between two peers. This paper suggest a strong bond between message, category and spam/ham decision, so that ite two messages can't be compared without putting them into a context first.

3.3.3.2 Architecture

"For every new spam message that is sent out to the network, there is a [email] user who first identifies it. Every subsequent [email] user [who] receives a suspicious message m, [...] can query his interest group corresponding to m's category to find if it has been already marked as spam." [3]

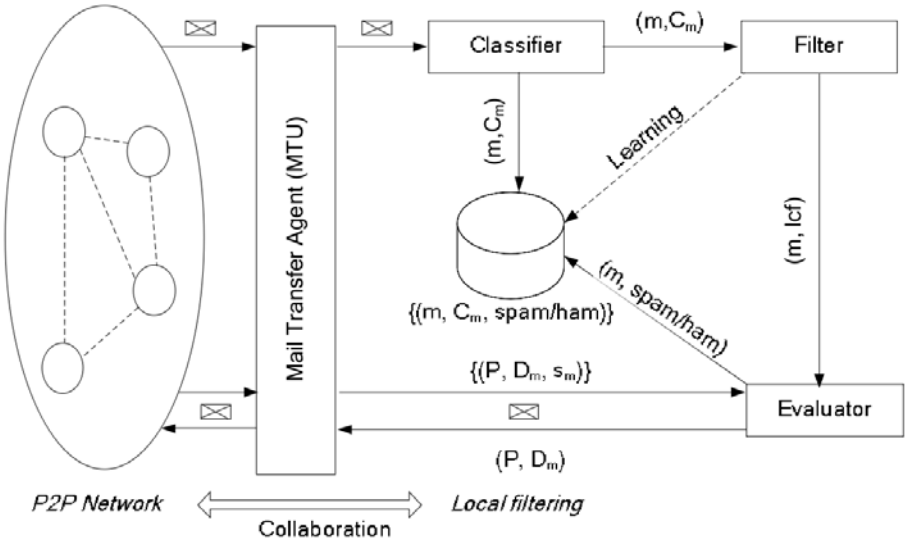

Figure 3.4: Interest Diversity Architecture

In order to understand the process a bit better, an overview of the architecture in Figure 3.4 is supplied. It begins by receiving an e-mail from a network:

1. Initial classifying
Before being submitted to the local spam filter, a message is classified into some category. It is proposed to implement a learning system, which can be trained so local categories match the P2P-network categories.

2. Local filtering
Messages are then being send through a filter, preferably a content filter like bayesian.

3. Signature indexing
This step shall make signatures out of the messages received, one signature for each message, which should uniquly identify the message. Since spam message can easily fool simple hash functions used for signatures, this step reminds the importance of a solid algorithm for the task. See Unique Identifiers

4. Final evaluation
This step is actually a combination of local filtering results and the results previously stored by others. In order to retrieve the results from others, the P2P network is queried, and uses the similarity calculation algorithm to combine all responses into a global score. then mix the two scores and value the local score more than the global score the more accurate the local score is. This usually happens if the bayesian filter has been trained a long enough time. After the evaluation, the local score will be stored and thus made available for the P2P network.

3.3.4 Anti-Voice-Spam

Currently, advertisement over telephony is a common thing to happen, very often there are several companies usually through call centers, call private persons and ask them to buy some of their products. In order to counter this, there are means to prevent them from calling. Switzerland has, as of writing this document, no law disallowing any unwanted advertisement calls to it's citizens, there merely exists a way to mark a person inside the national wide telephone book, so that call centers should not call that person. However this mark does not obligate them to not call that certan person, just suggest them to not call them. Germany is a lot more consumer friendly, as it is strictly forbidden to call and advertise over telephone, if not explicitly wanted by the callee.

There are several ways to counter advertisement calls:

- Block lists on the phone. Only works for repetitive callers, call-centers usually have whole ranges of numbers they can use to call.
- Change the phone number. One time permanent solution until the new number will be known again by spammers.
- Try to find out if a country/organisation has a list of call centers which can be used to block those numbers on the telephone provider side. Usually organizations defending consumer rights are putting such lists on their web sites.

SPIT means spam over Internet telephony. The following content is heavily based on paper [1]. While all of the above was written with the conventional telephone in mind, it still applies to new technology. Not only that, but there is almost zero call costs which allow spammers to call as many times as they want. Not counting skype, most protocols used for Internet telephony are open. This allows to write telephone applications which can directly connect to everyone's infrastructure and play advertisement. Skype goes around this problem by using a encrypted standard coupled with a user authorization system (only needed for the free Internet to Internet call system).

3.3.4.1 Anti-Voice-Spam Proposal

The main idea of the proposed Anti-Voice-Spam system [1] was to create an open distributed architecture to store and share feedback, just like all proposals before. The main

difference is that it tries to combat voice-spam instead of mail-spam, and since primarily differs from email that it only occurs interactively. Additionally, not the message will be blocked, the caller will. The caller will be identified through his number, which assumes that spam calls from the same organization or person will come primarily from the same number.

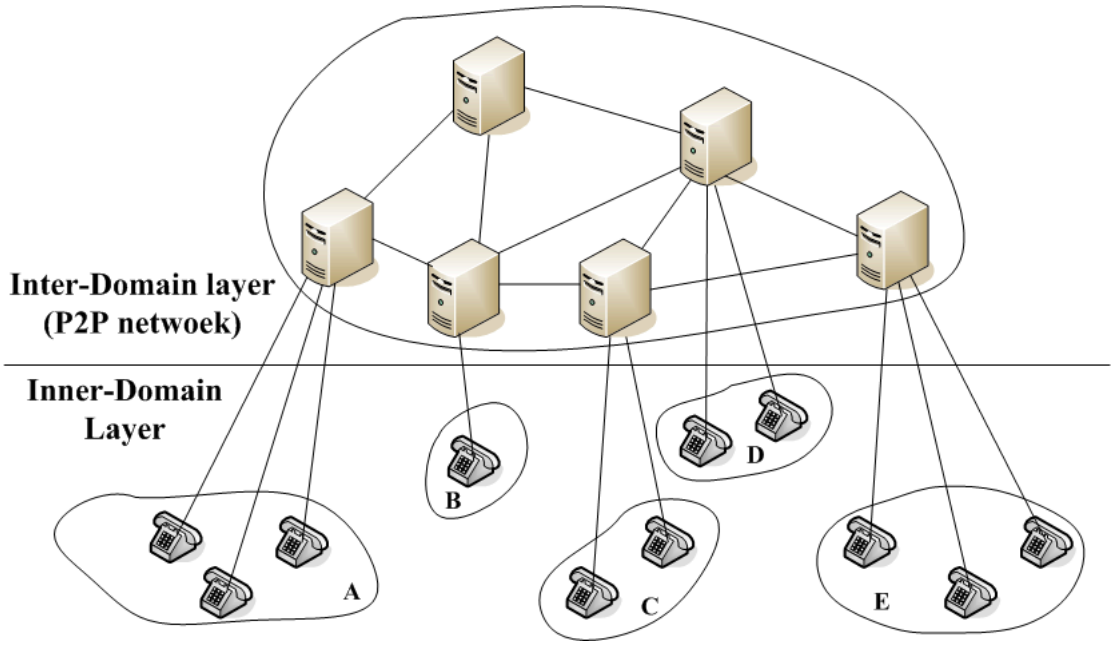

Figure 3.5: Peer-to-Peer Anti-Voice-System

Architecture wise (Figure 3.5) there are two parts, the inter-domain and the inner-domain layers. The inter-domain layer interconnects all voip-servers which have the P2P-AVS (system running aside. The inner-domain layer consists of all voip clients, usually voip-phones or similar devices which are directly connected to the voip-servers in a traditional client-server fashion. It is assumed that this allows the user-phones to stay untouched of the needed P2P-feedback system modifications, while only the servers have to be adapted. Although the user still needs a way to convey his feedback back to the server, the normal voip-phone has already enough functionality to accommodate this, like pressing a certain number after the call ended.

The paper further suggest using a DHT P2P system like Kademlia [8] or chord [7]. Non-feedback related information however will be directly stored on the server responsible for the user (and his number) and not on the phone or in the P2P network, this information includes the black and white lists and the user mode (see below for further explanation).

An interesting part is how this paper identifies and verifies the users for feedback inclusion into the paper. Each server will have a pair of unique public and private keys, each feedback given into the P2P-network will be signed with the private key. Signed information includes the callee and his evaluation about the call he received. The evaluation consists itself presumably of the caller number and if he is regarded as spammer or not by the callee.

To calculate if the caller is indeed a spammer, there are four steps which will lead to the reputation (in this paper reputation refers to the caller not callee, meaning the aledged spammer and not the reporter of spam):

1. The first step is to find the caller number on the black or white list, similar to black or white list for email.
2. Next step, and this is unique to the interactive form of communication, is the user mode evaluation. If the user is on "do not disturb", the call will be automatically blocked, like if the caller would be on the black list.
3. Then comes the call model into play. Previous calls can be used to get call intensity (how often) and call lengths and the paper suggest that these variables can indicate if a call can be spam or not.
4. At the end the subjective evaluation from the user is included.

3.4 Summary

Email is nowadays the most important communication method. However, since it is open to be used by everyone, some aspects like advertising became a nuisance. As everyone who uses that medium is similar affected and the wish to have less unwanted email (known as spam) is mutually shared between the users, collaborating, in order to fight the problem, is something people want to do. And indeed, alot of e-mail anti-spam systems work on the basis that a group of people flag emails as spam before they arrive at users' mailboxes. Known working systems are Razor or DCC, which run under the hood of or are at least plugged into an anti-email-spam system. However not every participant of such a system will be allowed to help identifying spam, as there are still central controls. In this seminar paper, we will discuss solutions, which work on peer-to-peer technology, to gather evaluations and feedbacks of every user. Most of the solutions are proposals constructed after conference gatherings, so they don't represent fully working systems.

There are four of peer-to-peer anti-spam proposals which all have unique properties, which could be combined to an advanced working system. One of them explains a three-tier system, and allows different points of inserting anti-spam methods. Anti-Voice-Spam explains the difference between email and voice, with the only difference being in the interactivity of telephone communication, which doesn't allow for content evaluation before the call. The Multi-Agent proposal explains how to attach P2P functionality by just adding two different "agents" onto an existing system. The last paper, Interest Diversity, then tries to solve the problem of different people having different opinions while choosing what email should be spam or ham.

Still, two key technologies were shared intensively between the proposals, one was the need of an algorithm for creating unique identifiers, the other are the peer-to-peer protocols. Unique identifiers are required to share email evaluation data between different users. Since spam has properties of not being exactly the same in between users, evaluation of

emails that differ only in small bits, still need to produce the same identifier. Of the peer-to-peer protocols, the generally most widely used is Kademlia, and although some proposals use Percolation or Chord, there is no difference as it is just required to store evaluation data in a network, and the currently DHT-based P2P protocols are able to do exactly that.

A fully working system has not field tested before, but every proposal has shown that peer-to-peer technology can be added to existing anti-spam technology to improved it for defeating spam. Accuracy and detection was raised and made better. However new problems of the technology emerged, privacy issues being the worst, which weren't discussed but would need more in depth research. Also how good everything would scale and how deceptable everything is to attacks wasn't discussed by the proposals either (Three Tier excluded).

Bibliography

[1] F.Wang, Y.Mo, B. Huang: *P2P-AVS: P2P Based Cooperative VoIP Spam Filtering* Wireless Communications and Networking Conference, 2007.WCNC 2007, pp.3547-3552, Hong Kon, 11-15 March 2007.

[2] E. Damiani, S. De Capitani di Vimercati, S. Paraboschi, P. Samarati: *P2P-based collaborative spam detection and filtering*, Fourth International Conference on Intelligent Agent Technology, 2006. IAT '06, pp.428-431, Hong-Kong, 18-22 Dec. 2006.

[3] F. Weidong, D. Shoubin: *Addressing Interest Diversity in P2P Based Collaborative Spam Filtering*, Fifth International Conference on Grid and Cooperative Computing Workshops, 2006. GCCW '06, pp.163-169, ChangSha, China, Oct. 2006.

[4] G. Mo, W.Zhao, H.Cao, J. Dong: *Multi-agent Interaction Based Collaborative P2P System for Fighting Spam*, IEEE/WIC/ACM International Conference on Intelligent Agent Technology, 2006. IAT '06, pp.428-431, Hong-Kong, 18-12 Dec. 2006.

[5] D. Dong, J. Zhang, Y. Wu, H. Lu, G. Wu: *A Spam Filter System Based on P2P Architecture*, International Conference on Networking, Architecture, and Storage, 2008. NAS '08, pp.155-156, Chong'Qing, China, 12-14 June 2008.

[6] E. Damiani, S. De Capitani di Vimercati, S. Paraboschi, P. Samarati: *An Open Digest-based Technique for Spam Detection*, Proc. of 2004 Int. Workshop on Security in Parallel and Distributed Systems, ISCA PDCS 2004, pp.559-564, San Francisco USA, 15-17 September 2004

[7] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, H. Balakrishnan: *Chord: A scalable peer-to-peer lookup service for Internet applications*, Proc. of ACM SIGCOMM 01' Conference, pp.149-160, San Diego California, August 2001

[8] P. Maymounkov, D. Mazieres: *Kademlia: A Peer-to-peer Information System Based on the XOR Metric*, Conference IPTPS 2002, 2002

[9] J. S. Kong, P. O. Boykiny, B. A. Rezaei, N. Sarshar, V. P. Roychowdhury: *Scalable and Reliable Collaborative Spam Filters: Harnessing the Global Social Email Networks*, Conference CEAS KongBRSR05, 2005

[10] Wikipedia: *E-Mail*, Found on 24.03.2009 under: http://en.wikipedia.org/wiki/E-Mail.

[11] Wikipedia: *Voice over Internet Protocol*, Found on 24.03.2009 under: http://en.wikipedia.org/wiki/Voice over IP.

[12] Wikipedia: *Peer-to-peer*, Found on 11.05.2009 under: http://en.wikipedia.org/wiki/Peer-to-peer.

[13] Information Technology Services: *Metrics Email spam statistics*, Yale University, Found on 17.04.2009 under: http://www.yale.edu/its/email/statistics.html.

[14] Burkhard Stiller: *P2P*, Lecture, University of Zürich, FS 2009.

[15] TMDA: Tagged message delivery agent. Found under http://tmda.net/.

3.5 Appendix

People/System	Fang Weidong Dong Shoubin	Guoqing Mo Wei Zhao	Ernesto Damiani Sabrina (..) Stefano Paraboschi Pierangela Samarati	Fei Wang Yijun Mo Benxiong Huang
Categories				
id	P2P Collaboration for E-Mail Anti-Spam			P2P-AVS
	#1	#2	#3	
base technologies	smtp, p2p	smtp, p2p	smtp, p2p	sip (p2p)
type of architecture	fully decentral p2p	p2p for spam info only		inner and inter domain
goals	addresses interest diversity to lower false positives	easy to understand system??	- compatibility for current infrastructure - false positive absence - collaborative identification of spam - robustness against attacks by spammers - protection of email confidentiality	builds on existing sip technology, find a way to block voice/real time spam
achieves the goal ... how?	uses categories to put users in virtual groups which represents his interessts the most	split of local and socialagent for easier understanding	- three-tier architecture (compatiblity, ~false postivies, p2p on super peer tier, thrust system on super-peer level intern and extern of that tier)	reputation module, historical data, user modes, statistical data about call
user actions	- categorize - probably final spam/ham info given to local content filter	- „user feedback“	- mark message as spam	- set user mode - user evaluation
client/peer actions	- local filter, e.g. content based - signature indexing - „final evaluation“	local agent (included into email app e.g.): -(pre) filtering - „user feedback“	mail server (peer): - query on copies threshold, and request spam info set (returns similar message digets with the originating servers ids) - weight that info for evaluation - take on user evaluation - add to own spam/contrary catalog - report to super seeders	- black/white list info sending based on evaluation or feedback - call model eval.
server/superpeer actions	- (not known, maybe not needed)	social agent: - digest agent - p2p agent(s)	- collect and allow polling - spam report processing - spam inquiry processing	- black/white list organisation and blocking of spam calls..
spam or ham decision intersections	- local filter - „final evaluation“, global and local points	- local filter (spamassassin) - p2p network info - user feedback (Rocchio algorithm works with baysian filter)	- spam check - mail server (threshold -> query) - super peer (info on all reports back) - user evaluation (spam or ham)	- black/white lists - user mode - call model - user evaluation
reporting system	- digest implantation at random walk size of L(N, t) - takes all clients/peers into account in range	- feedback from user goes to ?	- feedback to email client - or spam trap - goes to the „mailer“ / Peer Tier	- send white/blacklist info to servers - send eval of callee signed with private key
digest calculation	- (not SHA1, references indirectly nilsimsa)	- nilsimsa modified for chinese characters	- (says there exists one, talking in a technical paper unavailable to us)	- callee number
spam handling	- just calculate score, do not handle (but maybe give other filters)		- not specified what has to be done, we assume custom filter measurements will apply.	- create white/black list on demand
querying (after spreading)	- query implantation, random walk of L(N, t) + Bond perculation	- chord modified for inexact match (digest range)	- spam inquiry from mail server (peer) to super peer	spreading unknown domain servers responsible for blocking? (no queriyng needed?)
(p2p) lookup algorithm	perculation search	- chord	- something in direction of chord or kademlia	TODO
additional info	- hit rate and false positive reduction graphs	some graphs about results	rebustness because of inter super-peer connection and reputation (even though mail server has single connection)	- graphs for accuracy, learning rate, etc..
problems	doesn't take misuse of system into account	false positives can't be undone	- super peer not based on existing technology (well smtp technology that is)	- misuse system is assume to work on reputation of spaminfo providers
specials	dynamic group calculation	china modifications, simple and new naming	three tier, contrary reports	voip, simply callee based
p2p style	dht p2p	dht p2p	hybrid + dht p2p	hybrid + dht p2p

Chapter 4

Honeypots

Daniel Meier, Stefan Badertscher

The following paper gives an introduction to the topic of honeypots. Apart from the basic principles and the understanding of what honeypots are, it also covers the different shapes that honeypots can adopt. Honeypots can show up in many different types, such as an E-Mail or database entrys. There are some concrete examples, where honeypots are used to detect attacks from the outside and also from the inside of an organization. This paper covers the technical details only very briefly, because it should be considered more as an introduction to the topic of "honeypots" with the focus on ideas and strategies implementing various types of honeypots.

Contents

4.1 Introduction and Motivation

In this paper we will start with a short introduction followed by a definition of honeypots and shapes they can assume. After we work out the principles to understand the basic functionalities of honeypots, we will further have a look on some special aspects and features that honeypots can have. This means, that later we will discuss which tasks are adequate to be solved by honeypots and what makes this technique so powerful. We will demonstrate this on some concrete examples. In this paper we are not going much into technical details; instead we focus on how honeypots can be used to detect malicious activities. This also offers several aspects which can be used in further papers.

4.1.1 Motivation and why we have chosen this task

We have chosen this talk, because we both are interested in security aspects of networks. Also regarding that the number of malicious software such as viruses and worms is permanently increasing showed us that security aspects are becoming more and more important in the future. Nowadays signature-based as well as behavioral virus scanners are always in competition with developers of malicious software according to a classic cat-and-mouse game [4]. Unfortunately the producers of malware are always one step ahead of the anti-virus companies, because they have the initiative [13]. So it's unlikely that this cat-and-mouse game can be won by the camp who wants to make virtual world safer. This fact requires new techniques being applied. Honeypots are a relatively new technique, first serious research started in 1997 by Fred Cohen [10]. In this case Honeypots seem to be a really interesting and viable approach to us, because they can be used in various situations.

4.1.2 What are honeypots?

Let's start the main part of this paper with the following statement found on a security forum of over 5'000 security professionals [11], [10]. Later we will see how this definition makes honeypots to an extremely powerful technology.

> *"A honeypot is an information system resource whose value lies in unauthorized or illicit use of that resource."*

By definition, honeypots are decoy computer resources [6] with the main goal of attracting people who want to attack or compromise a computing resource or a network. Practically this means that no one is allowed to interact with it like reading entries or writing something on it [10]. Any interaction with it is per definition an unauthorized access and should be treated as suspicious. Basically when we talk about honeypots, they can be roughly partitioned into two categories: On one side we have physical honeypots. They are real computers, which have their own IP-address in a network [9]. On the other side there are virtual honeypots which are not real machines any more. In fact, it's a simulation of network resources which takes place on an existing machine. In this case realistic

networks are simulated including packet loss and latency. It's also possible to run several simulations of honeypots on the same machine at the same time. This is as a big advantage in comparison with physical honeypots, because you do not need to have for every honeypot a single computer system. The lower hardware requirements help to reduce the costs for the hardware investment and for the maintenance. If a large IP-address space is needed, it's nearly impossible to use physical honeypots, because the costs would increase dramatically. In this case it's necessary to choose a virtual honeypot [9]. For handling such tasks, there is a powerful framework called Honeyd to which we refer later in this paper.

However honeypots are much more, they can adopt various shapes and they do not need to be computers at all. In fact, they nearly can be everything [10], [11], [12]. In the listing below, we collected some shapes honeypots can take, but of course the list is open-ended:

- Windows programs
- Excel spreadsheets
- PowerPoint presentations
- Credit Card numbers
- Database entries
- Logins and passwords
- E-Mails
- and many more...

Later we will have a precise look on using such types of honeypots (called honeytoken) as a detection method of insider threats and also we will show some specific situations where honeypots can be utilized.

4.2 How honeypots can look like & their strengths and weaknesses

Further honeypots can also be roughly distinguished into low-interaction and high-interaction honeypots. High-interaction honeynets offer all the services, programs and operating systems in a real manner, meanwhile low-interaction honeynets offer them in a simulated way [13].

4.2.1 Honeynets

Honeynets can be defined as following [7]: "A Honeynet is a network, placed behind a reverse firewall that captures all inbound and outbound data. The reverse firewall limits the amount of malicious traffic that can leave the Honeynet. A Honeynet is a network that is intended to be compromised, to provide the system administrator with intelligence about vulnerabilities and compromises within the network." Networks such as from a university or a company have a large throughput. This fact makes them very attractive for attackers. "Owning" such a network offers a huge amount of criminal possibilities an attacker can utilize. The large throughput of the network makes it very difficult to collect and analyze the entire traffic passing through. So a honeynet can be a solution, because by its definition it should have no traffic. Every access form the enterprise network to a computer in the honeynet probably indicates a compromised enterprise system [7]. Honeynets can be separated into first generation and second generation Honeynets. The following table shows the differences between them. Table 4.1 shows some selected features of 1^{st} generation and 2^{nd} generation honeypots [7].

Table 4.1: Advantages of honeypots

Characteristics	1^{st} generation	2^{nd} generation
First developed in	1999	2002
Method of data control	limiting outbound connections	examining outbound data
Highly effective for	automated / beginner attacks	advanced target attacks
Weakness	fingerprinted as a honeynet by a hacker	more complex to install and maintain

It's important to mention, that you cannot state that a 1^{st} generation Honeynet is worse than a 2^{nd} generation Honeynet. It mainly depends on the context and the goal of its use. Another important influence is the cost factor for maintenance of a Honeynet.

Honeynets are an architecture. It's basically a network of two or more honeypots together, which are physical machines. Honeynets are high-interaction honeypots, because they consist of operating systems, applications and services which are ready to interact with an attacker [13]. More precisely, any operating system you want can be installed on a single honeypot in this network and be used as a decoy resource, no matter if you want to create an E-Commerce Site on a Windows XP System or maintain a database on Debian Linux. Because the Honeynet should not generate traffic (by its definition, a honeypot is only accessed in an unauthorized mode), any traffic is assumed being suspicious. So in comparison to traditional firewall logfiles, you do not have a large quantity of (quite often false) alerts. Traditional firewall solutions have also some serious weaknesses [7]:

- Inability to protect against attacks which bypass the firewall.
- A firewall does not protect against threats which are taking place on the inside of a network.
- When a virus is transferred, the stand-alone firewall solution will not detect it.

A great metaphor which is used often in literature is the comparison between a honeynet and a fishbowl [10], [13]. In a fishbowl one can add almost everything. It's possible to build a little world, which can be watched and observed from the outside. In a honeynet the same things are practicable. Instead of adding fishes and algae, you can add applications and devices to a honeynet. From now on, these things can be observed as just like you do with a fishbowl. To clarify the architecture of a Honeynet, Figure 4.1 is helpful.

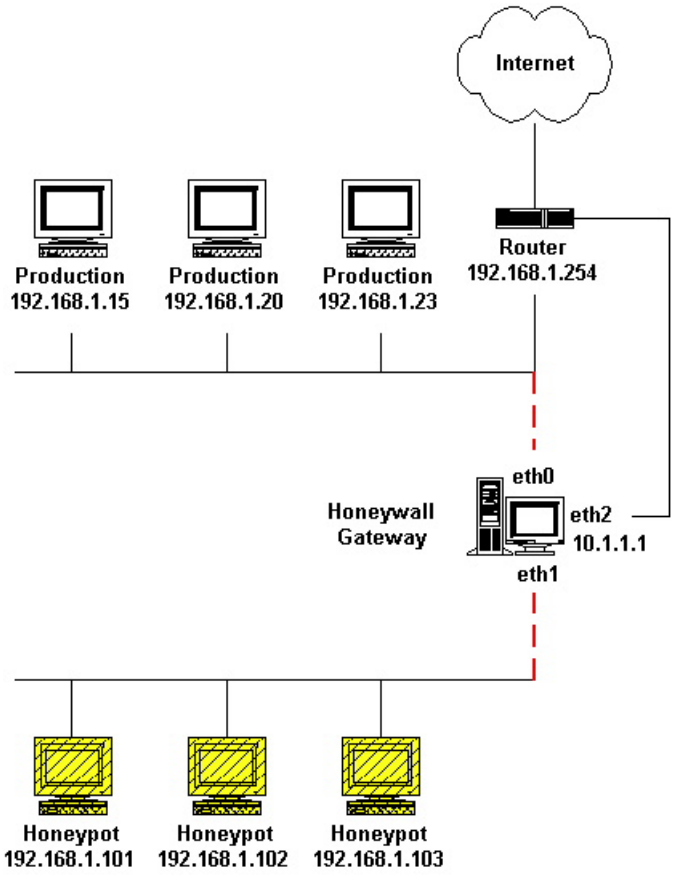

Figure 4.1: Typical topolgy of a 2^{nd} generation honeynet. Source [10], [13]

The key component in a Honeynet is the gateway device called Honeywall Gateway [13]. This gateway device separates the Honeynet from the other components of the network. The intent is, that any traffic has to pass the Honeywall Gateway, if its destination is a machine inside the Honeynet. The gateway is "traditionally a layer 2 bridging device, meaning the device is invisible to anyone interacting with the honeypots" [13]. In this example there are three network interfaces, namely: eth0, eth1 and eth2. The first two interfaces (eth0, eth1) separate the two pseudo subnets from each other. The third interface (eth2) can be used for remote administration. As we already have seen, the Honeywall Gateway is the most important component in the Honeynet. This is why it has to implement several important features; without them it's impossible to gain some benefit of such a configuration. Following features have to be implemented on the Honeywall Gateway [13], [7], [10]:

- **Data Control:** The purpose of data control is to minimize the risk that an attacker can infect other non honeypot systems; it also necessary to ensure that the honeynet

is not uncovered. In this context it's necessary to make a tradeoff between the possibilities one wants to give an attacker versus the restrictions in activities that an attacker is allowed to perform. There are some common options to do so like limiting the number of outgoing connections, bandwidth restrictions and an intrusion preventing gateway. These methods allows to adjust the risk an organization takes. Although these precautions can be made, it's impossible to eliminate the whole risk that an attacker can harm productive systems. It's evident that the risk level is always depending on the purpose of a Honeynet and it always should be observed by a network administrator.

- **Data Capture:** Data capture refers to the logging and monitoring facilities in a Honeynet with the goal collecting as much data as possible of a threat. These logging facilities have to be hidden carefully from an attacker. Some important precautions are that the logged data should not be stored on the honeypot itself, instead it should be transferred to a secure system like a remote log server. If the data is stored on the honeypot, it could be easily detected and manipulated by an attacker. Another importance is that the honeypot is not modified too much, because the more modifications are done, the higher the chance that a honeypot is uncovered.

- **Data Analysis:** Data analysis is the information one has gathered and generated from the collected data. It always depends on the needs of an organization, so it's evident that different organizations have different methods of data analysis.

- **Data Collection:** Data collection refers to the collected data of multiple distributed honeynets. These data can be combined at a central location which offers the possibility to do better and deeper threat analysis. The value of the data depends significantly on the way how it's combined and evaluated.

These requirements show that setting up a Honeynet is not such a simple affair. Mainly building a Honeywall is complex and time consuming. Luckily there is a bootable CD-ROM, called "Honeywall CD-ROM", where these features are already implemented. This reduces the time effort for setting up a Honeynet dramatically and so it should be widely applicable for organizations [14].

4.2.2 Honeytokens

In this section we discuss honeytokens. This term already appeared several times, now we are going to take a closer look at it. The term honeytoken can be roughly derived from the definition of a honeypot. In this case, the honeypot does not have to be a computing resource anymore. A honeytoken is just a specific resource, which is used as a decoy. Some well known shapes of honeytokens are the following ones [11]:

- Credit Card Number

- PowerPoint presentation

- Excel spreadsheet
- Word document
- Login and password data
- Database entries
- nearly everything...

The value of a honeytoken lies in the matter of fact that every access or interaction with it is per default suspicious and unauthorized. But honeytokens have extended capabilities. The main purpose of a honeytoken is to detect insider threats [10], [11]. This means, that its design is ideal to detect if a user inside a company violates for example data access restrictions. Honeytokens are not used for detecting attacks from the outside, because this would not be adequate and too dangerous in respect to security reasons. Honeytokens have a great potential to detect users who do not behave like they should [11]. With traditional security methods, it's nearly impossible to detect for example a system administrator who looks after restricted information which he is not allowed to see. System administrators have generally many security and access rights, because they have to guarantee that everything runs smooth on the network. In such cases access restriction and control is not effective, because a skilled system administrator could get around those restrictions or if that is impossible, he can ask the person in charge to allow him access due to some pretended system problems. So with traditional methods it's quite hard to detect such violations. If there are some honeytokens placed on those information systems, it's much more likely to detect such an unauthorized action. However like traditional honeypots, honeytokens are not made for solving specific problems like attack prevention [11]. They are designed to detect unauthorized access and malicious behaviors of "trusted" organization members. It's important that a honeytoken is placed within other information and that it is designed in a way that it attempts attraction without being suspicious. To prove a deliberate irregularity, you have to ensure that the honeytoken is not only accessed, but also that it has to be used. Accessing specific information can happen unintentionally if a confused employee is looking for things which he normally should not [11], but it is still a big difference if that particular employee just copied the information or is trying to make use of it. To be completely sure what kind of violation it is, there is the use of other methods to designate the real intent. The most important fact is, that the activity itself is detected and that it can be tracked from then on.

Surprisingly the underlying concept of honeytokens is nothing new. The principle was already used many years ago (sometimes still nowadays) for example implementing "Easter eggs" in programs [11]. The idea behind it was not to protect the source code from unauthorized resell or copying, instead the idea was to track it. If some undocumented special features were implemented in the code, the creator could easily verify if it was his own work or not in a case of a suspicion. It's just necessary to check if the program implements such a particular special feature. If so, it's proven that it was used in an illegitimate way. This idea is the core of the honeytokens functionality. Now it becomes very clear, why they can be used for numerous tasks. Later some specific examples

of placing a honeytoken will be shown and described. Beside from its high potential, honeytokens are also very cost effective, it's even possible to say that the costs are minimal [11]. There is no need in any extra investment or in buying some special security software. Just create a honeytoken, which can be for example a PowerPoint presentation [11], or an Excel spreadsheet. Just place within some pretended and faked confidential information, then just lean back and to wait until someone tries to use the contained information. The second big advantage of honeytokens is the high flexibility. Because they are information, there is no limitation of what a honeytoken can be. The only limit is one's imagination [11].

4.2.3 Virtual honeypots

As already mentioned in the introduction, this section is dedicated to Honeyd, a framework for creating virtual honeypots. The advantage of virtual honeypots is that they are not real machines, as the name tells us. It's obvious that hardware investments and maintenance costs are lower in comparison to physical honeypots [9]. Usually a network with various operating systems is simulated. Honeyd is also very powerful in detecting worms. Honeyd is a low-interaction honeypot, which can simulate ICMP, TCP and UDP services. One of the main features of Honeyd is the possibility to simulate thousands of virtual machines including their corresponding network services. This enables the possibility to utilize a large amount of IP-Address space [9]. Honeyd is also capable to simulate the network stack of many different operating systems. In the following Figure 4.2 demonstrates a graphic where the principle of a tree network topology is shown [9]:

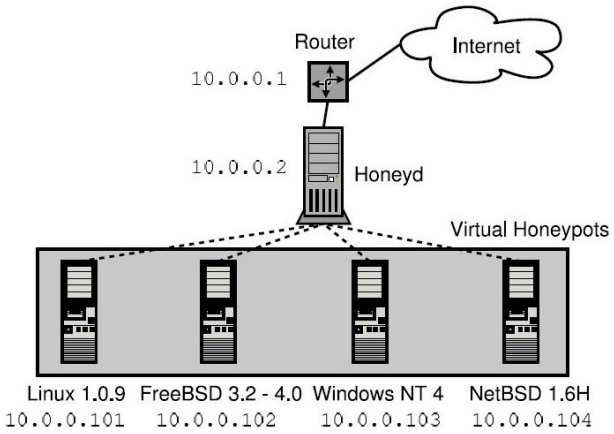

Figure 4.2: Overview of Honeyd simulating four virtual honeypots with different operating systems

Later we will see how this principle can be used in worm detection.

4.2.4 Strengths and weaknesses of honeypots

Table 4.2 and 4.3 show the advantages and disadvantages of honeypots, which have been summarized from the explanations in the preceding sections and also by using the article

of Spitzner [10]. In the first table the advantages are gathered with a short description of them and in the second table the disadvantages are also gathered in the same manner. That is followed by a discussion about these advantages and disadvantages:

Table 4.2: Advantages of honeypots

	Advantages	Reason
1.	Small data sets	Per definition the interaction with a honeypot is suspicious, so the collected data is of high value
2.	Easy change of its identity	Possibility to emulate different operating systems or services, and monitoring different IP-Addresses
3.	Highly flexible	Can be used in various situations, because it's easy to change their appearance
4.	A few false positives	Per definition every interaction with a honeypot is unauthorized, so one can be sure that in fact it is an attack on the system
5.	A few false negatives	Honeypots can detect unknown attacks, only due its definition
6.	Handling of encryption	Because the honeypot is an end point, the activity is decrypted. So it also can be used in encrypted environments (e.g. SSL, IPsec) of organizations
7.	Minimal ressources	Only little ressources are needed. Already an old computer can monitor millions of IP-Addresses

Table 4.3: Disadvantages of honeypots

	Disadvantages	Reason
1.	Risk of other attacks	An attacker could use a honeypot to gain access to other non-honeypots systems
2.	Limited field of view	Honeypots are only powerful in the direct interaction with them. Otherwise they only have a little value, because they do not monitor every traffic.
3.	Beeing discovered	If a honeypot is beeing discovered by an attacker, it can be used to generate false informations.

It's obvious from the list above and the short description of the advantages and disadvantages that there are much more advantages than disadvantages. Especially notable is, that honeypots only use little resources. It's said, that already an aging Pentium computer can handle a big amount of IP-Addresses [10]. As already mentioned in the introduction, this reason gets more and more important, as for example antivirus software is requesting more and more resources and hard disk space for signatures. It's evident, that a new technique for security is really to appreciate. The honeypot's definition that every interaction with it is suspicious and unauthorized by default, has an astonishing high value. The definition causes that the data generated by a honeypot has an extremely high quality, because only people who want to utilize confidential information or want to get access

to an unprotected system would even try to interact with a honeypot. "Normal users" who are not interested in illegal activities do not try to access them. So the data collected by honeypots does not contain a lot of so called false positive entries, which have to be filtered out before the data examination. Through evaluation of such data sets, it's possible to learn a lot about the behavior of attackers and the methods they're using. It's possible to track exactly which information is accessed in case of an insider threat or just which methods are used by trying to get access to an unprotected system. Another feature is the possibility to detect so far unknown styles of attack and exploitation. Nowadays companies use more and more encrypted connections in their organization [10]. As already mentioned before IPSec, SSL and SSH are used in these cases. Breaking into such connections and sniffing the traffic is a high effort and often takes long time to achieve or is even impossible. Using honeypots, this problem is solved in a very elegant manner. Due to the fact that the honeypot is the target or the end point, the traffic is always decrypted on it and can be easily evaluated by the honeypot operator. This advantage is very important when one has to catch an insider who is trying to get around data access restrictions and wants to get information which he is not allowed to access. Such offences are very difficult to detect by traditional inquiries. But by utilizing a specific resource like a prepared database entry (honeytoken), it's very comfortable just to wait until someone tries to access it. This principle of passivity is a powerful instrument in this contest. Further it's possible to change easily the identification of a honeypot [10]. It's not a big deal to just change the emulation of an Operating System to another (e.g. from Windows to Mac OS). Moreover it effortless to create another file or another database entry when using it as a honeytoken. Also new IP-Addresses can be allocated quickly and without much work. So the same infrastructure of a honeypot can be used several times, even if the honeypot was discovered. Being discovered is a possible weakness of a honeypot. In case of a honeytoken, the attacker can willingly create false information and so the honeypot records comprise useless data. Also the records about the behavior of an attacker are useless, because he behaves in a very different manner when he knows that he's interacting with a honeypot. It's pretty obvious that camouflage is essential and should be set up carefully and adequate to the applied situation where a honeypot is deployed. Unfortunately, honeypots have also some disadvantages as we can derive from the second table. A honeytoken used inside an organization has a total different purpose than a honeypot which has the goal of attracting attackers to break into it. Being aware of the different aims are the best assumptions for setting up the camouflage for the honeypot. It's better to use many different styles of honeypots for each little task, then only a few for some bigger tasks. This corresponds with the conclusion that a honeypot has a view like a microscope [10]. A honeypot cannot control the whole traffic running through a system, it only gets a very close view on interactions where it's involved. If nobody interacts with a honeypot, its existence is quite useless and the only information you can gather, is that it has to be presented in another way to potential attackers. For example in a honeynet there is a quite a high risk that the attacker can get access to other systems in the surrounding of a honeypot, if the environment is not secured enough. So it's important to check always what's going on in a honeypot and it has to be absolutely clear for what purposes it's being used. Also it's good to apply human monitoring and customization [13]. Human monitoring means that you have special trained staff that constantly checks that the system is running smooth. The comparison of the advantages and disadvantages shows a clear trend in favor of the positive aspects. The disadvantages

are of relatively little weight and their risks can be strongly reduced. Because attacks in a network are often caused intentionally and malicious, it's important to be cleverer than the attackers [15]. It turns out, that honeypots are a good concept for dealing with security aspects in a clever manner.

4.3 Honeypots in action

In this section we cover some examples of using the technique of honeypots and honeytokens. First we show two examples of the usage of honeytokens placed in an E-Mail or a database. After that we discuss the benefits of using honeypots and virtual honeypots in the area of worm detection.

4.3.1 Honeytoken placed in an E-Mail

The possibility to place a honeytoken in an E-Mail is important to catch an insider threat [10]. An insider threat means that one is interested in detecting a threat which does not come from the outside but from the inside, for example an employee who looks after information in a company network which he is not permitted to access. This kind of threat is very difficult to detect with traditional security solutions. System administrators or other persons who have to guarantee correct working computer systems, have often broad rights in accessing data and security settings. To check if a person behaves in the right manner, a lure can be placed in order to attract such a misbehaving individual. In this section the lure is a honeytoken placed in an E-Mail. Then, this E-Mail can be sent to an E-Mail account which is suspected to be unsafe. Such a "honeytoken E-Mail" could be formulated in the following way:

```
Date:       Mon, 07 Jun 2008 14:01:55
From:      ''Don Bowie, CFO''
           <bowied@sachsman-silver.com>
To:        ''Hon Eypotter, CEO''
           <eypotterh@sachsman-silver.com>
Subject:    Business forecast

Dear Mr. Hon Eypotter

I'm glad to inform you, that finally all the documents you requested from the
Administrative Board have been collected on a secure location. I just checked
the reports. The business forecast for the next six months seem to be really bad.
There are also some delicate issues documented which really could menace the
future of Sachsman-Silver.
You will find this information on the following secure location:

        URL:     ftp://secure.silversachs.com:21/
        Login:   secretDocs
        Pass:    d@TzC6Y{saP

With the best wishes
Don Bowie
```

If this E-Mail is read by an unauthorized individual, the probability is quite high that this login and password is used for the given FTP-Server. The clue here is, that the given

FTP-Server is working properly with this login and password combination. There are also some files placed that seem to be important, but the contents are freely faked. So the individual who accesses the files should think that he has found valuable information, but in fact the documents are futile. With this example the concept of a honeytoken becomes intuitively clear. In this example no one except the CEO and CFO should know the login data and therefore no one except these two should try to access the files. (It's possible that the CFO does not know everything about it and his name is used without his permission.) It's clearly noticeable that the whole traffic generated by this honeytoken is very suspicious. Because the honeypot (the FTP-Server) is always observed, it can be figured out who tried to access it and what actions were carried out. If a file access gets registered, it discloses that both the E-Mail account and the CEO's computer could be compromised by individuals with criminal intents. With some further checks it can be determined clearly, where the weak or breach point is and that something is not working as it is supposed to do. Maybe some delicate files were placed on the hard disk of the CEO's computer, so it can be checked if someone has even physical access to his computer.

This example demonstrates that honeytokens can be used effortless and without many costs and personal involvement. Also the logging activities can be reduced to a minimum, because every access to it is per se suspicious. In such cases insider threats would be nearly impossible to detect, because they take place on the highest levels in an organization. A honeytoken is here a very effective way to detect such insider threats.

4.3.2 Example of a honeytoken database entry

A very simple method of such a honeytoken database entry could be the following, suggested by Spitzner [10]:

> *"For example, the credit card number 4356974837584710 could be embedded into database, file server, or some other type of repository. The number is unique enough that there will be minimal, if any, false positives. An IDS tool, such as Snort, could be used to detect when that honeytoken is accessed. Such a simple signature could look as follows:*

```
alert ip any any
→ any any (msg:
''Honeytoken access - potential unauthorized activity'';
content:''4356974837584710'';)
```

This database entry is nothing else than a simple number, a credit card number. The number is a honeytoken, because it should not be used according to its definition. But this convenient way has some disadvantages. If the SQL query output is encrypted, then it does not work because a sniffing tool running on the same machine as a Database Management System could be suspicious to a attacker. Also the logs could be easily manipulated, because there are on the same machine [2].

It has to be the aim to insert some tables in the database with titles seeming attractive to malicious users. Because it's a whole table (and not just one single entry), a warning E-Mail can be generated and sent to the administrator by any interaction with the "honeytoken" table. Creating such a table requires a bit more work, because it should contain a large amount of real looking entries. This is necessary that it does not seem suspicious to anyone who access it [2]. For the control of the honeytoken module, a technique should be used which can monitor any type of database objects. In general, it can be set up by the following requirements to a honeytoken database entry [2]:

- **Database object:** The database object which is used as a honeytoken should contain real looking and "attractive" data. The database object can be a table, view or something else.
- **Program for monitoring access to the honeytoken:** A little program is required which monitors the access to the honeytoken module.
- **Program for the action after the honeytoken was accessed:** Further a program (it can be a little bash script) has to be created, which informs the database owner that the honeytoken module has been accessed (e.g. by sending an E-Mail to him).

There are a few alternatives for the implementation. The concept of inserting a honeytoken as a whole table is described in Figure 4.3, which shows this procedure on an Oracle Database Management System.

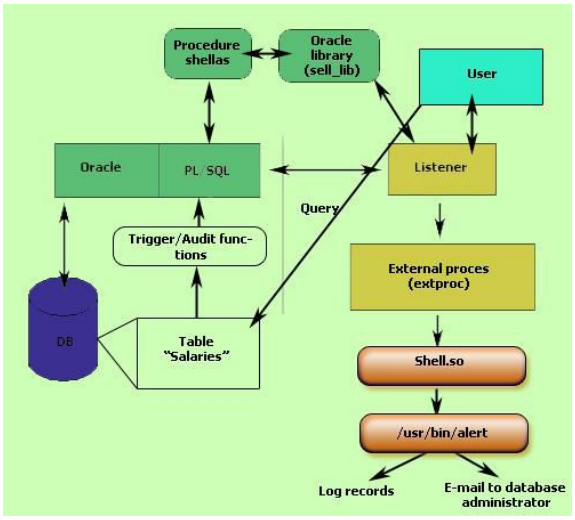

Figure 4.3: Scheme of a honeytoken database entry [2]

Let's start the description of Figure 4.3 when a user tries to interact with the honeytoken table "salaries". If someone does, a listener is ready to notice this action. This listener will start an internal Oracle procedure shell. The mentioned shell is going to run an external module which is responsible for the logging activities and for the notification of the database administrator. Here it's possible to use triggers running in the background and waiting for certain defined events [5]. The data table has to be considered as the

most important part of this undertaking. If the data table is not designed properly, then the honeytoken is quite useless. The honeytoken data table named "salaries" used in this example is shown in Table 4.4.

Table 4.4: Honeytoken table "salaries" (own design inspired by [2])

Salaries				
PersonalNumber	DateOfEntry	Name	Surname	SalaryUSD
CE00597546	22.10.1997	Miller	Joe	100'000
EX12099789	10.06.1995	Cooper	Michael	75'300
BC55781568	30.01.1999	Fisher	Gary	120'000
SA00501002	05.04.2001	Doe	Robert	60'000
...	...	...	...	...

An appropriate SQL Statement to create this table would be:

```
CREATE TABLE Salaries (
        PersonalNumber  VARCHAR(10) NOT NULL ,
        DateOfEntry     TIMESTAMP,
        Name            VARCHAR(20),
        Surname         VARCHAR(20),
        SalaryUSD       INTEGER(8),
        PRIMARY KEY(PersonalNumber)
);
```

The full power of the honeytoken strategy will appear when the database has a large amount of user and database entries. In that case it's very effective and also cost efficient and last but not least the setup is not so advanced at all.

4.3.3 Honeypots in the fight against worms

In this section some principles are shown, how honeypots can be used for worm detection. Internet worms can be roughly distinguished in two types regarding their spread behavior [8]:

- **Constant spread time** Due to the design of a network scanner the worm is implementing, it spreads on a constant time on all hosts.
- **Different spread time** Worms of this types do not spread constantly, the only limitations are in the CPU or in the Network bandwidth of hosts.

Due to the spread all over the globe, the need for global monitoring is becoming clear. Worm detection is covered actively in many research topics. Some topics are "worm spreading" and "early detection" where statistical models are used [3]. Monitoring large

IP-Space has (e.g. 2^{20} IP addresses) specific requirements, a famous architecture to handle this task are "Network telescopes" [3], [8]. A large space of IP-Addresses significantly increases the probability that a worm is detected. This can be shown with some calculus; for a deep explanation of the detailed mathematical and statistical derivation, please refer the paper of Moore [8]. In the following, an overview of detection times depending on different network sizes is presented. On one hand it shows the size of the network and on the other hand the time needed in order to detect a given percentage of attacks [8]. In the table lists given the values to detect 5% and 95% of the events, as well as the time for detecting the average and median of the events. In this table the / notation is used. "This notation is useful for the calculation of the probability of the telescope monitoring a target chosen by a host" [8]. This notation leads to the theme of "Subnetting". In this table is based on IPv4 (Internet Protocol Version 4), which is nowadays the most ubiquitous communication protocol. The IPv4 protocol specifies uniquely a machine with the use of 32 bits per address. So, according to this definition there are 2^{32} possible IP-Addresses [8]. The IPv4 protocol generates IP-Addresses in the shape xxx.xxx.xxx.xxx. There are four blocks where the numbers are written in. In each block there are 8 bits (this leads to the total of 32 bits by $4 \times 8\, bits = 32\, bits$). A binary representation of the "highest" possible IP-Address 255.255.255.255 (in decimal) would be: 11111111.11111111.11111111.11111111. (Each Block stands for 255). So a /8 describes the range of 2^{24} addresses which have the first 8 bits of their addresses in common [8], e.g. a /32 describes a single IP-Address like 89.206.64.93 and a /0 describes the whole range of the internet with the IPv4-Protocol. A general Formula for the maximum amount of addresses by /X can be given by:

$$2^{32-X} \Rightarrow 2^{32-X}\ possible\ adresses$$

Table 4.5: Detection time in relation to network size

Size of Network	5%	Average	Median	95%
/8	1.31 sec	25.6 sec	17.7 sec	1.3 min
...	...	...	...	...
/14	1.40 min	27.3 min	18.9 min	1.4 hours
/15	2.80 min	54.6 min	37.9 min	2.7 hours
/16	5.60 min	1.82 hours	1.26 hours	5.5 hours
...	...	...	...	...
/19	44.8 min	14.6 hours	10.1 hours	1.8 days
/20	1.49 hours	29.1 hours	20.8 hours	3.6 days
/21	2.99 hours	58.3 hours	40.4 hours	7.3 days
/22	5.98 hours	4.85 days	3.36 days	14.5 days
/23	12.0 hours	9.71 days	6.73 days	29.1 days
/24	23.9 hours	19.4 days	13.5 days	58.2 days

Table 4.5 shows that the more addresses are used in a telescope network, the faster a single packet can be detected with a given probability. Further it's important to mention that the detection time does not increase linearly with the amount of IP-Addresses, e.g. a /24 would need "65664 times as long as a /8 to detect at least one packet with the same probability" [8].

But in practice there occur some problems because the IP-Addresses are not chosen randomly. But having no bias or known discriminations of in the selection of data is one of the most important principles when dealing with statistics. For the following reasons the IPv4 protocol does not choose the IP-Addresses randomly form its whole space [8]: There are also shown up the statistical principles which are violated.

- **Narrowing of address space:** This means that certain regions are just left out. Due to statistical definitions, a sample has always to be drawn out of a whole population. It's not allowed that certain regions are left out.
- **Biasing:** Some regions are more used than others, for example a worm nearly always tries the addresses in its close range [1]. Due to statistical definitions a bias is not allowed. As long as the data are biased, every calculation made is valueless.
- **Bugs or bias in the underlying Pseudo-Random Number Generator:** The same principle is violated as in the second item.

But luckily there are some solutions which help in the setup of a telescope Network [8].

- **Distributed telescope network:** As it already can be derived from the name, there are some small networks in different regions of the whole network space. This networks are combined together in order to form a larger network based on these "pieces". With this method it's possible to achieve a quite homogeneous distribution of IP-Addresses.
- **Anycast network telescope:** "With an anycast telescope, multiple locations advertise routes for the same network address range prefix" [8]. In this type of network, which bases on the same principles as a distributed telescope, the traffic can be distributed in a more equilibrate way.
- **Honeyfarm telescope:** The key principle of a honeyfarm lies in the active response to some events. According to the definition of a honeypot, a honeyfarm does not passively observe the traffic. It's important that the extra traffic caused by the responses is handled correctly.
- **Transit network telescope:** The observation of the network traffic is done in the transit network instead at the network edges. So it's possible to observe traffic of a very high amount of IP-Addresses.

These structures are often used in universities for research in worm propagation. The architectures of large honeypot networks help the setup of honeypots in the practice. Unfortunately they are not very easy to create due to their large size and the maintenance of the telescope network can turn out as time consuming and exhausting. Also the enormous quantity of collected data has to be processed in a clever manner. A quite good approach seem to be the "distributed telescope networks". According to the principle of modularization is the utilization of some smaller networks from different locations combined to one global piece.

4.4 Summary and Conclusion

As it turns out, the concept of honeypots seems to be a very interesting and helpful one. Basically it's quite easy to create a honeypot by oneself. There are already many tools available to do so, such as the "honeywall CD-ROM", which should simplify the creation and maintenance of a virtual honeypot. Within an organization, a honeytoken is a very convenient way to detect insider threats. Honeytokens can be created with very little resources, but their power and effectiveness is quite impressive. Due to the fact that honeypots are very cheap to maintain they can be used in several environments, from private users to large corporate networks. They also can be used for a large variety of security and monitoring tasks. Given to the honeypots definition, that every access to it is per se an unauthorized access, there are big advantages in comparison to traditional security arrangements such as firewalls and/or antivirus scanners. With honeypots it's possible to choose a new way in detecting yet unknown attacks. Further, they also allow doing some deep research in worm detection and propagation. Unfortunately honeypots are not very well known. Traditional security measures are much more widespread, although it's known that firewalls have certain weaknesses and virus scanners often fail to detect new types of attacks. Using honeypots within an organization is a viable possibility, because some security problems are nearly impossible to handle in another way. Restricting data access is often not applicable to every person, because someone has to administrate the systems. Illegal data access can be discovered for example by a honeytoken database entry.

As expected the uses of honeypots show that security aspects in networks should not be a matter where the machines decide which actions are suspicious and which are not. Those decisions should be made by a human person. Unfortunately this is not possible with traditional security measures. A firewall creates many alerts in a short time, that even a large group of security specialists would need too much time to analyze all of these alerts. The alerts created by honeypots are much more serious, because they should not generate any. So in practice much less data has to be examined. Because of that smaller quantity it's possible for humans to check these alerts.

There are not many evidences against the use of honeypots and we hope that this paper will simplify the entrance to this topic.

Bibliography

[1] K. G. Anagnostakis, S. Sidiroglou, P. Akritidis, K. Xinidis, E. Markatos, A. D. Keromytis: *Detecting Targeted Attacks Using Shadow Honeypots*, In Proceedings of the 14th USENIX Security Symposium, Baltimore, Maryland, USA, pp. 129-144, August 2005.

[2] Antanas Čenys, Darius Rainys, Lukas Radvilavičius, Nikolaj Goranin: *Implementation of honeytoken module in DBMS Oracle 9iR2 enterprise edition for internal malicious activity detection*, Informtion Systems Laboratory, Semiconductor Physics Institute Lithuania, `http://www1.gi-ev.de/fachbereiche/sicherheit/fg/sidar/dimva/dimva2005/papers/cenys.pdf`

[3] D. Dagon, X. Qin, G. Gu, W. Lee, J. Grizzard, J. Levine, Henry Owen: *HoneyStat: Local Worm Detection Using Honeypots*, In Proceedings of the 7th International Symposium on Recent Advances in Intrusion Detection, Sophia Antipolis, France, pp. 39-58, September 2004.

[4] E. Kapersky: *Malware: Von Viren, Würmern, Hackern und Trojanern und wie man sich vor ihnen schützt*, Hanser Fachbuch; pp. 86-89, March 2008.

[5] A. Kemper, A. Eickler: *Datenbanksysteme: Eine Einführung*, Oldenbourg, pp. 163-164, March 2006.

[6] C. Kreibich, J. Crowcrof: *Honeycomb: creating intrusion detection signatures using honeypots*, ACM SIGCOMM Computer Communication Review, Volume 34, Issue 1, pp. 51-56, January 2004.

[7] J. Levine, R. LaBella, H. Owen, D. Contis, B. Culver: *The use of Honeynets to detect exploited systems across large enterprise networks*, In Proceedings of the 2003 IEEE Workshop on Information Assurance, West Point, New York, USA, pp. 92-99, June 2003.

[8] D. Moore, C. Shannon, G. M. Voelkery, S. Savagey: *Network Telescopes: Technical Report*, Cooperative Association for Internet Data Analysis - CAIDA, University of California San Diego Supercomputer Center, San Diego, California, USA, April 2004.

[9] N. Provos: *A Virtual Honeypot Framework*, In Proceedings of the 13th USENIX Security Symposium, San Diego, California, USA, pp. 1-14, August 2004.

[10] L. Spitzner: *Honeypots: catching the insider threat*, In Proceedings of the 19th Annual Computer Security Applications Conference, Las Vegas, Nevada, USA, pp. 170-179, December 2003.

[11] L. Spitzner: *Honeytokens: The Other Honeypot*, SecurityFocus, July 2003. `http://www.securityfocus.com/infocus/1713`

[12] L. Spitzner: *Know your enemy: Automated Credit Card Fraud*, honeynet Project, June 2003. `http://old.honeynet.org/papers/profiles/cc-fraud.pdf`

[13] L. Spitzner: *Know your enemy: Honeynets*, honeynet Project, May 2006. `http://www.honeynet.org/papers/honeynet`

[14] L. Spitzner: *Know your enemy: Honeywall CD-ROM*, honeynet Project, August 2005. `http://old.honeynet.org/papers/profiles/cc-fraud.pdf`

[15] A. S. Tanenbaum: *Computernetzwerke*, Pearson Studium, pp. 779-780, November 2000.

Kapitel 5

Inter-domain Routing

Lukas Keller, Sacha Gilgen

BGP ist das heute benutzte Routing-Protokol, um im Internet den Datenverkehr zu steuern. In diesem Bericht wird das Border Gateway Protokol in der aktuellen Version 4 beleuchten indem zuerst die Einsatz-Gebiete und der direkte Einsatz erklärt werden. Im zweiten Teil wird gezeigt, wie das Protokol technisch funtkioniert und es werden Designentscheidungen erklärt. Im letzten Teil des Berichts werden die ökonomischen Gesetze des Routings im Internet und wie die Internet Service Provider mit Hilfe der Policies von BGP mit diesen Gesetzen arbeiten erklärt. Am Ende wird dann noch auf die aktuellen Probleme von BGP und dem Internet-Routing eingegangen.

Inhaltsverzeichnis

5.1 Einleitung

Routing bedeutet das Steuern der Datenpakete auf ihrem Weg durch das Internet. Es ist ein grundlegendes Konzept des heutigen weltweiten Netzes und daher von grosser Wichtigkeit. Dabei wird unterschieden zwischen dem Intra-Domain Routing, was die Steuerung des Datenverkehrs innerhalb abgegrenzter Regionen (Autonome Systeme) bedeutet, und dem Inter-Domain Routing, mit dem die Verkehrsflusssteuerung zwischen diesen Autonomen Systemen gemeint ist. In dieser Seminararbeit wird vor allem das Inter-domain Routing beleuchtet, es wird aber auch kurz auf das Intra-Domain Routing eingegangen und die Unterschiede aufgezeigt. Beim Inter-Domain Routing ist ein wichtiger Punkt die angewendeten Metriken, mit denen Provider versuchen, den Datenstrom nach ihren Vorstellungen zu lenken. Dies und die Bedeutung dessen für das Inter-Domain Routing wird auch noch Rechnung getragen.

5.2 Grundlagen

Das Grundproblem, welches Routing löst, ist, wie Datenpakete schnellstmöglichst von ihrem Absender zu ihrem Empfänger kommen. Im Internet ist das besonders wichtig, weil das Netz eine weltumspannende Grösse hat, und daher die Effizienz der Datenübertragung eine zentrale Rolle spielt.

5.2.1 Internet und ISPs (Autonome Systeme)

Das Internet ist nicht, wie häufig angenommen, ein grosses Netz, in das sich jeder Nutzer einklinkt (s. Abb. 5.1), sondern es besteht aus verschiedenen kleineren und mittleren Netzen (s. Abb. 5.2). Was diese verschiedenen Netze zum Internet macht, ist die Verbindung dieser Netze untereinander. Dadurch können die Daten von einem Netz ins andere weiterreisen und in die ganze Welt gelangen. Diese Netze werden Autonome Systeme (im Folgenden nur noch AS) genannt und machen die eigentliche Architektur des Internets aus [4]. Jede Organisation jedwelcher Grösse kann ein AS unterhalten, es sind aber doch Internet Service Provider (in der Folge ISPs genannt), welche das Gros der ASe im heutigen Internet ausmachen. Diese haben ein Interesse daran, dass der Netzwerk-Verkehr durch ihre Netze geleitet wird, weil sie dadurch Geld verdienen können [4]. Dieser Teil wird später noch genauer behandelt.

Es gibt drei Klassen von ISPs. Tier-1, Tier-2 und Tier-3. Wenn ein Endnutzer sich ins Internet einwählt, dann verbindet er sich meist mit einem Tier-3 ISP, welcher nur lokalen oder regionalen Zugang bietet. Tier-2 ISPs haben grössere Netze, welche sich über mehrere Regionen oder Länder erstrecken. Und Tier-1 ISPs sind global tätig und haben in fast allen wichtigen Regionen der Welt Anschlüsse. Damit ein Tier-3 ISP Zugang zum weltweiten Netz bekommt, muss er sich einen Tier-2 oder Tier-1 ISP Anschluss verschaffen, indem er mit diesem ein Abkommen abschliesst. So schliessen sich die verschiedenen ASe der ISPs zu einem global vernetzten System zusammen. Damit kann schlussendlich über verschiedene Stationen jeder Teilnehmer des Internets erreicht werden und auch andere erreichen.

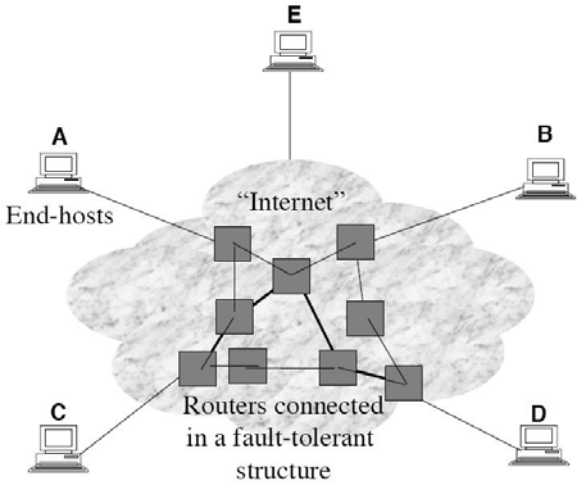

Abbildung 5.1: Abstrakte, aber irreführende Darstellung des Internets ([4], S. 2)

5.2.2 Intra- und Inter-Domain Routing

In den eben besprochenen ASen, wie auch zwischen diesen Systemen, muss der Netzwerk-Verkehr geleitet und gesteuert werden. Dies nennt man Routing. Es gibt dabei Unterschiede zwischen dem Routing innerhalb der ASe (Intra-domain Routing) und dem Routing zwischen den verschiedenen AS (Inter-domain Routing). Das Routing innerhalb eines AS wird vom Besitzer des AS selbst verwaltet. Es kann also von einer zentralen Stelle aus gemanagt werden. Beim Inter-domain Routing hingegen gibt es keine zentrale Instanz, weil hier verschiedene Organisationen beteiligt sind, welche zum Teil im Wettbewerb zueinander stehen. Es muss sich also quasi selbst verwalten. Dies ist einer der wichtigsten Unterschiede zwischen den beiden Routing-Arten. Ein anderer Unterschied ist, dass beim Inter-domain Routing die Skalierbarkeit des verwendeteten Routing-Protokolls viel wichtiger ist als beim Intra-domain Routing, da es eine riesige Anzahl AS gibt, welche alle irgendwie miteinander verbunden sind. Aus diesen Gründen gibt es unterschiedliche Protokolle für die beiden Routing Arten. Diese werden hier kurz vorgestellt.

5.2.2.1 Intra-domain Routing Protokolle

Die Intra-domain Routing Protokolle werden Interior Gateway Protocols (IGP) genannt. Sie werden verwendet, um den Datenverkehr innerhalb eines AS zu steuern. D.h. der Verkehr, welcher nur zwischen den Stationen des eigenen AS ausgetauscht wird, und der Verkehr, welcher von aussen in das AS reinkommt. Die Bekanntesten sind:

RIP (Routing Information Protokoll) ist ein Distanz-Vektor Protokoll und das simpelste heute eingesetzte IGP. Es wird wegen einiger Mängel aber nicht mehr häufig benutzt. Vor allem, weil es sich nicht gut in grossen Netzen einsetzen lässt (schlechte Skalierbarkeit) und bei Änderungen langsam reagiert [7].

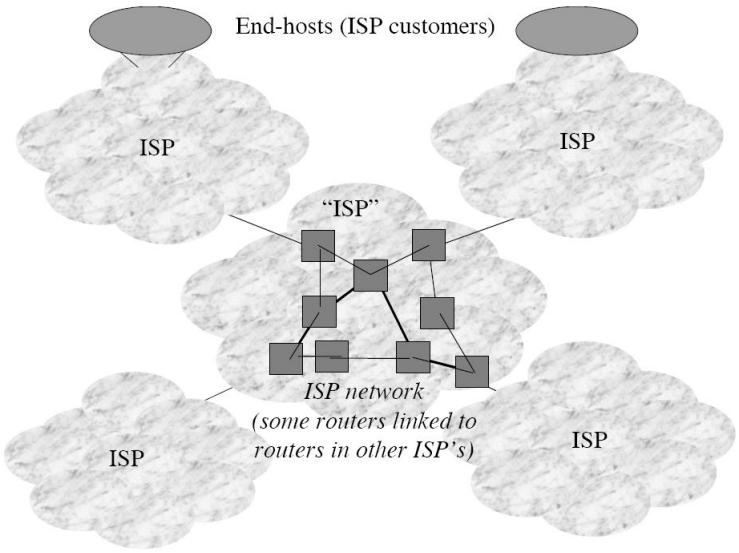

Abbildung 5.2: Struktur des Internets mit autonomen Systemen und Endnutzern ([4], S. 3)

OSPF (Open Shortest Path First) ist ein Link-State Protokoll und das heute am meisten benutzte IGP. Es wurde als direkter Ersatz für RIP entwickelt und behebt die wichtigsten Mängel von diesem [7].

5.2.2.2 Inter-domain Routing Protokolle

Die Inter-domain Protokolle werden Exterior Gateway Protocols (EGP) genannt. Mit ihnen wird, wie bereits geschrieben, der Datenverkehr zwischen den einzelnen ASs gesteuert. Dieses Protokoll ist dafür zuständig, dass ein Datenpaket, das von Europa aus geschickt wird, seinen Weg nach Nordamerika findet. Von diesen Protokollen wird heute eigenlich nur noch das Border Gateway Protocol (BGP), welches im zweiten Teil dieser Arbeit behandelt wird, eingesetzt [5]. Als das Internet entstand, bezeichnete EGP noch nicht die Klasse der Exterior Gateway Protokolle, sondern ein einzelnes Protokoll, das damals zum selben Zwecke eingesetzt wurde. Die Bezeichnung passte man aber an, da das ursprüngliche EGP wegen verschiedener Mängel (keine dynamische Anpassung an Änderungen im Netz, alle Router müssen alle anderen Router kennen, etc.) praktisch ganz von der Bildfläche verschwunden ist und nur noch in einigen amerikanischen Militär- oder staatlichen Netzen benutzt wird [5].

5.2.3 Ökonomischer Aspekt des Internetrouting

ISPs verdienen ihr Geld über die Abwicklung von Datenverkehr. Dies passiert auf jeder Ebene des Internets. Der Endbenutzer muss seinem ISP eine monatliche Gebühr dafür bezahlen, dass er ans Internet angeschlossen ist. Der lokale ISP hat monetäre Verträge mit grösseren ISPs, die ihm den Zugang zum weltweiten Netz ermöglichen. Und die ganz grossen Tier-1 Provider bekommen Einnahmen von verschiedenen Organisationen durch

die Weiterleitung der Daten über Kontinente und Ozeane hinweg [4]. In all diesen Bereichen spielen daher ökonomische Entscheidungen ein wichtige Rolle. Der weitaus grösste Teil dieser Entscheidungen betrifft die Zusammenarbeit zwischen ISPs. Dabei gibt es v.a. zwei Arten von Zusammenarbeit: Transit und Peering.

5.2.3.1 Transit

Von Transit spricht man, wenn ein kleiner ISP sich an einen grossen ISP anschliesst, damit er Zugang zum weltweiten Netz hat. Er bezahlt den grösseren Provider dafür, dass er Daten durch sein Netz schicken darf (s. Abb. 5.3: ISP X bezahlt P für den Transit seiner Daten und er Daten seiner Kunden C1-C3). Diese Verträge sind beliebt, da sie Einnahmen generieren [4].

5.2.3.2 Peering

Zwei ähnlich grosse ISPs schliessen sich zusammen, um direkten Zugang zu Adressen des jeweiligen anderen Netzes zu bekommen. Dies ist häufig der Fall, wenn ISPs merken, dass ein grosser Teil ihres eigenen Verkehrs zu Adressen im Netz des anderen ISPs geht, weil dort zum Beispiel eine beliebte Internetseite ist. Mit einem direkten Zugang muss der ISP diesen Verkehr nicht über eine teure Transitlinie eines grösseren ISPs leiten, sondern kann ihn direkt ins andere Netz steuern (s. Abb. 5.3: ISP X hat Peering-Abkommen mit Z). Der Vorteil ist, dass diese Peering-Möglichkeiten meist gratis sind, solange das ausgetauschte Datenvolumen in etwa gleich gross ist (bis zu Ratios von 4:1 sind dabei möglich) [4]. So kann ein ISP Geld sparen.

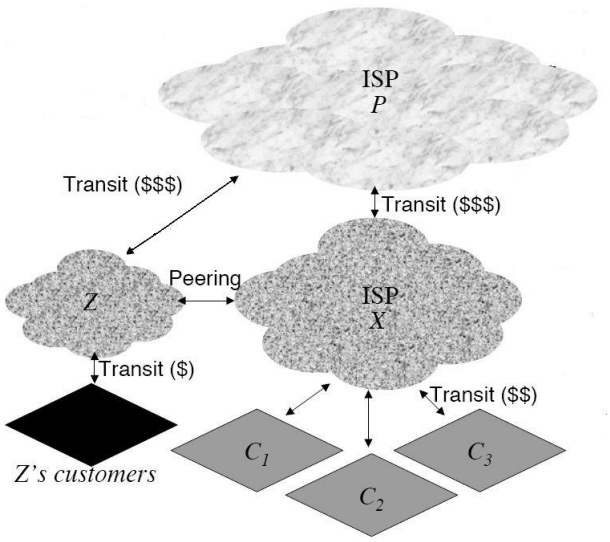

Abbildung 5.3: Beispiel von ISPs mit Transit- und Peering-Verträgen. Provider X leitet seinen Datenverkehr von seinen Kunden zum Kunden von Z direkt in Z's Netz, ohne den teuren Umweg über Provider P zu nehmen. (basierend auf [4], S. 5)

Durch diese beiden Möglichkeiten ergibt sich die Notwendigkeit, den Datenfluss mit dem verwendeten Routing-Protokoll steuern zu können. Denn wenn ein ISP einen Peering-Vertrag mit einem Nachbar-ISP hat, dann muss er seinen Datenverkehr auch zu diesem umleiten können, sonst geht er weiter über die teure Transit-Leitung des grösseren ISPs. Aus diesem Grunde hat das im zweiten Teil vorgestellte Border Gateway Protokoll die Möglichkeit, mit Policies die Routing-Entscheidungen der Router zu beeinflussen. Diese Entscheidungsmöglichkeiten werden im zweiten Teil noch näher erläutert.

5.3 Das Border Gateway Protokoll Version 4

Das Border Gateway Protokoll wurde im Jahre 1989 von der Internet Engineering Task Force (IETF) in erster Version (BGP-1) als RFC 1105 veröffentlicht und sollte das bis dahin verbreitete EGP-Protokoll ersetzen. Die Entwicklung von BGP durchlief seither mehrere Stufen in dem es schrittweise um weitere Mechanismen ergänzt wurde. Im Jahre 1995 wurde dann, die heute vorherrschende Version BGP-4, veröffentlicht. BGP-4 wird im RFC 4271 beschrieben. Nachfolgend sollen zunächst die Eigenschaften und die Funktionsweise von BGP-4 beschrieben werden bevor dann näher auf das Protokoll eingegangen wird.

5.3.1 Eigenschaften von BGP

BGP-4 ist ein Pfadvektorprotokoll das zum Austausch von Erreichbarkeitsinformationen zwischen ASen dient. Dieses Protokoll stellt dabei keine Anforderungen in Form von Restriktionen an die zugrunde liegende Netzwerktopologie, sondern geht von der Annahme aus, dass das Internet aus zufällig miteinander verbundenen ASen besteht [3] Damit grenzt sich BGP deutlich von seinem Vorgänger EGP ab, das als Grundlage eine Baumtopologie des Netzwerks veraussetzt und damit bezüglich der Skalierbarkeit sehr eingeschränkt war [2].

Die übermittelten Routing-Informationen zwischen BGP-Routern enthalten eine Liste von ASen, welche vom zu kommunizierenden Netzwerk aus, bisher durchlaufen wurden. Jeder BGP-Router ergänzt diese Liste um sein AS, bevor er die Informationen weiterleitet. Man spricht in diesem Zusammenhang von einem Pfad bestehend aus ASen. Mit Hilfe dieses Pfades können Routing-Schleifen leicht vermieden werden, da ein AS nicht zweimal im kommunizierten Pfad vorkommen darf. Die interne Struktur der ASe spielt beim Auffinden des Pfades zum kommunizierten Zielnetzwerk keine Rolle. Das Konzept der Aufteilung des Internets in ASe sorgt für die notwendige Abstraktion der internen Struktur dieser ASe.

Seit Version 4 unterstützt BGP das Classless Interdomain Routing (CIDR). Damit können mehrere Destinationen (Zielnetzwerke) unter einer IP-Prefix zusammengefasst werden. Nebst der unterstützung von CIDR beinhaltet BGP-4 auch Mechanismen um Routen zusammenzufassen, womit die Routing-Tabellen der BGP-Router verkleinert werden können.

Weiter unterstützt BGP die Definition von Routing-Policies. Dazu werden sogenannte Pfadattribute, welche Eingeschaften einer bestimmten Route darstellen, genutzt, um die Übermittlung der Routing-Informationen zu kontrollieren. Policies spielen im heutigen Internet eine wichtige Rolle, da damit die oben bereits beschriebenen ökonomisch geprägten Beziehungen zwischen ISP realisiert werden. Es gilt jedoch zu beachten, das Policies nicht Teil der Spezifikation von BGP sind. Wie Policies definiert werden und was für Policie-Strategien die ISPs verfolgen, wird im nachfolgenden Kapitel erläutert.

Ein weiteres wichtiges Merkmal von BGP ist, dass es für die Übermittlung der Informationen auf das Transportprotokoll TCP aufsetzt. TCP erlaubt dabei eine zuverlässige Verbindung zwischen den benachbarten Routern und garantiert eine "in-order delivery„ der gesendeten Nachrichten [4]. Damit wird sichergestellt, dass der Aufbau einer zuverlässigen Verbindung nicht durch das Border Gateway Protokoll zu erfolgen hat. Die Komplexität des Protokoll wird dadurch deutlich reduziert [3]. Zwei BGP-Router welche eine TCP-Verbindung untereinander aufbauen um Routing-Informationen (Erreichbarkeitsinformationen) auszutauschen, werden als Nachbarn oder auch als Peers bezeichnet.

5.3.2 Funktionsweise von BGP

Wenn zwei BGP-Router miteinander kommunizieren wollen, errichten diese zunächst eine TCP-Verbindung. Dazu lauscht BGP auf den wohlbekannten TCP-Port 179. Sobald die TCP-Verbindung zwischen den BGP-Routers steht, tauschen diese OPEN-Nachrichten aus, um die Parameter für die BGP-Session auszuhandeln (z.B. BGP-Version) [8]. Dieser Vorgang wird auch als "Neighbour Negotiation„ bezeichnet [3]. Werden die OPEN-Nachrichten von den BGP-Peers akzeptiert, senden diese eine KEEPALIVE-Nachricht als Bestätigung zurück und es besteht fortan eine BGP-Session zwischen den Peers.

Nachdem die BGP-Session erfogreich erstellt wurde, beginnen die benachbarten Router den gesamten Inhalt ihrer Routing-Tabellen auszutauschen. Dabei wird nur derjenige Teil der Routing-Tabelle kommuniziert, der nach Anwendung der Export-Policy verbleibt [1]. Nach diesem Initialaustausch erfolgen nur noch inkrementelle Aktualisierungen, um neu hinzugefügte, geänderte oder gelöschte Routen zu kommunizieren. Der Austausch von Routen erfolgt dabei anhand der Übermittlung von UPDATE-Nachrichten.

Liegen keine Aktualisierungen vor, werden KEEPALIVE-Nachrichten in einem vorher festgelegten Intervall zwischen den benachbarten BGP-Router gesendet. Mit diesen Nachrichten wird sichergestellt, dass der benachbarte Router, betriebsbereit ist und somit weiterhin erreicht werden kann.

Tritt ein Fehler bei der Kommunikation zwischen zwei BGP-Routern auf, wird eine NOTIFICATION-Nachricht gesendet und ansschliessend die Verbindung umgehend abgebaut. Die Nachricht beschreibt dem benachbarten Router unter welchen Bedigungen ein Fehler aufgetreten ist.

5.3.3 BGP-Nachrichten

Im vorherigen Unterkapitel wurde die Funktionsweise von BGP grob skizziert, indem beschrieben wurde, wann welche Nachrichtentypen zwischen den BGP-Peers ausgetauscht werden. In diesem Kapitel sollen nun diese Nachrichtentypen näher erläutert werden, um einen tieferen Einblick in das Border Gateway Protokoll zu gewähren. Besonderes Augenmerk gilt den UPDATE-Nachrichten, da diese die Routing-Informationen zwischen den ASen übermitteln und somit den Kern des Protokolls bilden.

5.3.3.1 BGP Message Header

Jede BGP-Nachricht enthält einen Header mit festgelegter Länge. Dieser dient dazu, die einzelnen Nachrichten im Datenfluss voneinander zu trennen. Der Header besitzt dazu ein Feld *Length*, welches die Länge der Nachricht angibt. Einem Header folgt dann, je nach Nachrichten-Typ, ein Datenteil oder auch nicht. Eine KEEPALIVE-Nachricht zum Beispiel enthält keine weiteren Daten, sondern besteht lediglich aus dem Header. Die Abbildung 5.4 zeigt das Format eines solchen Headers.

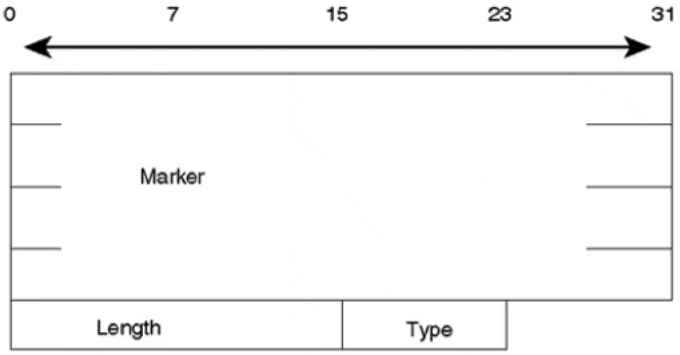

Abbildung 5.4: Format des Headers einer BGP-Nachricht [1]

Im folgenden sollen die drei Felder des Headers kurz erläutert werden:

Marker: Das 16 Byte grosse Feld dient entweder zur Authentifizierung eingehender BGP-Nachrichten oder um Synchronisationsverluste zwischen zwei BGP-Peers aufzudecken [3].

Length: Dieses 2 Byte grosses Feld dient dazu, die gesamte Länge der BGP-Nachricht anzugeben (den Header inbegriffen). Die kleinste mögliche BGP-Nachricht umfasst somit 19 Bytes, was der Grösse des Headers entspricht. Eine Nachricht darf zudem nicht grösser als 4096 Bytes sein [1].

Type: Das letzte 1 Byte grosse Feld gibt den Typ der Nachricht an. Folgende Nachrichtentypen sind möglich:

- OPEN-Nachricht
- KEEPALIVE-Nachricht
- NOTIFICATION-Nachricht
- UPDATE-Nachricht

5.3.3.2 OPEN-Nachricht

OPEN-Nachrichten werden übermittelt, um die Parameter einer BGP-Session auszuhandeln (z.B. Hold Time) und spezifische Informationen über den sendenden Router zu übermitteln (z.B. BGP-Identifier). Abbildung 5.5 veranschaulicht das Format einer solchen OPEN-Nachricht.

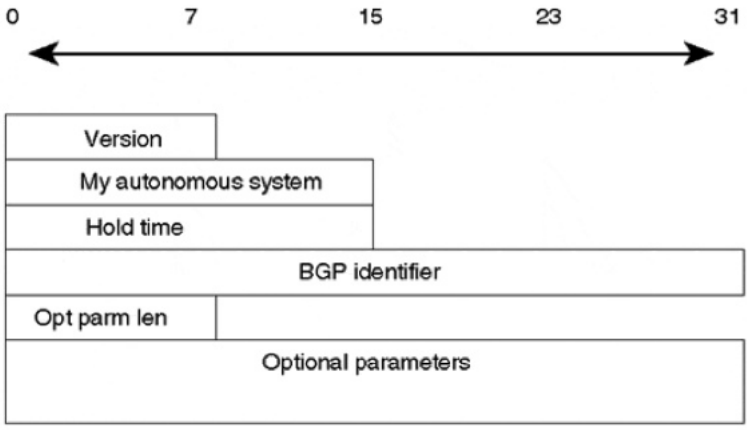

Abbildung 5.5: Format einer OPEN-Nachricht [1]

Die OPEN-Nachricht verfügt zusätzlich zum Header über folgende Felder:

Version: Dieses 1 Byte grosse Feld gibt die Version der BGP-Nachricht an (zum Beispiel BGP-3 oder die aktuelle Version BGP-4). Während der "Neighbor Negotiation„ versuchen die BGP-Peers sich auf die grösste, von beiden Routern unterstützte Version zu einigen [3].

My Autonomous System: Dieses 2 Byte grosse Feld gibt die Nummer des AS (AS-Nummer) an, in welchem sich der BGP-Router befindet [1].

Hold Time: Dieses 2 Byte grosse Feld gibt die maximal mögliche Anzahl von Sekunden an, die zwischen dem Erhalt von aufeinanderfolgenden KEEPALIVE oder UPDATE-Nachrichten verstreichen darf. Die Hold Time wird dabei zwischen den benachbarten BGP-Router ausgehandelt, indem jeder seine Hold Time dem benachbarten Router über die OPEN-Nachricht kommuniziert. Der niedrigere der beiden Werte wird dann von beiden Routern übernommen (bzw. beibehalten). Im jeweiligen Router zählt ein

Timer von 0 bis zum vereinbarten Hold Time. Erhält ein Router eine KEEPALIVE oder UPDATE-Nachricht innerhalb der Hold Time, beginnt der Timer wieder von Null an zu zählen. Wird jedoch die Hold-Time überschritten, geht der Router davon aus, dass sein Nachbarn nicht mehr aktiv ist und sendet eine NOTIFICATION-Nachricht, bevor er die Verbindung abbricht [1].

BGP Identifier: Dieses 4 Byte grosse Feld enthält eine IP-Schnittstellenadresse, welche dem BGP-Router zugewiesen wurde. Die IP-Adresse wird bei der Inbetriebsnahme des Routers festgelegt und für alle BGP-Aktionen verwendet, unabhängig von der Schnittstelle, welche die BGP-Pakete überträgt [2].

Optional Parameter Length (Opt Pram Len): Dieses 1 Byte grosse Feld gibt die gesamte Länge des *Optional Parameters* Feldes an. Bei einem Wert von Null enthält die OPEN-Nachricht keine optionale Parameter [1].

Optional Parameters: Dieses Feld hat eine variable Länge und enthält eine Liste von optionalen Parametern. Jeder Parameter wird als Triple der Form <Paramenter Type, Parameter Length, Parameter Value> angegeben. Mit *Parameter Type* wird ein Parameter eindeutig identifiziert. *Paramter Length* gibt die Länge des Parameter an und *Paramter Value* enthält den Wert des Parameter, der aufgrund des festgelegten Paramter Type interpretiert wird [1].

5.3.3.3 KEEPALIVE-Nachricht

KEEPALIVE-Nachrichten werden periodisch zwischen benachbarten BGP-Routern übermittelt und dienen dazu, die Erreichbarkeit des Nachbarn zu überprüfen. Diese Nachrichten werden in einer Periodizität gesendet, die sicherstellt, dass die vereinbarte Hold Time einer BGP-Session nicht überschritten wird. Als Daumenregel wird eine Periodizität von einem Drittel der Hold Time empfohlen. Die KEEPALIVE-Nachricht besteht einfach aus einem 19 Byte BGP-Header [1].

5.3.3.4 NOTIFICATION-Nachricht

Eine NOTIFICATION-Nachricht wird immer dann gesendet, wenn ein Fehler entdeckt wurde. Nachdem diese Nachricht von einem BGP-Router gesendet wurde, wird umgehend die TCP-Verbindung zum benachbarten BGP-Router geschlossen. Anhand der NOTIFICATION-Nachricht lässt sich dann die Fehlerart im Routing-Protokoll eingrenzen. Eine solche Nachricht wird beispielsweise gesendet, wenn die definierte Hold Time bei einem BGP-Router überschritten wurde. Die Abbildung 5.6 illustriert das Format einer solchen Nachricht.

Eine NOTIFICATION-Nachricht enthält somit zusätzlich zum Header folgende Felder:

Error: Dieses 1 Byte grosse Feld gibt den Typ der NOTIFICATION an [1].

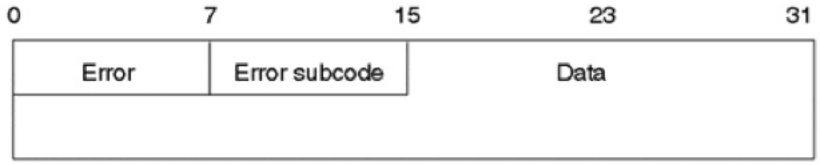

Abbildung 5.6: Format einer NOTIFICATION-Nachricht [1]

Error subcode: Dieses 1 Byte grosse Feld gibt spezifischere Informationen über die Art des Fehlers an. Jeder *Error Code* hat einen oder mehrere assoziierte *Error subcodes* [1]. Die Tabelle 5.1 zeigt alle möglichen Error Codes und die dazugehörigen Error subcodes.

Tabelle 5.1: Mögliche Error-Codes [3]

Error Code	Error Subcode
Message Header	1. Connection Not Synchronized 2. Bad Message Length 3. Bad Message Type
OPEN message error	1. Unsupposed Version Number 2. Bad Peer AS 3. Bad BGP Identifier 4. Unsupported Optional Parameter 5. Authentication Failure 6. Unacceptable Hold Time 7. Unsupported Capability
UPDATE message error	1. Malformed Attribute List 2. Unrecognized Well-Known Attribute 3. Missing Well-Known Attribute 4. Attribute Flags Error 5. Attribute Length Error 6. Invalid Origin Attribute 7. AS Routing Loop 8. Invalid NEXT_HOP Attribute 9. Optional Attribute Error 10. Invalid Network Field 11. Malformed AS_PATH
Hold Timer expired	N/A
Finite State Machine error	N/A
Cease (for fatal errors besides the ones already listed)	N/A

Data: Dieses variabel grosse Feld enthält fehlerrelevante Daten wie zum Beispiel eine illegale AS-Nummer oder einen fehlerhaften Header [3]

5.3.3.5 UPDATE-Nachricht

UPDATE-Nachrichten transportieren die Routing-Informationen zwischen BGP-Router und bilden somit den Kern des Inter-domain Routing. Eine UPDATE-Nachricht besteht zusätzlich zum Header aus drei Basisblöcken, einem Block aus Pfadattributen (Path Attributes), einem Block von unerreichbaren Routen (Unreachable Routes) und einem Block Network Layer Reachability Information [3]. Die Abbildung 5.7 gibt einen Überblick über die Basisblöcke einer UPDATE-Nachricht und deren jeweiliges Format.

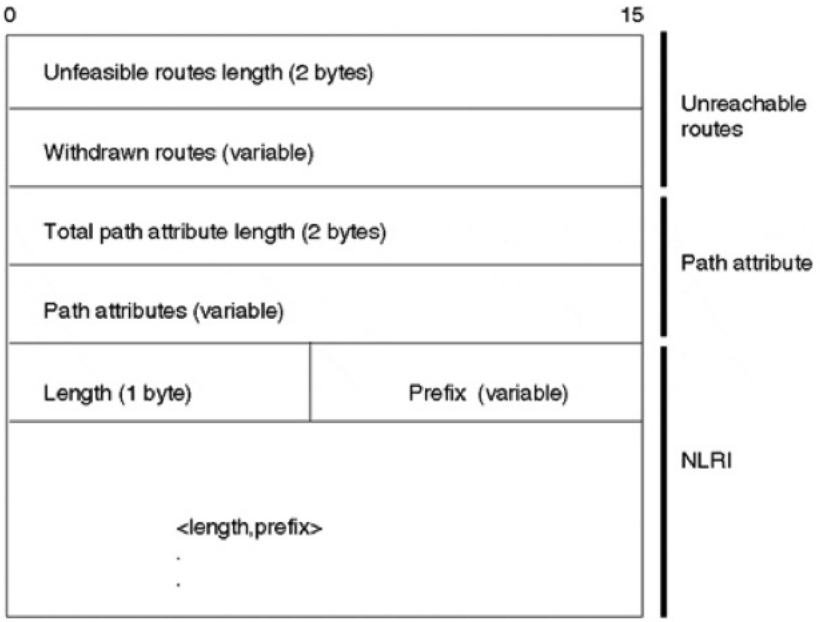

Abbildung 5.7: Format einer UPDATE-Nachricht [1]

Network Layer Reachability Information (NLRI): Mit dem NLRI-Block wird das Classless Inter Domain Routing durch BGP-4 unterstützt. Dieser Block beinhaltet eine Liste von IP-Prefixes zu Netzwerken, die dem benachbarten BGP-Router als „erreichbar“ kommuniziert werden sollen. Diese Erreichbarkeitsinfomationen werden als Tupel der Form <length, prefix> angegeben. Das *Lenght* Feld gibt die Länge der Maske eines IP-Prefixes in Bits an und das *Prefix* Feld beinhaltet eine IP-Netzwerkadresse, welche die Anzahl von Bits angibt, das die Netzwerknummer bildet [3]. Mit diesem Konzept können unter einer IP-Prefix mehrere Zielnetzwerke aggregiert werden. Mittel einer UPDATE-Nachricht kann nur eine Route kommuniziert werden.

Unreachable Routes: Dieser Block beinhaltet eine Liste von Routen zu Destinationen, die nicht mehr erreichbar sind und aus den Routing-Tabellen entfernt werden müssen. Die zu widerrufende Routen werden als Tupel der Form <length, prefix> im *Withdrawn Routes* Feld notiert. Das *Ufeasible Routes Legth* Feld gibt dabei die Gesamtlänge des *Withdrawn Routes* Feldes an. Es können somit mehrere Routen mittels einer UPDATE-Nachricht gleichzeitig widerrufen werden. Ein Wert von Null im *Unfeasible Routes Length* Feld bedeutet, dass keine Routen widerrufen werden.

Path attribute: Routes die mittels UPDATE-Nachrichten kommuniziert werden, haben assoziierte Pfad-Attribute. Diese Attribute werden dazu benutzt, um spezifische Informationen über eine Route zu übermitteln. Es handelt sich dabei um Informationen über den Pfad, den Degree of Preference einer Route, den NEXT_HOP Wert einer Route und Aggregationsinformationen [3]. BGP nutzt diese Attribute, um die beste Route zu einer bestimmten Destination auszuwählen, falls mehrere Pfade existieren. Diese Pfad-Attribute bilden dabei eine wichtige Grundlage für die Definition von Policies, welche im nächsten Kapitel näher erläutert werden.

Man unterscheidet vier Kategorien von Pfad-Attributen [3]:

Well-known mandatory: Dazu gehören Attribute, die in einer UPDATE-Nachricht enthalten sein müssen und von allen BGP-Implementationen anerkannt werden. Fehlt ein solches Attribut, wird eine NOTIFICATION-Nachricht erstellt und die Verbindung umgehend geschlossen. AS_PATH ist ein typisches Attribut, das in einer UPDATE-Nachricht enthalten sein muss.

Well-known discretionary: Darunter gehören Attribute die von allen BGP-Implementationen anerkannt, werden aber nicht zwingend in der UPDATE-Nachricht enthalten sein müssen. Das LOCAL_PREF Attribut gehört zum Beispiel zu dieser Kategorie von Attributen.

Zusätlich zu den Well-known Attributen kann eine Route auch optionale Attribute enthalten, welche entweder transitiv sind oder nicht. Diese Attribute müssen jedoch nicht von allen BGP Implementationen unterstützt werden.

Optional transitive: Wurde ein optionales Attribut von der BGP-Implementation nicht erkannt, wird dieses Attribut trotzdem akzeptiert und an weitere BGP-Peers weitergeleitet.

Optional nontransitive: Wurde das optionale Attribut von der BGP-Implementation nicht erkannt, wird das Attribut ignoriert und nicht an weitere BGP-Routers weitergeleitet.

Die Pfadattribute spielen für die Durchsetzung von Routing-Policies eine zentrale Rolle. Aus diesem Grund werden nachfolgend, die wichtigsten Attribute einer UPDATE-Nachricht beschrieben.

ORIGIN: Dieses Attribut gibt die Herkunkft der Pfad-Information an. Das ORIGIN-Attribut kann drei verschiedene Werte annehmen:

- IGP: Routes deren Herkunft innerhalb des AS stammt.
- EGP: Routes welche über das EGP erhalten wurden.
- Incomplete: Routes deren Herkunft unbekannt sind bzw. andersweitig erlernt wurden.

AS_PATH: Das AS_PATH-Attribut besteht aus einer Sequenz von AS-Nummern, welche den Pfad repräsentiert, den eine Route bisher durchlaufen hat. Dieses Attribut benutzt BGP, um eine schleifenfreie Topologie des Internets zu garantieren [3]. Jeder BGP-Router, der eine Route mittels einer UPDATE-Nachricht an einen benachbarten BGP-Router eines anderen AS übermittelt, erweitert zunächst den AS_PATH um die Nummer seines AS bevor er die Route weiterleitet. Das AS_PATH-Attribut ist somit ein Pfadvektor, der den gesamten Pfad von einem Ende der Route zum anderen Ende als eine geordnete Sequenz von AS-Nummern angibt [8]. Das letzte durchquerte AS steht dabei am Ende dieser Liste. Falls einem AS eine Route angekündigt wird, deren AS-Nummer bereits in AS_PATH vorhanden ist, wird die UPDATE-Nachricht ignoriert.

NEXT_HOP: Das NEXT_HOP Attribut beinhaltet die IP-Adresse des BGP-Routers, der als nächster Schritt genutzt werden soll um das angegebene Zielnetzwerk in der UPDATE-Nachricht zu erreichen [1]. NEXT_HOP gibt somit an, wohin ein ankommendes Packet als nächstes weitergeleitet werden soll.

LOCAL_PREF: Das LOCAL_PREF-Attribut ist ein well-known discretionary Attribut [1]. Es gibt den Grad der Präferenz von Routen an, welche dieselbe Destination beinhalten. Ein höherer LOCAL_PREF-Wert wird einem niedrigeren LOCAL_PREF-Wert bevorzugt. Dieses Attribut wird nur zwischen den Routern innerhalb eines AS ausgetauscht. Im Allgemeinen besitzt ein AS mehrere BGP-Router, welche Routen mit der gleichen Destination von verschiedenen ASen kommuniziert bekommen. Das LOCAL_PREF-Attribut wird dazu benutzt, um den bevorzugten Austrittspunkt zu einer bestimmten Destination festzulegen. Da das LOCAL_PREF-Attribut allen BGP-Router innerhalb eines AS bekannt ist, wird für eine bestimmte Route immer der gleiche Austrittspunkt gewählt. Das folgende Beispiel soll dies verdeutlichen [3]:

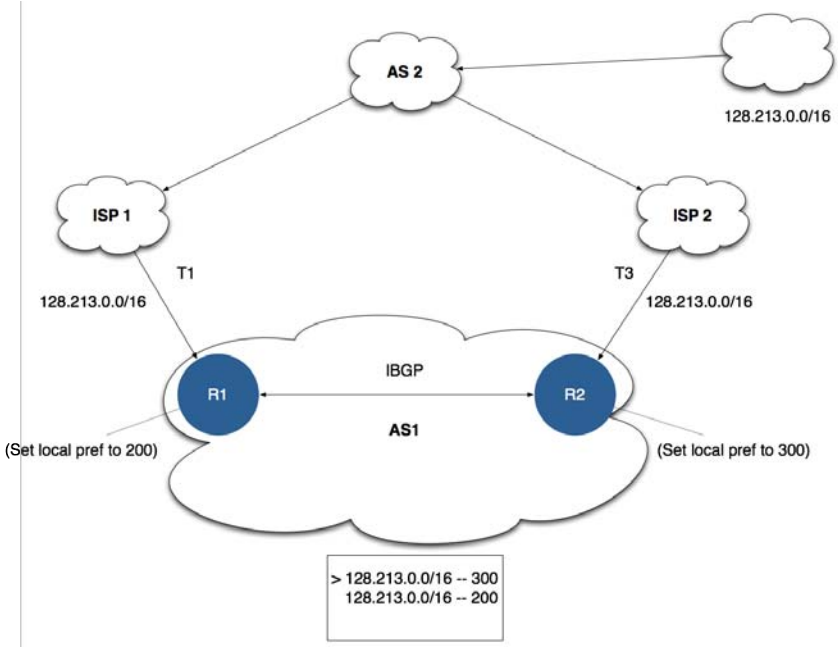

Abbildung 5.8: Beispiel zum LOCAL_PREF-Attributs [3]

In Abbildung 5.8 unterhält AS1 zwei Internet-Verbindungen zu verschiedenen ISPs (ISP1 und ISP2). Da die Verbindung über ISP 2 schneller ist als diejenige über ISP 1, möchte AS 1 den Verkehr bevorzugt über die Verbindung mit ISP 2 leiten. Um dies zu realisieren, wird das LOCAL_PREF-Attribut benutzt. Alle vom BGP-Router R1 erlernten Routen erhalten einen LOCAL_PREF-Wert von 200. Die erlernten Routen vom BGP-Router R2 erhalten hingegen einen höheren LOCAL_PREF-Wert von zum Beispiel 300. Da beide BGP-Router ihre von extern erlernten Routen miteinander austauschen, bevorzugen beide den Austrittspunkt über ISP 2 (aufgrund der höheren lokalen Präferenz). Wenn nun AS1 von beiden ISP eine Route mit derselben Destination (128.213.0.0/16) kommuniziert bekommt, sind sich beide Router einig, dass der Austrittspunktp über ISP 2 für die Destination 128.213.0.0/16 genutzt werden soll, da die lokale Präferenz höher ist.

MULTI_EXIT_DISC: Zwischen benachbarten ASen können mehrere Verbindungen bestehen. Der Multiexit Discriminator (MED) dient dazu, einem benachbarten AS die bevorzugte Route in das lokale AS vorzuschlagen. Dabei sollte ein tieferer MED-Wert einem grösseren MED-Wert bevorzugt werden. Das folgende Beispiel soll dies verdeutlichen [3]:

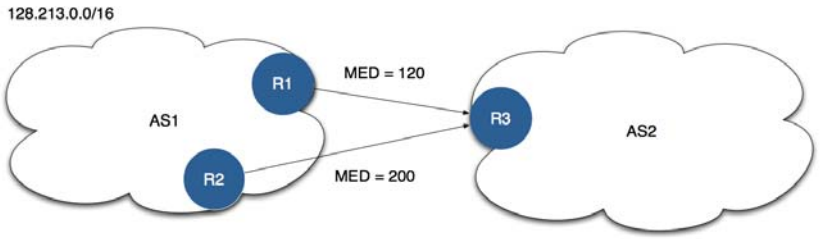

Abbildung 5.9: Beispiel zum LOCAL_PREF-Attributs [3]

In Abbildung 5.9 übermittelt AS1 dieselbe Route, mit dem IP-Prefix 128.213.0.0, über zwei verschiedene Router (R1 und R2) an AS2. Die über R2 kommunizierte Route enthält einen tieferen MED-Wert als diejenige über R1. In diesem Fall sollte das AS2 die Route über R2 bevorzugen, da sein MED-Wert kleiner ist. Es steht dem AS2 jedoch frei, auch die andere, über R1 kommunizierte Route, zu bevorzugen. Das MULIT_EXIT_DISC-Attribut stellt lediglich einen Vorschlag über den bevorzugten Eintrittspunkt eines AS dar.

Im Gegensatz zum LOCAL_PREF Attribut wird das MULTI_EXIT_DISC Attribut also zwischen ASen ausgetauscht, allerdings nur zwischen zwei direkt benachbarten Systemen. Erhält ein AS eine Route mit einem entsprechend MED-Wert, so wird dieser Wert wieder auf Null gesezt, falls die Route weitergeleitet wird. Beim MULTI_EXIT_DISC handelt es sich um ein optionales nicht transitives Attribut. Weiterhin gilt es zu beachten, dass MED-Werte für dieselbe Route, welche aber von unterschiedlichen ASen stammen, nicht verglichen werden dürfen.

COMMUNITY: Unter einer Community versteht man eine Gruppe von Destinationen, welche gemeinsame Eigenschaften besitzen. Diese Attribute beschränken sich nicht nur auf ein Netzwerk oder AS, sondern besitzen keine physischen Grenzen [3]. Communities dienen dazu die Routing-Policies zu vereinfachen, indem Routen

anhand ihrer logischen Eigenschaften statt einer IP-Prefix oder AS-Nummer identifiziert werden. BGP-Router können dieses Attribut zusammen mit anderen Attributen nutzen, um zu bestimmen, welche Routen akzeptiert, bevorzugt und an benachbarten Router übermittelt werden sollen [3]. Nachfolgend zwei wohlbekannte COMMUNITY-Attribute [3]:

- NO_EXPORT: Eine Route, die dieses Attribut besitzt, wird nicht an BGP-Peers ausserhalb des AS kommuniziert.
- NO_ADVERTISE: Eine angekündigte Route, die dieses Attribut besitzt, sollte keinem weiteren BGP-Peer kommuniziert werden.

Neben diesen wohlbekannten COMMUNITY-Attributen können auch private COMMUNITY-Attribute für spezielle Zwecke definiert werden. Eine Route kann dabei mehrere COMMUNITY-Attribute besitzen. Wurde einem BGP-Router eine Route mit mehreren COMMUNITY-Attributen angekündigt, kann dieser nur eine, ein paar, oder alle diese Attribute berücksichtigen [3]. Zudem bietet sich dem Router die Möglichkeit, die bestehenden COMMUNITY-Attribute zu verändern oder gar neue Attribute hinzuzufügen, bevor er die Route weiterleitet.

5.3.4 iBGP und eBGP

Man unterscheidet zwei Arten von BGP-Sessions zwischen benachbarten Router. Bei einer eBGP-Session werden Routing-Informationen zwischen BGP-Router verschiedener ASen ausgetauscht. Bei iBGP-Sessions hingegen werden die Routing-informationen zwischen BGP-Routern innerhalb desselben AS ausgetauscht. Beide nutzen dazu das selbe Protokoll, dienen aber unterschiedlichen Zwecken [4].

Im Allgemeinen besitzt ein AS mehr als einen Router der eine eBGP-Session zu benachbarten ASen unterhält. Jedem dieser eBGP-Router wird nur eine Teilmenge der Routing-Informationen übermittelt, die einem AS in der Gesamtheit kommuniziert wurden. Es besteht somit der Bedarf, dass jeder dieser eBGP-Router seine von extern erlernten Routen innerhalb des AS verteilt, um ein vollständiges Bild aller möglichen Routen innerhalb des AS zu schaffen. Die interne Verteilung von Routen muss dabei folgenden Anforderugen genügen [4]:

Schleifenlose Weiterleitung: Nach dem Austausch der von extern erlernten eBGP-Routen müssen die daraus resultierenden Routen, welche alle Router eines AS wählen, frei von Schleifen sein.

Komplette Transparenz: Die verschiedenen eBGP-Router innerhalb des AS müssen die externen Routing-Informationen so untereinander austauschen, dass sie ein vollständiges Bild aller externen Routen haben.

Die Verteilung der extern erlernten Routen innerhalb des AS erfolgt durch iBGP-Sessions. Die BGP-Spezifikation empfiehlt es, ein vollständiges Netz von iBGP-Session aufzubauen,

um Schleifen beim Austausch der Informationen zu verhindern. Jeder eBGP-Router unterhält demnach eine iBGP-Session mit jedem anderen BGP-Router innerhalb des AS. Ein eBGP-Router verteilt dann seine UPDATE-Nachrichten zu all seinen iBGP-Nachbarn. Die Empfehlung, dass iBGP-Routers miteinander ein vollstädiges Netz bilden, grenzt jedoch die Skalierbarkeit des AS stark ein [4].

Es gilt hervorzuheben, dass iBGP nicht mit einem Internal Gateway Protokol (IGP) wie RIP oder OSPF gleichzusetzen ist. iBGP dient lediglich dazu, externe Routen innerhalb des AS zu verteilen. Jeder Router im AS kombiniert dann die BGP- und IGP-Informationen, um eine Routing-Tabelle zu erstellen. Diese Tabelle assoziert zu jedem IP-Prefix ein oder mehrere ausgehende Verbindungen welche den kürzesten Pfad durchs Netzwerk zu einem eBGP-Router darstellen [9].

Die Konfiguration einer iBGP-Topologie, welche den oben genannten Anforderungen (Schleifenfrei und vollständige Transparenz) genügt, ist keine triviale Angelegenheit [4]. Eine falsche Konfiguration der Topologie kann leicht zu persistenten Schleifen und Oszillationen führen.

5.4 Policies

Seit das Internet kommerzialisiert wurde und die Aufrechterhaltung der Funktionalität von den staatlichen Organisationen zu privaten Unternehmen übergegangen ist, wurde es wichtig, dass die Betreiber das Routing beeinflussen können, um die Erreichung ihrer ökonomischen Ziele gewährleisten zu können. Mit Policies kann ein Betreiber eines AS bestimmen, wie mit Routing-Informationen (Advertisements) umgegangen wird. Indem er den Fluss der Daten indirekt beeinflusst, kann er seine ökonomischen Ziele besser erreichen. Man spricht dabei von Policy-Strategien, wobei mit Filtern und den BGP-eigenen Attributen Policies implementiert werden.

5.4.1 Definition von Policies

Policies sind im RFC 4271 definiert. Sie werden als Regeln angegeben, welche man darauf anwenden kann, in welcher Weise mit ankommenden und ausgehenden Routenankündigungen umgegangen wird, und wie sie intern behandelt werden [4]. Daraus entstehen drei Arten von Policies:

- Import Policies: Damit wird bestimmt, was mit den versch. reinkommenden Routen passiert [9].
- Export Policies: Damit wird bestimmt, was mit den versch. rausgehenden Routen passiert [9].
- Interne Policies: Damit wird bestimmt, wie innerhalb des AS mit den versch. Routen umgegangen wird [9].

5.4.2 Policy-Strategien

Typische Fälle von Policies sind, dass ein ISP billigere Routen bevorzugt, wenn ein Ziel über mehrere Routen erreichbar ist. Oder ein ISP möchte seinen eingehenden Netzwerkverkehr gleichmässig über seine Eingangsspunkte verteilen, um Stau auf bestimmten Routen zu verhindern [9]. Diese Beeinflussung des Netzwerkverkehrs, welcher sonst einfach die Route mit der kürzesten Distanz zum Ziel nehmen würde, wird über die BGP-Attribute gesteuert [9].

5.4.2.1 Geschäftsbeziehungen

ISPs haben das Ziel, möglichst viel Datenfluss über ihre Kunden und möglichst wenig Datenfluss über ihre Provider und Peering-Partner zu produzieren. Mit der Kontrolle, welche Routen weitergleitet werden, kann der Datenfluss indirekt gesteuert werden.

- Routen-Export kontrollieren: ISPs können kontrollieren, welche Routen weitergeleitet werden. Häufig möchten ISPs Routen, welche sie von ihrem Provider übermittelt bekommen, nicht an ihre Peering-Nachbarn weitergeben, weil sie kein Interesse daran haben, dass diese die Route über sie benutzen. Technisch wird das so implementiert, dass beim Empfang einer Route vom Provider ein Tag in Form eines Community-Attributes angehängt wird (Import Policy). So kann innerhalb des AS unterschieden werden, von wo die Routenankündigung gekommen ist und beim Export an den Peering-Nachbarn eventuell gefiltert werden (Export Policy).
- Kunden-Routen bevorzugen: Oft zählen ISPs auch andere ISPs zu ihren Kunden. Diese Kunden haben häufig weitere Verbindungen zu anderen Providern und erreichen dadurch weitere Adressen. Ist ein Ziel erreichbar über einen Provider und über einen Kunden, so wird meist die Kunden-Route bevorzugt, da sie Einnahmen generiert. Um das zu erreichen wird der LocalPref Wert für Routen zu Kunden höher gesetzt als für Routen zu Providern. So werden die Kunden-Routen im Normalfall bevorzugt behandelt.

5.4.2.2 Datenflusssteuerung

ISPs sind versucht, den Datenfluss zu steuern. Mit BGP kann dies nur indirekt via BGP-Attribute gemacht werden.

- Hot-Potato Routing: Ein ISP ist häufig versucht, seine Netzwerklast innerhalb seines AS tief zu halten. Ein Weg, dies zu erreichen, ist, nicht für das eigene AS bestimmte Datenpakete auf dem nächstkürzesten Weg wieder aus dem AS zu befördern. Dies ist zwar schlecht für die globale Routing-Performance, lindert aber dem ISP seine Netzwerklast-Probleme.

- Netzwerklast verteilen: Möchte ein ISP die Netzwerklast auf seinen Eingangsleitungen besser verteilen, macht er das mit dem MED-Attribut. Er kann mit diesem Attribut seinen angeschlossenen ASs angeben, auf welcher Leitung er mehr und auf welcher er weniger Daten bevorzugt.

5.4.2.3 Skalierbarkeit

Bei BGP ist die Skalierbarkeit ein Problem, da die Routing-Tabellen sehr gross werden können. Dies kommt daher, dass viele Ziele auf der Welt bekannt sind. Strategien dagegen sind:

- Routing-Tabellengrösse: Um die Routing-Tabellengrösse zu kontrollieren, gibt es zwei Strategien: a) Man macht aggressive Aggregation von IP-Prefixen und lässt nur noch Routen in der Routing-Tabelle zu, welche eine gewisse Prefix-Länge haben (z.B. /19), oder b) man filtert seine eigenen Routen und verschickt nur Routen-Ankündigungen mit einer bestimmten Prefix-Länge.
- Routing-Änderungen: Wenn eine Route instabil ist, gibt es häufige Routen-Änderungen (flapping). Dies ist unerwünscht, da es den Rechenaufwand der Systeme erhöht. Als Gegenmassnahme benutzen ISPs häufig „flap damping“. Dabei wird jeder Route ein Penalty-Wert zugeordnet, welcher hochzählt, wenn es eine Änderung im Zustand der Route gibt. Erreicht der Wert einen gewissen Schwellwert, dann wird die Route bei den weiteren Routing-Entscheidungen ignoriert.

5.5 Fundamentale Probleme im Inter-domain Routing

Nachdem in den vorherigen Kapiteln das Border Gateway Protokoll ausführlich erläutert wurde, sollen im Folgenden noch einige fundamentale Probleme von BGP aufgezeigt werden.

5.5.1 Skaliebarkeitsprobleme

Es gibt in der Welt immer mehr Destinationen, die ihre Erreichbarkeit mittels BGP in die Welt hinaus kommunizieren. Daher muss das Inter-domain-Routing bezüglich AS, Routern und Destinationen skalierbar sein. Um dies sicherzustellen, werden drei Methoden/Techniken verwendet:

- Repräsentation eines AS als Einheit: Jedes AS wird vom Rest der Welt als „ein“ Ziel angeschaut. Das heisst, alle Ziele die innerhalb des AS sind, interessieren den Rest der Welt nicht direkt. Sie schicken die Daten einfach ans AS und dieses schaut dann selbst wohin damit. Diese Technik wird ermöglicht, indem ASe nur den Bereich

der Adressen die in ihrem System sind weiterkommunizieren, und nicht jede einzelne Adresse selber (Mittels kürzeren Prefixen). Damit wird die Adressenmenge drastisch reduziert im weltweiten Netz.

- Filtering: Ein AS blockt mittels Import Policies gewisse Adressen schon bei ihrer Bekanntgabe ab und nimmt sie gar nicht erst in ihre Routing-Tabellen auf. Dies sind zum Beispiel Adressen, welche aus Ländern kommen, welche man boykottieren möchte.

- Prefix aggregation: Um die Adressen zu reduzieren, welche in den Routing-Tabellen stehen, versuchen ASe auch mehrere ähnliche Adressen, welche im gleichen AS sind, zu aggregieren und so zwar einen ungenaueren, aber dafür nur einen Eintrag in der Routing-Tabelle zu haben.

5.5.2 Instabilität von Routen

Im Internet ist es wichtig, dass Routen über längere Zeit sich nicht verändern, da jede Änderung an alle Internetteilnehmer kommuniziert werden muss. Passieren solche Änderungen zu häufig, kann das Netz instabil werden. Darum sind instabile Routen ein grosses Problem im Inter-domain Routing.

- Oszillieren: Bei BGP kann es unter gewissen Umständen vorkommen, dass zwei Routen zum gleichen Ziel beginnen zu oszillieren. D.h. eine Route wird unerreichbar, worauf die andere Route überlastet und auch unerreichbar wird. Kurz darauf wird die erste Route, und in der Folge auch die zweite Route, wieder erreichbar. Dies kann sich unendlich wiederholen. Dagegen gibt es aber Strategien, wie das Flap Damping [9].

- Falschkonfiguration: Durch die Komplexität von BGP ist es manchmal schwer vorauszusagen, was für Auswirkungen eine Änderung in einer Policy auf den Rest des Netzes hat. So kann ein lokaler Anbieter mit eine falschen Einstellung Probleme verursachen, welche in der ganzen Welt bemerkbar werden [10].

5.5.3 Sicherheit

Da BGP-Nachrichten weder verschlüsselt verschickt werden, noch auf Echtheit des Absenders überprüft werden können, gibt es Möglichkeiten die Nachrichten zu fälschen. Dadurch tun sich verschiedene Sicherheitslöcher auf. Ein AS kann zum Beispiel mit einer gefälschten Routenankündigung den Netzwerk-Verkehr einer bestimmten Destination über sich selber laufen lassen und dann zum Ziel weiterleiten und so den gesamten Verkehr abhören. Oder es kann den Verkehr auch gar nicht weiterleiten und so das Ziel unerreichbar machen bis geeignete Gegenmassnahmen ergriffen werden [10].

5.6 Fazit

Das Inter-domain Routing ist geprägt von den Geschäftsbeziehungen zwischen den Internet Service Providern, welche die autonomen Syteme verwalten. Die Service Provider versuchen über die Definition von Policies den Packetverkehr möglichst gewinnbringend durch ihr eigenes autonomens System zu lenken. Gleichzeitig müssen sie aber miteinander kooperieren, um eine globale Erreichbarkeit zu gewährleisten, was einen gewissen Wiederspruch in sich birgt.

Grundlage des Inter-domain Routings bildet dabei das Border Gateway Protokoll, ein in seinem Kern eher simples Protokoll, das im Praxiseinsatz extrem komplex ist [4]. Diese Komplexität von BGP rührt aus zwei Anforderungen an das Inter-domain Routing. Erstens besteht seitens der ISPs der Bedarf, Policies autonom definieren zu können, um den Packetverkehr zu lenken. Zweitens muss das Protokoll mit dem stetig wachsenden Internet skalierbar sein. Daraus ergeben sich eine Vielzahl von Problemen des Border Gateway Protokolls, wie zum Beispiel durch Policies verursachte Oszillationen, Falschkonfigurationen oder Skalierbarkeitsprobleme aufgrund der Aggregation von Routing-Informationen. Obwohl auf diesem Gebiet intensiv geforscht wird und in den letzten Jahren auch einige Fortschritte erreicht wurden, bleibt das Inter-domain Routing immer noch schwer zu verstehen und zu modellieren [4].

Literaturverzeichnis

[1] Y. Rekhter, T. Li, S. Hares: *A Border Gateway Protocol 4 (BGP-4)*. IETF RFC 4271, January 2006.

[2] C. Huitema: *Routing im Internet*. Prentice-Hall, 1996.

[3] S. Halabi, D. McPherson: *Internet Routing Architectures*. Cisco Press, 2000.

[4] H. Balakrishnan, N. Feamster: *Interdomain Internet Routing, Lecture4*, MIT Lecture Notes, 2005. `http://pages.cs.wisc.edu/~akella/CS740/F08/740-Papers/hari-bgp-notes.pdf`

[5] J. Doyle, J. DH. Carrol: *Routing TCP/IP*, CCIE Professional Development, Vol. 2, Indianapolis, Cisco Press, 2007.

[6] N. Feamster, H. Balakrishnan, J. Rexford: *Some Foundational Problems in Interdomain Routing*, 3rd ACM SIGCOMM Workshop on Hot Topics in Networks (HotNets), San Diego, CA, USA, November 2004.

[7] A. S. Tannenbaum: *Computer Networks*, Fourth Edition, New Jersey, Pearson Education International, 2003.

[8] C. Hunt: *TCP/IP Network Administration*, Third Edition, O'Reilly, April 2002.

[9] M. Caesar, J. Rexford: *BGP routing policies in ISP networks*, California Univ., Berkeley, CA, USA, December 2005.

[10] YouTube Hijacking: A RIPE NCC RIS case study; `http://www.ripe.net/news/study-youtube-hijacking.html`, May, 2009.

Kapitel 6

Information and Communication Technology for the Needy

Rahel Jerjen, Sarah Schneiter

Informations- und Kommunikationstechnologien sind ein wichtiger Entwicklungsfaktor eines Landes. Von Interesse ist deshalb zu wissen, wie man Menschen in Ländern, die noch nicht so weit entwickelt sind, erreichen kann, wie man mit IKT ihren Bedürfnissen gerecht wird, einen Nutzen für die Personen vor Ort generiert, die Situation verbessert, die Wirtschaft stärkt, die Entwicklung in diesem Bereich fördert.
In Europa ist der Entwicklungsstand unterschiedlich. In Dänemark, einem der reichsten Länder Europas, ist die Entwicklung bereits weit fortgeschritten und läuft stets weiter. In Litauen, einem eher ärmeren Land Europas, ist die Entwicklung noch nicht so weit fortgeschritten, doch schreitet sie in rasantem Tempo voran.
Ausserhalb Europas ist die Lage noch unterschiedlicher. In Afrika ist die Abdeckung mit IKT verhältnismässig gering und in Asien differiert der Entwicklungsstand stark nach Land, Region. In den ärmeren Regionen wurden und werden jedoch verschiedene Projekte durchgeführt, wie das Dorftelefon, Telezentren und 'One Laptop Per Child (OLPC)'.
IKT sind in der heutigen Zeit sehr wichtig für die Entwicklung eines Landes und deshalb ist es zentral, dass auch ärmere Menschen, Regionen, Länder nicht von den Technologien abgeschnitten und den Möglichkeiten, die jene Technologien eröffnen, beraubt werden. Gerade in diesen Gebieten lässt sich mit Hilfe neuer Technologien viel erreichen. Menschen, die kaum gebildet sind, können sich übers Radio, Telefon, Internet informieren und weiterbilden. Regionen können durch verschiedene Technologien ihre Situation verbessern und Länder können über verschiedene IKT ihre Entwicklung nachhaltig positiv beeinflussen.

Inhaltsverzeichnis

6.1 Einleitung und Fragestellung

6.1.1 Einführung in die Thematik

Informations- und Kommunikationstechnologien, kurz IKT, sind in den letzten Jahren für die Wirtschaft und die Entwicklung eines Landes immer bedeutender geworden. Die Menschen in entwickelten Ländern mit einer guten IKT-Abdeckung wachsen heute bereits von Kindesalter an in die Nutzung von IKT hinein. Dies ist jedoch in weniger entwickelten Ländern nicht der Fall. Diese Länder sind noch auf dem Stand, auf dem auch die entwickelten Länder vor noch nicht allzu langer Zeit waren, als erste IKT aufkamen, sich rasant entwickelten und verbreiteten.

Die Entwicklung im Bereich von IKT ist sehr wichtig. Länder mit genügend Ressourcen entwickeln IKT fortwährend weiter und schreiten dadurch auch in ihrer Gesamtentwicklung rasant voran. Wenn weniger entwickelte Länder den Anschluss verpassen und sich nicht darum bemühen in diesen Technologien Fuss zu fassen, werden sie mit den fortgeschrittenen Ländern nicht mehr mithalten können. Natürlich ist es schwierig für weniger entwickelte Länder die nötigen Ressourcen aufzubringen, um den Einstieg zu schaffen. Eine gewisse Infrastruktur ist von Nöten, die Technologien an sich sind zu erwerben oder zu produzieren, Fachwissen ist zentral und für eine nutzbringende Einführung der Technologien ist auch die Schulung der Bevölkerung wichtig. All dies und die weiteren benötigten Ressourcen kosten Geld, welches weniger entwickelte Länder meist nicht oder nur teilweise aufbringen können.

6.1.2 Ziele und Aufgabenstellung

Die zentrale Frage, welche man sich zur Förderung von IKT in Entwicklungs- und Schwellenländern stellen muss, ist: Inwiefern sind die Projekte, wie Dorftelefone, Telezentren und 'One Laptop per Child', längerfristig von Nutzen für alle Beteiligten in der jeweils anzutreffenden Situation? Es ist wichtig, sich auch technischen, wirtschaftlichen und sozialen Grenzen bewusst zu sein. Deshalb werden die Betrachtungen in der vorliegenden Arbeit auf diese Punkte gerichtet und unterschiedliche Vertiefungen gewählt, da die Situation geographisch und situationsbedingt abweichen kann. Einerseits wird die Situation innerhalb Europas exemplarisch an den Ländern Dänemark und Litauen erläutert. Andererseits wird auch die Situation ausserhalb Europas, in Afrika und Asien (Indien) betrachtet. Schliesslich werden auch noch erwartete Auswirkungen diskutiert. Doch zuerst wird nun der Blick auf die betrachteten IKT gerichtet, damit die Grundlagen für die danach folgenden Ausführungen vorhanden sind.

6.2 Betrachtete Informationstechnologie

Unter dem Begriff IKT werden Technologien im Bereich von Information und Kommunikation zusammengefasst. Darunter fallen unter anderem das Radio, das Telefon, das

Mobiltelefon und das Internet. Dies sind nur einige vieler IKT. Diese wurden im Rahmen der Arbeit ausgewählt, weil sie für das Thema und die Fragestellung besonders relevant sind. Relevant sind sie aus verschiedenen Gründen: das Radio, weil es meist eines der ersten IKT ist, welches sich in ärmeren Gebieten verbreitet; das Telefon, weil man damit nicht nur einseitig sondern auch zweiseitigen kommunizieren kann; das Mobiltelefon, weil es im Gegensatz zum Telefon nicht von physikalischen Leitungen und Kabeln abhängig ist; das Interent, weil es die vielseitigste IKT darstellt, welche nahezu unbegrenzte Nutzungsmöglichkeiten bietet. Im folgenden werden diese Technologien gemäss World Encyclopedia [8] definiert.

Das Radio ermöglicht das Senden und Empfangen von elektromagnetischer Strahlung in Form von Radiowellen. Ein Transmitter generiert ein Radiosignal mit fester Frequenz (Trägerwelle), ein Mikrofon konvertiert das Geräusch in ein elektrisches Signal, das mit Hilfe einer Modulation mit der Trägerwelle verknüpft wird. Die modulierte Trägerwelle wird zu einer Antenne übermittelt, welche das Signal in die Atmosphäre sendet. Der Empfänger, eine andere Antenne, fängt das Signal ab und dieses wird demoduliert, um das Radiosignal abzurufen.

Das Telefon ist ein Kommunikationsmittel, welches Sprachlaute über Distanzen mit Hilfe von Kabeln überträgt. Der Prototyp hatte eine Membran aus Weicheisen, welche zu den Schallwellen vibrierte. Diese Vibrationen verursachten Störungen im magnetischen Feld des Magnetstabes, welche wiederum einen elektrischen Strom von schwankender Intensität im dünnen Kupferdraht, welcher um den Magnet gewickelt war, verursachte. Der Strom konnte mittels Kabeln zu einem entfernten Apparat gesendet werden, welcher den Prozess umkehrte und so den Hörschall reproduzierte. In Weiterentwicklungen wurde der Sender vom Empfänger separiert und an Stelle des Magneten traten Batterien.

Das Mobiltelefon ist ein tragbares Telefon, welches den Benutzer mit dem öffentlichen Telefonsystem verbindet. Mobiltelefone arbeiten innerhalb eines Netzwerkes von Funkzellen. Die erste Generation arbeitete mit analogen Signalen, die zweite Generation mit digitalen Signalen und die dritte Generation beinhaltet Hochgeschwindigkeitsdatentransfer und kann sich mit dem Internet verbinden.

Das Internet ist ein weltweites Kommunikationssystem, welches eine Vielzahl heterogener Computernetzwerke verbindet. Es ist ein Netzwerk von Netzwerken in denen Nachrichten und Daten über kurze lokale Verbindungen von Ort zu Ort gesendet werden. Das World Wide Web besteht aus einer Reihe von Computernetzwerken, auf die über lokale Server zugegriffen werden kann. Das World Wide Web beinhaltet Seiten, welche Benutzer übers Internet abrufen können, um Daten zu erhalten oder zu senden.

Diese IKT ermöglichen die Kommunikation und die Beschaffung, Verarbeitung und Bereitstellung von Informationen in neuen Ausmassen. Doch zu ihrer Implementierung müssen zuerst die nötigen Voraussetzungen vorhanden sein. Ressourcen wie die Finanzen, die Infrastruktur, die technischen Mittel, die Fähigkeiten etc. müssen gegeben sein.

Im Folgenden werden die eben definierten IKT im Zusammenhang mit Europa, spezifisch mit Dänemark und Litauen und anschliessend im Zusammenhang mit Afrika und Asien, respektive verschiedenen Gebieten der genannten Kontinente, diskutiert.

6.3 Bedeutung der Informationstechnologie in Europa

IKT sind in den letzen Jahren immer verbreiteter und auch bedeutender geworden. Um im Bereich von IKT in Europa einen Einblick zu gewinnen, wird zuerst die IKT-Situation in Dänemark, einem der reichsten Länder Europas, und anschliessend die IKT-Situation in Litauen, einem der ärmeren Länder Europas, genauer erläutert. Wichtig ist jedoch nicht nur die IKT-Situation, sondern auch die Bedeutung der verschiedenen IKT für das jeweilige Land und seine Bewohner.

6.3.1 Dänemark

Die folgenden Angaben zu Dänemark sind, falls nicht anders angezeigt, aus dem Bericht 'Key Figures on the Danish Information Society 2006 - Danish Figures' [9] entnommen.

Der IKT-Sektor nimmt in Dänemark eine wichtige Stellung ein. Der Export von IKT-Produkten beträgt 8 % aller Exporte im Jahr 2005 und nimmt die gleiche Position ein wie der Export von medizinischen Produkten und Landwirtschaftsprodukten.

Der Zugang zum Internet ist gut abgedeckt, so haben im Jahr 2005 bereits 79 % der Bevölkerung einen eigenen Internetanschluss zu Hause. Die meisten davon, 78 %, haben Zugang über einen PC, 4 % über einen Organizer und 13 % über ihr Mobiltelefon.

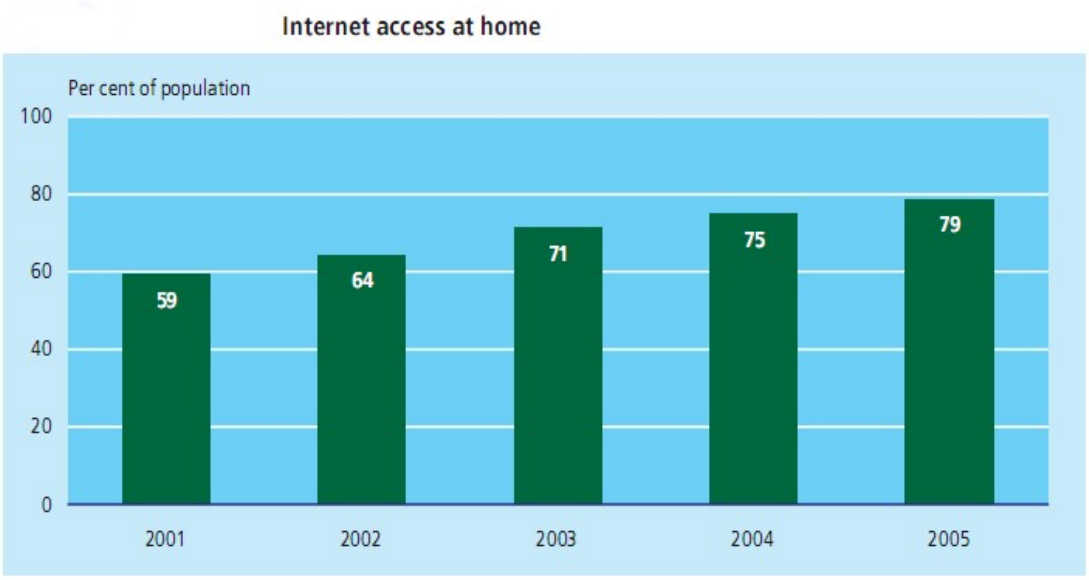

Abbildung 6.1: Internetzugang zu Hause [9]

Das Internet wird am häufigsten zur Kommunikation, Informationssuche und Nutzung von Onlinediensten genutzt. Einen Breitbandanschluss haben 55 % der Gesamtbevölkerung, im Jahr 2001 sind es hingegen erst 7 %. Die Preise für ADSL 512 kbit/s und 2 Mbit/s

fallen von 2001 bis 2005 um 60 %. Und auch die Preise für andere Breitbandverbindungen, z.B. Breitband übers Festnetz und übers Mobiltelefon, nehmen ab. Von den 79 % der Gesamtbevölkerung, die zu Hause einen Internetanschluss haben, sind es unter den 16-39jährigen 87 %, unter den 40-59jährigen 82 % und unter den 60-74jährigen 52 %.

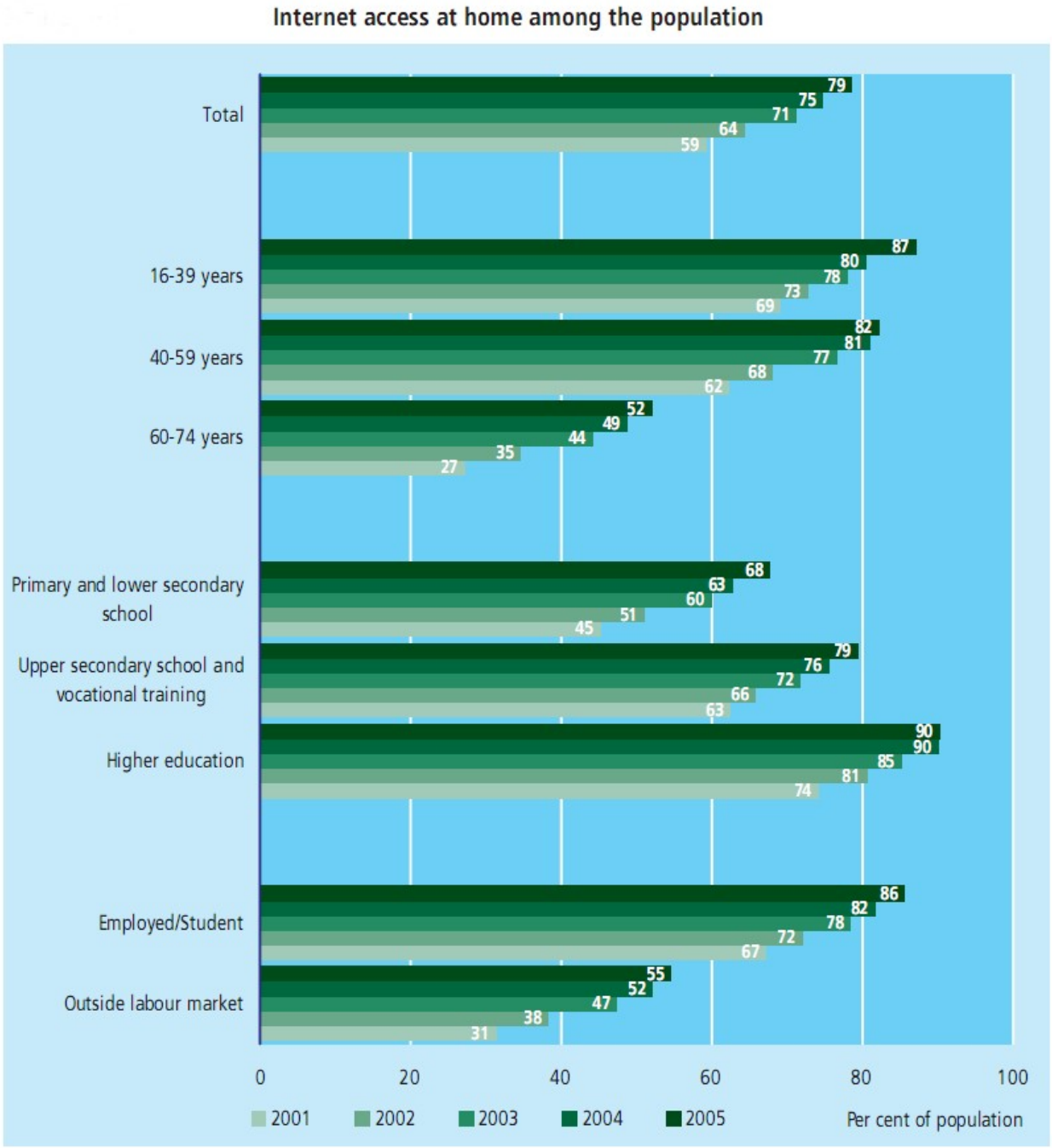

Abbildung 6.2: Internetzugang nach Alter und Bildung [9]

Alle Altersgruppen erhalten von 2001 bis 2005 vermehrten Zugang zum Internet, der höchste Anstieg erfolgt im Bereich der 60-74jährigen. Personen mit einer höheren Bildung, in Unternehmen und Studenten haben den höchsten Anteil an Internetzugängen zu Hause.

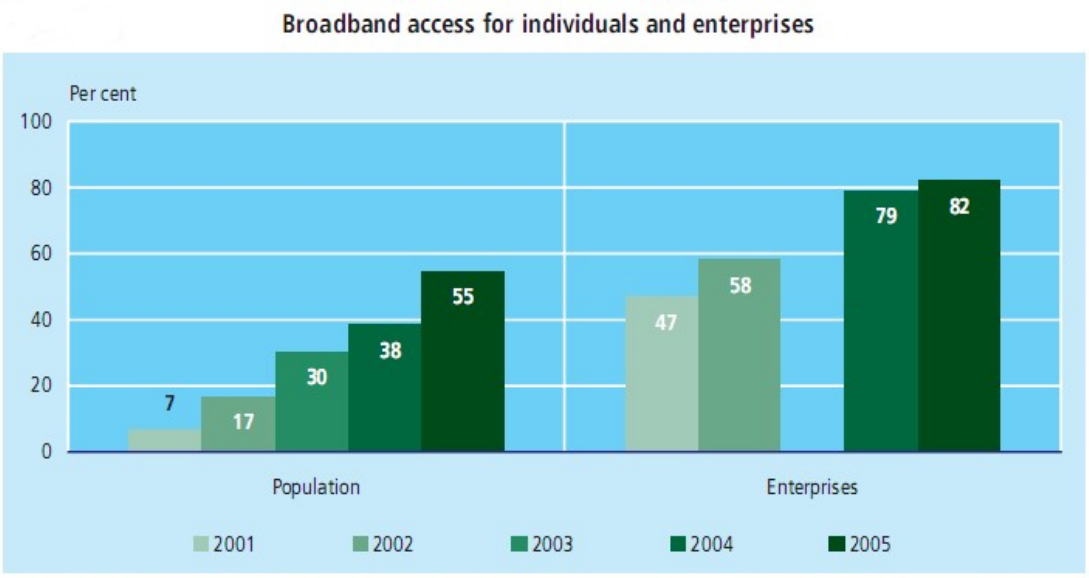

Abbildung 6.3: Breitbandzugang der Bevölkerung und Unternehmen[9]

82 % aller Firmen haben im Jahr 2005 einen Breitbandzugang zum Internet, im Jahr 2001 sind es hingegen nur 47 %. Firmen nutzen das Internet zur Abfertigung von Aufträgen und zum Austausch von Daten. Auch andere IKT-Systeme werden in Firmen eingesetzt. 69 % der Firmen fertigen im Jahr 2005 Aufträge über IKT-Systeme ab. Sie nutzen IKT-Systeme zudem häufig um ihre Geschäftsprozesse zu unterhalten. 92 % der Firmen nutzen die Systeme zur Finanzverwaltung, 69 % zur Auftragsabwicklung, 52 % fürs Management von Produktion, Logistik und Dienstleistungen, 47 % fürs Personalmanagement und 24 % fürs Marketingmanagement. Das grösste Hindernis für Firmen sind Fehler und Defekte in der Software, fehlende Flexibilität der IKT-Anbieter und höhere Kosten als erwartet. Im Jahr 2005 nutzen 87 % der Firmen öffentliche technische Dienste um Informationen auf Webseiten zu suchen, Formulare herunterzuladen und auch zu übertragen. 52 % der Firmen im IKT-Sektor sind im Jahr 2003 aktiv um Innovation bemüht. Die Ausgaben für die Forschung und Entwicklung im IKT-Sektor sind pro Angestellten viel höher als im Businesssektor insgesamt. Der IKT-Sektor finanziert 68 % seiner Forschung und Entwicklung selbst, 3 % stammen von anderen dänischen Quellen aus dem privaten und öffentlichen Bereich und 29 % werden von privaten und öffentlichen ausländischen Quellen bezahlt.

Die öffentlichen Behörden haben zu 91 % elektronische Systeme zum Erfassen von Dokumenten und Fallakten, zu 74 % ein elektronisches Dokumentenmanagement und zu 52 % ein elektronisches Fallmanagement. Zudem sind interne Arbeitsabläufe zur Bescheinigung und Bezahlung in 68 % vollkommen digitalisiert. Die fünf zentralsten Hindernisse fürs E-Government sind Schwierigkeiten im Freigeben von Ressourcen, fehlende Lösungen im öffentlichen Sektor, Schwierigkeiten bestehende Systeme zu integrieren, fehlende allgemeine Standards und höhere IKT-Kosten als erwartet.

Im Jahr 2005 nutzen 17 % der Bevölkerung das Internet für Zwecke der Bildung und des Trainings und 13 % der Unternehmen nutzen das Internet zum Trainieren ihres Personals.

24 % der Gesamtbevölkerung interagiert mit öffentlichen Behörden, 10 % der Gesamtbevölkerung interagiert übers Internet. Studenten und Angestellte interagieren zu 26 % mit öffentlichen Behörden und 11 % aller Studenten und Angestellten interagieren übers Internet.

Eines der höchsten IKT-Sicherheitsprobleme sind Virusattacken. Während Virusattacken von 2002 bis 2005 unter Bürgern etwas ansteigen, gehen die Attacken im Unternehmensbereich und im öffentlichen Sektor zurück. Die Unternehmen erleben im Jahr 2005 zu 13 % Denial of Service-Attacken und zu 5 % unbefugten Zugang zum System und zu Daten des Unternehmens. Der öffentliche Sektor erlebt zu 17 % Denial of Service-Attacken und zu 10 % unbefugten Zugang zum System und zu Daten. Die zentrale Regierung besitzt 2005 zu 91 % eine bewährte, erprobte IKT-Sicherheits-Strategie, regionale Behörden ebenfalls zu 91 % und lokale Behörden zu 88 %.

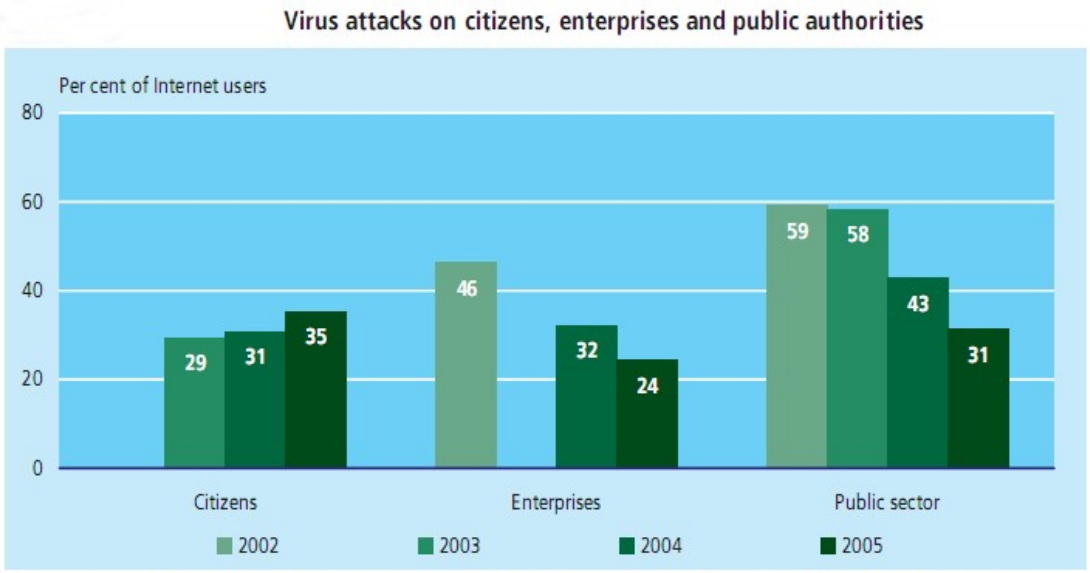

Abbildung 6.4: Virusattacken bei der Bevölkerung, den Unternehmen und den öffentlichen Behörden [9]

Gerade im Unternehmensbereich und für die Regierung kann es beträchtliche Folgen haben, wenn IKT von Viren attackiert werden. Denial of Service-Attacken können das System für eine gewisse Zeit lahmlegen, und um das System wieder in den normalen Betrieb überzuführen, werden Ressourcen wie Zeit und Geld benötigt. Dies bedeutet nicht nur einen Mehraufwand auf Seiten der betroffenen Stelle sondern auch auf Seiten jener Nutzer, die in dieser Zeit auf das System zugreifen wollen. Ihre Anfragen werden kaum durchkommen und bearbeitet, so dass sie die Stelle später noch einmal kontaktieren müssen. Noch grösser sind jedoch die Risiken, wenn unbefugt auf das System und die Daten zugegriffen wird. Die gestohlenen Informationen, wie z.B. vertrauliche oder geheime Daten werden vielleicht veröffentlicht. Die Folgen daraus können zahlreich sein, so leidet beispielsweise das Image der betroffenen Stelle darunter, oder Geheimhaltungsgesetze oder -vereinbarungen werden verletzt, oder die Zeit und das Geld, welches in die Entwicklung bestimmter Produkte investiert wurde, verlieren teilweise ihren Wert. Die

Informationen können der betroffenen Stelle jedoch auch schaden, indem sie zur Erpressung missbraucht werden, oder um die Entwicklung eines eigenen Produkts voranzutreiben u.v.m. Virenattacken und die daraus folgenden Konsequenzen schaden dem entsprechenden Unternehmen, der Wirtschaft, dem Land und den Bewohnern. Daher ist es wichtig IKT entsprechend gut zu sichern.

6.3.2 Litauen

Die folgenden Angaben zu Litauen stammen, falls nicht anders referenziert wird, aus dem Bericht 'Information Technologies in Lithuania' [10].

In Litauen beläuft sich der Export von IKT-Produkten im Jahr 2006 auf 5 % des Gesamtexports und der Import beläuft sich auf 6 % des Gesamtimports. Der grösste Anteil des Imports machen Computer und dazugehörige Ausrüstung (34 %) und Telekommunikationsausrüstung (24 %) aus. Den grössten Exportanteil haben Audio- und Videoausrüstung (33 %). IKT-Produkte machen 2 % aller Industrieprodukte aus. Die Erträge sind von 2005 bis 2006 um 25 % gestiegen. IKT-Unternehmen machen im Jahr 2006 4 % des Businesssektors aus. Insgesamt sind 4 % der Bevölkerung Litauens in diesen Unternehmen angestellt.

Im Jahr 2007 haben 42 % der Haushalte einen PC zu Hause, 53 % in städtischen Gebieten und 24 % in ländlichen Gegenden. Einen Desktopcomputer besitzen 38 %, einen Laptop besitzen 10 % und 78 % haben ein Mobiltelefon.

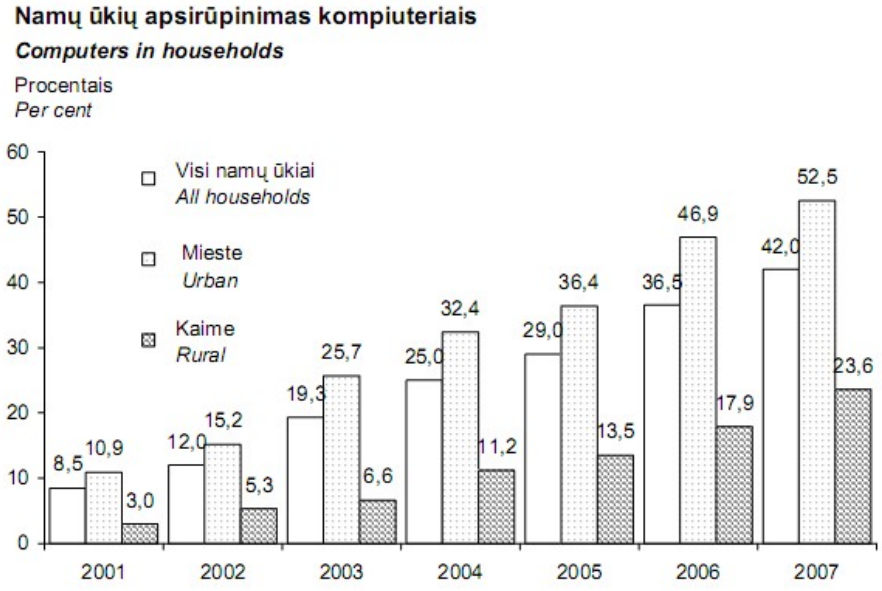

Abbildung 6.5: Computerzugang zu Hause [10]

Internetzugang zu Hause haben im Jahr 2007 40 % der Haushalte, 50 % in städtischen Gebieten und 23 % in ländlichen Gegenden. 73 % der Haushalte mit Internetzugang

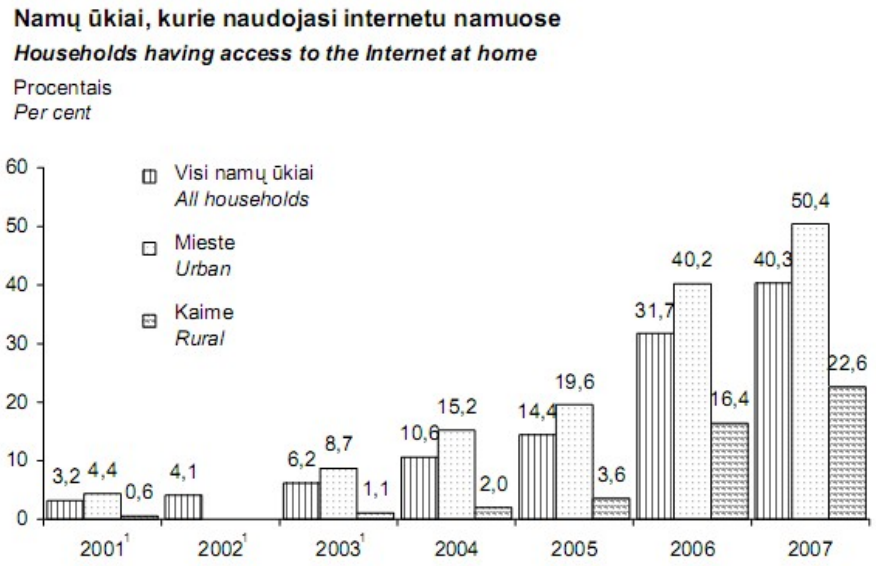

Abbildung 6.6: Internetzugang zu Hause [10]

erhalten diesen über einen Desktopcomputer, 19 % über einen Laptop und 52 % über ein Mobiltelefon.

23 % haben eine Schmalbandverbindung, 47 % eine Breitbandverbindung und 30 % haben beide Verbindungsmöglichkeiten. Die Gründe keine Breitbandverbindung zu unterhalten, sind zu 38 % die hohen Kosten, zu 31 % keinen Verwendungszweck zu haben, zu 9 % die Nichtverfügbarkeit, zu 21 % die Möglichkeit eine Breitbandverbindung an einem anderen Ort nutzen zu können, und zu 16 % andere Gründe. Im Jahr 2007 nutzen 52 % einen Computer und 49 % das Internet.

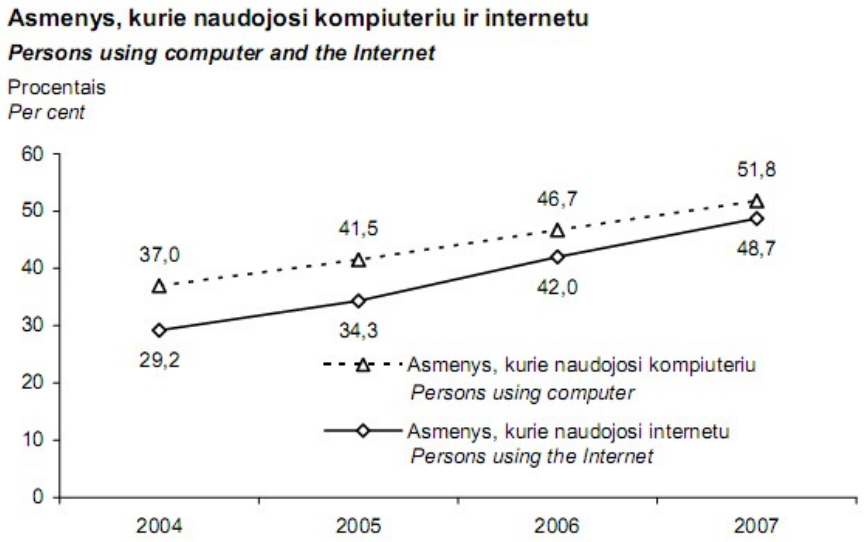

Abbildung 6.7: Personen die den Computer, das Internet in den Jahren 2004-2007 nutzen [10]

Im Alter zwischen 16-24 nutzen 91 % einen Computer und 90 % das Internet, zwischen 25-34 nutzen 71 % einen Computer und 67 % das Internet, zwischen 35-44 nutzen 59 % einen Computer und 54 % das Internet, zwischen 45-54 nutzen 41 % einen Computer und 38 % das Internet, zwischen 55-64 nutzen 19 % einen Computer und 16 % das Internet und zwischen 65-74 nutzen 5 % einen Computer und 4 % das Internet.

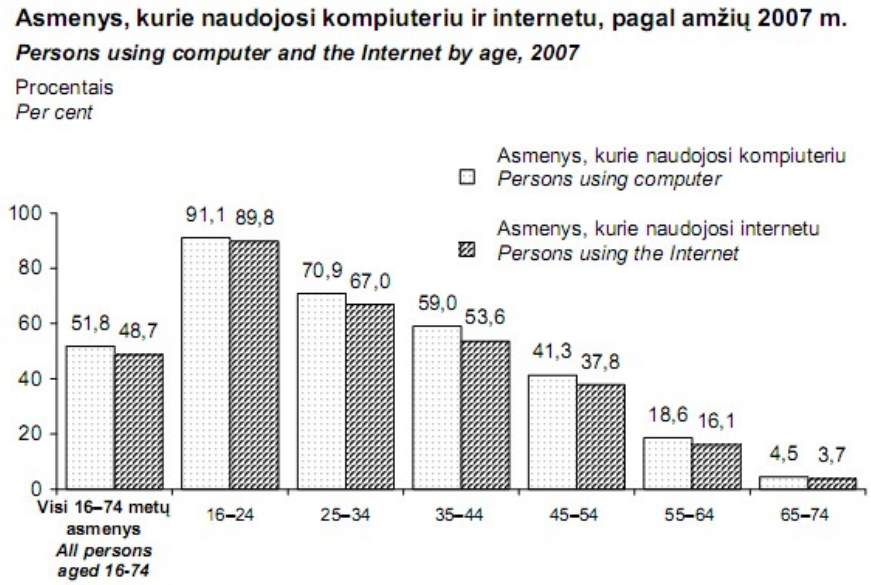

Abbildung 6.8: Personen die den Computer, das Internet nutzen nach Alter [10]

Schüler und Studenten nutzen den Computer am meisten, nämlich zu 99.8 %, gefolgt von Arbeitnehmern zu 62 %, Pensionären zu 4 % und anderen zu 27 %. Auch in Bezug auf das Internet sind die Schüler und Studenten mit 98.8 % die grössten Nutzer, gefolgt von den Arbeitnehmern mit 58 %, den Pensionären mit 3 % und anderen mit 24 %.

Das Internet wird im Bereich der Kommunikation genutzt um E-Mails zu verschicken und zu empfangen (39 %), um zu telefonieren und Videokonferenzen abzuhalten (19 %) und anderes (19 %). Im Bereich der Informationen und Dienstleistungen wird das Internet genutzt um Informationen über Güter und Dienstleistungen zu suchen (36 %), um Dienstleistungen im Bereich Reisen und Unterbringungsmöglichkeiten zu nutzen (14 %), um Webradio zu hören oder Webfernsehen zu sehen (20 %), um zu spielen und Spiele und Musik herunterzuladen (27 %), zum Softwaredownload (16 %), zum lesen und herunterladen von Onlinezeitschriften und -magazinen (32 %), zur Jobsuche und Stellenausschreibung (10 %) und zur Suche nach Informationen zum Thema Gesundheit. Im Bereich der Interaktion mit Behörden wird das Internet genutzt um Informationen von Webseiten der öffentlichen Behörden zu erhalten (18 %), um offizielle Formulare herunterzuladen (12 %) und um ausgefüllte Formulare zu senden (11 %). Der Bereich des E-Banking wird zu 21 % genutzt. Hindernisse zum erweiterten Gebrauch des Internets jener Menschen, die das Internet öfters nutzen möchten, sind zu 31 % fehlende Sprachfähigkeiten, zu 73 % Zeitmangel, zu 14 % zu langsame Verbindungen, zu 10 % zusätzliche Verbindungs- oder Download-Kosten, zu 12 % Kosten der Online-Inhalte, zu 3 % nicht genug interessante Inhalte, zu 20 % Mangel an Fähigkeiten und Wissen, zu 8 % Sicherheitsgründe oder pri-

vate Gründe und zu 15 % andere Gründe. 41 % der Internetnutzer haben Probleme mit Computerviren.

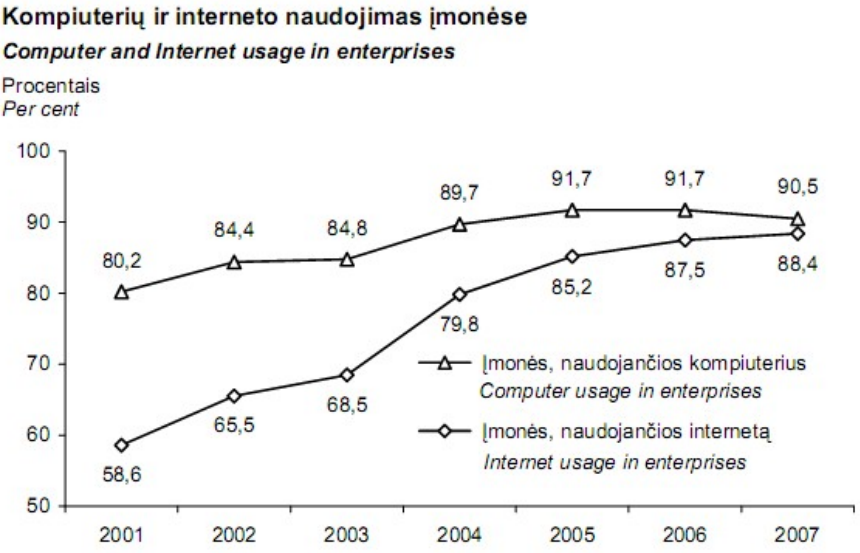

Abbildung 6.9: Computer- und Internetgebrauch der Unternehmen in den Jahren 2001-2007 [10]

Im Jahr 2007 nutzen 91 % der Unternehmen Computer und 88 % das Internet, über eine Breitbandverbindung verfügen jedoch nur 53 % der Unternehmen. 55 % der Firmen unterhalten ein Intranet, 47 % verfügen über LAN, 12 % über Wireless LAN und 8 % unterhalten ein Extranet. Zugang zum Internet haben 37 % über ein traditionelles Modem, 15 % über ISDN, 40 % über DSL, 20 % über andere feste Verbindungen und 25 % über drahtlose Verbindungen. Das Internet wird im Jahr 2006 zu 50 % zum Training und zur Bildung genutzt, zu 56 % zur Marktüberwachung, zu 83 % zur Erledigung von Bank- und Finanzdiensten und zu 76 % zur Interaktion mit öffentlichen Behörden. Webseiten von öffentlichen Behörden werden von Firmen zu 70 % zur Informationsbeschaffung, zu 74 % zur Formularbeschaffung, zu 60 % zur Retournierung ausgefüllter Formulare, zu 36 % zur vollelektronischen Fallbehandlung und zu 16 % zur Übertragung eines Angebots in ein elektronisches Angebotssystem genutzt. 47 % der Unternehmen verfügen auch über eine eigene Firmenwebseite. 95 % der Unternehmen, welche im Jahr 2007 über eine Internetverbindung verfügen, nutzen elektronische Sicherheitswerkzeuge. 90 % nutzen Antivirussoftware, 37 % eine Firewall, 29 % sichere Server, 47 % externe Datenbackups, 27 % Authentisierungsmechanismen und 12 % Datenverschlüsselung. Während im Jahr 2006 noch 40 % der Unternehmen mit Sicherheitsproblemen konfrontiert sind, sind es im Jahr 2007 keine Unternehmen mehr.

Im Jahr 2006 nutzen 70 % aller Angestellten in öffentlichen Administrationsbehörden einen Computer und 58 % das Internet. Die öffentlichen Stellen verfügen zu 22 % über ein traditionelles Modem, zu 14 % über ISDN, zu 47 % über DSL, zu 57 % über andere feste Verbindungen und zu 25 % über drahtlose Verbindungen. 56 % der Institutionen unterhalten eine Webseite und 39 % bieten elektronische Dienste an.

Im Bereich der Bildung wird im Jahr 2006 viel Geld ausgegeben, um IKT-Systeme im Bildungssystem einzuführen. Im Bereich des Gesundheitswesens besitzen 98 % der Insti-

tutionen Computer und 94 % einen Internetzugang, 35 % besitzen eine eigene Webseite. Im Bereich der Forschung und Entwicklung werden 47 % der im Businesssektor zur Verfügung stehenden Ressourcen in den IKT-Sektor investiert. 23 % der Unternehmen bemühen sich in den Jahren 2002 bis 2004 um Innovation.

6.3.3 Bedeutung

Für die Entwicklung einer Wissensgesellschaft sind zwei Dinge nötig, einerseits eine geeignete Infrastruktur und andererseits geeignete Qualifikationen. Eine gute IKT-Infrastruktur beschleunigt Prozesse der Kenntnisspeicherung, der Verarbeitung und der Anwendung. Dies allein reicht jedoch nicht, ohne die Qualifikation, diese Kenntnisse effektiv und sinnvoll anwenden zu können, verliert auch die Infrastruktur an Bedeutung. In Dänemark sind diese beiden Dinge bereits vorhanden und in Litauen ist die Politik auf die Entwicklung dieser beiden Komponenten gerichtet [14].

Litauen hat durchaus gute Bedingungen für die Entwicklung einer Wissensgesellschaft, denn Litauen ist eine der am schnellsten wachsenden Wirtschaften der EU. Das Bruttoinlandsprodukt hat im Jahr 2004 um 7 % zugenommen, im Jahr 2005 um 7.4 % und im Jahr 2006 um 7.5 %. Zudem verfügt Litauen über ein Bildungssystem, das das gleiche Recht, Kenntnisse zu bekommen, zu nutzen und zu teilen, für alle Bürger sichert. Und auch die Informationsinfrastruktur in Litauen ist weit entwickelt und ermöglicht eine Kommunikation, Verbreitung und Verarbeitung von Informationen [14].

Litauen belegt Platz 33 des ‚Networked Readiness Index 2007-2008'-Rankings des Global Information Technology Report 2007-2008 [7]. Die am stärksten vernetzte Wirtschaft der Welt und somit auf Platz eins des Rankings ist Dänemark, gefolgt von Schweden und der Schweiz. Der Bericht erfasst weltweit 127 Nationen und ist mittlerweile international zum massgebenden Bewertungsinstrument der Auswirkungen von IKT auf den Entwicklungsprozess und die Wettbewerbsfähigkeit von Ländern geworden. Der ‚Networked Readiness Index' prüft, inwieweit Länder bereit sind IKT effektiv einzusetzen. Dazu werden das allgemeine geschäftliche, regulatorische und infrastukturelle Umfeld für IKT, die Bereitschaft der wichtigsten beteiligten Gruppen (Einzelpersonen, Geschäfte, Regierungen) IKT gewinnbringend einzusetzen und ihre tatsächliche Nutzung der neusten verfügbaren IKT, geprüft [16].

Dänemark hat eine führende Stellung inne, und auch andere weit entwickelte Länder Europas sind in ihrer Entwicklung fortgeschritten. Daher sind auch die üblichen Probleme, die sich im Zuge der Entwicklung von einer Industriegesellschaft hin zu einer Inforamtions-, respektive Wissensgesellschaft ergeben, in Hinblick auf Dänemark kaum mehr von Bedeutung. Litauen jedoch hat trotz seines bereits beachtlichen Entwicklungsstandes mit einigen Problemen zu kämpfen. Noch längst nicht alle Menschen haben in Litauen Zugang zum Internet. Der Unterschied zwischen städtischen und ländlichen Gebieten, der Unterschied zwischen den Jungen und den Älteren und auch Unterschiede im Einkommen spielen eine Rolle. Wenn man jung ist, eine höhere Ausbildung und mehr Einkommen hat, in einer höheren sozialen Lage ist und wenn man in einer Stadt, respektive Grossstadt lebt, desto besser sind die Computerkenntnisse und desto mehr besteht ein Interesse daran Computerkenntnisse zu erlangen. Der grösste Bevölkerungsanteil Litauens lebt jedoch auf dem

Land und kann die Dienstleistungen des Bildungssystems nicht nutzen, wohnt weit weg von den IKT-Hauptzentren und hat nicht nur eine kleinere Möglichkeit IKT-Kenntnisse zu erwerben, sondern verspürt auch weniger das Bedürfnis dazu [15].

Litauen entwickelt sich jedoch schnell weiter und geht die vorhandenen Probleme an. Der Zugang zum Internet wird ausgebaut, Bildungsangebote im IKT-Bereich werden zur Verfügung gestellt, die Entwicklung und Forschung gefördert.

Schliesslich lässt sich sagen, dass die Situation der 'Needy' und ihre Bedürfnisse in jedem Land Europas unterschiedlich ist und dass sich alle Länder jetzt und in Zukunft damit beschäftigen werden den IKT-Bereich zu fördern, auch für die Ärmsten.

6.4 Bedeutung der Informationstechnologie ausserhalb Europas

In diesem Kapitel wird die Bedeutung von IKT ausserhalb Europas betrachtet; hierzu werden zwei Beispiele herangezogen. Einerseits ist dies die Situation auf dem afrikanischen Kontinent, wobei jedoch der Fokus der Betrachtung auf ländliche arme Regionen gerichtet ist. Andererseits werden IKT in Asien genauer betrachtet. Da hier jedoch viele Unterschiede bestehen, wird Indien als Beispiel für ein Schwellenland herausgegriffen und analysiert.

Die Definition von Armut - also den Needy im Titel dieser Arbeit - ist laut der Literatur in den meisten Entwicklungs- und Schwellenländern bei einem Verdienst von weniger als einem US-Dollar pro Tag angesetzt [4],[5]. Diese Definiton trifft auf mehr als 1.2 Milliarden Menschen weltweit zu, wobei jedoch bei anderen Ländern die Grenze höher angesetzt werden müsste, da Währungen und Kosten unterschiedlich sind. Von einer solchen auf Richtwerten basierenden Definition kann in den folgenden Beispielen ausgegangen werden.

6.4.1 Afrika

In Afrika ist das zentrale Problem die beinahe über den ganzen Kontinent verbreitete Armut, sowie der Mangel an ausreichender Infrastruktur und technischen Möglichkeiten. Jedoch sind hier starke Gefälle zwischen Stadt und Land sowie arm und reich festzustellen. Insbesondere in städtischen Gebieten geht beim wirtschaftlichen Mittelstand die Entwicklung erstaunlich rasch voran. Zwar hat weiterhin die Minderheit einen Telefon- oder Internetanschluss, doch ein Mobiltelefon ist erschwinglich geworden. Dies ist auch eine häufig praktizierte Lösung von kleineren Unternehmen, welche im urbanen Raum angesiedelt sind. Zudem ist in den Städten die Netzabdeckung für kabelgebundene und mobile Geräte wesentlich besser als auf dem Land, wo sie oft gar nicht existiert.

Ein weiterer Faktor ist der Analphabetismus, welcher in ländlichen Regionen bedeutend höher ist als in Städten und ein weiteres Hindernis bei der Benutzung vieler IKT darstellt [4]. Schliesslich gibt es ein weiteres soziales Phänomen, welches Afrika vor weitere

Probleme stellt, den sogenannten 'Brain Drain', die Abwanderung von gut ausgebildeten Personen. Meist sehen sie in anderen Ländern (insbesondere den USA und Europa) bessere berufliche Chancen, sodass sie dort Fuss zu fassen versuchen. Das grösste Problem hierbei ist der Verlust an Wissen, sowie die Tatsache, dass damit eine rasche wissenschaftliche und somit auch wirtschaftliche Entwicklung von vornherein verhindert wird. Gefördert wird diese Abwanderung nicht zuletzt dadurch, dass Forschungen und Arbeiten von afrikanischen Universitäten nur selten Anerkennung in internationalen Publikationen finden, und es kaum afrikanische Fachmagazine gibt. Hinzu kommt ein Mangel an der gesamten Infrastruktur des Kontinents, insbesondere den Verkehrsverbindungen. Zwar ist der physische Austausch zwischen unterschiedlichen Orten nicht eine direkte Voraussetzung für Informationsaustausch, meist jedoch trotzdem von nicht vernachlässigbarer Bedeutung. Schliesslich muss zum Austausch von Wissen zumindest ein minimaler Austausch an Material erfolgen. Ein Bauer profitiert nichts davon, zu wissen, dass das Saatgut einige Dörfer entfernt billiger erhältlich ist, wenn er nicht die Möglichkeit hat, es zu beschaffen.

Obwohl bereits mehrere Projekte zur Verbesserung der Telekom-Infrastruktur lanciert wurden, hatten bisher die wenigsten Erfolg [4]. Nicht zuletzt lag dies an der Problematik, dass der afrikanische Kontinent viele Bürgerkriegsregionen aufweist und in nicht wenigen Staaten die politische Lage instabil ist. Auch Hilfswerke unterstützen mehrere Projekte zur Vernetzung, einige werden später kurz vorgestellt. Jedoch muss man sich auch hier der Prioritäten bewusst sein. Ein Internetanschluss in einem Dorf, wo es an Nahrung und sauberem Trinkwasser mangelt, sowie niemand das Lesen beherrscht, ist faktisch nutzlos für die Bedürfnisse der Bevölkerung. Die Internetbenutzung Afrikas macht gerade mal 3 % jener der gesamten Welt aus [11]. Eine Abdeckung von über 10 % der Bevölkerung ist eine wahre Rarität und nur den reichsten Ländern vorbehalten, der Schnitt liegt bei einer Abdeckung von 6 %. Jedoch hat auf dem gesamten Kontinent innerhalb der letzten acht Jahre die Abdeckung um 1100 % zugenommen. Bei Radioempfängern wird in ländlichen Regionen von Entwicklungsländern allgemein von einem Gerät in 40 % der Haushalte ausgegangen [5]. Jedoch muss hier angemerkt werden, dass dies, wenn man eine Streuung über die Dörfer annimmt und auch das Stattfinden eines sozialen Austausches, bereits über 90 % der Bevölkerung einen Zugang zu Radiosendungen ermöglicht. Zur Verwendung von Telefonen sind kaum dokumentierte Zahlen vorhanden, insbesondere weil sich die Angaben und die Art des Zuganges (Mobil/Festnetz) rasch verändern und viele Provider keine genauen Dokumentationen durchführen.

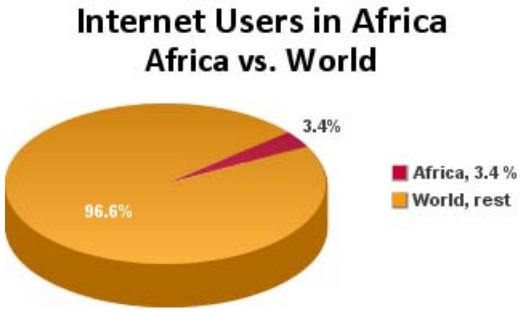

Abbildung 6.10: Internetbenutzung in Afrika [11]

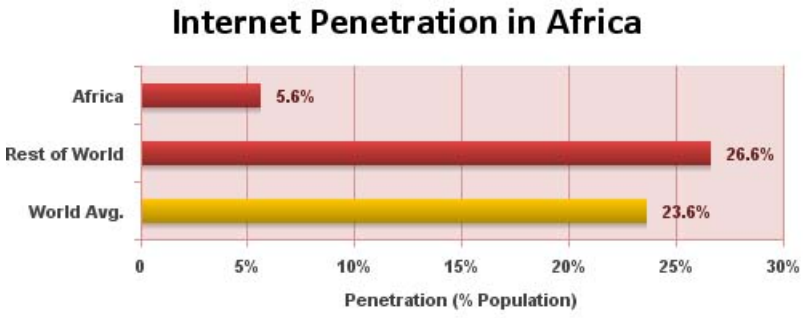

Abbildung 6.11: Internetabdeckung in Afrika [11]

In den Dörfern kann das Radio als das mit Abstand am weitesten verbreitete Medium angesehen werden [4]. Die Vorteile hierbei liegen auf der Hand; der Betrieb eines Empfängers ist mit sehr niederen Kosten verbunden und deshalb quasi für jedes Dorf, wenn auch nicht jeden Bewohner, erschwinglich. Die Netzabdeckung ist verglichen mit anderen Medien sehr hoch und aufgrund einer Vielzahl lokaler Stationen sind die gelieferten Informationen für die Bevölkerung von Bedeutung, da sie oft lokaler Natur sind. Hinzu kommt noch der Vorteil, dass Verständnisprobleme vermieden werden können, da die lokale Sprache verwendet wird und die Information so auch Analphabeten problemlos zugänglich ist.

Telefonie, sowohl auf Festnetz wie auch auf Funk basierend, ist wesentlich weniger verbreitet. Zwar gibt es in den zentraleren oder nahe urbanen Regionen gelegenen Dörfern oft Netzabdeckung und mehr als ein Telefon im Dorf, doch in entlegenen Regionen ist dies kaum immer der Fall. Es gibt einzelne Projekte zur Förderung der Kommunikation, meist indem eine besser gebildete Person im Dorf mit der Bedienung eines Telefons vertraut gemacht und das nötige Gerät bereitgestellt wird.

Der Computer und damit auch das Internet ist das mit Abstand am wenigsten verbreitete Medium. Einerseits liegt dies an einfachen technischen Gründen: eine Telefonleitung sowie Stromversorgung sind Voraussetzungen für den Betrieb der Geräte und der Kostenpunkt für Beschaffung und Betrieb liegt bei weitem über jenem der anderen IKT. Hinzu kommen soziale Faktoren, welche die Benutzung weiter erschweren. Es werden zumindest grundlegende Computerkenntnisse vorausgesetzt, zudem muss man des Lesens kundig sein. Ein weiteres Problem ist die Sprache. Der Grossteil des Internets ist in Englisch oder anderen westlichen Sprachen verfasst, sodass Fremdsprachenkenntnisse unerlässlich sind. Schliesslich sollten die verfügbaren Informationen auch für den lokalen Kontext relevant sein. Dies führt dazu, dass Netzzugang oft mit der Erstellung eines lokalen Informationsnetzwerkes verbunden werden muss, was weitere Kosten mit sich bringt.

Oft werden auch Medien kombiniert verwendet. Beispielsweise kann sich jemand telefonisch, schriftlich oder direkt bei einer Radiostation melden, um eine für ihn wichtige Information zu erhalten. Die Mitarbeiter beschaffen diese dann meist mittels Internetsuche und teilen die wichtigen Ergebnisse in einer Sendung mit.

Eine häufig angewendete Lösung für Netzzugang sind sogenannte Telezentren, vom Prinzip her ähnlich aufgebaut wie Internetcafés. Jedoch bieten sie nicht bloss Zugang zu vernetzten

Computern, sondern meist auch Hilfe für den Fall, dass eine Suche aufgrund mangelnder Kenntnisse nicht selbst durchgeführt werden kann. Hier ist jedoch meist ein Betreiber notwendig, welcher den Einheimischen bei der Bedienung der Geräte hilft oder falls diese Analphabeten sind, auch die Recherche übernimmt.

Eines der in der aktuellen Zeit am häufigsten erwähnten Projekte ist 'One Laptop Per Child (OLPC)', auch bekannt als der 'Einhundert-Dollar-Laptop', obwohl sich die Kosten pro Gerät auf ungefähr 200 Dollar belaufen [12]. Das Ziel ist es, an afrikanische Kinder günstige portable Computer abzugeben, und damit deren Lernen sowie Studium zu fördern. Zudem soll so ein reger Austausch zwischen den Lernenden stattfinden und die Kommunikation sowie das Wissen um neue Technologien gefördert werden. Zudem kann so eine der wichtigsten Ressourcen gefördert werden, welche zu einer besseren wirtschaftlichen und sozialen Entwicklung führt: gebildeter Nachwuchs. Die Initianten des Projekts standen vor einigen technischen Herausforderungen; die Geräte mussten möglichst günstig und langlebig sein, wenig Energie verbrauchen und zudem leicht zu handhaben sein.

Dies wurde mit einem einfachen Prozessor, günstigen Teilen, freier Software sowie einem kleinen Flash-Speicher anstelle einer Festplatte erreicht; trotzdem ist ein Austausch über eine WLAN-Verbindung möglich. Die Initianten bevorzugen das Modell von einem Laptop pro Kind und somit auch Familie vor den klassischen Telezentren, weil so eine weitere Verbreitung des Wissens und ein grösserer Austausch stattfinden können. Bisher sind Verbindungen mit dem Internet nur bei Zugangsstellen, wie Schulen, verbreitet, doch es gibt die Möglichkeit einer Peer-to-Peer Verbindung zwischen den einzelnen Geräten. Die Vermarktung erfolgt in grossen Stückzahlen insbesondere an Regierungen, Verwaltungen und Hilfsorganisationen.

Da das Projekt erst 2007 wirklich ins Laufen gekommen ist, sind bisher keine abschliessenden Zahlen sowie genaue Aussagen zum längerfristigen Erfolg machbar. Bisher wurden nach Angaben des Herstellers über 23'000 Laptops in Afrika verkauft, ganze 10'000 davon in Ruanda. Obwohl mittlerweile auch eine wesentlich grössere Verbreitung in anderen Kontinenten stattfindet (Abb. 2.12), insbesondere Südamerika, und insgesamt schon weit über eine Million Laptops verteilt wurden, scheint es angebracht, das Projekt im Kontext mit Afrika zu nennen, da hier die Verbreitung solcher Geräte zur stärksten Entwicklung führt; Afrika ist der bisher durch IKT am schlechtesten erschlossene Kontinent und ein Projekt dieser Grössenordnung ein wahres Novum. Ob jedoch auch Afrika die Voraussetzungen bietet, um von OLPC einen grösstmöglichen Nutzen zu ziehen, ist bisher nicht bekannt.

6.4.2 Asien

Sämtliche in diesem Abschnitt verwendeten Angaben stammen, wenn nicht anders vermerkt, aus Quibria/Tschang 2001 [3].

In Asien ist die Situation stark abhängig von der Region. Es gibt drastische Unterschiede zwischen technologisch hoch entwickelten Ländern wie beispielsweise Japan, welches in der Forschung und Entwicklung von Technologie weit vorne steht und sehr armen Ländern, beispielsweise Tibet. Hinzu kommen viele Schwellenländer, welche innerhalb der letzten

Abbildung 6.12: Geographische Verteilung der OLPC-Laptops [12]

Jahre eine grosse Entwicklung durchlaufen haben. Hier sind vor allem China und Indien zu nennen.

Indien ist für eine genauere Betrachtung ein guter Kandidat, insbesondere da hier die Streuung des Zugangs zu Technologien sowie deren Verwendung stark variiert. Indien ist zudem in einer guten Position, weil seine Bevölkerung auch die Chance hat, indirekt von IKT zu profitieren, dies hauptsächlich aufgrund einer guten Software-Industrie sowie Outsourcing von einfachen Operationen wie Datenverarbeitung und Call-Center.

Bereits mehrere asiatische Länder haben sich in der technologischen Entwicklung aufgeschwungen, insbesondere mit der Herstellung von Elektronik oder deren Programmierung. Trotzdem ist auch in diesem Fall noch eine grosse Armut, insbesondere in ländlichen Regionen oder Slums, vorhanden. Hier ist jedoch in einigen Ländern der Aufbau von neuen Technologien und Netzwerken weniger schwer, da im Schnitt etwas mehr Wissen vorhanden ist und die Infrastruktur statt neu erstellt nur ergänzt sowie verbessert werden muss. Deshalb ist auch die Einführung von Internetverbindungen teils wesentlich erleichtert. Hinzu kommt zumindest bei grösseren Ländern wie Indien ein kultureller Vorteil. Hier sind zwar, wie auch in Afrika, viele unterschiedliche Sprachen auf engem Raum vorhanden, doch die meisten Bewohner sprechen auch eine der gängigen Sprachen wie Englisch oder Hindi, sodass es wesentlich effizienter und einfacher ist, Internetseiten mit lokalen Inhalten anzubieten; zudem existieren auch in diesem Fall bereits viele Ressourcen. Jedoch ist auch hier Analphabetismus sowie eine sehr schlechte oder gar keine Schulbildung ein weit verbreitetes Problem. Die Infrastruktur hat jedoch durchaus einige Vorteile zu bieten, sind doch meist auch entlegene Dörfer mit Strassen erschlossen und haben auch oft Bus- oder Bahnverbindungen.

Zudem hat Indien überdurchschnittlich viele Projekte zur Verbreitung von IKT in der Bevölkerung vorzuweisen, weshalb der Fokus der Beispiele in dieser Arbeit auf dieses Land gerichtet ist. Dabei darf jedoch die Tatsache nicht vernachlässigt werden, dass andere Länder in wesentlich schwierigeren Situationen sind. Asien hat eine ziemlich gute Internet-Abdeckung erreicht, welche mit 17 % der Bevölkerung im Schnitt drei Vierteln

des weltweiten Mittels entspricht [11]. Insbesondere die aktuellen Schwellenländer haben innerhalb der letzten acht Jahre ein drastisches Wachstum von teils weit über 1000 % verzeichnet, das asiatische Mittel liegt bei 469 %. Interessant ist, dass auch arme Länder und Krisenregionen wie beispielsweise Afghanistan ein sehr grosses Wachstum hier ca. 49'000 %, wenn auch eine verhältnismässig noch immer niedere Abdeckung (2 %), aufweisen. Die Länder mit den höchsten Anteilen sind Japan (74 %), Hong Kong (70 %), Singapur (67 %) sowie Taiwan (66 %). Indien als typisches Schwellenland weist eine Abdeckung mit Internetanschlüssen von 7 % und 38 % bei Telefonen auf.

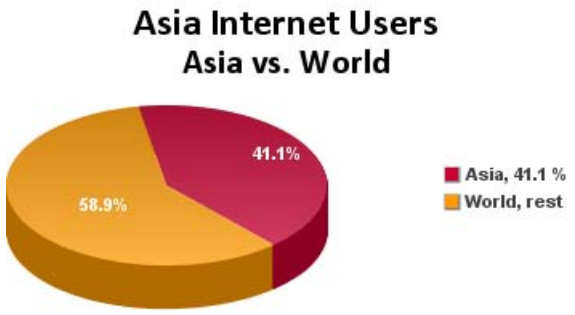

Abbildung 6.13: Internetbenutzung in Asien [11]

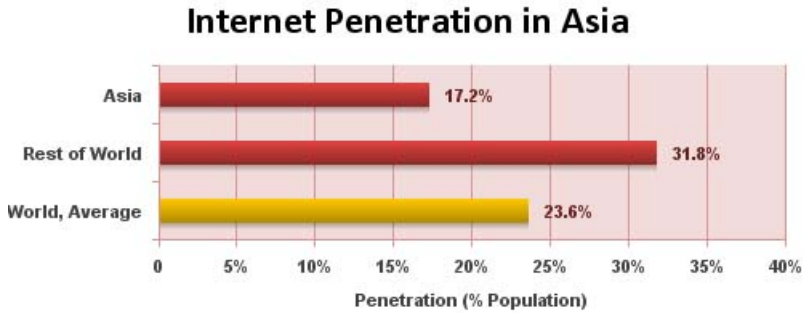

Abbildung 6.14: Internetabdeckung in Asien [11]

Ein insbesondere in Indien weitverbreitetes Prinzip bei Entwicklungsprojekten ist die Einführung sogenannter Dorftelefone. Diese Projekte machen jedoch nur in Dörfern Sinn, welche bisher noch keinen eigenen Telefonanschluss hatten und meist ziemlich arm sind. Hierbei wird an eine Person im Dorf, meist eine besser gebildete Hausfrau, ein Mikrokredit vergeben, mit welchem sie sich ein günstiges Mobiltelefon mit der nötigen Ausrüstung erwerben kann. Meist wird sie von einer Person mit den nötigen Kenntnissen kurz in die Benützung eingeweiht und hat danach die Möglichkeit, das Telefon gegen Bezahlung anderen Dorfbewohnern zur Verfügung zu stellen. Obwohl diese arm sind, nutzen sie das Angebot, da ihnen nicht genügend Geld für einen eigenen Anschluss zur Verfügung steht, sie sich jedoch die vergleichsweise günstige Benützung des Dorftelefons leisten können. In den meisten Fällen fällt es den Frauen leicht, den Kredit innerhalb einiger Zeit zurückzuerstatten.

Ein weiterer Bereich, welcher jedoch bereits im Zusammenhang mit Afrika diskutiert wurde, ist der Aufbau von Telezentren. Hier besteht jedoch teils ein Problem mit der Finanzierung, da Telezentren in den meisten Fällen nach dem Aufbau nur bedingt unterstützt

werden und folglich selbsterhaltend sein müssen. Im Gegensatz zu Afrika besteht jedoch in Indien der Vorteil, dass es viele Programme der Regierung zur Vernetzung und dem Aufbau lokaler Telezentren gibt und diese auch so gestaltet werden, dass sie sich gut erhalten. Meist werden auch hier einige Betreiber eingestellt, welche die Bewohner ohne Technikkenntnisse bei der Verwendung der Computer unterstützen. Ein Problem der Telezentren ist ihre Verfügbarkeit. Da meist bloss eine sehr kleine Anzahl an Computern für eine ziemlich grosse Zahl von Personen vorhanden ist, muss ein Weg gefunden werden, die vorhandenen Ressourcen fair zur Verwendung bereitzustellen.

E-Government, der Zugriff auf Daten und Dienstleistungen der Regierung, ist in Asien ebenfalls von grosser Bedeutung. Da viele Beamte der lokalen Behörden als korrupt gelten und das Ausstellen teils wichtiger Dokumente sehr teuer sowie mit Zeit- und Kostenaufwand für entsprechende Reisen verbunden ist, ist der Zugriff auf Behörden über das Internet eine gute und günstige Alternative [3]. Hier werden ebenfalls oft die Telezentren genutzt, da diese einen bezahlbaren Zugriff für Dorfbewohner bieten. Mittels elektronischer Kommunikation werden hier sowohl Kosten als auch Aufwand reduziert und damit das Leben der Einheimischen vereinfacht. Zudem kann so auch ein Weg zur Kommunikation mit den Behörden gefunden werden, welcher Verbesserungen ermöglicht. Korruption ist hier ebenfalls beinahe unmöglich, da alles über ein zentrales Netzwerk abgewickelt wird und geprüft werden kann. Zudem haben die elektronisch deponierten Anregungen aus der Bevölkerung bereits zu vielen Verbesserungen der staatlich angebotenen Dienstleistungen, wenn auch meist nur in lokalem Kontext, geführt.

Insbesondere in den städtischen Regionen von Indien ist zudem vor einigen Jahren erstmals ein Phänomen aufgetreten, welches sich seitdem stetig verbreitet hat: Die Auslagerung von Dienstleistungen. Meist sind es grosse westliche Firmen, welche aus finanziellen Gründen Datenverarbeitung und Kundendienst an indische Firmen auslagern. Hierbei werden oft Angestellte gesucht, welche fliessend Englisch sprechen und über grundlegende Computerkenntnisse verfügen, jedoch keine Spezialisten. So werden insbesondere Stellen für Bewerber aus der unteren Mittelschicht frei, welche für lokale Verhältnisse gut entlohnt werden. Da es sich bei den Arbeiten meist um Routinearbeit handelt, wird eine erstaunlich grosse Menge an Stellen angeboten, für welche nicht viele Kenntnisse vorausgesetzt werden.

Jedoch sind auch immer mehr asiatische Länder an der Forschung und Entwicklung sowie der Industrie beteiligt. Während die Forschung noch immer grösstenteils auf die entwickelten Länder beschränkt ist, haben neuerdings auch Schwellenländer ihre Beteiligung zu erweitern versucht. Indien ist mittlerweile ziemlich bekannt für gute Software, während Taiwan einen grossen Anteil an Elektronikindustrie aufweist.

Natürlich muss hier die Betrachtung gemacht werden, dass dadurch die ärmste Bevölkerung kaum direkt profitiert, sondern die Wirtschaft des Landes sowie der Mittelstand. Jedoch ist ein indirekter Profit meist möglich, da sich die Einnahmen des Staates erhöhen und dieser damit auch die Möglichkeit hat, mehr in eine Verbesserung der Lebensbedingungen der Armen zu investieren.

6.5 Erwartete Auswirkungen

Die Auswirkungen sind je nach Projekt und Situation sehr unterschiedlich. Nebst technologischen und sozialen Aspekten können auch kulturelle nicht vernachlässigt werden. Die Unterschiede zwischen den Kontinenten und auch deren einzelnen Ländern sind zu gross, um eine einzelne abschliessende Aussage zu treffen [2]. Schliesslich muss auch noch zwischen langfristigem und kurzfristigem Gewinn unterschieden, sowie direktes und indirektes Wachstum betrachtet werden.

Ein langfristiger Gewinn für alle beteiligten Parteien, beispielsweise Investoren und Betroffene, kann nur erreicht werden, wenn die Projekte auf lange Sicht selbsterhaltend und -finanzierend sind und unabhängig betrieben werden können. Eine weiter anhaltende Abhängigkeit von Hilfswerken sowie eine gleichbleibende Abhängigkeit von anderen Finanzierungsmöglichkeiten führen kaum zu einem Abbau der Armut. Zudem müssen einige Grundvoraussetzungen gegeben sein, um mit einer Entwicklung rechnen zu können. IKT sind kaum hilfreich, wenn ein Mangel an ausreichender Nahrung, sauberem Wasser und Hygiene gegeben ist. Dies lässt sich besonders gut an der Maslow'schen Hierarchie der Bedürfnisse aufzeigen [13]. Während IKT kaum dazu dienen können, physiologische Bedürfnisse auf der untersten Stufe zu erfüllen, ermöglichen sie jedoch eine Vebesserung auf allen darüberliegenden Stufen. Auf der Stufe der Sicherheit handelt es sich grösstenteils um indirekten Nutzen, wie erleichterter Zugang zu Wetterberichten, Stellenmärkten, Regierung etc., während insbesondere der soziale Austausch auch über weitere Distanzen sowie der erleichterte Zugang zu Wissen auf den darüberliegenden Stufen eine Rolle spielen.

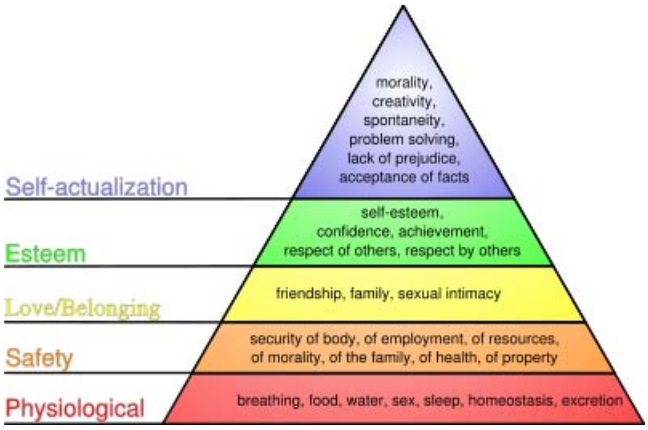

Abbildung 6.15: Maslows Pyramide der Bedürfnisse [13]

In der Bildung bieten IKT, insbesondere Computer und Internet, weitere Optionen, können jedoch nicht traditionelle Mittel wie den Unterricht gänzlich ersetzen. Je nach Situation kann eine technische Entwicklung auch die beruflichen Möglichkeiten verbessern, da neue Stellen geschaffen werden und bei einer guten Vernetzung mehr Zugriffsmöglichkeiten auf Angebote gegeben sind. Insbesondere bei der Beschaffung und dem Austausch von Informationen haben IKT ein grosses Potential zur Verbesserung der Situation. Das bereits vorgestellte E-Government kann eine grosse Effizienzsteigerung von Abläufen sowie ein

beträchtliches finanzielles Ersparnis für Arme bieten, setzt jedoch Telezentren voraus. Zudem sind so für Händler und Bauern Marktinformationen, Wetterprognosen und ähnliche Daten leichter zugänglich, welche zu einer Verbesserung von Abläufen sowie günstigeren Einkäufen führen können. Zudem eröffnen sich so zumindest für Einzelpersonen oder kleine Gruppen weitere Einkommensmöglichkeiten wie den Verkauf von Handarbeiten an entferntere Städte oder westliche Länder.

Das Kombinieren von IKT ist ein substantieller Faktor für eine kluge und effiziente Nutzung. Zudem müssen die Informationen an den lokalen Kontext angepasst sein und für die Zielgruppe verständlich vermittelt werden, was einen weiteren Aufwand bedeutet.

Vergleicht man die Wirkung unterschiedlicher IKT, so kann man das Radio als das noch immer wichtigste und bedeutendste Medium bezeichnen. Hier besteht bereits eine grosse Abdeckung und zudem ist aufgrund der vielen Stationen kaum eine Investition notwendig. Lokale Informationen können gut in der jeweiligen Sprache verbreitet werden und bieten so der Bevölkerung eine solide Grundlage, von welcher sie profitieren kann. Zudem sind die Kosten auch für die Ärmsten relativ gering, sodass in den meisten Fällen mindestens ein Radioempfänger pro Dorf vorhanden ist. Der Nachteil des Radios ist jedoch, dass es bloss einseitige Kommunikation ermöglicht und meist nicht zur direkten Suche von Informationen geeignet ist. Trotzdem bietet es eine stabile Grundlage, ist teils ein sehr grosser Gewinn für die Gemeinde und kann ohne grosse Investitionen betrieben werden.

Das Telefon hat zwar eine wesentlich geringere Verbreitung als das Radio, dies kann jedoch mit den höheren Betriebskosten erklärt werden. Diese ergeben sich primär durch mehr Voraussetzungen bei der Technik, es ist entweder eine Leitung oder ein wesentlich teureres Mobilgerät sowie Netzabdeckung notwendig, um telefonieren zu können. Jedoch kann hier auch gegenseitig kommuniziert werden, was die Nutzungsmöglichkeiten erweitert. Ein Modell, welches auch wirtschaftlichen Profit bringt, ist das in Indien angewandte System mit den Mikrokrediten. Die Bank verdient zwar kaum, aber wirtschaftliches Wachstum sowie ein freier Markt wird auch für die Ärmeren gefördert. Zudem werden auch finanzielle Anreize für private Telefongesellschaften geschaffen, auch entlegenere Regionen zu vernetzen und so ihr Einzugsgebiet zu erweitern.

Das Internet schliesslich ist die am schwersten zu verbreitende Technologie, da sie die grössten Voraussetzungen bei der Infrastruktur mit sich bringt. Folglich sind verglichen mit den anderen Technologien die notwendigen Investitionen deutlich höher. Zudem sind mehr Voraussetzungen bei der Bevölkerung notwendig, wie beispielsweise die Fähigkeit Lesen zu können, die Beherrschung des Umgangs mit dem Gerät und je nach Recherche auch das Sprechen weiterer Sprachen. Zudem müssen hier lokale Informationen mit mehr Aufwand aufbereitet werden.

Das grösste Wachstumspotential im Informatiksektor ist jedoch in der Wirtschaft zu finden, und nicht bei der Zivilbevölkerung. Hier ist kein direkter Profit für die Armen auszumachen, sondern ein indirekter aufgrund des wirtschaftlichen Aufschwungs. Umgekehrt kann mehr Wohlstand auch zu einem indirekten Profit für die Wirtschaft werden, da so mehr Investoren angezogen werden.

6.6 Zusammenfassung und Schlussfolgerung

IKT sind alle Technologien im Bereich von Information und Kommunikation, wie z.B. das Radio, das Telefon und das Mobiltelefon, das Internet und damit das World Wide Web.

In Europa ist der Stand im Bereich der IKT bereits sehr gut. Die Bevölkerung Dänemarks, eines der reicheren Länder Europas, hat bereits im Jahr 2005 zu 79 % einen Internetanschluss, die Unternehmen haben zu 82 % sogar einen Breitbandanschluss und auch andere IKT-Systeme werden von Unternehmen und auch den Behörden genutzt. Die Bevölkerung Litauens, einem ärmeren Land Europas, hat im Jahr 2007 zu 40 % einen Internetzugang, die Unternehmen zu 88 % und die Behörden zu 58 %. Litauen liegt also klar hinter Dänemark zurück, die Entwicklung schreitet jedoch nicht nur in Dänemark sonder auch in Litauen schnell voran.

In Schwellen- und Entwicklungsländern geht die technologische Entwicklung unterschiedlich rasch voran. In Afrika haben gerade 6 % der Bevölkerung einen Internetanschluss, in Asien sind es 17 % [11]. Es kann eine auffallende Konzentration in den Städten und den Schwellenländern und kaum eine Verbreitung auf dem Land gefunden werden. Jedoch haben weit mehr als 90 % der Bevölkerung die Möglichkeit, Nachrichten via Radio zu empfangen. Das Telefon ist ein weiterer wichtiger Faktor in der Entwicklung, da es die Möglichkeit bietet, Informationen beidseitig auszutauschen und zudem ist ein Anschluss Voraussetzung für eine Internetverbindung. Internetanschlüsse werden oft an zentralen öffentlichen Orten angeboten, um möglichst viele Bewohner profitieren zu lassen.

Wie sinnvoll IKT-Projekte sind, hängt von der jeweiligen Situation, Kultur, Infrastruktur, Finanzierung und Bildung ab. Es lassen sich jedoch in den meisten Fällen mögliche Entwicklungen prognostizieren und die Chancen bis zu einem gewissen Grad abschätzen. Es gilt hierbei zu beachten, dass wirtschaftliche Investitionen zwar notwendig sind, jedoch wenn irgendwie möglich eine langfristige Abhängigkeit vermieden werden sollte. Zudem ist zwischen einem direkten Nutzen für die Armen und einem indirekten Nutzen - meist durch Wirtschaftswachstum - zu unterscheiden. Um IKT sinnvoll implementieren zu können, muss zudem ein entsprechendes Bedürfnis sowie ein Nutzen für die Betroffenen gegeben sein - sonst enden Projekte meist ohne sichtbaren Erfolg. Ein weiterer wichtiger Faktor ist jener der Kosten - je komplexer eine Technologie ist, desto teurer wird sie - wenn sie auch nicht zwingend lohnender ist. Hier lässt sich eine Abstufung vom günstigen Radio, über das Telefon hin zum verhältnismässig teuren Computer oder Internetanschluss erkennen. Die jeweiligen Technologien sind auch entsprechend ihres Preises verbreitet, je kostengünstiger desto weiter verbreitet. Die häufigsten Verbesserungen durch IKT treten bei der Verwendung zur Bildung, zum Zugang zu lokalen Informationen, sowie bei Dorftelefonen auf. Situationsbedingt bieten auch erweiterte Möglichkeiten wie beispielsweise E-Government ein Entwicklungspotential.

Literaturverzeichnis

[1] Pedro Puga, Gustavo Cardoso, Rita Esphana, Sandor Mendonca: *Telecommunications for the needy: an exploratory approach from Portugal*; OberCom, Lisboa, http://www.lini-research.org/np4/?newsId=5&fileName=Apresenta__o_Obercom_verde___Corrigido_.pdf, Februar 2009.

[2] L-F Pau: *Mobile Service Affordability for the Needy, Addiction, and ICT Policy Implications*; Erasmus Research Institute of Management, Erasmus Universität Rotterdam, April 2008, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1131032.

[3] M.G. Quibria, Ted Tschang: *Information and Communication Technology and Poverty: An Asian Perspective*; Asian Development Bank Institute, Tokyo, Januar 2001, http://www.adbi.org/research%20paper/2001/01/01/152.communications.technology/.

[4] J.J. Britz, P.J. Lor, I.E.M. Coetzee, B.C. Bester: *Afirca as a knowledge society: A reality chek*; The International Information and Library Review 38, 2006, S.25-40, http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6WGP-4J4HH6J-1-1&_cdi=6828&_user=5294990&_orig=search&_coverDate=03%2F31%2F2006&_sk=999619998&view=c&wchp=dGLbVlz-zSkzS&md5=15f6355ed6d1dbc6d5e10cedfd4f879f&ie=/sdarticle.pdf.

[5] Charles Kenny: *Information and Communication Technologies for Direct Poverty Alleviation: Costs and Benefits*; Development Policy Review 20(2), 2002, S.141-157, http://www3.interscience.wiley.com/cgi-bin/fulltext/118955768/PDFSTART.

[6] *Ihr Schlüssel zur europäischen Statistik*; http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1090,1&_dad=portal&_schema=PORTAL, Februar 2009.

[7] *The Global Information Technology Report 2007-2008*; World Economic Forum, 2008, http://www.wefforum.org/en/initiatives/gcp/Global%20Information%20Technology%20Report/index.htm und http://www.insead.edu/v1/gitr/wef/main/fullreport/index.html.

[8] *World Encyclopedia. Oxford Reference Online*; Oxford University Press, 2008, http://www.oxfordreference.com/views/GLOBAL.html.

[9] *Key Figures on the Danish Information Society 2006 - Danish Figures*; Statistics Denmark and Ministry of Science, Technology and Innovation, 2006, `http://www.dst.dk/HomeUK/Statistics/ofs/Publications/KFICT_total/KFICT_DK2006.aspx`.

[10] *Information Technologies in Lithuania*; Statistics Lithuania, Technology and Innovation, 2007, `http://www.infobalt.lt/sl/index_de.php`.

[11] *Internet World Stats. Usage and Population Statistics*; `http://www.internetworldstats.com`, April 2009.

[12] *Give a Laptop, Change the World*; `http://laptop.org`, April 2009.

[13] A.H. Maslow: *A Theory of Human Motivation*; Psychological Review 50(4), 1943, S.370-96.

[14] Edgaras Leichteris: *Richtlinien der Konkurrenzfähigkeit von Litauen*; Asociation Knowledge Economy Forum, `http://www.infobalt.lt/sl/index_de.php?t=konkurenc`, März 2009.

[15] *Damit Computer Peters Leben erleichtert*; `http://www.infobalt.lt/sl/index_de.php?t=langas`, März 2009.

[16] Matthias Lüfkens: *Dänemark, Schweden und die Schweiz führen im Global Information Technology Report 2007-2008*; World Economic Forum, Genf, April 2008, `http://www.weforum.org/pdf/gitr/2008/press/german.pdf`.