D//D DÜSSELDORF INSTITUTE FOR INTERNET AND DEMOCRACY

Précis [pʁe'si:] // MAR 2017

Compliance Management bei zensurresistenten Netzwerken im Internet

Alexander Schneider
Kontakt: aschneider@cs.uni-duesseldorf.de

ABSTRACT

So genannte Peer-to-Peer Netzwerke werden häufig verwendet, um eine mögliche Zensur des Internets zu umgehen, z.B. in totalitären oder unfreien Staaten. Ein Nachteil dabei ist, dass diese Netzwerke von jedem Nutzer erfordern sämtliche Inhalte des Netzwerks weiterzuleitet und auch zu speichert. Diese Inhalte können teilweise vom Benutzer unerwünscht oder ethisch fragwürdig sein, weshalb viele Nutzer von der Benutzung solcher Netzwerke abgeschreckt werden. Hier wird ein Compliance Management für Peer-to-Peer Netzwerke beschrieben, welches aktuell an der Heinrich-Heine-Universität entwickelt wird. Bei diesem System können die Benutzer selber festlegen, welche Arten von Inhalten sie behandeln möchten. Gleichzeitig ist das Netzwerk auf eine Weise aufgebaut, bei der die Funktionalität des Netzwerks nicht merklich nachlässt.

Alexander Schneider
ist wissenschaftlicher Mitarbeiter am
Lehrstuhl für Rechnernetze an der
Heinrich-Heine-Universität Düsseldorf.

Er studierte Informatik in
Düsseldorf (M.Sc.).
In seiner Forschung
konzentrierte er sich auf
Computernetzwerke, Online-Partizipation
und elektronische Wahlen
und deren Sicherheit.

Einleitung

Nicht jedes Land ermöglicht seinen Bürgern einen freien und unzensierten Zugang zu Inhalten im Internet. Die klassische Architektur des Internets macht eine Zensur einfach. So liegen die Inhalte (z.B. einer Webseite) in der Regel auf einem bzw. wenigen Rechnern (sogenannten Servern) im Netz, auf die die einzelnen Nutzer (die Clients) zugreifen. Für die zensierende Instanz bedeutet dies, dass die Zensur nur an einer Stelle, nämlich dem Server, notwendig ist. Peer-to-Peer Netzwerke sind Netzwerke, bei denen keine zentralen Instanzen existieren. Stattdessen werden die Daten bei den am Netzwerk teilnehmenden Endbenutzern gespeichert und von diesen weitergeleitet. So liegen die Inhalte nicht nur einmal bei einem einzelnen Server, sondern vielfach vor, da immer mehrere Nutzer eine Kopie des gleichen Inhalts zwischenspeichern. Eine Zensur ist in einem Peer-to-Peer Netzwerk de facto unmöglich, da dazu alle Kopien des Inhalts gelöscht werden müssten. Im Gegensatz zum eingangs dargestellten Server-Client-Modell benötigen Peer-to-Peer Netzwerke aber die aktive Beteiligung eines Großteils ihrer Teilnehmerinnen und Teilnehmer, um ihre Funktionalität aufrechterhalten zu können. Nur, wenn genügend Personen Inhalte zwischenspeichern und weiterleiten, ist ein Datenaustausch für alle Teilnehmenden gewährleistet.

Aus diesem Grund hat man als Benutzer im Normalfall keine oder nur sehr wenig Kontrolle über die Daten, welche man weiterleitet und für andere speichert. Diese Daten können ethisch fragwürdig und im Extremfall auch illegal sein, da zensurresistente Netzwerke nicht nur für z.B. politische Dissidenten, sondern auch für kriminelle Aktivitäten von Interesse sind. Dies sorgt unter Umständen dafür, dass viele Menschen vom Gebrauch von Peer-to-Peer Netzwerken absehen.

Aus diesem Grund wird am Institut für Informatik der Heinrich-Heine-Universität Düsseldorf und im Rahmen des Düsseldorfer Instituts für Internet und Demokratie aktuell ein „Compliance Management“ für Peer-to-Peer Netzwerke entwickelt, welches genau dieses Problem lösen soll. Das Ziel des Compliance Managements ist es, den Nutzenden zu erlauben selbst fest zu legen, welche Arten von Inhalten automatisiert verarbeitet werden, während das System technisch so gestaltet ist, dass es keine bedeutenden Einbußen in seiner Funktionsfähigkeit erfährt.

Im Folgenden geben wir einen kurzen Überblick über die technischen Grundlagen, die für ein Verständnis unseres vorgeschlagenen Systems nötig sind. Anschließend beschreiben wir die Systemarchitektur sowie Implikationen eines funktionierenden Compliance Management Netzwerks und geben einen Ausblick auf die noch ausstehenden Probleme, welche es zu lösen gilt.

Technische Grundlagen

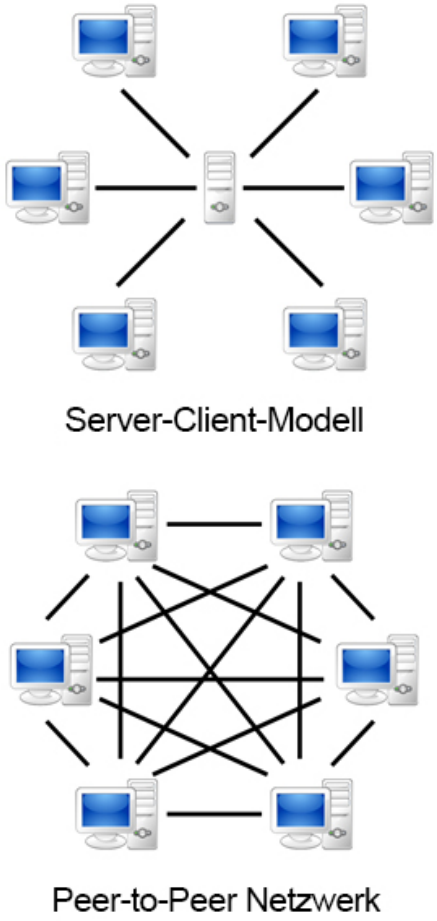

Abb. 1: Der Unterschied zwischen Server-Client und Peer-to-Peer Modellen.[1]

Um die Implikationen von einem Compliance Management Netzwerk korrekt einordnen zu können, müssen einige technische Grundlagen gelegt

1 Bildrechte liegen bei Saul Albert. www.saulalbert.net

werden. Die Hauptarchitektur, auf der das gesamte System aufbaut, ist die Peer-to-Peer Architektur. Die meisten Benutzer eines Netzwerks, wie des Internets, sind das Client-Server Modell gewöhnt, bei dem eine zentrale Instanz (Server) einen Dienst bereitstellt, der von den Endanwendern (Clients) abgerufen wird (siehe Abbildung 1 oben). Die zentrale Kontrolle über den angebotenen Dienst ist hierbei Vor- und Nachteil zugleich. Während die Zentralisierung eine höhere Kontrolle über den Dienst und dessen Güte ermöglicht, kommt es im Falle von Störungen am Server zu einem Zusammenbruch des Dienstes für alle Beteiligten. Auch Angreifer können diesen Schwachpunkt ausnutzen. Im Gegensatz dazu steht das Peer-to-Peer Modell, bei dem die Endanwender gleichzeitig auch Anbieter von Diensten sind (siehe Abbildung 1 unten). Dazu verbinden sich die Clients für gewöhnlich untereinander und die Grenze zwischen Clients und Servern verschwimmt. Damit trotz dem Fehlen einer zentralen Instanz und damit auch dem Fehlen von zentraler Koordination, ein geregelter Betrieb des Netzwerks möglich ist, werden für gewöhnlich so genannte Overlays benutzt (siehe Abbildung 2). Diese abstrahieren von den physikalischen Verbindungen der Geräte und verwenden Algorithmen, die alle Clients innerhalb des Netzwerks strukturieren (z.B. durch Vergabe abstrakter Adressen, die nicht an den physikalischen Anschluss gebunden sind) und den Fluss von Daten und Verbindungen zwischen den Clients regeln. Bekannte Overlays, die häufig verwendet werden sind, zum Beispiel Kademlia (Maymounkov & Mazieres, 2002), Chord (Stoica et al., 2001) oder das BitTorrent Protokoll (Izal et al., 2004).

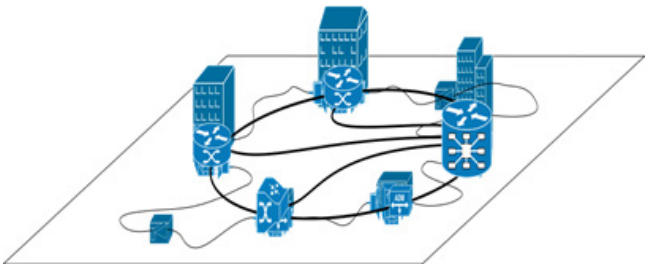

Abb. 2: Ein Beispiel für ein Overlay Netzwerk. Die physikalischen Verbindungen sind mit dünnen Kabeln gekennzeichnet, die logischen Verbindungen mit dicken. [2]

Systemarchitektur

Comademlia, das Compliance Management System, welches sich aktuell in Entwicklung befindet, besteht aus zwei Teilen. Der erste Teil ist ein Peer-to-Peer Overlay (Lua et al., 2005), welches sich um die Weiterleitung und die Verteilung von Daten innerhalb des angestrebten Netzwerkes kümmert. Der zweite Teil ist ein System, welches die Daten in Kategorien unterteilt und dafür sorgt, dass diese sachgerecht verwaltet werden.

Für das Peer-to-Peer Overlay wurde ein Programm entwickelt, welches auf dem Kademlia Overlay basiert. Der Grundpfeiler des Overlays sind die Knoten (Endbenutzer), die in einer listenähnlichen „Policy“ (Regeln/Regelwerk) einstellen können, welche Arten von Inhalten sie zur Weiterleitung und Speicherung erlauben. Jeder Knoten annonciert seine Policy an seine Nachbarknoten, woraufhin diese ihre verschiedenen Routingtabellen anhand ihrer Umgebung befüllen. Im Detail wird auf jedem Knoten für jede bekannte Datenkategorie eine Routingtabelle angelegt, die angibt an welche Nachbarknoten Daten einer bestimmten Kategorie weitergeleitet werden dürfen (siehe Abbildung 3). Eine Simulation des Overlays hat gezeigt, dass das Netzwerk trotz Compliance Management Auflagen weiterhin zufriedenstellend funktioniert.[3] Es gibt zumindest eine probabilistische Garantie, dass selbst Inhalte aus selten vorkommenden Kategorien gefunden und abgerufen werden können.

Der Ansatz, mehrere Routingtabellen zu benutzen, führt natürlich zu einem Anstieg an versendeten Nachrichten und einem Mehraufwand an Speicherplatz. Dies lässt sich allerdings mit weiteren Optimierungen wie dem Zusammenlegen von redundanten Einträgen verringern und wurde für die Simulation nicht berücksichtigt, um die Komplexität der Implementierung möglichst niedrig zu halten.

Damit die Policies der Benutzer ihre Wirkung entfalten können müssen die Daten kategorisiert werden. Auf welche Weise diese Kategorisierung vorgenommen wird, z.B. ob automatisiert durch

2 Bildrechte liegen bei wikimedia.org

3 Dafür wurde der PeerfactSIM.KOM Simulator genutzt (Graffi, 2011; Stingl et al., 2011)

Algorithmen oder manuell durch die Herausgebenden der Daten, ist dabei unerheblich. Die entscheidende Herausforderung ist, wie sichergestellt werden kann, dass die Kategorisierung korrekt ist. Es stellt sich heraus, dass dieses Problem im Rahmen eines dezentral organisierten Netzwerkes nicht lösbar ist. Zumindest nicht, wenn man eine korrekte Kategorisierung bei der Einspeisung in das Netzwerk erwartet.

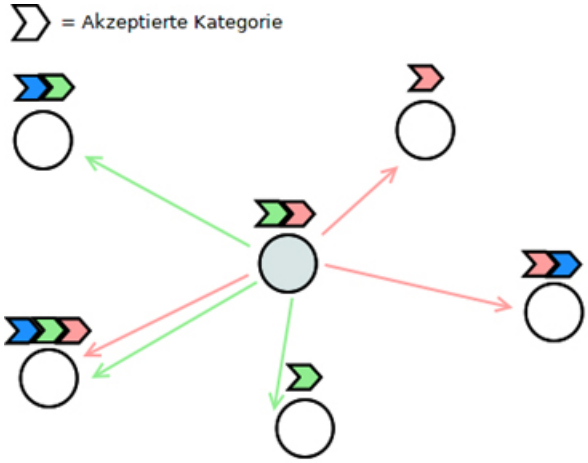

Abb. 3: Illustration von möglichen Verbindungen des grauen Knotens. Die Pfeile stellen Verbindungen zu Daten einer bestimmten Kategorie dar.

Da eine vollständig korrekte Markierung der Daten nicht realistisch erscheint konzentrieren wir uns auf den entgegengesetzten Weg und suchen eine Methode, wie inkorrekt kategorisierte Daten markiert werden können. Auf diese Weise kann zwar nicht sichergestellt werden, dass alle im Netzwerk vorhandenen Daten korrekt markiert sind, aber alle Daten werden systematisch und kontinuierlich überprüft.

Diese Konsenshaltung lässt sich auf das „byzantinische Generäleproblem“ (Treaster, 2005) reduzieren. Bei diesem versucht man in einem dezentralen System mit teilweise böswilligen, verdeckt schädlich agierenden Teilnehmern einen globalen Konsens zu finden. Dies trifft auf das Comademlia-Netzwerk zu, da auch in diesem Fall ein globaler Konsensus gefunden werden soll nämlich der, welche Daten als inkorrekt markiert werden müssen. Ebenfalls analog zu dem Generäleproblem können wir davon ausgehen, dass einige der Netzwerkteilnehmer bösartig handeln werden, indem sie z.B. korrekt kategorisierte Daten als inkorrekt markieren möchten.

Eine Lösung dieses Problems für das Comademlia-Netzwerk wird auf Basis der Blockchain Technologie (Swan, 2015) angestrebt. Blockchains werden unter anderem von der Bitcoin Währung (Nakamoto, 2008) genutzt und für Comademlia entsprechend extensiv angepasst und verändert. Als Grundprinzip geht man davon aus, dass ein Knoten bestimmte Daten nur dann als inkorrekt markieren darf, wenn dieser vorher eine bestimmte Menge an für das Netzwerk nützlichen Handlungen und Ressourcen investiert hat. Wie Inhalte markiert werden, ist den konkreten Implementierungen der Endsysteme überlassen.[4]

Die benötigten positiven Handlungen haben zur Folge, dass die schädlichen Aktionen von böswilligen Knoten nur dann durchgeführt werden können, wenn diese vorher mit genügend positiven Handlungen aufgewogen werden. Die technische Umsetzung beruht auf verschiedenen kryptographischen Primitiven, wie etwa kryptographischen „Quittungen“ für erbrachte Leistungen für das Netzwerk.

Zusammenfassung & Implikationen

In diesem Précis wurde das Konzept eines Compliance Management für Peer-to-Peer Netzwerke beschrieben mit dem, Nutzerinnen und Nutzer die Kontrolle über die Art der von ihnen zur Verfügung gestellten Daten ermöglicht wird. Damit soll ein zentrales Motivationsproblem für die Beteiligung an solchen Netzwerken gelöst werden, die eine wichtige Rolle für die Möglichkeit zur unzensierten Kommunikation spielen können. Des Weiteren wurde ein konkreter Ansatz für ein Peer-to-Peer Compliance Management Netzwerk vorgestellte, welches sich aktuell in der Entwicklung befindet.

Der Anwendungsbereich von Compliance Management Netzwerken ist nicht auf die Verhinderung von Internetzensur beschränkt. In der hier vorgestellten Form können Compliance Management Netzwerke auch innerhalb von beliebigen Inhaltsnetzwerken genutzt werden, um automatisiert regional angepasste Inhalte nur an bestimmten Standorten zu speichern und auszuliefern.

4 Ein Beispiel dafür wäre der Facebook „Beitrag melden“ Button, wenn Facebook das Endsystem ist, welches das Comademlia Netzwerk implementiert.

In zukünftiger Forschung liegt der Schwerpunkt darauf Compliance Management Netzwerke auch außerhalb von Simulationen in Echtweltszenarien einsetzbar zu machen. Dazu müssen allerdings noch verschiedene offene Fragestellungen technischer Natur geklärt werden, wie zum Beispiel, ob das für Peer-to-Peer Netzwerke typische stetige Beitreten und Verlassen des Netzwerks zu Problemen führt oder, ob der Blockchain-basierte Mechanismus zur Kategorisierung auch in kleinen Netzwerken funktioniert.

Literatur

Graffi, K. (2011). PeerfactSim. KOM: A P2P system simulator—Experiences and lessons learned. In *Peer-to-Peer Computing (P2P), 2011 IEEE International Conference on* (pp. 154-155). IEEE.

Izal, M., Urvoy-Keller, G., Biersack, E. W., Felber, P. A., Al Hamra, A., & Garces-Erice, L. (2004, April). Dissecting bittorrent: Five months in a torrent's lifetime. In *International Workshop on Passive and Active Network Measurement* (pp. 1-11). Springer: Berlin/Heidelberg.

Lua, E. K., Crowcroft, J., Pias, M., Sharma, R., & Lim, S. (2005). A survey and comparison of peer-to-peer overlay network schemes. *IEEE Communications Surveys & Tutorials*, 7(2), 72-93.

Maymounkov, P. &Mazieres, D. (2002). Kademlia: A peer-to-peer information system based on the xor metric. *International Workshop on Peer-to-Peer Systems*. Springer: Berlin/Heidelberg.

Nakamoto, S. (2008). Bitcoin: ***A** peer-to-peer electronic cash system*. Online abrufbar am 21.03.2017 unter: https://bitcoin.org/bitcoin.pdf.

Stingl, D., Gross, C., Rückert, J., Nobach, L., Kovacevic, A., & Steinmetz, R. (2011). Peerfactsim. kom: A simulation framework for peer-to-peer systems. In W. W. Smari (Ed.). Proceedings of the *2011 International Conference on Hih Performance Computing & Simulation* (HPCS), (pp. 577-584). IEEE.

Stoica, I., Morris, R., Karger, D., Kaashoek, M. F., & Balakrishnan, H. (2001). Chord: A scalable peer-to-peer lookup service for internet applications. *ACM SIGCOMM Computer Communication Review*, 31(4), 149-160.

Swan, M. (2015). *Blockchain: Blueprint for a new economy*. Sebastopol: O'Reilly Media.

Treaster, M. (2005). A survey of fault-tolerance and fault-recovery techniques in parallel systems. *arXiv preprint* cs/0501002