

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
20. September 2007 (20.09.2007)

PCT

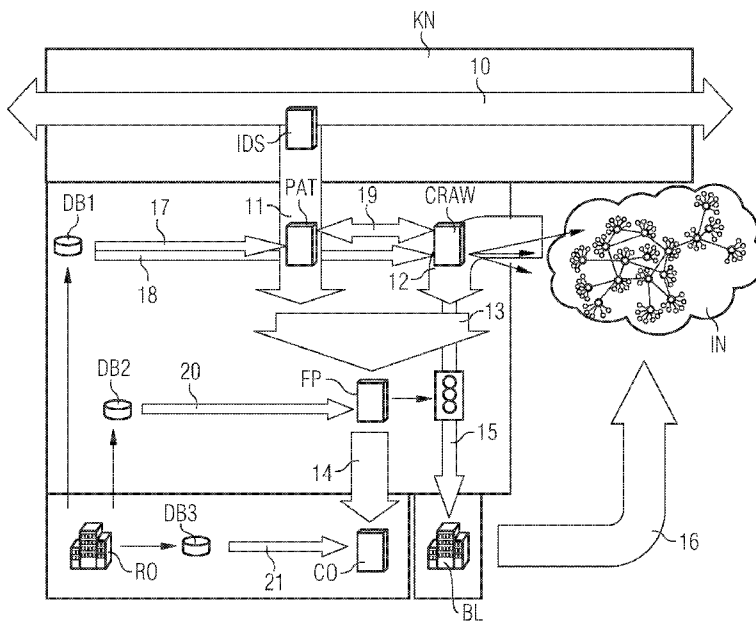
(10) Internationale Veröffentlichungsnummer
WO 2007/104691 A2

- (51) Internationale Patentklassifikation: Nicht klassifiziert
- (72) Erfinder; und
(75) Erfinder/Anmelder (nur für US): BAUSCHERT, Thomas [DE/DE]; Olivierstr. 18, 81477 München (DE). BÄSE, Gero [DE/DE]; Thalkirchner Str. 184, 81371 München (DE). FINKENZELLER, Michael [DE/DE]; Welslerstr. 3, 81373 München (DE). WINTER, Martin [DE/DE]; Hugo-Wolf-Str. 1, 83024 Rosenheim (DE).
- (21) Internationales Aktenzeichen: PCT/EP2007/052161
- (22) Internationales Anmeldedatum: 8. März 2007 (08.03.2007)
- (25) Einreichungssprache: Deutsch
- (74) Gemeinsamer Vertreter: SIEMENS AKTIENGESELLSCHAFT; Postfach 22 16 34, 80506 München (DE).
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität: 10 2006 011 294.6 10. März 2006 (10.03.2006) DE
- (81) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR,

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD AND COMMUNICATION SYSTEM FOR THE COMPUTER-AIDED DETECTION AND IDENTIFICATION OF COPYRIGHTED CONTENTS

(54) Bezeichnung: VERFAHREN UND KOMMUNIKATIONSSYSTEM ZUM RECHNERGESTÜTZTEN AUFFINDEN UND IDENTIFIZIEREN VON URHEBERRECHTLICH GESCHÜTZTEN INHALTEN



(57) Abstract: Disclosed is a method for the computer-aided detection and identification of copyrighted contents that are exchanged between at least two computers in a communication network, especially in peer-to-peer networks. Said method comprises the following steps: - first data packets that are specified according to an execute command and are analyzed regarding at least one first criterion are fed to a first computer (PAT), first and second parameters being determined from the data packets meeting the at least one first criterion; - the first computer (PAT) determines the first data packets encompassing the second parameter from all first data packets that are fed to the first computer (PAT) and transmits said data packets to a second computer (FP); - a third computer (CRAW) sends at least one inquiry message for detecting data with copyrighted contents to the communication network, said third computer (CRAW)

receives reply messages in reaction to the at least one inquiry message and requests second data packets meeting at least one second criterion from the communication network and analyzes the same, third and fourth parameters being determined from the data packets meeting the at least one second criterion; - the third computer (CRAW) determines the second data packets encompassing the fourth parameter from all second data packets that are fed to the third computer (CRAW) and transmits said data packets to the second computer (FP); - the first computer (PAT) transmits the first parameters to the third computer (CRAW) in order for said first parameters to be used in the second criteria; and - the third computer (CRAW) transmits the third parameters to the second computer (PAT) in order for said third parameters to be used in the first criteria.

[Fortsetzung auf der nächsten Seite]

WO 2007/104691 A2



LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF,

CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

- ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(57) **Zusammenfassung:** Es wird ein Verfahren zum rechnergestützten Auffinden und Identifizieren von urheberrechtlich geschützten Inhalten beschrieben, welche in einem Kommunikationsnetzwerk, insbesondere in Peer-to-Peer-Netzwerken, zwischen zumindest zwei Rechnern ausgetauscht werden. Dieses weist die folgenden Schritte auf: - es werden einem ersten Rechner (PAT) gemäß einer Ausführungs Vorschrift spezifizierte erste Datenpakete zugeführt, die hinsichtlich zumindest eines ersten Kriteriums analysiert werden, wobei aus den das zumindest eine erste Kriterium erfüllenden Datenpaketen erste und zweite Parameter ermittelt werden; - der erste Rechner (PAT) ermittelt aus sämtlichen ihm zugeführten ersten Datenpaketen diejenigen ersten Datenpakete, die den zweiten Parameter umfassen und übermittelt diese Datenpakete an einen zweiten Rechner (FP); - ein dritter Rechner (CRAW) sendet zumindest eine Anfragenachricht zum Auffinden von Daten mit urheberrechtlich geschützten Inhalten an das Kommunikationsnetzwerk, wobei der dritte Rechner (CRAW) in Reaktion auf die zumindest eine Anfragenachricht Antwortnachrichten erhält und zumindest ein zweites Kriterium erfüllende zweite Datenpakete aus dem Kommunikationsnetzwerk anfordert und analysiert, wobei aus den das zumindest eine zweite Kriterium erfüllenden Datenpaketen dritte und vierte Parameter ermittelt werden; - der dritte Rechner (CRAW) ermittelt aus sämtlichen ihm zugeführten zweiten Datenpaketen diejenigen zweiten Datenpakete, die den vierten Parameter umfassen und übermittelt diese Datenpakete an den zweiten Rechner (FP); - der erste Rechner (PAT) übermittelt die ersten Parameter an den dritten Rechner (CRAW) zur Verwendung in den zweiten Kriterien; und - der dritte Rechner (CRAW) übermittelt die dritten Parameter an den zweiten Rechner (PAT) zur Verwendung in den ersten Kriterien.

Beschreibung

Verfahren und Kommunikationssystem zum rechnergestützten Auf-
finden und Identifizieren von urheberrechtlich geschützten
5 Inhalten

Die Erfindung betrifft ein Verfahren und ein Kommunikations-
system zum rechnergestützten Auffinden und Identifizieren von
urheberrechtlich geschützten Inhalten, welche in einem Kommu-
10 nikationsnetzwerk, insbesondere in Peer-to-Peer-Netzwerken,
zwischen zumindest zwei Rechnern ausgetauscht werden.

Die Ausbreitung von digitalen Formaten und Kompressions-
Technologien für Audio- und Videodaten hat Kommunikations-
15 netzwerke, wie das Internet, als Leitungen für den weltweiten
Tausch von Musik, Videos und Kinofilmen, Software und sonsti-
gen digitalen Informationen stark beeinflusst. Dank Digitali-
sierung und Codier-Techniken enthalten Dateien vollständige
Songs oder auch Filme, die auf einfache Weise über das Inter-
20 net verbreitet und getauscht werden können. Die Dateien kön-
nen unter Verwendung herkömmlicher Browser üblicherweise über
das World Wide Web (www) auf einen Rechner geladen werden.
Dabei existieren spezielle Applikationen, wie z. B. KaZaA,
Bittorrent, eMule und andere, welche im Rahmen von Peer-to-
25 Peer-Netzwerken die Suche und den Austausch von urheberrecht-
lich geschützten Daten auf einfache Weise ermöglichen. Die
Urheber der Inhalte, wie z. B. die Musik- und Filmindustrie,
erleiden durch derartige Piraterie-Netzwerke große Umsatzver-
luste. Aufgrund der zunehmenden Bandbreite bei der Übertra-
30 gung von Daten in den Kommunikationsnetzwerken wird auch der
Tausch von großen Dateien, wie z. B. Filmen, auf immer einfa-
chere Weise möglich.

Um den Tausch von Daten mit urheberrechtlich geschützten In-
35 halten zu verhindern oder einzudämmen, sind aus dem Stand der
Technik verschiedene Möglichkeiten aufgezeigt worden. Dabei
werden im Wesentlichen zwei Techniken verwendet, die in Fach-

kreisen als "Fingerprinting" und "Watermarking-Technology" bekannt sind.

5 Beim "Fingerprinting" wird ein Fingerabdruck einer Datei bzw. eines Datenpakets mit Audio- und/oder Videodaten ermittelt. Dabei werden die in einem Datenpaket vorhandenen Bits analysiert und ein Fingerabdruck, z. B. eine Identifikationszeichenfolge, errechnet und mit in einer Datenbank hinterlegten Identifikationszeichenfolgen verglichen, um die Identität
10 oder Gleichheit der Daten festzustellen.

Beim so genannten "Watermarking" wird vom Inhaber der urheberrechtlich geschützten Inhalte ein Wasserzeichen in die Datenpakete einer Datei eingearbeitet, welches den Inhalt und den Empfänger der Datei beschreibt. Diese in die Dateien eingearbeiteten Wasserzeichen können extrahiert und mit in einer Datenbank hinterlegten Wasserzeichen zur Identitätsprüfung verglichen werden.

20 Grundsätzlich können durch Fingerabdrücke und Wasserzeichen markierte Daten, die in Peer-to-Peer-Netzwerken getauscht werden, anhand der Fingerabdrücke und Wasserzeichen aufgefunden und identifiziert werden. Da dieses Vorgehen jedoch mit einem großen Zeitaufwand verbunden ist, werden zum Auffinden
25 von urheberrechtlich geschützten Inhalten in Peer-to-Peer-Netzwerken üblicherweise Schlüsselwörter, auch als Keywords bezeichnet, verwendet. Der Nachteil dieser Vorgehensweise besteht darin, dass eine Suche nach Schlüsselwörtern eine große Anzahl an dieses Kriterium erfüllende Daten gibt, wobei diese
30 nur zum Teil in Peer-to-Peer-Netzwerken unberechtigtweise ausgetauschte Inhalte betreffen.

Die in Peer-to-Peer-Netzwerken oder -Tauschbörsen verfügbaren Medieninhalte, was stellvertretend für Audio- und/oder Video-
35 Inhalte stehen soll, sind üblicherweise mit einer eindeutigen Kennzeichnung versehen, anhand der ein so genannter Peer-to-Peer-Client-Rechner ein Laden des gewünschten Inhaltes vornehmen kann. Die eindeutige Kennzeichnung ermöglicht es, dass

die Vielzahl an Datenpaketen, welche den gesamten Medieninhalt beschreibt, von unterschiedlichen Peer-to-Peer-Hosts geladen werden kann.

5 Das Ausfindigmachen von urheberrechtlich geschützten Inhalten (verkörpert in Form einer Datei, welche als Vielzahl von Datenpaketen in einem Kommunikationsnetzwerk übertragbar ist) in Peer-to-Peer-Netzwerken kann auf verschiedenen Schichten des Kommunikationsnetzwerks erfolgen. So kann dies beispielsweise durch Analyse eines Datenpakets, einschließlich Header und Nutzdaten, erfolgen. Das Auffinden kann aber auch ausschließlich auf Basis der Analyse der Nutzdaten erfolgen, indem z. B. nach dem oben beschriebenen Fingerabdrücken oder Wasserzeichen gesucht wird. Die Suche kann alternativ anhand 10 der oben genannten Schlüsselwörter oder sonstigen Inhalte, welche durch das Peer-to-Peer-Netzwerk von sich bereitgestellt werden, erfolgen. 15

Um das Tauschen von urheberrechtlich geschützten Inhalten in Peer-to-Peer-Netzwerken eindämmen zu können, sind verschiedene Mechanismen bekannt. So kann beispielsweise das Blockieren von Datenpaketen oder das Einschränken der Bandbreite eines Peer-to-Peer-Teilnehmerrechners (Host und/oder Client) erfolgen. Es kann ein Umlenken oder Zwischenspeichern (zur Erlangung einer zeitlichen Verzögerung) von Peer-to-Peer-Datenpaketen erfolgen. Ebenso ist es bekannt, die in einem Peer-to-Peer-Netzwerk ausgetauschten Dateien mit "Dummy-Daten" anzureichern, um so beim Empfänger einer über eine Peer-to-Peer-Tauschbörse geladenen Datei eine Verfälschung, d.h. Verschlechterung des Inhalts, herbeizuführen. 20 25 30

Es ist Aufgabe der vorliegenden Erfindung, ein Verfahren und ein Kommunikationssystem zum rechnergestützten Auffinden und Identifizieren von urheberrechtlich geschützten Inhalten anzugeben, mit welchen der Austausch von Dateien in Tauschbörsen verhindert oder zumindest erschwert wird. 35

Diese Aufgaben werden mit einem Verfahren zum rechnergestützten Auffinden und Identifizieren von urheberrechtlich geschützten Inhalten gemäß Patentanspruch 1 sowie durch ein Kommunikationssystem mit den Merkmalen des Patentanspruches
5 13 gelöst. Vorteilhafte Ausgestaltungen finden sich jeweils in den abhängigen Patentansprüchen.

Bei dem erfindungsgemäßen Verfahren zum rechnergestützten Auffinden und Identifizieren von urheberrechtlich geschützten
10 Inhalten, welche in einem Kommunikationsnetzwerk, insbesondere in Peer-to-Peer-Netzwerken, zwischen zumindest zwei Rechnern ausgetauscht werden, werden folgende Schritte durchlaufen: Es werden einem ersten Rechner gemäß einer Ausführungsvorschrift spezifizierte erste Datenpakete zugeführt, die
15 hinsichtlich zumindest eines ersten Kriteriums analysiert werden, wobei aus den das zumindest eine erste Kriterium erfüllenden Datenpaketen erste und zweite Parameter ermittelt werden. Der erste Rechner ermittelt aus sämtlichen ihm zugeführten ersten Datenpaketen diejenigen ersten Datenpakete,
20 die den zweiten Parameter umfassen und übermittelt diese Datenpakete an einen zweiten Rechner. Ein dritter Rechner sendet zumindest eine Anfragenachricht zum Auffinden von Daten mit urheberrechtlich geschützten Inhalten an das Kommunikationsnetzwerk, wobei der dritte Rechner in Reaktion auf die zu-
25 mindest eine Anfragenachricht Antwortnachrichten erhält und zumindest ein zweites Kriterium erfüllende zweite Datenpakete aus dem Kommunikationsnetz anfordert und analysiert, wobei aus den das zumindest eine zweite Kriterium erfüllenden Datenpaketen dritte und vierte Parameter ermittelt werden. Der
30 dritte Rechner ermittelt aus sämtlichen ihm zugeführten zweiten Datenpaketen diejenigen zweiten Datenpakete, die den vierten Parameter umfassen und übermittelt diese Datenpakete an den zweiten Rechner. Der erste Rechner übermittelt die ersten Parameter an den dritten Rechner zur Verwendung in den
35 zweiten Kriterien. Der dritte Rechner übermittelt die dritten Parameter an den zweiten Rechner zur Verwendung in den ersten Kriterien.

Durch den Einsatz zweier Rechner, den ersten und den dritten Rechner, zum Auffinden von urheberrechtlich geschützten Inhalten können verschiedene Arten des Ausfilterns von relevanten Datenpaketen vorgenommen werden. Die dabei jeweils gewonnenen Erkenntnisse werden zwischen dem ersten und dem dritten Rechner ausgetauscht, so dass deren Suche mit zunehmender Zeit immer zielgerichteter wird. Das Auffinden von urheberrechtlich geschützten Inhalten ist dadurch in sehr kurzer Zeit möglich. Die als relevant erachteten Datenpakete werden einem zweiten Rechner zur genaueren Analyse zugeführt, welcher mit großer Zuverlässigkeit die Entscheidung treffen kann, ob es sich bei den ausgefilterten Datenpaketen um solche mit urheberrechtlich geschützten Inhalten handelt oder nicht.

15

Der erste Rechner analysiert die ihm zugeführten ersten Datenpakete hinsichtlich zumindest eines ersten Kriteriums, wobei der erste Rechner im Wesentlichen überprüft, ob das bzw. die ihm zugeführten ersten Datenpakete eine so genannte Anfragenachricht darstellen. Ist dies der Fall, so ermittelt der erste Rechner erste und zweite Parameter, wobei die ersten Parameter beispielsweise Schlüsselwörter, welche als Keywords bezeichnet werden, und die zweiten Parameter Peer-to-Peer-Metadaten, wie z. B. Hash-Keys, verifizierte Schlüsselwörter (d.h. solche Schlüsselwörter, die mit hoher Wahrscheinlichkeit oder sogar Sicherheit Peer-to-Peer-Daten kennzeichnen) oder inhaltsbezogene Daten, darstellen. In gleicher Weise analysiert der dritte Rechner die ihm zugeführten zweiten Datenpakete hinsichtlich eines zweiten Kriteriums. Der dritte Rechner prüft im Wesentlichen, ob die ihm auf eine Anfragenachricht gelieferten Ergebnisse Peer-to-Peer-Tauschbörsen zuzuordnen sind. Ist dies der Fall, so ermittelt der dritte Rechner dritte und vierte Parameter, wobei die dritten Parameter beispielsweise Schlüsselwörter (Keywords) und die zweiten Parameter Peer-to-Peer-Metadaten, insbesondere Hash-Keys, darstellen. Durch das wechselweise zur Verfügung Stellen des ersten und des vierten Parameters ergibt sich ein Selbstlernmechanismus, durch den das Auffinden und

35

Identifizieren von urheberrechtlich geschützten Daten in sehr kurzer Zeit durchführbar ist. Darüber hinaus ist es möglich, innerhalb eines kurzen Zeitraumes eine solch große Datenmenge mit Datenpaketen mit urheberrechtlich geschützten Inhalten aufzufinden, um den Nachweis zu erbringen, dass tatsächlich ein Urheberrecht verletzt wird.

In einer Ausführungsform werden die den zweiten Parameter umfassenden ersten Datenpakete und die den vierten Parameter umfassenden zweiten Datenpakete zur weiteren Analyse in einer Datengesamtheit zusammengeführt, wenn der zweite und der vierte Parameter übereinstimmen. Die Auswahl, welche der zweiten und vierten Parameter zur Weiterleitung der Daten an den zweiten Rechner führen, kann beispielsweise mit Hilfe eines selbstlernenden Verfahrens durchgeführt werden. Zur Analyse, ob Datenpakete urheberrechtlich geschützte Inhalte aufweisen, wird eine Datenmenge gebildet, die sowohl erste als auch zweite Datenpakete umfasst, welche von dem ersten Rechner bzw. dem dritten Rechner ermittelt wurden. Um eine zielgerichtete Auswertung vornehmen zu können, werden dabei jeweils solche erste und zweite Datenpakete zur weiteren Verarbeitung in einer Datengesamtheit zusammengeführt, bei welchen der zweite und der vierte Parameter, z. B. ein Schlüsselwort oder bevorzugt ein Hash-Key, übereinstimmen. Hierdurch ist auf einfache Weise die Überprüfung möglich, ob ein bestimmter urheberrechtlich geschützter Inhalt im Rahmen der Peer-to-Peer-Tauschbörsen ausgetauscht bzw. von einem Teilnehmer der Peer-to-Peer-Tauschbörse heruntergeladen wird.

Im Weiteren wird zumindest eines der Datenpakete jeder der Datengesamtheit einer Fingerabdruck-Analyse unterzogen, indem aus dem zumindest einen der Datenpakete jeder der Datengesamtheit eine Identifikationszeichenfolge ermittelt wird und mit Referenz-Identifikationszeichenfolgen verglichen wird. Die Fingerabdruck-Analyse ist, wie einleitend bereits erwähnt, in Fachkreisen auch als Fingerprint-Analyse bekannt, bei der das zumindest eine Datenpaket hinsichtlich einer bestimmten Bitfolge untersucht wird. Die als Fingerprint be-

zeichnete Bitfolge wird mit Referenz-Identifikationszeichenfolgen verglichen. Bei einer Übereinstimmung kann davon ausgegangen werden, dass das Datenpaket urheberrechtlich geschützten Inhalt umfasst. Bevorzugt wird im Rahmen der Analyse jedes der Datenpakete jeder der Datengesamtheit einer Fingerabdruck-Analyse unterzogen. Anhand dessen kann beispielsweise mit großer Zuverlässigkeit eine Unterscheidung getroffen werden, ob verbotenerweise ein Song oder ein Film oder ein legal zu ladender Trailer über die Peer-to-Peer-Tauschbörse ausgetauscht wird. Diese Unterscheidung ist für Frage, ob und welche Mittel zum Unterbinden des unerlaubten Tausches derartiger Daten eingesetzt werden, von Bedeutung.

Gemäß einer weiteren Ausführungsform sind die Referenz-Identifikationszeichenfolgen von dem oder den Urhebern der geschützten Inhalte bereitgestellt.

Es ist in einer Ausführungsform vorgesehen, dass bei übereinstimmenden Identifikationszeichenfolgen einer Datengesamtheit der zweite bzw. vierte Parameter an einen vierten Rechner übertragen werden, welcher anhand des zweiten bzw. vierten Parameters eine Beeinflussung von solchen Datenpaketen in dem Kommunikationsnetz vornehmen kann, welche den zweiten bzw. vierten Parameter aufweisen. Die Beeinflussung ist in Fachkreisen auch unter dem Begriff des "Policing" bekannt. Die Beeinflussung von solchen Datenpaketen in dem Kommunikationsnetz, welche den zweiten bzw. vierten Parameter aufweisen, kann einen oder mehrere der folgenden Schritte umfassen:

- Blockieren der Datenpakete,
- Umleiten der Datenpakete zu einem von dem in dem Datenpaket angegebenen Zielrechner verschiedenen Rechner,
- Verwerfen der Datenpakete,
- Verändern der Datenpakete.

In einer weiteren Ausführungsform ist vorgesehen, dass bei übereinstimmenden Identifikationszeichenfolgen einer Datengesamtheit der zweite bzw. vierte Parameter sowie die Datengesamtheit an einen fünften Rechner übertragen wird, welcher

anhand dieser Daten eine Wasserzeichen-Analyse vornehmen kann. Die Wasserzeichen-Analyse stellt die eingangs erwähnte "Watermarking-Technology" dar, anhand welcher nicht nur eine Überprüfung der Datenpakete vorgenommen werden kann, ob es
5 sich dabei um urheberrechtlich geschütztes Datenmaterial handelt, sondern auch, wer der Empfänger des bzw. der Datenpakete ist. Diese Vorgehensweise ist insbesondere dazu bestimmt, eine Rechtsverfolgung des unerlaubten Datentausches vornehmen zu können.

10

Gemäß einer weiteren Ausführungsform werden die ersten und die dritten Parameter aus einer Datenbank ausgelesen, wobei die in der Datenbank vorgehaltenen Daten von einer den fünften Rechner verwaltenden Organisation bereitgestellt werden.

15

Bei der den fünften Rechner verwaltenden Organisation kann es sich beispielsweise um den Inhaber oder Urheber des urheberrechtlich geschützten Inhalts handeln. Die ersten und die dritten Parameter umfassen insbesondere Schlüsselwörter, welche den urheberrechtlich geschützten Inhalt charakterisieren
20 und kennzeichnen. Darüber hinaus können die ersten und die dritten Parameter jedoch durch solche Inhalte ergänzt werden, welche im Rahmen der Analyse der Datenpakete durch den ersten und den dritten Rechner ermittelt werden.

25

Gemäß einer weiteren Ausführungsform analysiert ein Filter-Rechner die in einem ersten Kommunikationsnetz übertragenen Datenpakete und führt die, die Ausführungsvorschrift erfüllenden, Datenpakete als erste Datenpakete dem ersten Rechner zur weiteren Verarbeitung zu.

30

Bei dem Filter-Rechner kann es sich beispielsweise um einen Netzwerk-Zugangsknoten- oder einen Aggregationspunkt-Knotenrechner handeln. Die Aufgabe des Filter-Rechners besteht darin, die in einem ersten Kommunikationsnetz übertragenen Datenpakete dahingehend zu analysieren, ob das Datenpaket ein "Peer-to-Peer-Datenpaket" ist. Diese Analyse
35 kann auf unterschiedlichsten Weisen erfolgen. Es ist eine Analyse möglich, welche das gesamte Datenpaket, das heißt so-

wohl Header als auch Nutzdaten, betrachtet. Die Analyse kann sich jedoch auch ausschließlich auf die Analyse der Headerdaten oder der Nutzdaten beziehen. Schließlich kann auch eine Analyse anhand eines bekannten Kontextes erfolgen. Die Art und Weise, wie die, die ersten Ausführungsvorschrift erfüllenden, Datenpakete ermittelt werden, ist prinzipiell beliebig.

Ein erfindungsgemäßes Computerprogrammprodukt kann direkt in den internen Speicher eines digitalen Computers geladen werden und umfasst Softwarecodeabschnitte, mit denen die Schritte gemäß einem der vorhergehenden Ansprüche ausgeführt werden, wenn das Produkt auf einem Computer läuft.

Ein erfindungsgemäßes Kommunikationssystem zum rechnergestützten Auffinden und Identifizieren von urheberrechtlich geschützten Inhalten, welche in einem Kommunikationsnetzwerk, insbesondere in Peer-to-Peer-Netzwerken, zwischen zumindest zwei Rechnern ausgetauscht werden, umfasst einen ersten, einen zweiten und einen dritten Rechner. Der erste Rechner, dem gemäß einer Ausführungsvorschrift spezifizierte erste Datenpakete zuführbar sind, ist dazu ausgebildet:

- die ersten Datenpakete hinsichtlich zumindest eines ersten Kriteriums zu analysieren;
- aus den das zumindest eine erste Kriterium erfüllenden Datenpaketen erste und zweite Parameter zu ermitteln;
- aus sämtlichen ihm zugeführten ersten Datenpaketen diejenigen ersten Datenpakete zu ermitteln, die den zweiten Parameter umfassen und diese Datenpakete an einen zweiten Rechner zu übertragen;
- die ersten Parameter an den dritten Rechner zur Verwendung in den zweiten Kriterien zu übertragen.

Der dritte Rechner ist dazu ausgebildet,

- zumindest eine Anfragenachricht zum Auffinden von Daten mit urheberrechtlich geschützten Inhalten an das Kommunikationsnetzwerk zu senden und in Reaktion auf die zumindest eine Anfragenachricht Antwortnachrichten zu empfangen;

- zumindest ein zweites Kriterium erfüllende zweite Datenpakete aus dem Kommunikationsnetz anzufordern und zu analysieren, und aus den das zumindest eine zweite Kriterium erfüllenden Datenpaketen dritte und vierte Parameter zu ermitteln;
5
- aus sämtlichen ihm zugeführten zweiten Datenpaketen diejenigen zweiten Datenpakete zu ermitteln, die den vierten Parameter umfassen und diese Datenpakete an den zweiten Rechner zu übermitteln;
- 10 - die dritten Parameter an den zweiten Rechner zur Verwendung in den ersten Kriterien zu übermitteln.

Mit dem erfindungsgemäßen Kommunikationssystem sind dabei die gleichen Vorteile verbunden, wie sie vorstehend in Verbindung
15 mit dem erfindungsgemäßen Verfahren erläutert wurden.

In einer Ausführungsform ist der zweite Rechner dazu ausgebildet, die den zweiten Parameter umfassenden ersten Datenpakete und die den vierten Parameter umfassenden zweiten Datenpakete zur weiteren Analyse in einer Datengesamtheit zusammenzuführen, wenn der zweite und der vierte Parameter übereinstimmen.
20

Gemäß einer weiteren Ausbildung ist der zweite Rechner weiter dazu ausgebildet, zumindest eines der Datenpakete jeder der Datengesamtheit einer Fingerabdruck-Analyse zu unterziehen, indem aus dem zumindest einen der Datenpakete jeder der Datengesamtheit eine Identifikationszeichenfolge ermittelt wird und mit Referenz-Identifikationszeichenfolgen verglichen
25 wird.
30

Es ist gemäß einer weiteren Ausführungsform ein vierter Rechner vorgesehen, dem bei übereinstimmenden Identifikationszeichenfolgen einer Datengesamtheit der zweite bzw. vierte Parameter zuführbar sind, wobei der vierte Rechner dazu ausgebildet ist, anhand des zweiten bzw. vierten Parameters eine Beeinflussung von solchen Datenpaketen in dem Kommunikations-
35

netz vorzunehmen, welche den zweiten bzw. vierten Parameter aufweisen.

In einer weiteren Ausbildung ist ein fünfter Rechner vorgesehen, dem bei übereinstimmenden Identifikationszeichenfolgen einer Datengesamtheit der zweite bzw. vierte Parameter sowie die Datengesamtheit zuführbar sind, wobei der fünfte Rechner dazu ausgebildet ist, anhand dieser Daten eine Wasserzeichen-Analyse vorzunehmen.

10

Dabei ist es vorteilhaft, wenn der vierte und/oder der fünfte Rechner von einem von dem Kommunikationssystem unterschiedlichen Anbieter verwaltet werden. Insbesondere kann der fünfte Rechner in der Einflussosphäre der Rechteinhaber der urheberrechtlich geschützten Inhalte vorgesehen sein. Der vierte Rechner, mit dem geeignete Maßnahmen zum Unterbinden oder Erschweren des Austauschs der urheberrechtlich identifizierten Inhalte vorgenommen werden, kann z.B. einer weiteren, dritten Organisation zugeordnet sein, welche von dem Rechteinhaber beauftragt ist, eine solche Beeinflussung der Datenpakete vorzunehmen.

20

Das erfindungsgemäße Kommunikationssystem umfasst weiter eine erste Datenbank, welche die ersten und die dritten Parameter umfasst, wobei die in der Datenbank vorgehaltenen Daten von einer den fünften Rechner verwaltenden Organisation bereitgestellt sind. Das Kommunikationssystem kann eine zweite Datenbank umfassen, welche die Identifikationszeichenfolgen für die Fingerabdruck-Analyse umfasst, wobei die in der Datenbank vorgehaltenen Daten von einer den fünften Rechner verwaltenden Organisation bereitgestellt sind. Die in der ersten und der zweiten Datenbank enthaltenen Daten bilden die Grundlage für das Auffinden und Identifizieren urheberrechtlich geschützter Daten bzw. Datenpakete. Insbesondere die darin vorgehaltenen Parameter ermöglichen eine zielgerichtete und damit zeit-effiziente Suche nach solchen Inhalten.

35

Es ist ferner zumindest ein Filter-Rechner vorgesehen, der dazu ausgebildet ist, die, in einem ersten Kommunikationsnetzwerk übertragenen, Datenpakete zu analysieren und die, die Ausführungsvorschrift erfüllenden, Datenpakete als erste
5 Datenpakete dem ersten Rechner zur weiteren Verarbeitung zuzuführen.

Wie vorstehend bereits ausgeführt besteht die Aufgabe des Filter-Rechners darin, aus den ihm zugeführten Datenpaketen
10 solche Datenpakete auszufiltern, die Peer-to-Peer-Tauschbörsen zugeordnet sind. Zweckmäßigerweise ist der zumindest eine Filter-Rechner an einem Netzwerkzugangsknoten und/oder an einem Aggregationsknoten des ersten Kommunikationsnetzwerks angeordnet. Die Anordnung des Filter-Rechners an
15 derartigen Netzwerkknoten weist den Vorteil auf, dass ein Großteil der über das erste Kommunikationsnetz übertragenen Datenpakete durch diese Netzwerkknoten geleitet wird.

Die Erfindung wird nachfolgend anhand einer Figur näher erläutert. Die einzige Figur zeigt in schematischer Darstellung ein erfindungsgemäßes Kommunikationssystem zum rechnergestützten Auffinden und Identifizieren von urheberrechtlich geschützten Inhalten.

25 Mit IN ist in der Figur ein Kommunikationsnetzwerk bezeichnet, wie z. B. dem Internet. Das Kommunikationsnetzwerk IN kann eine Vielzahl von Kommunikationsnetzwerken aufweisen, welche von jeweiligen Providern verwaltet werden. Das Kommunikationsnetzwerk IN beherbergt Peer-to-Peer-Tauschbörsen,
30 mit einer Vielzahl von Nutzern. Beispiele für derartige Tauschbörsen sind beispielsweise KaZaA, Bittorrent, eMule und viele andere. In diesen Tauschbörsen werden in digitaler Form gespeicherte Inhalte, wie z.B. Songs und Filme, zwischen den einzelnen Mitgliedern der Tauschbörsen getauscht. Die in digitalisierter Form vorliegenden Daten umfassen dabei oftmals
35 urheberrechtlich geschützten Inhalt.

Das mit KN bezeichnete Kommunikationsnetzwerk ist eines der Vielzahl von unterschiedlichen Providern verwalteten Kommunikationsnetzwerken des Internets (Kommunikationsnetzwerk IN). Mit dem Bezugszeichen 10 ist ein durch das Kommunikations-

5 netzwerk KN übertragener Datenstrom gekennzeichnet, der durch einen Netzwerkknoten-Zugangsrechner IDS geleitet wird. Der Rechner IDS könnte auch in einem Aggregationsknoten des Kommunikationsnetzwerkes KN angeordnet sein. Der Rechner IDS ist dazu ausgebildet, jedes Datenpaket des Datenstroms 10 zu ana-

10 lysieren. Die Analyse erfolgt dabei dergestalt, dass der Rechner IDS eine Unterscheidung von solchen Datenpaketen vornimmt, welche Peer-to-Peer-Tauschbörsen zugeordnet werden können und welche nicht. Diejenigen Datenpakete, die keinen im Peer-to-Peer-Kontext aufweisen, werden durch den Rechner

15 IDS ohne weitere Aktion zu dem gewünschten Zielknoten weitergeleitet. Diejenigen Datenpakete, welche einen Peer-to-Peer-Kontext aufweisen, werden jedoch ausgefiltert und als Datenstrom 11 einem Rechner PAT zugeführt. Entgegen der zeichnerischen Darstellung und der nachfolgenden Beschreibung kann eine Mehrzahl an Rechnern IDS, z.B. an jedem Netzübergangsknoten, vorgesehen sein.

Die Analyse, ob ein Datenpaket einen Peer-to-Peer-Kontext aufweist oder nicht, kann auf prinzipiell beliebige Weise erfolgen. Eine Zuordnung zu einer Peer-to-Peer-Tauschbörse kann

25 beispielsweise anhand der Auswertung der Header-Daten erfolgen. So weisen im Rahmen von Peer-to-Peer-Tauschbörsen ausgetauschte Datenpakete beispielsweise spezielle Codierungen in den Header-Daten auf, welche durch den Rechner IDS erkannt

30 werden können. Eine Erkennung ist jedoch auch anhand einer Analyse des Nutzdatenteils eines Datenpakets möglich. Im Rahmen der Analyse, ob ein Datenpaket einen Peer-to-Peer-Kontext aufweist oder nicht, kann auch die Betrachtung eines kompletten Datenpakets, d.h. sowohl des Headers als auch der Nutzdaten,

35 erfolgen. Dies bietet sich insbesondere dann an, wenn innerhalb der Datenpakete nach Hash-Keys und Schlüsselwörtern gesucht wird, was unter Verwendung von Signaturen erfolgt. Dabei wird nach einem bestimmten Bytemuster, wie dies bei Vi-

russcannern der Fall ist, gesucht, die Teil der Medieninhalte sind. Eine andere Möglichkeit besteht darin, nach bestimmten Verkehrsprofilen, d.h. nach bestimmten Mustern des Datenaustausches zu suchen. Durch die Analyse, welche Rechner mit welchem anderen Rechner innerhalb welchen Zeitraums wie viel Daten austauscht, lässt sich feststellen, welche Rechner Partner eines File-Sharings sind.

Um eine gute Filter-Effizienz des Rechners IDS zu erreichen, ist es zweckmäßig, wenn der Rechner regelmäßig mit neuen, Peer-to-Peer-Datenpakete kennzeichnenden Signaturen und Datenmustern aktualisiert wird.

Die Aufgabe des Rechners PAT, dem Datenpakete mit dem Datenstrom 11 vom Rechner IDS zugeführt werden, besteht darin, eine Analyse der Protokoll-Semantik vorzunehmen. Dazu verfügt der Rechner PAT über Informationen der Protokoll-Semantik zumindest der am meisten populären Peer-to-Peer-Netzwerke. Die von dem Rechner PAT durchgeführte Aufgabe besteht darin, aus den Datenpaketen solche Datenpakete zu identifizieren, welche eine Suchanfrage an ein Peer-to-Peer-Tauschnetzwerk beinhalten, um daraus Schlüsselwörter und Metadaten, wie z. B. Hash-Keys (HK) oder Inhaltsbeschreibungen zu extrahieren. Zur Durchführung dieser Aufgabe kann sich der Rechner PAT bereits der Suche von Schlüsselwörtern oder anderer Parameter bedienen, welche in einer Datenbank DB1 vorgehalten sind. Die in der Datenbank DB1 enthaltenen Parameter werden dem Rechner PAT als Datenstrom 17 zur Verfügung gestellt.

Die Inhalte der Datenbank DB1 werden von dem Rechteinhaber der urheberrechtlich geschützten Inhalte bereitgestellt. Dieser ist mit dem Bezugszeichen RO gekennzeichnet.

Die von dem Rechner PAT durchzuführende Aufgabe ist hinsichtlich der Effizienz des vorliegenden Kommunikationssystems von großer Bedeutung. Es ist zu berücksichtigen, dass das Laden eines über Peer-to-Peer-Tauschbörsen geladenen Inhalts binnen einer bestimmten Zeit beendet ist. Innerhalb dieses Zeitrau-

mes muss der Prozess des Auffindens und Verifizierens (ob die
aufgefundenen Inhalte ein Urheberrecht verletzen) sowie gege-
benenfalls das Beeinflussen des Ladens des Datenstroms vorge-
nommen sein. In Anbetracht der immer größeren zur Verfügung
5 stehenden Bandbreiten eines Downloads kann das Laden großer
Dateien in immer kürzerer Zeit erfolgen. In der Praxis kann
die typische Download-Zeit eines neuen und begehrten Medien-
Inhaltes von Peer-to-Peer-Netzwerken mehrere Stunden oder so-
gar Tage aufgrund der begrenzten Upload-Ressourcen und den
10 großen Download-Anfragen betragen. Dieser Umstand wird sich
im Rahmen der vorliegenden Erfindung zunutze gemacht.

Die Aufgabe des Rechners PAT besteht damit im Wesentlichen
darin, aus den ihm zugeführten Datenpaketen Parameter zu er-
15 mitteln, anhand derer eine gezielte Suche nach Peer-to-Peer-
Inhalten möglich ist.

Ein dritter Rechner CRAW ist vorgesehen, um Suchanfragen und
Lade-Anfragen bei mehreren Peer-to-Peer-Netzwerken parallel
20 durchzuführen. Die Suchbegriffe werden ihm hierzu von der Da-
tenbank DB1 und dem Rechner PAT zur Verfügung gestellt. Dies
ist durch die mit den Bezugszeichen 18 und 19 gekennzeichneten
Pfeile dargestellt. Aus der Analyse der aus den Peer-to-
Peer-Tauschbörsen heruntergeladenen Daten (Bezugszeichen 12)
25 ist der Rechner CRAW in der Lage, Hash-Keys zu extrahieren.
Hash-Keys werden in Tauschbörsen üblicherweise dazu verwen-
det, um einen bestimmten Inhalt eindeutig zu kennzeichnen.
Mit anderen Worten bedeutet dies, dass jeder Medieninhalt,
sei es ein Song oder ein Film, einen eindeutigen Hash-Key
30 aufweist. Der Hash-Key wird von den Clients der Peer-to-Peer-
Tauschbörsen verwendet, um das Laden eines gewünschten Me-
dieninhalts vorzunehmen.

Die von dem Rechner CRAW aufgefundenen Hash-Keys werden des-
halb dazu verwendet, um aus dem Kommunikationsnetz IN Daten-
35 pakete mit einem oder mehreren Hash-Keys zu laden. Die Hash-
Keys werden weiterhin von dem Rechner CRAW auch dem Rechner
PAT zur Verfügung gestellt (Bezugszeichen 19), damit dieser

zielgerichtet Datenpakete mit den entsprechenden Hash-Keys
ausfindig machen kann. Die von den Rechnern PAT und CRAW ge-
ladenen Datenpakete werden einem Rechner FP zugeführt (Be-
zugszeichen 14). Durch den wechselweisen Austausch von
5 Schlüsselwörtern und Hash-Keys zwischen den Rechnern PAT und
CRAW wird die Suche nach Datenpaketen mit Peer-to-Peer-
Kontext wesentlich beschleunigt. Das Laden von Datenpaketen,
welche einen bestimmten Hash-Key aufweisen, durch den Rechner
PAT ist deshalb hilfreich, da aufgrund der Anordnung des
10 Rechners IDS an einem Netzwerkzugangsknoten zum Netzwerk KN
ein beachtlicher Datenstrom 10 durch den Rechner IDS geleitet
wird. Die Wahrscheinlichkeit, dass deshalb auch eine große
Anzahl an Datenpaketen mit Peer-to-Peer-Kontext und gegebe-
nenfalls den gesuchten Hash-Keys durchgeleitet wird, ist des-
15 halb groß.

Der Rechner FP führt die von den Rechnern PAT und CRAW zuge-
führten Datenpakete einer genauen Analyse durch. Zu diesem
Zweck bildet der Rechner FP jeweils eine Datenmenge mit Da-
20 tenpaketen identischen Hash-Keys. Jedes der Datenpakete ist
mit einem Fingerabdruck versehen, welcher durch den Rechner
FP ausfindig gemacht werden kann. Von einer Datenbank DB2,
die von dem Rechteinhaber RO gespeist wird, werden dem Rech-
ner FP Referenz-Fingerabdrücke bzw. -Identifikationszeichen-
25 folgen zur Verfügung gestellt. Durch einen Vergleich der Re-
ferenz-Identifikationszeichenfolgen mit den aus den Datenpa-
keten identifizierten Zeichenfolgen ist der Rechner FP im-
stande, festzustellen, ob es sich um Datenpakete mit urheber-
rechtlich geschütztem Inhalt handelt oder nicht. Insbesondere
30 ist der Rechner FP in der Lage, illegal getauschte Medienin-
halte von beispielsweise Trailern, die legal ausgetauscht
werden dürfen, zu unterscheiden. Dies ist deshalb möglich, da
der Rechner FP eine vergleichsweise große Datenmenge zur Ana-
lyse zur Verfügung gestellt bekommt, wobei vorzugsweise jedes
35 Datenpaket der Datenmenge einer Fingerabdruck-Analyse unter-
zogen wird.

Sofern durch den Rechner FP festgestellt wurde, dass es sich bei den ausgefilterten Datenpaketen um einen urheberrechtlich geschützten und verbotenerweise ausgetauschten Dateninhalt handelt, überträgt der Rechner FP Schlüsselwörter, Hash-Keys und die Datengesamtheit an einen Rechner CO (Bezugszeichen 14) sowie die Schlüsselwörter und Hash-Keys an ein Rechner BL (Bezugszeichen 15).

Der Rechner CO befindet sich bevorzugt in der Einflussosphäre des Rechteinhabers. Der Rechteinhaber ist aufgrund in einer Datenbank DB3 gespeicherten Daten in der Lage, die Datenmenge einer Wasserzeichen-Analyse zu unterziehen. Zu diesem Zweck werden die in der Datenbank gespeicherten Daten an den Rechner CO übertragen (Bezugszeichen 21). Anhand des Wasserzeichens ist der Rechteinhaber RO weiterhin in der Lage, denjenigen Datenpaket zu ermitteln, der die Daten in das Kommunikationsnetz eingespeist hat. Bei diesem handelt es sich um einen Teilnehmer des Peer-to-Peer-Netzwerks, welcher verbotenerweise den urheberrechtlich geschützten Inhalt heruntergeladen hat. Der Rechteinhaber RO ist damit in die Lage versetzt, den Peer-to-Peer-Nutzer ausfindig zu machen und gegebenenfalls weitere Schritte gegen ihn einzuleiten.

Der Rechner BL befindet sich bevorzugt bei einem von dem Betreiber des erfindungsgemäßen Kommunikationssystems und dem Rechteinhaber unabhängigen dritten Betreiber, z.B. einem Dienstleister. Der Betreiber des Rechners BL ist damit in der Lage, eine Beeinflussung der in dem Internet getauschten Datenpakete vorzunehmen, indem er beispielsweise Datenpakete mit beliebigem Inhalt und gleichem Hash-Key in das Internet einspeist, so dass für einen Empfänger eines heruntergeladenen Dateninhalts ein nicht-sinnvoller Datenstrom ankommt (Bezugszeichen 16). Die Beeinflussung des Datenstroms kann prinzipiell beliebig und z.B. im Verbund mit einem Internet Service Provider vorgenommen werden. So könnten Datenpakete eines bestimmten Hash-Keys verworfen oder verändert werden. Es könnten ferner die Quellen der Datenpakete blockiert oder deren Bandbreite einschränkt werden.

Das Anordnen der Datenbanken DB1 und DB2 und das Bereitstellen der darin gespeicherten Schlüsselwörter und Fingerprints weist den Vorteil auf, dass eine Analyse und Identifikation von urheberrechtlich geschütztem Inhalt mittels des erfindungsgemäßen Kommunikationssystems vorgenommen werden kann. Die Datenbanken DB1 und DB2 können dabei von einem Provider verwaltet werden, der nicht mit dem Rechteinhaber RO identisch ist. Andererseits ist der Rechteinhaber RO nicht gezwungen, die Originaldaten des zu schützenden Inhaltes zur Verfügung zu stellen, so dass der Provider nicht selbst die Quelle für ein Peer-to-Peer-Tauschnetzwerk darstellen kann.

Das erfindungsgemäße Kommunikationssystem weist eine Reihe von Vorteilen auf, welche durch die Analyse von Daten auf unterschiedlichen Schichten herkommen. Die Erfindung kombiniert Tracking-Lösungen auf unterschiedlichen Schichten mit einem extern (durch den Rechner IDS) durchgeführten Tracking. Der Datenaustausch zwischen mehreren Tracking-Rechnern basiert auf einem Selbstlern-Mechanismus.

Das erfindungsgemäße Kommunikationssystem arbeitet innerhalb des Netzwerks eines Internetservice-Providers und eines Netzwerkbetreibers. Damit ist ein direkter Zugang zu Daten möglich, welcher zwischen Nutzern ausgetauscht wird. Die Erfindung kombiniert verschiedene Stufen von spezialisierten Filterungen und Umlenkungen, um die Gesamtleistungsfähigkeit zu erhöhen. Dabei können existierende IDS-Systeme (Intrusion Detection System) und Protokoll-Analysatoren verwendet werden. Es ist dadurch möglich, in kürzerer Zeit eine kritische Menge an Inhalten zwecks weiterer Analyse zu sammeln. Dies erfolgt aufgrund des Ladens von Daten einer so genannter Crawler-Komponente und eines Paket-Filters. Ein weiterer Vorteil besteht darin, dass kein zusätzlicher Netzwerkverkehr durch die Erfindung verursacht wird. Wesentlich ist dabei der Selbstlernerneffekt durch den Austausch von Schlüsselwörtern und zugeordneten Hash-Keys zwischen einem Paketfilter und einer Crawler-Komponente. Der Selbstlernmechanismus kann durch künstli-

che Intelligenz unterstützt sein. Die Erfindung ermöglicht die zuverlässige Identifikation von unerlaubt ausgetauschten Inhalten, verglichen mit dem blinden Blockieren von Peer-to-Peer-File-Sharing. Die vorgeschlagene Lösung ist deshalb
5 nicht für rechtliche Angriffe der Nutzer von Tauschbörsen anfällig.

Patentansprüche

1. Verfahren zum rechnergestützten Auffinden und Identifizieren von urheberrechtlich geschützten Inhalten, welche in einem Kommunikationsnetzwerk, insbesondere in Peer-to-Peer-Netzwerken, zwischen zumindest zwei Rechnern ausgetauscht werden:

- 5 - es werden einem ersten Rechner (PAT) gemäß einer Ausführungsvorschrift spezifizierte erste Datenpakete zugeführt, die hinsichtlich zumindest eines ersten Kriteriums analysiert werden, wobei aus den das zumindest eine erste Kriterium erfüllenden Datenpaketen erste und zweite Parameter ermittelt werden;
- 10 - der erste Rechner (PAT) ermittelt aus sämtlichen ihm zugeführten ersten Datenpaketen diejenigen ersten Datenpakete, die den zweiten Parameter umfassen und übermittelt diese Datenpakete an einen zweiten Rechner (FP);
- 15 - ein dritter Rechner (CRAW) sendet zumindest eine Anfragenachricht zum Auffinden von Daten mit urheberrechtlich geschützten Inhalten an das Kommunikationsnetzwerk, wobei der dritte Rechner (CRAW) in Reaktion auf die zumindest eine Anfragenachricht Antwortnachrichten erhält und zumindest ein zweites Kriterium erfüllende zweite Datenpakete aus dem Kommunikationsnetz anfordert und analysiert, wobei aus den das zumindest eine zweite Kriterium erfüllenden Datenpaketen dritte und vierte Parameter ermittelt werden;
- 20 - der dritte Rechner (CRAW) ermittelt aus sämtlichen ihm zugeführten zweiten Datenpaketen diejenigen zweiten Datenpakete, die den vierten Parameter umfassen und übermittelt diese Datenpakete an den zweiten Rechner (FP);
- 25 - der erste Rechner (PAT) übermittelt die ersten Parameter an den dritten Rechner (CRAW) zur Verwendung in den zweiten Kriterien; und
- 30 - der dritte Rechner (CRAW) übermittelt die dritten Parameter an den zweiten Rechner (PAT) zur Verwendung in den ersten Kriterien.
- 35

2. Verfahren nach Anspruch 1,
dadurch gekennzeichnet, dass
die den zweiten Parameter umfassenden ersten Datenpakete und
die den vierten Parameter umfassenden zweiten Datenpakete zur
5 weiteren Analyse in einer Datengesamtheit zusammengeführt
werden, wenn der zweite und der vierte Parameter übereinstim-
men.

3. Verfahren nach Anspruch 2,
10 dadurch gekennzeichnet, dass
zumindest eines der Datenpakete jeder der Datengesamtheit ei-
ner Fingerabdruck-Analyse unterzogen wird, indem aus dem zu-
mindest einen der Datenpakete jeder der Datengesamtheit eine
Identifikationszeichenfolge ermittelt wird und mit Referenz-
15 Identifikationszeichenfolgen verglichen wird.

4. Verfahren nach Anspruch 3,
dadurch gekennzeichnet, dass
jedes der Datenpakete jeder der Datengesamtheit einer Finger-
20 abdruck-Analyse unterzogen wird.

5. Verfahren nach Anspruch 3 oder 4,
dadurch gekennzeichnet, dass
die Referenz-Identifikationszeichenfolgen von dem oder den
25 Urhebern der geschützten Inhalte bereitgestellt sind.

6. verfahren nach einem der Ansprüche 3 bis 5,
dadurch gekennzeichnet, dass
bei übereinstimmenden Identifikationszeichenfolgen einer Da-
30 tengesamtheit der zweite bzw. vierte Parameter an einen vier-
ten Rechner (BL) übertragen werden, welcher anhand des zwei-
ten bzw. vierten Parameters eine Beeinflussung von solchen
Datenpaketen in dem Kommunikationsnetz vornehmen kann, welche
den zweiten bzw. vierten Parameter aufweisen.

35

7. Verfahren nach Anspruch 6,
dadurch gekennzeichnet, dass

die Beeinflussung von solchen Datenpaketen in dem Kommunikationsnetz, welche den zweiten bzw. vierten Parameter aufweisen, einen oder mehrere der folgenden Schritte umfassen kann:

- Blockieren der Datenpakete,
- 5 - Umleiten der Datenpakete zu einem von dem in dem Datenpaket angegebenen Zielrechner verschiedenen Rechner,
- Verwerfen der Datenpakete,
- Verändern der Datenpakete.

10 8. Verfahren nach einem der Ansprüche 3 bis 7, dadurch gekennzeichnet, dass bei übereinstimmenden Identifikationszeichenfolgen einer Datengesamtheit der zweite bzw. vierte Parameter sowie die Datengesamtheit an einen fünften Rechner übertragen wird, welcher anhand dieser Daten eine Wasserzeichen-Analyse vornehmen
15 kann.

9. Verfahren nach einem der vorherigen Ansprüche, dadurch gekennzeichnet, dass
20 die ersten und die dritten Parameter aus einer Datenbank ausgelesen werden, wobei die in der Datenbank vorgehaltenen Daten von einer den fünften Rechner verwaltenden Organisation bereitgestellt werden.

25 10. Verfahren nach einem der vorherigen Ansprüche, dadurch gekennzeichnet, dass ein Filter-Rechner (IDS) die in einem ersten Kommunikationsnetz übertragenen Datenpakete analysiert und die, die Ausführungsvorschrift erfüllenden, Datenpakete als erste Datenpakete
30 dem ersten Rechner (PAT) zur weiteren Verarbeitung zuführt.

11. Verfahren nach Anspruch 10, dadurch gekennzeichnet, dass
35 die Ausführungsvorschrift dann erfüllt ist, wenn das Datenpaket ein Peer-to-Peer-Datenpaket ist.

12. Computerprogrammprodukt das direkt in den internen Speicher eines digitalen Computers geladen werden kann und Softwarecodeabschnitte umfasst, mit denen die Schritte gemäß einem der vorhergehenden Ansprüche ausgeführt werden, wenn das
5 Produkt auf einem Computer läuft.

13. Kommunikationssystem zum rechnergestützten Auffinden und Identifizieren von urheberrechtlich geschützten Inhalten, welche in einem Kommunikationsnetzwerk, insbesondere in Peer-
10 to-Peer-Netzwerken, zwischen zumindest zwei Rechnern ausgetauscht werden, mit einem ersten (PAT), einem zweiten (FP) und einem dritten (CRAW) Rechner, wobei:

- der erste Rechner (PAT), dem gemäß einer Ausführungsvorschrift spezifizierte erste Datenpakete zuführbar sind,
15 dazu ausgebildet ist,
 - die ersten Datenpakete hinsichtlich zumindest eines ersten Kriteriums zu analysieren;
 - aus den das zumindest eine erste Kriterium erfüllenden Datenpaketen erste (KW) und zweite Parameter zu
20 ermitteln;
 - aus sämtlichen ihm zugeführten ersten Datenpaketen diejenigen ersten Datenpakete zu ermitteln, die den zweiten Parameter umfassen und diese Datenpakete an einen zweiten Rechner (FP) zu übertragen;
 - die ersten Parameter an den dritten Rechner (CRAW)
25 zur Verwendung in den zweiten Kriterien zu übertragen;
- der dritte Rechner (CRAW) dazu ausgebildet ist,
 - zumindest eine Anfragenachricht zum Auffinden von
30 Daten mit urheberrechtlich geschützten Inhalten an das Kommunikationsnetzwerk zu senden und in Reaktion auf die zumindest eine Anfragenachricht Antwortnachrichten zu empfangen;
 - zumindest ein zweites Kriterium erfüllende zweite Datenpakete aus dem Kommunikationsnetz anzufordern und
35 zu analysieren, und aus den das zumindest eine zweite Kriterium erfüllenden Datenpaketen dritte und vierte Parameter zu ermitteln;
 - aus sämtlichen ihm zugeführten zweiten Datenpaketen

diejenigen zweiten Datenpakete zu ermitteln, die den vierten Parameter umfassen und diese Datenpakete an den zweiten Rechner (FP) zu übermitteln;

5 -- die dritten Parameter an den zweiten Rechner (PAT) zur Verwendung in den ersten Kriterien zu übermitteln.

14. Kommunikationssystem nach Anspruch 13, dadurch gekennzeichnet, dass
10 der zweite Rechner (FP) dazu ausgebildet ist, die den zweiten Parameter umfassenden ersten Datenpakete und die den vierten Parameter umfassenden zweiten Datenpakete zur weiteren Analyse in einer Datengesamtheit zusammenzuführen, wenn der zweite und der vierte Parameter übereinstimmen.

15 15. Kommunikationssystem nach Anspruch 14, dadurch gekennzeichnet, dass
der zweite Rechner (FP) dazu ausgebildet ist, zumindest eines der Datenpakete jeder der Datengesamtheit einer Fingerabdruck-Analyse zu unterziehen, indem aus dem zumindest einen
20 der Datenpakete jeder der Datengesamtheit eine Identifikationszeichenfolge ermittelt wird und mit Referenz-Identifikationszeichenfolgen verglichen wird.

16. Kommunikationssystem nach Anspruch 15,
25 dadurch gekennzeichnet, dass
ein vierter Rechner (BL) vorgesehen ist, dem bei übereinstimmenden Identifikationszeichenfolgen einer Datengesamtheit der zweite bzw. vierte Parameter zuführbar sind, wobei der vierte Rechner dazu ausgebildet ist, anhand des zweiten bzw. vierten
30 Parameters eine Beeinflussung von solchen Datenpaketen in dem Kommunikationsnetz vorzunehmen, welche den zweiten bzw. vierten Parameter aufweisen.

17. Kommunikationssystem nach Anspruch 15 oder 16,
35 dadurch gekennzeichnet, dass
ein fünfter Rechner (CO) vorgesehen ist, dem bei übereinstimmenden Identifikationszeichenfolgen einer Datengesamtheit der zweite bzw. vierte Parameter sowie die Datengesamtheit zu-

föhrbar sind, wobei der fönfte Rechner dazu ausgebildet ist, anhand dieser Daten eine Wasserzeichen-Analyse vorzunehmen.

18. Kommunikationssystem nach Anspruch 16 oder 17,

5 dadurch gekennzeichnet, dass

der vierte und/oder der fönfte Rechner (BL, CO) von einem von dem Kommunikationssystem unterschiedlichen Anbieter verwaltet werden.

10 19. Kommunikationssystem nach einem der vorherigen Anspröche, dadurch gekennzeichnet, dass

dieses eine erste Datenbank (DB1) umfasst, welche die ersten und die dritten Parameter umfasst, wobei die in der Datenbank vorgehaltenen Daten von einer den fönften Rechner verwaltenden Organisation bereitgestellt sind.

15

20. Kommunikationssystem nach einem der vorherigen Anspröche, dadurch gekennzeichnet, dass

20 dieses eine zweite Datenbank (DB2) umfasst, welche die Identifikationszeichenfolgen für die Fingerabdruck-Analyse umfasst, wobei die in der Datenbank vorgehaltenen Daten von einer den fönften Rechner verwaltenden Organisation bereitgestellt sind.

25 21. Kommunikationssystem nach einem der vorherigen Anspröche, dadurch gekennzeichnet, dass

30 zumindest ein Filter-Rechner (IDS) vorgesehen ist, der dazu ausgebildet ist, die, in einem ersten Kommunikationsnetzwerk übertragenen, Datenpakete zu analysieren und die, die Ausführungsvorschrift erfüllenden, Datenpakete als erste Datenpakete dem ersten Rechner (PAT) zur weiteren Verarbeitung zuzuföhren.

22. Kommunikationssystem nach Anspruch 21,

35 dadurch gekennzeichnet, dass

der zumindest eine Filter-Rechner (IDS) an einem Netzwerkszugangsknoten und/oder einem Aggregationsknoten des ersten Kommunikationsnetzwerks angeordnet ist.

23. Kommunikationssystem nach Anspruch 21 oder 22,
dadurch gekennzeichnet, dass
der zumindest eine Filter-Rechner (IDS) dazu ausgebildet ist,
5 Peer-to-Peer-Datenpakete zu erkennen.

