(19) Deutsches Patent- und Markenamt

(10) **DE 10 2017 125 655 A1** 2018.07.05

(12) # Offenlegungsschrift

(21) Aktenzeichen: **10 2017 125 655.5**
(22) Anmeldetag: **02.11.2017**
(43) Offenlegungstag: **05.07.2018**

(51) Int Cl.: ***G06F 21/00*** (2006.01)

(30) Unionspriorität:
15/344220 **04.11.2016** **US**

(71) Anmelder:
General Motors LLC, Detroit, Mich., US

(74) Vertreter:
Manitz Finsterwald Patentanwälte PartmbB, 80336 München, DE

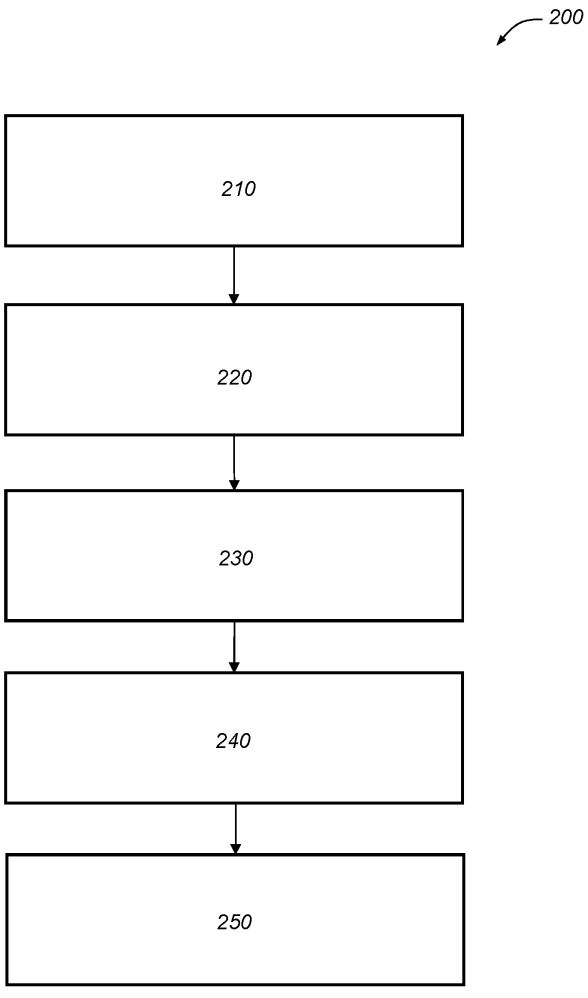
(72) Erfinder:
Ramamurthy, Kannan, Detroit, Mich., US;
Morampudi, Srinivas R., Detroit, Mich., US

Prüfungsantrag gemäß § 44 PatG ist gestellt.

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **ANONYMISIERUNG VON STREAMING-DATEN**

(57) Zusammenfassung: Ein System und Verfahren zur Anonymisierung von Streaming-Datensätzen beinhaltet: das Verarbeiten eines oder mehrerer Datensätze in einem oder mehreren anonymen Vektordarstellungen dieses Datensatzes (der Datensätze); das Zugreifen auf einen verallgemeinerten Vektor, der eine gewünschte Datenanonymisierungsstufe beinhaltet; das Vergleichen der anonymen Vektordarstellungen des Datensatzes/ der Datensätze mit dem verallgemeinerten Vektor; das Ermitteln, ob die anonymen Vektordarstellungen des Datensatzes/ der Datensätze basierend auf dem Vergleich ausreichend anonym sind; das Identifizieren eines Zeitintervalls zum Senden von Datensätzen, die ausreichend anonym sind; und das Erhöhen oder Verringern einer Menge der Rechnerressourcen, die bestimmen, ob die Vektordarstellungen des Datensatzes/ der Datensätze basierend auf einer Menge der Zeitspanne, die im Zeitintervall verbleibt, ausreichend anonym sind.



Beschreibung

TECHNISCHES GEBIET

[0001] Die vorliegende Erfindung betrifft das Verarbeiten von Daten und insbesondere das Anonymisieren von Datensätzen.

HINTERGRUND

[0002] Datensätze können erstellt werden, die jeweils mehrere Datenwerte enthalten. Diese Datensätze können eine Reihe von Phänomenen beschreiben. Die Datensätze können beispielsweise mehrere Datenwerte enthalten, die einer Person zugeordnet sind, die diese Person beschreibt oder mit ihr in Beziehung steht. Das heißt, die Person könnte mit Daten in Verbindung gebracht werden, die angeben, wo diese Person lebt, wie etwa Postleitzahl, Geschlecht oder Alter der Person. Datensätze können alternativ andere Themen beschreiben, wie etwa Fahrzeug- und Fahrzeugfunktionen, die mit diesem Fahrzeug verbunden sind. Da die Datensätze Identifizierungsinformationen beinhalten können, werden sie im Allgemeinen anonymisiert, sodass der Empfänger solcher Informationen bestimmte Datenwerte (beispielsweise Personen oder Fahrzeuge), die in dem Datensatz beinhaltet sind, nicht gesondert identifizieren können.

[0003] Die Anonymisierung kann mithilfe von k-Anonymitätstechniken erfolgen, um Datensätze zu verarbeiten und anonym zu rendern. K-Anonymität wird jedoch im Allgemeinen auf statische Datensätze angewendet. Im Gegensatz dazu werden viele Datensätze nun als Streaming-Daten gesendet, und die Anonymisierung dieser Streaming-Datensätze kann eine Herausforderung darstellen. Bei der Anwendung bestehender k-Anonymitätstechniken auf Streaming-Daten kann es schwierig sein, die optimale Anonymität für die Datensätze zu ermitteln, um sicherzustellen, dass die Daten ausreichend anonym sind und die Datensätze innerhalb eines definierten Zeitbedarfs verarbeiten.

ZUSAMMENFASSUNG

[0004] Gemäß einer Ausführungsform ist ein Verfahren zur Anonymisierung von Streaming-Datensätzen vorgesehen. Das Verfahren beinhaltet das Verarbeiten eines oder mehrerer Datensätze in einem oder mehreren anonymen Vektordarstellungen dieses Datensatzes (der Datensätze); das Zugreifen auf einen verallgemeinerten Vektor, der eine gewünschte Datenanonymisierungsstufe beinhaltet; das Vergleichen der anonymen Vektordarstellungen des Datensatzes/ der Datensätze mit dem verallgemeinerten Vektor; das Ermitteln, ob die anonymen Vektordarstellungen des Datensatzes/ der Datensätze basierend auf dem Vergleich ausreichend anonym sind; das Identifizieren eines Zeitintervalls zum Senden von Datensätzen, die ausreichend anonym sind; und das Erhöhen oder Verringern einer Menge der Rechnerressourcen, die bestimmen, ob die Vektordarstellungen des Datensatzes/ der Datensätze basierend auf einer Menge der Zeitspanne, die im Zeitintervall verbleibt, ausreichend anonym sind.

[0005] Gemäß einer weiteren Ausführungsform ist ein Verfahren zur Anonymisierung von Streaming-Datensätzen vorgesehen. Das Verfahren beinhaltet das Verarbeiten eines oder mehrerer Datensätze in einem oder mehreren anonymen Vektordarstellungen dieses Datensatzes (der Datensätze); das Zugreifen auf einen verallgemeinerten Vektor, der eine gewünschte Datenanonymisierungsstufe beinhaltet; das Vergleichen der anonymen Vektordarstellungen des Datensatzes/ der Datensätze mit dem verallgemeinerten Vektor; das Ermitteln, ob die anonymen Vektordarstellungen des Datensatzes/ der Datensätze basierend auf dem Vergleich ausreichend anonym sind; das Berechnen einer Rate, mit der die eingehenden Datensätze empfangen werden; das Erhöhen oder Verringern einer Menge der Rechnerressourcen, die bestimmen, ob die Vektordarstellungen des Datensatzes/ der Datensätze basierend auf der Rate, mit der Datensätze empfangen werden, ausreichend anonym sind; und das Übertragen von Datensätzen, die ausreichend anonymisiert sind, bevor ein Zeitintervall an eine dritte Partei abläuft.

Figurenliste

[0006] Eine oder mehrere Ausführungsformen der Erfindung werden im Folgenden in Verbindung mit den beigefügten Zeichnungen beschrieben, worin gleiche Bezeichnungen gleiche Elemente bezeichnen, und worin:

Fig. 1 ein Blockdiagramm ist, das eine Ausführungsform eines Kommunikationssystems darstellt, das fähig ist, das hierin offenbarte Verfahren zu verwenden; und

Fig. 2 ein Flussdiagramm ist, das eine Ausführungsform eines Verfahrens zur Anonymisierung von Streaming-Datensätzen darstellt.

AUSFÜHRLICHE BESCHREIBUNG DER VERANSCHAULICHTEN AUSFÜHRUNGSFORM(EN)

[0007] Das unten beschriebene System und Verfahren verwendet K-Anonymitätstechniken, um Streaming-Datensätze zu anonymisieren und sicherzustellen, dass eine ausreichende Anonymität in Bezug auf eine optimierte Anonymitätsstufe besteht. Die optimierte Anonymitätsstufe kann mithilfe eines verallgemeinerten Vektors festgelegt werden. Wenn Streaming-Datensätze anonymisiert wurden und durch Vektoren dargestellt werden, können diese Vektoren mit dem verallgemeinerten Vektor verglichen werden, um eine berechnete Entfernung zu bestimmen, der zwischen ihnen existiert. Wenn die berechnete Entfernung innerhalb eines akzeptablen Wertebereichs liegt, kann der Datensatz ausreichend anonymisiert sein. Der Bereich kann eine obere und eine untere Grenze aufweisen, um nicht nur sicherzustellen, dass der Datensatz ausreichend anonymisiert ist, sondern auch, dass das System nicht zu viele Daten entfernt hat, wodurch der Datensatz weniger aussagekräftig werden kann, als er sein könnte. Der Bereich kann durch die Anzahl bestimmt werden, wie oft ein Schlüssel aus dem Datensatz, wie etwa eine Fahrzeugkennung, in einem bestimmten Datensatz wiederholt wird. Wenn der gesamte Schlüssel bereitgestellt wird, besteht keine Anonymität. Umgekehrt, wenn der Schlüssel redigiert oder verallgemeinert ist, dann wird der Schlüssel anonym. Auf diese Weise kann der verallgemeinerte Vektor gesteuert werden, um eine optimale Anonymitätsstufe festzulegen, die die Geheimhaltung aufrechterhält, aber auch sicherstellt, dass die Datensätze so viele Informationen wie möglich beinhalten, ohne diese Geheimhaltung zu beeinträchtigen.

[0008] Das System überwacht auch die verbleibende Zeitdauer in einem Zeitintervall, die zur Verarbeitung einer definierten Menge an anonymen Datensätzen vorgesehen ist, und erhöht oder verringert die Menge an Computerressourcen, die für die Verarbeitung der Datensätze bestimmt sind, basierend auf der Menge an Zeit und der Verarbeitung von Datenmengen. Bestimmte Softwareanwendungen können eine Anforderung beinhalten, dass eine bestimmte Anzahl von anonymisierten Datensätzen für jeden definierten Zeitraum erstellt und/oder übertragen werden. Während dieses Zeitraums kann die Menge an anonymisierten Datensätzen, die von einer Fernvorrichtung empfangen werden, über einen Zeitraum hinweg schwanken und die Menge an Computerressourcen kann in Abhängigkeit von der Rate variiert werden, mit der Datenmengen empfangen werden. Ausführungsformen des offenbarten Verfahrens bzw. der offenbarten Verfahren werden nachstehend in Bezug auf eine Fernvorrichtung beschrieben, die drahtlos Datensätze von mehreren Fahrzeugen empfängt und an Dritte anonymisierte Daten liefert. Es sollte jedoch beachtet werden, dass das oder die Verfahren mithilfe anderer Systeme als das Fahrzeug/ die Fernvorrichtung implementiert werden könnten. Die Verfahren können beispielsweise alternativ mithilfe von mehreren Computern implementiert werden, die Datenpakete über das Internet übertragen.

Kommunikationssystem -

[0009] Mit Bezug auf **Fig. 1** ist eine Betriebsumgebung dargestellt, die ein mobiles Fahrzeugkommunikationssystem **10** umfasst, das verwendet werden kann, um das hierin offenbarte Verfahren zu implementieren. Das Kommunikationssystem **10** umfasst im Allgemeinen ein Fahrzeug **12**, ein oder mehrere Drahtlosträgersysteme **14**, ein Festnetz **16**, einen Computer **18** und ein Call-Center **20**. Es versteht sich, dass das offenbarte Verfahren mit einer beliebigen Anzahl von unterschiedlichen Systemen verwendet werden kann und nicht speziell auf die hierin gezeigte Betriebsumgebung einschränkt ist. Auch die Architektur, Konstruktion, Konfiguration und der Betrieb des Systems **10** und seiner einzelnen Komponenten sind in der Technik allgemein bekannt. Somit stellen die folgenden Absätze lediglich einen kurzen Überblick über ein solches Kommunikationssystem **10** bereit; aber auch andere, hierin nicht dargestellte Systeme könnten die offenbarten Verfahren einsetzen.

[0010] Fahrzeug **12** ist in der veranschaulichten Ausführungsform als ein Personenkraftwagen dargestellt, es sollte jedoch beachtet werden, dass jedes andere Fahrzeug, einschließlich Motorräder, Lastwagen, Geländewagen (SUV), Campingfahrzeuge (RV), Wasserfahrzeuge, Flugzeuge usw. ebenfalls verwendet werden kann. Ein Teil der Fahrzeugelektronik **28** wird im Allgemeinen in **Fig. 1** gezeigt und beinhaltet eine Telematikeinheit **30**, ein Mikrofon **32**, eine oder mehrere Tasten oder andere Steuereingänge 34, ein Audiosystem **36**, eine optische Anzeige **38**, ein GPS-Modul **40** sowie eine Anzahl von Fahrzeugsystemmodulen (VSMs) **42**. Einige dieser Vorrichtungen können direkt mit der Telematikeinheit, wie z. B. dem Mikrofon **32** und der/den Taste(n) 34, verbunden sein, während andere indirekt unter Verwendung einer oder mehrerer Netzwerkverbindungen, wie einem Kommunikationsbus **44** oder einem Entertainmentbus **46**, verbunden sind. Beispiele geeigneter

Netzwerkverbindungen beinhalten ein Controller Area Network (CAN), einen medienorientierten Systemtransfer (MOST), ein lokales Kopplungsstrukturnetzwerk (LIN), ein lokales Netzwerk (LAN) und andere geeignete Verbindungen, wie z. B. Ethernet oder andere, die u. a. den bekannten ISO-, SAE- und IEEE-Standards und -Spezifikationen entsprechen.

[0011] Die Telematikeinheit **30** selbst ist ein Fahrzeugsystemmodul (VSM) und kann eine OEM-installierte (eingebettete) oder eine Aftermarketvorrichtung sein, die in dem Fahrzeug installiert ist und eine drahtlose Sprach- und/oder Datenkommunikation über das drahtlose Trägersystem **14** und über eine drahtlose Vernetzung ermöglicht. Dies ermöglicht, dass das Fahrzeug mit Call-Center **20**, anderen telematikfähigen Fahrzeugen oder einer anderen Entität oder Vorrichtung kommunizieren kann. Die Telematikeinheit verwendet vorzugsweise Funkübertragungen, um einen Kommunikationskanal (einen Sprachkanal und/oder einen Datenkanal) mit dem Drahtlosträgersystem **14** herzustellen, sodass Sprach- und/oder Datenübertragungen über den Kanal gesendet und erhalten werden können. Durch Bereitstellen von sowohl Sprach- als auch Datenkommunikation ermöglicht die Telematikeinheit **30** dem Fahrzeug das Anbieten einer Anzahl von unterschiedlichen Diensten, darunter solche, die sich mit Navigation, Fernsprechen, Nothilfe, Diagnose, Infotainment usw. befassen. Daten können entweder über eine Datenverbindung, z.B. die Paketdatenübertragung über einen Datenkanal oder über einen Sprachkanal unter Verwendung von auf dem Fachgebiet bekannten Techniken gesendet werden. Für kombinierte Dienste, die sowohl Sprachkommunikation (z. B. mit einem Live-Berater oder einer Sprachdialogeinheit im Call-Center **20**) als auch Datenkommunikation (z. B. für die Bereitstellung von GPS-Ortsdaten oder Fahrzeugdiagnosedaten im Call-Center **20**) einschließen, kann das System einen einzelnen Anruf über einen Sprachkanal verwenden und nach Bedarf zwischen Sprach- und Datenübertragung über den Sprachkanal umschalten, und dies kann unter Verwendung von Techniken erfolgen, die dem Fachmann bekannt sind.

[0012] Gemäß einer Ausführungsform verwendet die Telematikeinheit **30** Mobilfunkkommunikation gemäß entweder den GSM-, CDMA- oder LTE-Standards und beinhaltet daher einen Mobilfunkstandardchipsatz **50** für die Sprachkommunikation, wie Freisprechen, ein drahtloses Modem für die Datenübertragung, ein elektronisches Verarbeitungsgerät **52**, eine oder mehrere Digitalspeichervorrichtungen **54** und eine Dual-Antenne **56**. Es versteht sich, dass das Modem entweder durch Software implementiert sein kann, die in der Telematikeinheit gespeichert und durch den Prozessor **52** ausgeführt wird, oder es kann eine separate Hardwarekomponente sein, die sich innerhalb oder außerhalb der Telematikeinheit **30** befinden kann. Das Modem kann mithilfe einer beliebigen Anzahl unterschiedlicher Standards oder Protokolle, wie z. B. LTE, EVDO, CDMA, GPRS und EDGE, betrieben werden. Die drahtlose Vernetzung zwischen dem Fahrzeug und den anderen vernetzten Vorrichtungen kann auch unter Verwendung der Telematikeinheit **30** erfolgen. Für diesen Zweck kann die Telematikeinheit **30** konfiguriert sein, gemäß einem oder mehreren Protokollen drahtlos zu kommunizieren einschließlich drahtloser Nahbereichskommunikation (SRWC), wie irgendwelche von den IEEE 802.11-Protokollen, WiMAX, ZigBee™, Wi-Fi direct, Bluetooth™ oder Nahfeldkommunikation (NFC). Wenn die Telematikeinheit für paketvermittelte Datenkommunikation, wie z. B. TCP/IP verwendet wird, kann sie mit einer statischen IP-Adresse konfiguriert oder eingerichtet werden, um eine zugewiesene IP-Adresse von einer anderen Vorrichtung im Netzwerk, wie z. B. einem Router oder einem Netzwerkadressenserver, automatisch zu empfangen.

[0013] Der Prozessor **52** kann jede Geräteart sein, die fähig ist elektronische Befehle zu verarbeiten, einschließlich Mikroprozessoren, Mikrocontrollern, Hostprozessoren, Steuerungen, Fahrzeugkommunikationsprozessoren und anwendungsspezifische integrierte Schaltungen (ASICs). Er kann ein speziell dafür vorgesehener Prozessor sein, der nur für die Telematikeinheit **30** verwendet wird, oder er kann mit anderen Fahrzeugsystemen geteilt werden. Der Prozessor **52** führt verschiedene Arten von digital gespeicherten Befehlen aus, wie Software oder Firmwareprogramme, die im Speicher **54** gespeichert sind, welche der Telematikeinheit ermöglichen, eine große Vielfalt von Diensten bereitzustellen. Zum Beispiel kann der Prozessor **52** Programme ausführen oder Daten verarbeiten, um mindestens einen Teil des Verfahrens auszuführen, das hierin beschrieben ist.

[0014] Die Telematikeinheit **30** kann verwendet werden, um eine vielfältige Palette von Fahrzeugdiensten bereitzustellen, die drahtlose Kommunikation zu und/oder vom Fahrzeug beinhalten. Derartige Dienste beinhalten: Wegbeschreibungen und andere navigationsbezogene Dienste, die in Verbindung mit dem GPS-basierten Fahrzeugnavigationsmodul **40** bereitgestellt sind; Benachrichtigung über die Airbagauslösung und andere mit Notruf oder Pannendienst verbundene Dienste, die in Verbindung mit einem oder mehreren Crashsensor-Schnittstellenmodulen, wie einem Fahrzeugbeherrschbarkeitsmodul (nicht gezeigt), bereitgestellt sind; Diagnosemeldungen unter Verwendung von einem oder mehreren Diagnosemodulen; und mit Infotainment verbundene Dienste, wobei Musik, Internetseiten, Filme, Fernsehprogramme, Videospiele und/oder andere Informationen durch ein Infotainmentmodul (nicht gezeigt) heruntergeladen und für die gegenwärtige oder spätere

Wiedergabe gespeichert werden. Die vorstehend aufgelisteten Dienste sind keineswegs eine vollständige Liste aller Fähigkeiten der Telematikeinheit **30**, sondern sie sind einfach eine Aufzählung von einigen der Dienste, welche die Telematikeinheit anbieten kann. Des Weiteren versteht es sich, dass mindestens einige der vorstehend genannten Module in der Form von Softwarebefehlen implementiert sein könnten, die innerhalb oder außerhalb der Telematikeinheit **30** gespeichert sind, sie könnten Hardwarekomponenten sein, die sich innerhalb oder außerhalb der Telematikeinheit 30 befinden, oder sie könnten integriert sein und/oder miteinander oder mit anderen Systemen geteilt zu sein, die sich im Fahrzeug befinden, um nur einige Möglichkeiten zu nennen. Für den Fall, dass die Module als VSM **42** implementiert sind, die sich außerhalb der Telematikeinheit **30** befinden, könnten sie den Fahrzeugbus **44** verwenden, um Daten und Befehle mit der Telematikeinheit auszutauschen.

[0015] Das GPS-Modul **40** empfängt Funksignale von einer Konstellation **60** von GPS-Satelliten. Von diesen Signalen kann das Modul **40** die Fahrzeugposition ermitteln, die verwendet wird, um Navigation und andere mit der Position verbundene Dienste an den Fahrzeugführer bereitzustellen. Navigationsinformationen können auf der Anzeige **38** (oder einer anderen Anzeige innerhalb des Fahrzeugs) dargestellt oder in verbaler Form präsentiert werden, wie es beispielsweise bei der Wegbeschreibungsnavigation der Fall ist. Die Navigationsdienste können unter Verwendung von einem zugehörigen Fahrzeugnavigationsmodul (das Teil des GPS-Moduls **40** sein kann) bereitgestellt werden, oder einige oder alle Navigationsdienste können über die Telematikeinheit **30** erfolgen, wobei die Positionsinformationen zum Zweck des Ausstattens des Fahrzeugs mit Navigationskarten, Kartenanmerkungen (Sehenswürdigkeiten, Restaurants usw.), Routenberechnungen und dergleichen zu einem entfernten Standort gesendet werden. Die Positionsinformationen können an das Call-Center **20** oder ein anderes Remotecomputersystem, wie Computer **18**, für andere Zwecke, wie Flottenmanagement, bereitgestellt werden. Außerdem können neue oder aktualisierte Kartendaten zum GPS-Modul **40** vom Call-Center **20** über die Telematikeinheit **30** heruntergeladen werden.

[0016] Abgesehen von der Telematikeinheit **30**, dem Audiosystem **36** und dem GPS-Modul 40 kann das Fahrzeug **12** andere Fahrzeugsystemmodule (VSMs) **42** in Form von elektronischen Hardwarekomponenten beinhalten, die sich im Fahrzeug befinden und typischerweise eine Eingabe von einem oder mehreren Sensoren erhalten und die erfassten Eingaben verwenden, um die Diagnose, Überwachung, Steuerung, Berichterstattung und/oder andere Funktionen auszuführen. Jedes der VSMs **42** ist vorzugsweise durch den Kommunikationsbus **44** mit den anderen VSM sowie der Telematikeinheit **30** verbunden und kann programmiert werden, Fahrzeugsystem- und Subsystemdiagnosetests auszuführen. So kann beispielsweise ein VSM **42** ein Motorsteuergerät (ECM) sein, das verschiedene Aspekte des Motorbetriebs, wie z. B. Kraftstoffzündung und Zündzeitpunkt, steuert, ein weiteres VSM **42** kann ein Antriebsstrangsteuermodul sein, das den Betrieb von einer oder mehreren Komponenten des Fahrzeugantriebsstrangs reguliert, und ein weiteres VSM **42** kann ein Chassis-Steuermodul sein, das verschiedene im Fahrzeug befindliche elektrische Komponente, wie beispielsweise die Zentralverriegelung des Fahrzeugs und die Scheinwerfer, verwaltet. Gemäß einer Ausführungsform ist das Motorsteuergerät mit integrierten Diagnose (OBD)-Funktionen ausgestattet, die unzählige Echtzeitdaten, wie z. B. die von verschiedenen Sensoren, einschließlich Fahrzeugemissionssensoren, erhaltenen Daten bereitstellen und eine standardisierte Reihe von Diagnosefehlercodes (DTCs) liefern, die einem Techniker ermöglichen, Fehlfunktionen innerhalb des Fahrzeugs schnell zu identifizieren und zu beheben. Fachleute auf dem Fachgebiet werden erkennen, dass es sich bei den vorgenannten VSMs nur um Beispiele von einigen der Module handelt, die im Fahrzeug **12** verwendet werden können, zahlreiche andere Module jedoch ebenfalls möglich sind.

[0017] Die Fahrzeugelektronik **28** beinhaltet auch eine Anzahl von Fahrzeugbenutzeroberflächen, die Fahrzeuginsassen mit einem Mittel zum Bereitstellen und/oder Empfangen von Informationen ausstattet, einschließlich Mikrofon **32**, Taste(n) 34, Audiosystem **36**, und optischer Anzeige **38**. Wie hierin verwendet, beinhaltet der Begriff „Fahrzeugbenutzeroberfläche“ weitgehend jede geeignete Form von elektronischer Vorrichtung, die sowohl die im Fahrzeug befindlichen Hardware- als auch Softwarekomponenten beinhaltet und einem Fahrzeugbenutzer ermöglicht, mit einer oder durch eine Komponente des Fahrzeugs zu kommunizieren. Das Mikrofon **32** stellt eine Audioeingabe an die Telematikeinheit bereit, um dem Fahrer oder anderen Insassen zu ermöglichen, Sprachsteuerungen bereitzustellen und Freisprechen über das Drahtlosträgersystem **14** auszuführen. Für diesen Zweck kann es mit einer integrierten automatischen Sprachverarbeitungseinheit verbunden sein, welche die unter Fachleuten auf dem Gebiet bekannte Mensch-Maschinen-Schnittstellen (HMI)-Technologie verwendet. Die Taste(n) 34 ermöglichen eine manuelle Benutzereingabe in die Telematikeinheit **30**, um drahtlose Telefonanrufe zu initiieren und andere Daten, Antworten oder eine Steuereingabe bereitzustellen. Separate Tasten können zum Initiieren von Notrufen gegenüber regulären Dienstunterstützungsanrufen beim Call-Center **20** verwendet werden. Das Audiosystem **36** stellt eine Audioausgabe an einen Fahrzeuginsassen bereit und kann ein zugehöriges selbstständiges System oder Teil des primären Fahrzeugaudiosystems sein.

Gemäß der bestimmten Ausführungsform, die hierin gezeigt ist, ist das Audiosystem **36** operativ sowohl mit dem Fahrzeugbus **44** als auch mit dem Entertainmentbus **46** gekoppelt und kann AM-, FM- und Satellitenradio, CD-, DVD- und andere Multimediafunktionalität bereitstellen. Diese Funktionalität kann in Verbindung mit dem vorstehend beschriebenen Infotainmentmodul oder davon unabhängig bereitgestellt werden. Die optische Anzeige **38** ist vorzugsweise eine Grafikanzeige, wie z. B. ein Touchscreen am Armaturenbrett oder eine Warnanzeige, die von der Frontscheibe reflektiert wird, und verwendet werden kann, um eine Vielzahl von Eingabe- und Ausgabefunktionen bereitzustellen. Verschiedene andere Fahrzeugbenutzeroberflächen können ebenfalls verwendet werden, denn die Schnittstellen von **Fig. 1** dienen lediglich als Beispiel für eine bestimmte Implementierung.

[0018] Das Drahtlosträgersystem **14** ist vorzugsweise ein Smartphonesystem, das eine Vielzahl von Mobilfunkmasten **70** (nur einer gezeigt), eine oder mehrere mobile Vermittlungszentrale Einrichtung (MSC) **72** sowie irgendwelche anderen Netzwerkkomponenten beinhaltet, die erforderlich sind, um das Drahtlosträgersystem **14** mit dem Festnetz **16** zu verbinden. Jeder Mobilfunkturm **70** beinhaltet Sende- und Empfangsantennen und eine Basisstation, wobei die Basisstationen von unterschiedlichen Mobilfunktürmen mit der MSC **72** entweder direkt oder über zwischengeschaltete Geräte, wie z. B. eine Basisstationssteuereinheit, verbunden sind. Das Mobilfunksystem **14** kann jede geeignete Kommunikationstechnik implementieren, die beispielsweise, analoge Technologien, wie AMPS oder die neueren Digitaltechnologien, wie CDMA (z. B. CDMA2000) oder GSM/GPRS, beinhaltet. Der Fachmann wird erkennen, dass verschiedene Zellenturm-/Basisstation/MSC-Anordnungen möglich sind und mit dem drahtlosen System 14 verwendet werden könnten. Zum Beispiel könnten sich Basisstation und Zellentürme an derselben Stelle oder entfernt voneinander befinden, jede Basisstation könnte für einen einzelnen Zellenturm zuständig sein oder eine einzelne Basisstation könnte verschiedene Zellentürme bedienen und verschiedene Basisstationen könnten mit einer einzigen MSC gekoppelt werden, um nur einige der möglichen Anordnungen zu nennen.

[0019] Abgesehen vom Verwenden des Drahtlosträgersystems **14** kann ein unterschiedliches Drahtlosträgersystem in der Form von Satellitenkommunikation verwendet werden, um unidirektionale oder bidirektionale Kommunikation mit dem Fahrzeug bereitzustellen. Dies kann unter Verwendung von einem oder mehreren Fernmeldesatelliten **62** und einer aufwärtsgerichteten Sendestation **64** erfolgen. Bei der unidirektionalen Kommunikation kann es sich beispielsweise um Satellitenradiodienste handeln, worin programmierte Inhaltsdaten (Nachrichten, Musik usw.) von der Sendestation **64** erhalten werden, für das Hochladen gepackt und anschließend zum Satelliten **62** gesendet werden, der die Programmierung an die Teilnehmer sendet. Bidirektionale Kommunikation kann beispielsweise Satellitentelefoniedienste unter Verwendung der Satelliten **62** sein, um Telefonkommunikationen zwischen dem Fahrzeug **12** und der Station **64** weiterzugeben. Bei Verwendung kann dieses Satellitenfernsprechen entweder zusätzlich zum oder anstatt des Drahtlosträgersystems **14** verwendet werden.

[0020] Das Festnetz **16** kann ein konventionelles landgebundenes Telekommunikationsnetzwerk sein, das mit einem oder mehreren Festnetztelefonen verbunden ist und das Drahtlosträgersystem **14** mit dem Call-Center **20** verbindet. So kann beispielsweise das Festnetz **16** ein Fernsprechnetz (PSTN) wie jenes sein, das verwendet wird, um festverdrahtetes Fernsprechen, paketvermittelte Datenkommunikationen und die Internetinfrastruktur bereitzustellen. Ein oder mehrere Segmente des Festnetzes **16** könnten durch Verwenden eines normalen drahtgebundenen Netzwerks, eines Lichtleiter- oder eines anderen optischen Netzwerks, eines Kabelnetzes, von Stromleitungen, anderen drahtlosen Netzwerken, wie drahtlose lokale Netzwerke (WLANs) oder Netzwerke, die drahtlosen Breitbandzugang (BWA) bereitstellen oder jeder Kombination davon, implementiert sein. Des Weiteren muss das Call-Center **20** nicht über das Festnetz **16** verbunden sein, sondern könnte Funktelefonieausrüstung beinhalten, sodass direkt mit einem drahtlosen Netzwerk, wie dem Drahtlosträgersystem **14**, kommuniziert werden kann.

[0021] Der Computer **18** kann einer von einer Anzahl von Computern sein, die über ein privates oder öffentliches Netzwerk, wie das Internet, zugänglich sind. Jeder dieser Computer 18 kann für einen oder mehrere Zwecke, wie einen Webserver verwendet werden, der vom Fahrzeug über die Telematikeinheit **30** und das Drahtlosträgersystem **14** zugänglich ist. Andere derartige zugängliche Computer **18** können beispielsweise sein: ein Kundendienstzentrumcomputer, wo Diagnoseinformationen und andere Fahrzeugdaten vom Fahrzeug über die Telematikeinheit **30** hochgeladen werden können; ein Clientcomputer, der vom Fahrzeugbesitzer oder einem anderen Teilnehmer für solche Zwecke, wie das Zugreifen auf oder das Erhalten von Fahrzeugdaten oder zum Einstellen oder Konfigurieren von Teilnehmerpräferenzen oder Steuern von Fahrzeugfunktionen, verwendet wird; oder ein Drittparteispeicherort, zu dem oder von dem Fahrzeugdaten oder andere Informationen entweder durch Kommunizieren mit dem Fahrzeug **12** oder dem Call-Center **20** oder beiden bereitgestellt werden. Ein Computer **18** kann auch für das Bereitstellen von Internetkonnektivität, wie DNS-

Dienste oder als ein Netzwerkadressenserver, verwendet werden, der DHCP oder ein anderes geeignetes Protokoll verwendet, um dem Fahrzeug **12** eine IP-Adresse zuzuweisen. Der Computer **18** beinhaltet Rechenressourcen in der Form von einem oder mehreren Computerprozessoren, die die Verarbeitungslast, die sie als Reaktion auf Softwareanweisungen tragen, selektiv erhöhen oder verringern können. Der Computer **18** kann auch einen Arbeitsspeicher (RAM) beinhalten, der einen Datenpuffer oder Zwischenspeicher zum Speichern von Datensätzen hostet, bevor er sie an eine dritte Partei sendet.

[0022] Das Call-Center **20** ist konzipiert, die Fahrzeugelektronik **28** mit einer Anzahl von unterschiedlichen System-Back-End-Funktionen bereitzustellen, und beinhaltet nach dem hierin gezeigten Ausführungsbeispiel im Allgemeinen einen oder mehrere Switches **80**, Server **82**, Datenbanken **84**, Live-Berater **86** sowie ein automatisiertes Sprachausgabesystem (VRS) **88**, die alle auf dem Fachgebiet bekannt sind. Diese verschiedenen Komponenten des Call-Centers sind vorzugsweise miteinander über ein verdrahtetes oder drahtloses lokales Netzwerk **90** gekoppelt. Der Switch **80**, der ein Nebenstellenanlagen (PBX)-Switch sein kann, leitet eingehende Signale weiter, sodass Sprachübertragungen gewöhnlich entweder zum Live-Berater **86** über das reguläre Telefon oder automatisiert zum Sprachdialogsystem 88 unter Verwendung von VoIP gesendet werden. Das Live-Berater-Telefon kann auch VoIP verwenden, wie durch die gestrichelte Linie in **Fig. 1** angezeigt. VoIP und andere Datenkommunikation durch den Switch **80** werden über ein Modem (nicht gezeigt) implementiert, das zwischen dem Switch **80** und Netzwerk **90** verbunden ist. Datenübertragungen werden über das Modem an den Server **82** und/oder die Datenbank **84** weitergegeben. Die Datenbank **84** kann Kontoinformationen, wie Teilnehmerauthentisierungsinformationen, Fahrzeugbezeichner, Profilaufzeichnungen, Verhaltensmuster und andere entsprechende Teilnehmerinformationen, speichern. Datenübertragungen können zudem durch drahtlose Systeme, wie z. B. 802.11x, GPRS und dergleichen, erfolgen. Obwohl die veranschaulichte Ausführungsform beschrieben wurde, als ob sie in Verbindung mit einem bemannten Call-Center **20** verwendet werden würde, das den Live-Berater **86** einsetzt, ist es offensichtlich, dass das Call-Center stattdessen VRS **88** als einen automatisierten Berater verwenden kann, oder eine Kombination von VRS **88** und dem Live-Berater **86** verwendet werden kann.

Verfahren -

[0023] Unter Bezugnahme auf **Fig. 2** ist nun eine Ausführungsform eines Verfahrens (**200**) zur Anonymisierung von Streaming-Datensätzen dargestellt. Das Verfahren **200** beginnt im Schritt **210** durch das Verarbeiten eines oder mehrerer Datensätze in einem oder mehreren anonymen Vektordarstellungen dieser Datensätze. In dieser Implementierung kann das Fahrzeug **12** Datensätze erzeugen und die Datensätze drahtlos an eine Fernvorrichtung, wie etwa den Computer **18** oder das Call-Center **20**, übertragen, wo die Datensätze verarbeitet werden. Datensätze umfassen mehrere Dateneinträge, die personenbezogene Daten (PID - Personally Identifying Data) enthalten. Die Datensätze können einen oder mehrere Kennungs- oder Quasi-Kennungs-Dateneinträge und einen oder mehrere Inhalts-Dateneinträge beinhalten. Diese Informationen können in Tabellen- oder Matrixform organisiert werden. In einem Beispiel kann jeder Datensatz eine Fahrzeug-Identifizierungsnummer (VIN) des Fahrzeugs **12** als Kennungs-Dateneintrag sowie Kilometerzähler- und Ölstandswerte für Inhalts-Dateneinträge, die dem Fahrzeug **12** zugeordnet sind, beinhalten. Zur Veranschaulichung dieser Datensätze sind drei beispielhafte Datensätze in der folgenden Tabelle dargestellt:

VIN	KILOMETERZÄHLER	ÖLWECHSELINTERVALL
2GABCDEFGHJKLM234	1000	10%
2GABCDEFGHJKLM234	1500	45 %
2GABCDEFGHJKLM234	2000	22%

[0024] Die Fernvorrichtung kann große Mengen von Datensätzen empfangen, die von vielen Fahrzeugen erstellt wurden. Einzelne Fahrzeuge, wie das Fahrzeug **12**, können jeweils mehrere Datensätze erzeugen und an die zentrale Einrichtung senden. Die zentrale Einrichtung kann diese Datensätze vom Fahrzeug **12** gleichzeitig empfangen, während die zentrale Einrichtung Datensätze von einer großen Anzahl von anderen Fahrzeugen empfängt.

[0025] Die Datensätze können mit den k-Anonymitätstechniken der Unterdrückung und Verallgemeinerung anonymisiert werden. Die Unterdrückung beinhaltet das Ersetzen von Kennungs-Dateneinträgen durch Sternchen oder Leerzeichen, während die Verallgemeinerung ausgeführt werden kann, indem bestimmte Dateneinträge durch breitere Bereiche oder Kategorien ersetzt werden. Beide dieser Techniken können durch eine

zweite Tabelle, die unten gezeigt ist, betrachtet werden, die die Kennungs-Dateneinträge unterdrückt und die Inhalts-Dateneinträge verallgemeinert.

VIN	KILOMETERZÄHLER	ÖLWECHSELINTERVALL
*	1000-1500	10-20 %
*	1000-1500	40-50 %
*	1500-2000	20-30 %

[0026] Die in der zweiten Tabelle dargestellten Dateneinträge, die verallgemeinert wurden, können jeweils durch einen Punkt im Raum mit Koordinaten dargestellt werden. Es kann ein Vektor erstellt werden, der alle Punkte beinhaltet, die Dateneinträge in einer Tabelle darstellen. Dieser anonymisierte Vektor, der den Datensatz darstellt, kann die Anonymitätsstufe des bestimmten Datensatzes angeben. Das Verfahren **200** fährt dann mit Schritt **220** fort.

[0027] Bei Schritt **220** wird auf einen verallgemeinerten Vektor zugegriffen, der eine gewünschte Datenanonymisierungsstufe beinhaltet. Der verallgemeinerte Vektor kann basierend auf einer Benutzereingabe definiert werden, die eine ideale oder gewünschte Anonymisierungsstufe festlegt. Diese gewünschte Anonymisierungsstufe repräsentiert, welcher Teil des Datensatzes, den der Benutzer als geeignet betrachtet, veröffentlicht wird. Es ist möglich, den verallgemeinerten Vektor als einen Startpunktvektor zu betrachten, sodass die anonymisierten Vektordarstellungen von Datensätzen, die denselben Vektor wie der verallgemeinerte Vektor aufweisen, keinen Informationsverlust bezüglich der gewünschten Datenanonymisierungsstufe aufweisen würden. Es sollte beachtet werden, dass eine Anonymisierung von Daten entfernt von Fahrzeug **12** stattfindet. Das Verfahren **200** fährt dann mit Schritt **230** fort.

[0028] In Schritt **230** werden die anonymen Vektordarstellungen des Datensatzes/ der Datensätze mit dem verallgemeinerten Vektor verglichen und es wird basierend auf dem Vergleich bestimmt, ob die anonymen Vektordarstellungen der Datensätze ausreichend anonym sind. Da Datensätze in der zentralen Einrichtung empfangen werden, können die Datensätze mithilfe von k-Anonymitätstechniken anonymisiert und in Form von Matrizen in Vektoren umgewandelt werden. Eine lineare Transformation kann dann an den Matrizen bezüglich des verallgemeinerten Vektors durchgeführt werden. Die zentrale Einrichtung kann dann den Kosinus des Winkels zwischen dem anonymen Vektor und dem verallgemeinerten Vektor bestimmen. Wenn der Kosinus des Winkels kleiner als ein definierter Schwellenwert ist, kann die zentrale Einrichtung bestimmen, dass der anonymisierte Datensatz ausreichend anonym ist. In einer Implementierung beträgt der Schwellenwert 0,3, was anzeigt, dass ein Datenverlust von 30 % akzeptabel ist. Wenn jedoch der Kosinus des Winkels zwischen dem anonymen Vektor und dem verallgemeinerten Vektor größer als der definierte Schwellwert ist, kann der Datensatz als nicht ausreichend anonym oder zu anonym abgewiesen werden, indem eine größere Menge der Dateneinträge im Datensatz unterdrückt oder verallgemeinert wurde. Die zentrale Einrichtung kann die anonymen Vektordarstellungen modifizieren und dann den modifizierten anonymen Vektor mit dem verallgemeinerten Vektor vergleichen, um sicherzustellen, dass der Kosinus des Winkels kleiner als der definierte Schwellenwert ist. Da anonymisierte Datensätze in der zentralen Einrichtung erzeugt werden oder bereits vom Fahrzeug **12** anonymisiert empfangen wurden, kann die zentrale Einrichtung weiterhin eine lineare Transformation dieser neuen anonymisierten Datensätze in Bezug auf den verallgemeinerten Vektor durchführen. Anonymisierte Datensätze, die als ausreichend anonym identifiziert wurden, können von der zentralen Einrichtung in einem Zwischenspeicher oder Datencache gespeichert werden, bis sich genügend Datensätze zum Senden an oder zum Zugreifen durch eine dritte Partei angesammelt haben. Derzeit besteht eine Anzahl an Parteien von Dritten, die die Datensätze empfangen können. Ein Beispiel für eine solche dritte Partei ist Telogis™. Das Verfahren **200** fährt dann mit Schritt **240** fort.

[0029] Bei Schritt **240** wird ein Zeitintervall zum Senden von Datensätzen bestimmt, die ausreichend anonym sind. Die zentrale Einrichtung kann angewiesen werden, während einem definierten Zeitintervall eine definierte Anzahl von Datensätzen an eine dritte Partei zu senden. Die zentrale Einrichtung kann beispielsweise angewiesen werden, 1000 Datensätze pro Sekunde zu anonymisieren. Zu Beginn eines Zeitintervalls von einer Sekunde kann die zentrale Einrichtung beauftragt werden, eingehende Datensätze zu anonymisieren, diese anonymisierten Datensätze mit dem verallgemeinerten Vektor zu vergleichen, die ausreichend anonymisierten Datensätze zu identifizieren und mindestens 1000 der ausreichend anonymisierten Datensätze im Speicher zur Übertragung an die dritte Partei zu speichern. Es sollte beachtet werden, dass das Zeitintervall, die definierte Anzahl der Datensätze oder beide vom vorgesehenen Beispiel abweichen können. Das Verfahren **200** fährt dann mit Schritt **250** fort

[0030] Bei Schritt **250** wird eine Menge von Rechenressourcen, die bestimmen, ob die Vektordarstellungen des Datensatzes/ der Datensätze ausreichend anonym sind, basierend auf einer Zeitdauer, die im Zeitintervall verbleibt, erhöht oder verringert. Die zentrale Einrichtung kann die Menge an Datensätzen überwachen, die sie pro Zeiteinheit empfängt, und die Menge an Rechenressourcen basierend auf dieser Rate steuern. Wenn die eingehenden Datensätze in einem Beispiel 21 Kilobytes (KB) groß sind und während einer ersten Stunde **30 000** Fahrzeuge jeweils einen Datensatz senden, wird die zentrale Einrichtung **630** Gigabyte (GB) an Daten empfangen. Mit der Zeit kann die Anzahl der Fahrzeuge, die Datensätze senden, variieren. Die zentrale Einrichtung kann beispielsweise in den nächsten sieben Stunden Datensätze von einer zunehmenden Anzahl von Fahrzeugen empfangen. Dieses Beispiel wird im Folgenden in einer Tabelle gezeigt, die die Größe der Datensätze, die Anzahl der Fahrzeuge, die Datensätze senden und die Gesamtmenge der empfangenen Daten (in KB) pro Stunde angibt.

Stunde	Bytes in KB	Fahrzeuge	Datengröße
0	21	30,000	630,000
1	21	40,000	840,000
2	21	50,000	1.050,000
3	21	60,000	1.260,000
4	21	70,000	1.470,000
5	21	80,000	1.680,000
6	21	90,000	1.890,000
7	21	100,000	2.100,000
8	21	200,000	4.200,000
9	21	175,000	3.675,000
10	21	150,000	3.150,000
11	21	140,000	2.940,000
12	21	155,000	3.255,000
13	21	150,000	3.150,000
14	21	150,000	3.150,000
15	21	150,000	3.150,000
16	21	170,000	3.570,000
17	21	200,000	4.200,000
18	21	190,000	3.990,000
19	21	170,000	3.570,000
20	21	160,000	3.360,000
21	21	100,000	2.100,000
22	21	70,000	1.470,000
23	21	40,000	840,000

[0031] Im Laufe der Zeit schwankt die Anzahl der Fahrzeuge, die Datensätze an die zentrale Einrichtung senden, und die Menge an Computerverarbeitungsressourcen, die benötigt wird, um die Datensätze zu verarbeiten, variiert basierend auf dem Zustrom von Datensätzen. Die zentrale Einrichtung kann die pro Zeiteinheit empfangene Datenmenge darstellen und die Steigung des Polynoms finden, die eine Kurve darstellt, die den gezeichneten Datenpunkten angenähert ist. Dieses Polynom kann verwendet werden, um die Durchflussrate der von der zentralen Einrichtung empfangenen Datensätze darzustellen. Die zentrale Einrichtung kann die Menge an Datensätzen überwachen, die innerhalb des Zeitintervalls anonymisiert wurden, und abhängig von der Menge an Datensätzen, die noch anonymisiert werden müssen und der Durchflussrate der Datensätze, die Rechenressourcen, die zum Verarbeiten oder Anonymisieren der Datensätze verwendet werden, erhöhen oder verringern. Erhöhungen der Durchflussrate von Datensätzen können eine Zunahme der Menge an Re-

chenressourcen veranlassen, die verwendet werden, um die Datensätze zu verarbeiten, während eine Abnahme der Durchflussrate die Menge an Rechenressourcen verringern kann.

[0032] Die berechnete Durchflussrate kann verwendet werden, um die Größe eines Zwischenspeichers zu messen, der zum Speichern der eingehenden Datensätze in der zentralen Einrichtung im Speicher benötigt wird. Der Zwischenspeicher kann basierend auf einem Schlüssel, wie etwa einem Kennungs-Dateneintrag, mithilfe eines Schlüssel- und Hash-Knoten-Kombinationsalgorithmus geteilt werden. Es ist möglich, eine erforderliche Kraft zu bestimmen, die einen anonymisierten Vektor aus dem Zwischenspeicher schieben wird. Die Größe (M) dieser Kraft (F) um den Ursprung (O) des Vektors kann als M = Fd definiert werden, wobei d der kürzeste Abstand von O zu einer Kraftlinie ist, unter der Annahme, dass d der benötigte Abstand ist, um den anonymisierten Vektor aus dem Zwischenspeicher zu schieben. Der Abstand d kann basierend auf der Strömungsgeschwindigkeit (v) und der Beschleunigung (a) zu einem Zeitpunkt berechnet werden, der durch $d=vt + 1/2at^2$ dargestellt ist, wobei v und a aus der Durchflussrate bestimmt werden.

[0033] Die Zwischenspeichergröße kann von Faktoren wie RAM abhängig sein und verwendet werden, um die Datensätze vorübergehend zu verschieben, bevor sie an einen Ausgabepunkt verschoben werden, wie etwa eine Nachrichtenwarteschlange, die die Daten auf die Festplatte schreibt. Es ist hilfreich, die Zwischenspeichergröße so zu steuern, dass die Datensätze aus einem Cache-Zwischenspeicher in einen Zwischenspeicher mit anonymen Daten verschoben werden. Dieser Abstand kann anhand der Geschwindigkeit des eingehenden Datenstroms berechnet werden. Die obige Formel zeigt die Abstandsberechnung. Da die Daten aus dem Zwischenspeicher-Cache mit Rohdaten in den Zwischenspeicher-Cache mit anonymisierten Daten verschoben werden, wird die Kraft, die dem Verarbeitungsdurchsatz entspricht, mithilfe der Momententheorie berechnet. Das Verfahren **200** endet dann.

[0034] Es versteht sich, dass das Vorstehende eine Beschreibung einer oder mehrerer Ausführungsformen der Erfindung ist. Die Erfindung ist nicht auf die besondere(n) hierin offenbarte(n) Ausführungsform(en) beschränkt, sondern ausschließlich durch die folgenden Patentansprüche definiert. Darüber hinaus beziehen sich die in der vorstehenden Beschreibung gemachten Aussagen auf bestimmte Ausführungsformen und sind nicht als Einschränkungen des Umfangs der Erfindung oder der Definition der in den Patentansprüchen verwendeten Begriffe zu verstehen, außer dort, wo ein Begriff oder Ausdruck ausdrücklich vorstehend definiert wurde. Verschiedene andere Ausführungsformen und verschiedene Änderungen und Modifikationen an der/den ausgewiesenen Ausführungsform(en) sind für Fachleute offensichtlich. Alle diese anderen Ausführungsformen, Änderungen und Modifikationen sollten im Geltungsbereich der angehängten Patentansprüche verstanden werden.

[0035] Wie in dieser Beschreibung und den Ansprüchen verwendet, sind die Begriffe „zum Beispiel“, „beispielsweise“, „z. B.“, „wie“ und „gleich“ und die Verben „umfassen“, „aufweisen“, „beinhalten“ und ihre anderen Verbformen, wenn sie in Verbindung mit einer Auflistung einer oder mehrerer Komponenten oder anderen Gegenständen verwendet werden, jeweils als offen auszulegen, was bedeutet, dass die Auflistung nicht so berücksichtigt wird, als dass sie andere, zusätzliche Komponenten oder Elemente ausschließt. Andere Begriffe sind in deren weitesten vernünftigen Sinn auszulegen, es sei denn, diese werden in einem Kontext verwendet, der eine andere Auslegung erfordert.

Patentansprüche

1. Verfahren zur Anonymisierung von Streaming-Datensätzen, umfassend die Schritte:
(a) Verarbeiten eines oder mehrerer Datensätze in einem oder mehreren anonymen Vektordarstellungen dieses Datensatzes/dieser Datensätze;
(b) Zugreifen auf einen verallgemeinerten Vektor, der eine gewünschte Datenanonymisierungsstufe beinhaltet;
(c) Vergleichen der anonymen Vektordarstellungen des Datensatzes/ der Datensätze mit dem verallgemeinerten Vektor;
(d) Ermitteln, ob die anonymen Vektordarstellungen des Datensatzes/ der Datensätze basierend auf dem Vergleich ausreichend anonym sind;
(e) Identifizieren eines Zeitintervalls zum Senden von Datensätzen, die ausreichend anonym sind; und
(f) Erhöhen oder Verringern einer Menge der Rechnerressourcen, die bestimmen, ob die Vektordarstellungen des Datensatzes/ der Datensätze basierend auf einer Menge der Zeitspanne, die im Zeitintervall verbleibt, ausreichend anonym sind.

2. Verfahren nach Anspruch 1, worin der eine oder die mehreren Datensätze mindestens ein Kennungs-Dateneintrag und mindestens ein Inhalts-Dateneintrag umfassen.

3. Verfahren nach Anspruch 1, worin jeder Dateneintrag, der in dem einen oder den mehreren Datensätzen beinhaltet ist, einen Punkt auf der einen oder den mehreren anonymen Vektordarstellungen dargestellt

4. Verfahren nach Anspruch 1, worin Schritt (c) ferner den Schritt des Ermittelns des Kosinus eines Winkels zwischen der anonymen Vektordarstellung und dem verallgemeinerten Vektor umfasst.

5. Verfahren nach Anspruch 1, worin Schritt (d) ferner den Schritt des Vergleichens des Kosinus des Winkels mit einem Schwellenwert umfasst

6. Verfahren nach Anspruch 1, ferner umfassend den Schritt der Bestimmung der Durchflussrate der Datensätze, die an einer zentralen Einrichtung empfangen werden.

7. Verfahren nach Anspruch 4, ferner umfassend den Schritt des Ermittelns einer Größe eines Zwischenspeichers basierend auf der Durchflussrate der Datensätze.

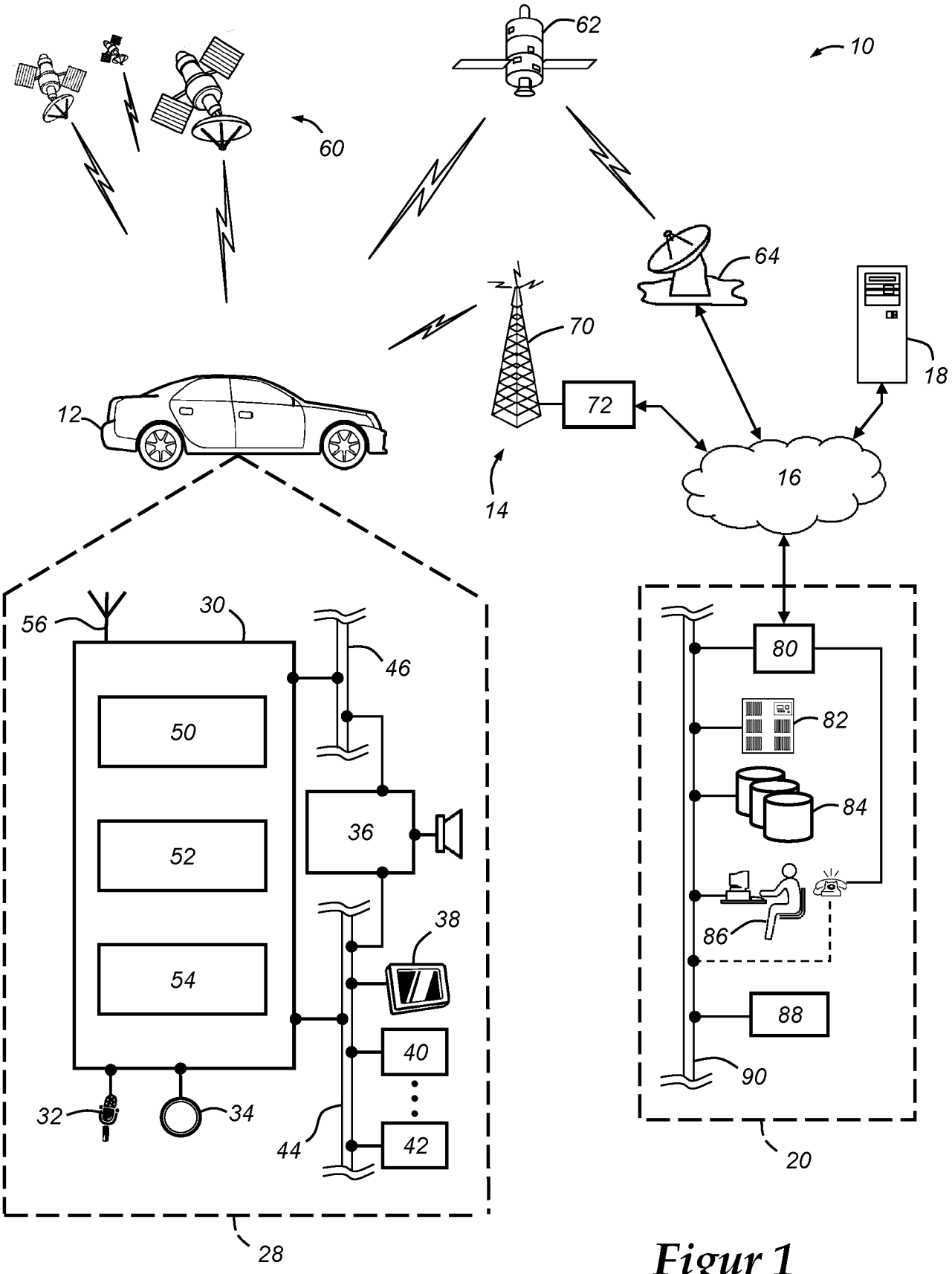
8. Verfahren nach Anspruch 1, worin die Datensätze in anonymen Vektordarstellungen in einem Fahrzeug verarbeitet werden.

9. Verfahren zur Anonymisierung von Streaming-Datensätzen, umfassend die Schritte:
(a) das Verarbeiten eines oder mehrerer Datensätze in einem oder mehreren anonymen Vektordarstellungen dieses Datensatzes/dieser Datensätze;
(b) das Zugreifen auf einen verallgemeinerten Vektor, der eine gewünschte Datenanonymisierungsstufe beinhaltet;
(c) das Vergleichen der anonymen Vektordarstellungen des Datensatzes/ der Datensätze mit dem verallgemeinerten Vektor;
(d) das Ermitteln, ob die anonymen Vektordarstellungen des Datensatzes/ der Datensätze basierend auf dem Vergleich ausreichend anonym sind;
(e) das Berechnen einer Rate, mit der die eingehenden Datensätze empfangen werden;
(f) das Erhöhen oder Verringern einer Menge der Rechnerressourcen, die bestimmen, ob die Vektordarstellungen des Datensatzes/ der Datensätze basierend auf der Rate, mit der Datensätze empfangen werden, ausreichend anonym sind; und
(g) das Übertragen von Datensätzen, die ausreichend anonymisiert sind, bevor ein Zeitintervall an eine dritte Partei abläuft.

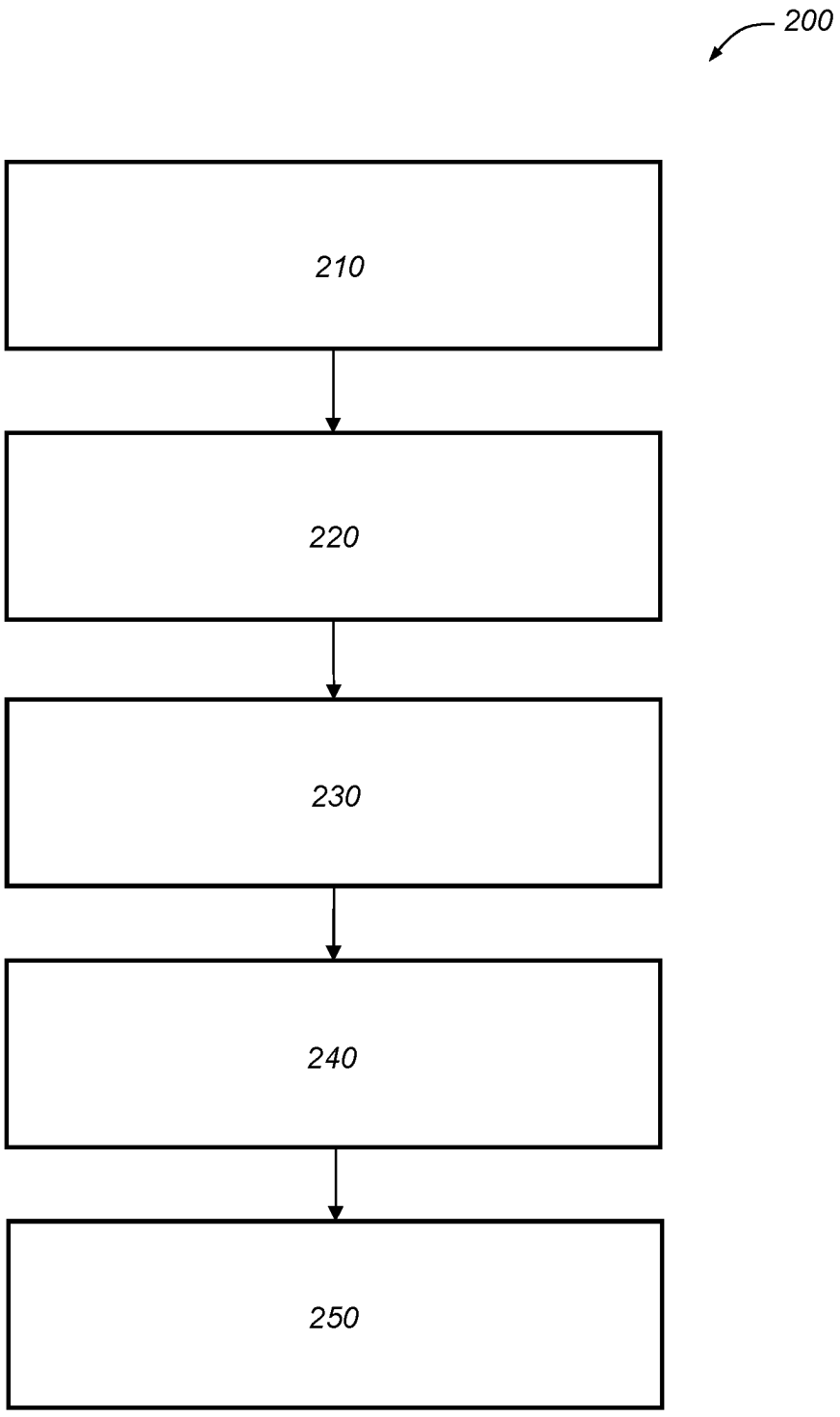
10. Verfahren nach Anspruch 9, worin Schritt (c) ferner den Schritt des Ermittelns des Kosinus eines Winkels zwischen der anonymen Vektordarstellung und dem verallgemeinerten Vektor umfasst.

Es folgen 2 Seiten Zeichnungen

Anhängende Zeichnungen



Figur 1



Figur 2