

(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
14. Dezember 2023 (14.12.2023)



(10) Internationale Veröffentlichungsnummer
WO 2023/237205 A1

(51) Internationale Patentklassifikation:

G06F 9/50 (2006.01) G06N 3/088 (2023.01)
G06N 3/0464 (2023.01)

(21) Internationales Aktenzeichen: PCT/EP2022/065756

(22) Internationales Anmeldedatum:
09. Juni 2022 (09.06.2022)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(71) Anmelder: SIEMENS AG ÖSTERREICH [AT/AT]; Sie-
mensstraße 90, 1210 Wien (AT).

(72) Erfinder: SCHALL, Daniel; Schabelgasse 1, 2020 Holla-
brunn (AT).

(74) Anwalt: MAIER, Daniel; Postfach 22 16 34, 80506 Mün-
chen (DE).

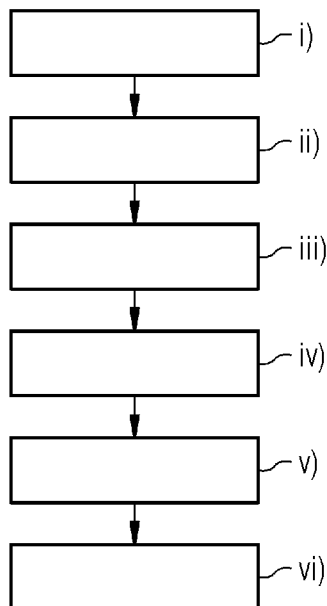
(81) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare nationale Schutzrechtsart): AE, AG, AL,
AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE,
KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU,
LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA,
NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO,
RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,
TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare regionale Schutzrechtsart): ARIPO (BW,

(54) Title: COMPUTER-IMPLEMENTED METHOD AND SYSTEM FOR DETECTING ANOMALIES DURING THE OPER-
ATION OF A TECHNICAL DEVICE

(54) Bezeichnung: COMPUTER-IMPLEMENTIERTES VERFAHREN UND SYSTEM ZUR ANOMALIE-ERKENNUNG BEIM
BETRIEB EINES TECHNISCHEN GERÄTS

FIG 1



(57) Abstract: Computer-implemented method for processing a set of computing tasks by means of a system, comprising at least one first computing apparatus (ELC1-ELCN) and at least two second computing apparatuses (EAC1-EACN), wherein the following steps are carried out: i) determining a weight function (w11-w13, wN1-wN3) for each computing task from the set of computing tasks between the at least one first computing apparatus (ELC1-ELCN) and the at least two second computing apparatuses (EAC1-EACN), ii) determining the execution capacity with regard to communication, storage and execution properties of the at least two second computing apparatuses (EAC1-EACN), iii) sorting the previously determined weight functions (w11-w13, wN1-wN3), and determining a subset of the sorted weight functions that are within a predetermined range of values of the determined execution capacity, iv) distributing the subset of computing tasks assigned to the subset of sorted weight functions to the at least two second computing apparatuses (EAC1-EACN), executing the subset of computing tasks and deleting the associated weight function, v) increasing the sequence factor for the computing tasks which were not distributed in the preceding step on account of the execution capacity, vi) continuing with step ii) as long as there are weight functions, and otherwise terminating the method.

(57) Zusammenfassung: Computer-implementiertes Verfahren zur Bearbeitung einer Menge an Rechenaufgaben durch ein System, umfassend zumindest eine erste Rechenvorrichtung (ELC1-ELCN) und zumindest zwei zweite Rechenvorrichtungen (EAC1-EACN), wobei folgende Schritte ausgeführt werden: i) Bestimmen jeweils einer Gewichtsfunktion (w11-w13, wN1-wN3) für jede Rechenaufgabe aus der Menge an Rechenaufgaben zwischen der zumindest einen ersten Rechenvorrichtung (ELC1-ELCN) und den zumindest zwei zweiten Rechenvorrichtungen (EAC1-EACN), ii) Ermitteln der Ausführungskapazität hinsichtlich Kommunikations-, Speicher- und Ausführungs-Eigenschaften der zumindest zwei zweiten Rechenvorrichtungen (EAC1-EACN), iii) Sortieren der zuvor bestimmten Gewichtsfunktionen (w11-w13, wN1-wN3), und Festlegen einer Teilmenge der sortierten Gewichtsfunktionen, welche innerhalb eines vorbestimmten Wertebereichs der ermittelten Ausführungskapazität liegen, iv) Verteilen der Teilmenge der Rechenaufgaben, welche der Teilmenge der sortierten Gewichtsfunktionen zugeordnet ist, an die zumindest zwei zweiten Rechenvorrichtungen

WO 2023/237205 A1

GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, RU, TJ, TM), europäisches (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— mit internationalem Recherchenbericht (Artikel 21 Absatz 3)

(EAC1-EACN), Ausführen der Teilmenge an Rechenaufgaben und Löschen der zugehörigen Gewichtsfunktion, v) Erhöhen des Reihungsfaktors für die Rechenaufgaben, die im vorhergehenden Schritt aufgrund der Ausführungskapazität nicht verteilt wurden, vi) Solange Gewichtsfunktionen vorliegen, Fortsetzen mit Schritt ii), sonst Beenden.

Computer-Implementiertes Verfahren und System zur Anomalie-Erkennung beim Betrieb eines technischen Geräts

Die Erfindung betrifft ein computer-implementiertes Verfahren und ein Rechen-System zur Bearbeitung einer Menge an Rechen-
5 aufgaben durch ein System, umfassend zumindest eine erste Rechen-
vorrichtung und zumindest zwei zweite Rechenvorrichtungen,
jeweils mit einem Prozessor und einem Speicher, wobei
die zumindest eine erste Rechenvorrichtung mit den zumindest
zwei zweiten Rechenvorrichtungen durch jeweilige Kommunikati-
10 onsverbindungen verbunden sind.

Die Erfindung betrifft auch eine Verwendung des Verfahrens zur Anomalie-Erkennung beim Betrieb eines technischen Geräts.
Ferner betrifft die Erfindung ein Computerprogramm, einen elektronisch lesbaren Datenträger und ein Datenträgersignal.
15 Typische Systeme künstlicher Intelligenz (kurz „KI“) passen Modelle des maschinellen Lernens (kurz „ML“) nur sehr langsam an neue Daten oder Umgebungsänderungen an. Der typische ML-Lebenszyklus umfasst drei Phasen:

- i) Datenerfassung und Datenaufbereitung,
- 20 ii) Modellerstellung und
- iii) Modellbereitstellung und Inferenz.

Das Schließen von Daten auf (hypothetische) Modelle wird als statistische Inferenz bezeichnet.

ML-Modelle können sich jedoch häufig nicht schnell an Live-
25 Daten oder neue Eingabedaten anpassen.

Ferner sind nicht alle Daten in einem Cloud-Edge-Kontext in der Cloud verfügbar. Das Senden aller Daten an die Cloud umgeht den Vorteil des Edge-Computing.

In einem Klienten-Server-System liegt eine Rechenvorrichtung
30 mit einem Prozessor und einem Speicher an der Edge.

Außerdem verfügen Edge-Geräte häufig nicht über ausreichende Fähigkeiten, um ein vollständiges Training von ML-Modellen durchzuführen.

5 In modernen Cloud-Edge-Umgebungen nach dem Stand der Technik wird beispielsweise das Training und die Inferenz nur in der Cloud durchgeführt. Dies hat den Vorteil, dass alle Daten an einem Ort verarbeitet und gespeichert werden. Klienten rufen diese Daten von einer zentralen Stelle ab. Jedoch ist keine Echtzeitinteraktion mit einem Prozess, welcher an einem Kli-
10 enten ausgeführt wird, möglich.

Alternativ wird im Stand der Technik Training in der Cloud und die Inferenz an der Edge durchgeführt. Dies führt dazu, dass performantes Training in der Cloud und Prozessinteraktion in Echtzeit an der Edge ausgeführt wird. Jedoch werden da-
15 bei Daten über die Cloud und die Edge verstreut verarbeitet oder gespeichert, was nachteilig ist, da das Aktualisieren von Modellen typischerweise nur in der Cloud möglich ist.

Soll ein ML-Verfahren jedoch auf einem verteilten System ausgeführt werden, kann sich zusätzlich eine höhere Komplexität
20 ergeben, nämlich insofern, als ein initiales Modell zur Anomalie-Erkennung in der Cloud trainiert werden muss, und anschließend an diverse Edge-Vorrichtungen verteilt werden muss.

Dabei ist es oft schwierig, eine geeignete Verteilung von Rechen-Aufgaben, wie ML-Modell-Aktualisierungen durch erneutes
25 Training, an mehrere performante Rechen-Vorrichtungen durchzuführen und eine vorteilhafte Auslastung des gesamten Rechen-Systems an der Edge zu erreichen.

Im Kontext von maschinellem Lernen wird häufig ein Initiales Modell in der Cloud erzeugt, trainiert und verteilten Klien-
30 ten bereitgestellt.

Eine Berechnung and er Edge kann aus Datenschutz-Gründen bevorzugt werden.

Eine Anwendung dieses ML-Modells erfolgt daher vorteilhaft an der Edge, wobei ML-Modell-Aktualisierungen aus dem Betrieb beziehungsweise der Anwendung des ML-Modells aus denselben Gründen ebenfalls an der Edge durchgeführt werden, was jedoch
5 eine sehr hohe Rechen-Kapazität, wie beispielsweise mit einem hoch parallelisierten Rechen-Kern wie einer GPU (engl. „graphic processing unit“), erfordern kann.

Die reine Anwendung eines ML-Modells kann meistens auf Rechenvorrichtungen mit einfacherer Komplexität und Kapazität
10 ausgeführt werden. Dies umfasst beispielsweise auch Datenerfassung durch Sensoren wie Kameras oder von einer PLC (engl. „programmable logic controller“) oder anderen Sensor-Mitteln.

Es ist daher Aufgabe der Erfindung eine Lösung bereitzustellen, mittels welcher eine vorteilhafte Verteilung von Rechenaufgaben auf mehrere Rechenvorrichtungen durchgeführt werden
15 kann.

Die erfindungsgemäße Aufgabe wird durch ein Verfahren eingangs genannter Art gelöst, wobei folgende Schritte ausgeführt werden:

20 i) Bestimmen jeweils einer Gewichtsfunktion für jede Rechenaufgabe aus der Menge an Rechenaufgaben zwischen der zumindest einen ersten Rechenvorrichtung und den zumindest zwei zweiten Rechenvorrichtungen, wobei die Gewichtsfunktion aufweist:

- 25 • eine Priorität der Rechenaufgabe,
- einen Kommunikations-Parameter bezüglich der Kommunikationsverbindung,
- einen Reihungsfaktor für die Bearbeitung der Rechenaufgabe, vorzugsweise beginnend bei null,
- 30 • einen Speicher-Parameter für den benötigten Speicherbedarf bei der Ausführung der Rechenaufgabe,

- einen Ausführungs-Parameter für den benötigten Rechenbedarf bei der Ausführung der Rechenaufgabe,
- ii) Ermitteln der Ausführungskapazität hinsichtlich Kommunikations-, Speicher- und Ausführungs-Eigenschaften der zumindest zwei zweiten Rechenvorrichtungen,
- iii) Sortieren der zuvor bestimmten Gewichtsfunktionen, und Festlegen einer Teilmenge der sortierten Gewichtsfunktionen, welche innerhalb eines vorbestimmten Wertebereichs der ermittelten Ausführungskapazität liegen,
- iv) Verteilen der Teilmenge der Rechenaufgaben, welche der Teilmenge der sortierten Gewichtsfunktionen zugeordnet ist, an die zumindest zwei zweiten Rechenvorrichtungen, Ausführen der Teilmenge an Rechenaufgaben und Löschen der zugehörigen Gewichtsfunktion,
- v) Erhöhen des Reihungsfaktors für die Rechenaufgaben, die im vorhergehenden Schritt aufgrund der Ausführungskapazität nicht verteilt wurden,
- vi) Solange Gewichtsfunktionen vorliegen, Fortsetzen mit Schritt ii), sonst Beenden.

Dadurch wird erreicht, dass die Rechen-Aufgaben auf besonders effiziente Weise im Rechen-System abgearbeitet werden können und eine Edge-Vorrichtung, welche die Rechen-Vorrichtungen umfasst, für die entsprechende Verwendung optimal angepasst ausgelegt werden kann.

In einer Weiterbildung der Erfindung ist es vorgesehen, dass jeweils die Gewichtsfunktion bestimmt wird, indem die Priorität, der Kommunikations-Parameter und der Reihungsfaktor jeweils in einem Wertebereich zwischen 0 und 1 liegen und deren Summe einen Rechen-Parameter bilden, und der Speicher-Parameter und der Ausführungs-Parameter jeweils in einem Wertebereich zwischen 0 und 1 liegen und deren Summe einen Berechnungs-Parameter bilden, und die Differenz von Rechen-

Parameter und Berechnungs-Parameter die Gewichtsfunktion bilden.

Dadurch wird erreicht, dass die Gewichtsfunktion auf eine einfache Weise bestimmt werden kann und welche besonders re-
5 präsентative Parameter zur Lösung der erfindungsgemäßen Aufgabe verwendet.

In einer Weiterbildung der Erfindung ist es vorgesehen, dass nach Schritt i) geprüft wird, ob die jeweilige Gewichtsfunktion gleich null ist, und falls ja, die zugehörige Rechenauf-
10 gabe in Teil-Rechenaufgaben unterteilt wird und die Teil-Rechenaufgaben der Menge an Rechenaufgaben hinzugefügt wird und die aktuelle Rechenaufgabe aus der Menge an Rechenaufgaben entfernt wird und das Verfahren mit Schritt i) fortgesetzt wird.

15 Dadurch wird erreicht, dass auch große Rechenaufgaben beziehungsweise deren zugeordneten Gewichtsfunktionen - beispielsweise mit hohem Rechenbedarf oder hohem Speicherbedarf - durch das Verfahren effizient abgearbeitet werden können und zugleich andere Rechenaufgaben nicht zu lange unbearbeitet
20 bleiben.

Eine Rechenaufgabe ist einer Gewichtsfunktion zugeordnet und durch diese abgebildet, wonach eine Teilmenge einer Gewichtsfunktion auch einer Teilmenge einer Rechenaufgabe entspricht. Dementsprechend können Teil-Rechenaufgaben auch Teil-
25 Gewichtsfunktionen zugeordnet werden.

In einer Weiterbildung der Erfindung ist es vorgesehen, dass das Verfahren eine Verwendung zur Anomalie-Erkennung beim Betrieb eines technischen Geräts findet, welches von einem Klienten-Server-System mit einem Server und einen Klienten mit
30 der zumindest einen ersten Rechenvorrichtung betrieben wird und zusätzlich folgende Verfahrensschritte ausgeführt werden:

- a) Erfassen aus dem Betrieb des technischen Geräts und Bereitstellen von ersten Trainings-Daten vom Klienten an den Server, Erzeugen und Trainieren eines ersten Modells in Form eines Auto-Encoders durch den Server, und Bereitstellen des ersten Modells in Form von ersten Gewichten vom Server an den Klienten, sowie Laden und Speichern des ersten Modells in der zumindest einen ersten Rechenvorrichtung,
- b) Erfassen von Betriebs-Daten des technischen Geräts und Bestimmen einer Anomalie-Kenngröße hinsichtlich der Übereinstimmung mit dem ersten Modell durch die zumindest eine erste Rechenvorrichtung,
- c) Prüfen, ob die Anomalie-Kenngröße innerhalb eines vorbestimmten Werte-Bereichs liegt, wenn ja, fortfahren mit Schritt b), sonst fortfahren mit Schritt d),
- d) Bereitstellen der ersten Modell-Gewichte des Auto-Encoders und Speichern mit der Anomalie-Kenngröße im Speicher durch die zumindest eine erste Rechenvorrichtung,
- e) Prüfen, ob eine vorbestimmte Anzahl an gespeicherten ersten Modell-Gewichte des Auto-Encoders für jeweilige Anomalie-Kenngrößen erreicht ist, wenn ja, dann Fortfahren mit Schritt b), sonst Fortfahren mit Schritt f),
- f) Übermitteln der gespeicherten ersten Modell-Gewichte für jeweilige Anomalie-Kenngrößen an die zumindest zwei zweiten Rechenvorrichtungen, welche vom Klienten umfasst ist und Berechnen eines zweiten Modells in Form eines Auto-Encoders mit zweiten Modell-Gewichten durch Bestimmen der Inferenz durch eine der zumindest zwei zweiten Rechenvorrichtungen,
- g) Übermitteln der zweiten Modell-Gewichte des zuvor berechneten zweiten Modells an die zumindest eine erste Rechenvorrichtung, und Löschen der im Schritt f) übermittelten ersten Modell-Gewichte aus dem Speicher der zumindest ei-

nen ersten Rechenvorrichtung, und Fortfahrrten mit Schritt
b).

Dadurch ist ein Modell-Training in der Cloud vorgesehen, die
Inferenz wird jedoch an der Edge berechnet, außerdem erfolgt
5 ein erneutes Trainieren (engl. „re-training“) an der Edge.

Dadurch ist es möglich ein performantes Training in der Cloud
auszuführen, sowie eine Prozess-Interaktion in Echtzeit zu
erlauben, und ferner ein Re-Training von ML-Modellen wiederum
an der Edge durchzuführen, um beispielsweise die Modellgenau-
10 igkeit aufgrund neuer Trainings-Daten, welche aus dem vorher-
gehenden Betrieb eines technischen Geräts mit dem vorherge-
henden ML-Modell gewonnen wurden, zu verbessern.

Die Daten liegen zwar über Klienten und Server verteilt, je-
doch kann durch Anwendung föderierten Lernens das neue ML-
15 Modell an andere Klienten verteilt werden, ohne die Rohdaten
selbst im verteilten Klienten-Server-System zu verteilen, was
zu einer verbesserten Datensicherheit und Privatsphäre führt.

In einer Weiterbildung der Erfindung ist es vorgesehen, dass
die Anomalie-Kenngröße der Rekonstruktions-Fehler des Auto-
20 Encoders ist.

Dadurch kann auf besonders einfache Weise eine besonders aus-
sagekräftige Anomalie-Kenngröße ermittelt werden.

In einer Weiterbildung der Erfindung ist es vorgesehen, dass
die Größe der Speicherbelegung des Speichers der zumindest
25 einen ersten Rechenvorrichtung zum Speichern der im Schritt
e) vorbestimmte Anzahl an gespeicherten ersten Modell-
Gewichte des Auto-Encoders proportional zur Bandbreite der
Kommunikations-Verbindung zwischen der zumindest einen ersten
Rechenvorrichtung und der zumindest zwei zweiten Rechenvor-
30 richtungen vorgesehen ist.

Dadurch wird erreicht, dass erst bei Überschreiten einer vorgegebenen Anzahl an Rekonstruktions-Fehlern eine Aktualisierung des Modells erfolgt.

In einer Weiterbildung der Erfindung ist es vorgesehen, dass
5 die im Schritt f) berechneten zweiten Modell-Gewichte vom Klienten dem Server bereitgestellt werden, und das Klienten-Server-System vorzugsweise einen weiteren Klienten mit einem weiteren technischen Gerät aufweist und der Server dazu eingerichtet ist, die zweiten Modell-Gewichte dem weiteren Klienten durch föderiertes Lernen bereitzustellen.
10

Dadurch kann erreicht werden, dass das aktualisierte zweite Modell auf einfache Weise von weiteren Klienten genutzt werden können, ohne selbst dafür nötige Rechenoperationen auszuführen. Somit ist für weitere Klienten eine Bauart mit reduzierter Komplexität und geringeren Kosten möglich, und es
15 können insbesondere Prinzipien föderierten Lernens angewendet werden.

Die erfindungsgemäße Aufgabe wird durch ein Rechen-System eingangs genannter Art gelöst, umfassend zumindest eine erste
20 Rechenvorrichtung und zumindest zwei zweite Rechenvorrichtungen, jeweils mit einem Prozessor und einem Speicher, wobei die zumindest eine erste Rechenvorrichtung jeweils mit den zumindest zwei zweiten Rechenvorrichtungen durch eine Kommunikationsverbindung verbunden ist und das Rechen-System dazu
25 eingerichtet ist, das erfindungsgemäße Verfahren auszuführen.

Die erfindungsgemäße Aufgabe wird durch ein Computerprogramm gelöst, umfassend Befehle, welche bei deren Ausführung durch einen Computer diesen veranlassen, das erfindungsgemäße Verfahren auszuführen.

30 Die erfindungsgemäße Aufgabe wird durch einen elektronisch lesbaren Datenträger mit darauf gespeicherten lesbaren Steuereinformatoren gelöst, welche zumindest das erfindungsgemäße

Computerprogramm umfassen und derart ausgestaltet sind, dass sie bei Verwendung des Datenträgers in einer Recheneinrichtung das erfindungsgemäße Verfahren durchführen.

Die erfindungsgemäße Aufgabe wird durch ein Datenträgersignal
5 gelöst, welches das erfindungsgemäße Computerprogramm überträgt.

Die Erfindung wird nachfolgend anhand eines in den beigefügten Zeichnungen dargestellten Ausführungsbeispiels näher erläutert. In den Zeichnungen zeigt:

- 10 Fig. 1 ein Ausführungsbeispiel des erfindungsgemäßen Verfahrens als Ablaufdiagramm,
- Fig. 2 ein Ausführungsbeispiel des erfindungsgemäßen Systems als Blockschaltbild,
- Fig. 3 ein Ausführungsbeispiel einer Verwendung des erfindungsgemäßen Verfahrens als Ablaufdiagramm,
- 15 Fig. 4 ein Ausführungsbeispiel einer Verwendung des erfindungsgemäßen Systems als Blockschaltbild.

Fig. 1 stellt ein Ausführungsbeispiel des erfindungsgemäßen Verfahrens als Ablaufdiagramm dar, wobei das Verfahren auf
20 einem in der **Fig. 2** dargestellten System ablaufen kann.

Das Verfahren zur Bearbeitung einer Menge an Rechenaufgaben durch ein System ist zumindest in Teilen computerimplementiert.

Ein Rechen-System umfasst zumindest eine erste Rechenvorrichtung ELC1-ELCN und zumindest zwei zweite Rechenvorrichtungen EAC1-EACN, jeweils mit einem Prozessor und einem Speicher.
25

Die zumindest eine erste Rechenvorrichtung ELC1-ELCN ist mit den zumindest zwei zweiten Rechenvorrichtungen EAC1-EACN
30 durch jeweilige Kommunikationsverbindungen verbunden.

Folgende Verfahrensschritte werden ausgeführt:

- 5 i) Bestimmen jeweils einer Gewichtsfunktion w_{l1} - w_{l3} , w_{N1} - w_{N3} für jede Rechenaufgabe aus der Menge an Rechenaufgaben zwischen der zumindest einen ersten Rechenvorrichtung ELC1-ELCN und den zumindest zwei zweiten Rechenvorrichtungen EAC1-EACN, wobei die Gewichtsfunktion aufweist:
- eine Priorität der Rechenaufgabe,
 - einen Kommunikations-Parameter bezüglich der Kommunikationsverbindung,
 - 10 • einen Reihungsfaktor für die Bearbeitung der Rechenaufgabe, vorzugsweise beginnend bei null,
 - einen Speicher-Parameter für den benötigten Speicherbedarf bei der Ausführung der Rechenaufgabe,
 - einen Ausführungs-Parameter für den benötigten Rechenbedarf bei der Ausführung der Rechenaufgabe,
 - 15
- ii) Ermitteln der Ausführungskapazität hinsichtlich Kommunikations-, Speicher- und Ausführungs-Eigenschaften der zumindest zwei zweiten Rechenvorrichtungen EAC1-EACN,
- 20 iii) Sortieren der zuvor bestimmten Gewichtsfunktionen w_{l1} - w_{l3} , w_{N1} - w_{N3} , und Festlegen einer Teilmenge der sortierten Gewichtsfunktionen, welche innerhalb eines vorbestimmten Wertebereichs der ermittelten Ausführungskapazität liegen,
- 25 iv) Verteilen der Teilmenge der Rechenaufgaben, welche der Teilmenge der sortierten Gewichtsfunktionen zugeordnet ist, an die zumindest zwei zweiten Rechenvorrichtungen EAC1-EACN, Ausführen der Teilmenge an Rechenaufgaben und Löschen der zugehörigen Gewichtsfunktion,
- 30 v) Erhöhen des Reihungsfaktors für die Rechenaufgaben, die im vorhergehenden Schritt aufgrund der Ausführungskapazität nicht verteilt wurden,

vi) Solange Gewichtsfunktionen vorliegen, Fortsetzen mit Schritt ii), sonst Beenden.

Optional kann die Gewichtsfunktion jeweils bestimmt werden, indem die Priorität, der Kommunikations-Parameter und der Reihungsfaktor jeweils in einem Wertebereich zwischen 0 und 1 liegen und deren Summe einen Rechen-Parameter bilden, und der Speicher-Parameter und der Ausführungs-Parameter jeweils in einem Wertebereich zwischen 0 und 1 liegen und deren Summe einen Berechnungs-Parameter bilden, und die Differenz von Rechen-Parameter und Berechnungs-Parameter die Gewichtsfunktion bilden.

Beispielsweise kann die Priorität P (engl. „priority factor“) für eine aktuelle Rechenaufgabe, welche durch eine erste Rechenvorrichtung ausgewählt ist, im Werte-Bereich 0% ... 100% liegen.

Ein hoher Wert wie 100% bedeutet, dass das Modell eine dringende Aktualisierung benötigt, und ein niedriger Wert wie 1% bedeutet, dass eine Aktualisierung nicht notwendig ist.

Der Kommunikations-Parameter E (engl. „communication efficiency factor“) kann ebenso im Werte-Bereich 0% ... 100% liegen, wobei hoher Wert wie 100% bedeutet, dass eine sehr gute Netzwerk-Qualität vorliegt, und ein niedriger Wert wie 1% bedeutet, dass die Verbindung eine schlechte Übertragungs-Qualität aufweist.

Der Reihungsfaktor S (engl. „starvation factor“) kann ebenso im Werte-Bereich 0% ... 100% liegen, wobei der Reihungsfaktor bei jedem Iterationsdurchgang erhöht wird, beispielsweise um 10%. Dies dient dazu, die Dringlichkeit der Abarbeitung einer Rechenaufgabe schrittweise zu erhöhen, um die Rechenaufgabe nicht „verhungern“ zu lassen.

Der Speicher-Parameter U (engl. „utilization factor“) kann auch im Werte-Bereich 0% ... 100% liegen, wobei der Speicher-

Parameter eine Speicher-Belegung des Arbeitsspeichers und/oder des Festplattenspeichers abbilden kann. Der Speicher kann benötigt werden, um beispielsweise temporär oder dauerhaft einen Datensatz auf einer zweiten Rechenvorrichtung zu speichern.

Ein hoher Wert wie 100% bedeutet, dass das Modell einen hohen Speicherbedarf aufweist, und ein niedriger Wert wie 1% bedeutet, dass das Modell einen niedrigen Speicherbedarf aufweist.

Der Ausführungs-Parameter C (engl. „computational factor“) kann auch im Werte-Bereich 0% ... 100% liegen, wobei der Ausführungs-Parameter eine Rechenkapazität abbilden kann. Die Rechenkapazität kann benötigt werden, um beispielsweise eine Rechenaufgabe auf einer zweiten Rechenvorrichtung auszuführen.

Ein hoher Wert wie 100% bedeutet, dass das Modell einen hohen Rechenbedarf aufweist - beispielsweise durch eine hohe Anzahl parallel organisierter Rechenkern, wie bei einer GPU -, und ein niedriger Wert wie 1% bedeutet, dass das Modell einen niedrigen Rechenbedarf aufweist

Der Rechen-Parameter wird durch die Summe der Priorität, des Kommunikations-Parameters und des Reihungsfaktors gebildet.

Der Berechnungs-Parameter wird durch die Summe des Speicher-Parameters und des Ausführungs-Parameters gebildet.

Wird der Berechnungs-Parameter von dem Rechen-Parameter subtrahiert, erhält man die Gewichtsfunktion

$$w_{AB} = (P + E + S) - (U + C)$$

wobei A eine erste Rechenvorrichtung und B eine zweite Rechenvorrichtung ist.

Für jede Rechenaufgabe mit einer Gewichtsfunktion von mehr als 100% können diese „ausgehende“ Kanten sortiert werden,

wenn zur Veranschaulichung eine Analogie zu einem bipartiten Graphen wie nach Fig. 2 hergestellt wird.

Nach der Neusortierung werden all jene Kanten entfernt, welche im aktuellen Durchlauf nicht abgearbeitet werden können

5 und der verbleibende Graph wird zur weiteren Bearbeitung herangezogen.

Anschließend werden die noch nicht abgearbeiteten Rechenaufgaben erfasst und als neue Iterationsschleife wie zuvor ausgeführt bearbeitet, bis alle Rechenaufgaben erledigt sind.

10 Beispielsweise können $P=100\%$, $E=80\%$, $S=0\%$ sowie $U=50\%$, $C=50\%$ gewählt werden, sodass die Gewichtsfunktion $w=80\%$ beträgt.

In diesem Beispiel bedeutet das für eine Rechenaufgabe, dass sie eine hohe Priorität aufweist, eine gute Netzwerk-

Bandbreite vorliegt, die Rechenaufgabe erstmalig zur Abarbeitung vorgesehen ist, eine moderate Größe des Datensatzes vor-

15 liegt und ein moderater Rechenbedarf vorgesehen ist. Daher ist es wahrscheinlich, dass eine Anforderung zur Abarbeitung der Rechenaufgabe erfolgreich erledigt wird, auch wenn aktuell eine hohe Last bei einer zweiten Rechenvorrichtung vor-

20 liegt. Es kann ferner erwogen werden, ob zur Abarbeitung der Rechenaufgabe durch die zweite Rechenvorrichtung eine Modell-Aktualisierung angefordert wird, um die Parameter zur Bestimmung der Gewichtsfunktion zu verbessern.

Fernern kann optional nach Schritt i) geprüft werden, ob die

25 jeweilige Gewichtsfunktion gleich null ist, und falls ja, die zugehörige Rechenaufgabe in Teil-Rechenaufgaben unterteilt

wird und die Teil-Rechenaufgaben der Menge an Rechenaufgaben hinzugefügt wird und die aktuelle Rechenaufgabe aus der Menge an Rechenaufgaben entfernt wird und das Verfahren mit

30 Schritt i) fortgesetzt wird.

Beispielsweise können $P=80\%$, $E=50\%$, $S=0\%$ sowie $U=80\%$, $C=0\%$ gewählt werden, sodass die Gewichtsfunktion $w=-30\%$ beträgt.

In diesem Beispiel bedeutet das für eine Rechenaufgabe, dass sie eine hohe Priorität aufweist, eine eher moderate Netzwerk-Bandbreite vorliegt, die Rechenaufgabe erstmalig zur Abarbeitung vorgesehen ist, eine große Größe des Datensatzes vorliegt und ein hoher Rechenbedarf vorgesehen ist. Daher ist es unwahrscheinlich, dass eine Anforderung zur Abarbeitung der Rechenaufgabe erfolgreich erledigt wird, wenn aktuell eine hohe Last bei einer zweiten Rechenvorrichtung vorliegt, solange der Reihungsfaktor S erhöht wird oder eine erste Rechenvorrichtung die Rechenaufgabe in kleinere Teil-Rechenaufgaben mit kleineren Datensätzen unterteilt.

Mit anderen Worten kann anhand des Vorliegens einer negativen Gewichtsfunktion erkannt werden, dass eine Rechenaufgabe vorteilhafterweise in kleinere Teilrechenaufgaben unterteilt werden soll, um eine effiziente Abarbeitung zu erreichen.

Eine vorteilhafte Verwendung des Verfahrens zur Anomalie-Erkennung beim Betrieb eines technischen Geräts ist in **Fig. 3** dargestellt, wofür das System der **Fig. 4** verwendet werden kann.

Ein Klienten-Server-System CSS zur Anomalie-Erkennung beim Betrieb eines technischen Geräts TD1 umfasst einen Server S und einen Klienten $K1$ mit einer ersten Rechenvorrichtung 1 und einen weiteren, in diesem Beispiel baugleichen Klienten $K2$ mit einem weiteren technischen Gerät TD2.

Es können jedoch auch weitere Klienten im System eingesetzt werden, welche eine einfachere Bauart vorweisen und das das Modell mithilfe föderierten Lernens beziehen.

Die erste Rechenvorrichtung 1 weist einen Prozessor und einen Speicher auf.

Der Klient $K1$ umfasst ferner eine zweite Rechenvorrichtung 2, welche mit der ersten Rechenvorrichtung 1 über eine Kommuni-

kations-Verbindung C verbunden ist, wobei die Kommunikations-Verbindung C eine vordefinierte Bandbreite aufweist.

Der Klient K1 erfasst vorzugsweise mithilfe der ersten Rechenvorrichtung 1 während eines Referenzbetriebs des technischen Geräts TD1 Testdaten DT, welche einen zulässigen, gültigen Betrieb technischen Geräts TD1 repräsentieren.

Die Datenerfassung kann beispielsweise mit einem Sensor-Mittel wie einer Kamera erfolgen und die Testdaten DT können Kamera-Aufnahmen sein.

Mithilfe dieser Testdaten DT kann im Server durch Anwendung eines Auto-Encoders ein globales Modell auf Basis maschinellen Lernens für das technische Geräts TD1 erzeugt und trainiert werden.

In weiterer Folge kann das globale Modell an Klienten K1, K2 zum Betrieb des technischen Geräts TD1 in Form von Gewichten des globalen Modells bereitgestellt werden.

Das Verfahren ist computer-implementiert, das heißt ein oder mehrere Schritte werden von einem Computer ausgeführt.

Das Verfahren kann besonders vorteilhaft zur Anomalie-Erkennung beim Betrieb eines technischen Geräts TD1, welches von einem Klienten-Server-System CSS mit einem Server S und einem Klienten K1 mit der zumindest einen ersten Rechenvorrichtung ELC1-ELCN betrieben wird, verwendet werden, und zusätzlich folgende Verfahrensschritte ausgeführt werden:

a) Erfassen aus dem Betrieb des technischen Geräts TD1 und Bereitstellen von ersten Trainings-Daten DT vom Klienten K1 an den Server S, Erzeugen und Trainieren eines ersten Modells in Form eines Auto-Encoders durch den Server, und Bereitstellen des ersten Modells in Form von ersten Gewichten vom Server S an den Klienten K1, sowie Laden und Speichern des ersten Modells in der zumindest einen ersten Rechenvorrichtung ELC1-ELCN,

- b) Erfassen von Betriebs-Daten des technischen Geräts TD1 und Bestimmen einer Anomalie-Kenngröße hinsichtlich der Übereinstimmung mit dem ersten Modell durch die zumindest eine erste Rechenvorrichtung ELC1-ELCN,
- 5 c) Prüfen, ob die Anomalie-Kenngröße innerhalb eines vorbestimmten Werte-Bereichs liegt, wenn ja, fortfahren mit Schritt b), sonst fortfahren mit Schritt d),
- d) Bereitstellen der ersten Modell-Gewichte des Auto-Encoders und Speichern mit der Anomalie-Kenngröße im Speicher durch die zumindest eine erste Rechenvorrichtung ELC1-ELCN,
- 10 e) Prüfen, ob eine vorbestimmte Anzahl an gespeicherten ersten Modell-Gewichte des Auto-Encoders für jeweilige Anomalie-Kenngrößen erreicht ist, wenn ja, dann Fortfahren mit Schritt b), sonst Fortfahren mit Schritt f),
- 15 f) Übermitteln der gespeicherten ersten Modell-Gewichte für jeweilige Anomalie-Kenngrößen an die zumindest zwei zweiten Rechenvorrichtungen EAC1-EACN, welche vom Klienten K1 umfasst ist und Berechnen eines zweiten Modells in Form eines Auto-Encoders mit zweiten Modell-Gewichten durch Bestimmen der Inferenz durch eine der zumindest zwei zweiten Rechenvorrichtungen EAC1-EACN,
- 20 g) Übermitteln der zweiten Modell-Gewichte des zuvor berechneten zweiten Modells an die zumindest eine erste Rechenvorrichtung ELC1-ELCN, und Löschen der im Schritt f) übermittelten ersten Modell-Gewichte aus dem Speicher der zumindest einen ersten Rechenvorrichtung ELC1-ELCN, und Fortfahren mit Schritt b),
- 25

Die Größe der Speicherbelegung des Speichers der zumindest einen ersten Rechenvorrichtung ELC1-ELCN zum Speichern der im Schritt e) vorbestimmte Anzahl an gespeicherten ersten Modell-Gewichte des Auto-Encoders ist proportional zur Bandbreite der Kommunikations-Verbindung C zwischen der zumindest

30

einen ersten Rechenvorrichtung ELC1-ELCN und der zumindest zwei zweiten Rechenvorrichtungen EAC1-EACN vorgesehen.

Die zumindest zwei zweiten Rechenvorrichtungen EAC1-EACN sind vorzugsweise Grafikprozessoren (engl. „graphics processing unit“, kurz GPU), welche auf effiziente Berechnung spezialisierte und optimierte Prozessoren sind und für künstliche Intelligenz und maschinelles Lernen verwendet werden kann. Weil GPUs außerordentlich viel Rechenleistung bieten, können sie beispielsweise aufgrund von Parallelverarbeitung eine enorme Beschleunigung bei Rechenaufgaben erzielen.

Ein Grafik-Prozessor kann integrierter Teil der ersten Rechenvorrichtung sein, wie eine integrierte Grafikkarte eines Edge-Computers.

Das Bereitstellen der ersten Modell-Gewichte des Auto-Encoders im Schritt d) wird auch als „embedding“ komprimierter Merkmale beziehungsweise eines komprimierten Merkmalsraums, welcher durch die Modell-Gewichte gebildet wird, bezeichnet.

Es ist klar, dass im gesamten Verfahren ein oder mehrere Anomalie-Kenngrößen angewendet werden können.

Die im Schritt f) berechneten zweiten Modell-Gewichte können vom Klienten K1 dem Server S bereitgestellt werden.

Der Server S kann sein globales ML-Modell nun aktualisieren.

Wenn das Klienten-Server-System CSS zumindest einen weiteren Klienten K2 mit einem weiteren technischen Gerät TD2 aufweist, kann das System dem weiteren Klienten K2 durch föderiertes Lernen die zweiten Modell-Gewichte bereitstellen.

Bezugszeichenliste:

	C	Datenverbindung mit Übertragungskapazität, (engl. „communication line“)
5	CSS	Klienten-Server-System (engl. „client-server-system“)
	D, DT	Daten, Trainings-Daten
	EAC	Rechenvorrichtung mit Prozessor und Speicher, Klient mit hoher Verarbeitungs-Kapazität (engl. „edge advanced compute“)
10	ELC	Rechenvorrichtung mit Prozessor und Speicher , Klient mit geringer Verarbeitungs-Kapazität (engl. „edge limited computing“)
	K1, K1	Klient, Edge-Vorrichtung
	S	Server
15	TD1, TD2	technisches Gerät, z.B. Motor einer Produktionsanlage
	w11-w13,	
	wN1-wN3	Gewicht, Gewichtsfunktion

Patentansprüche

1. Computer-implementiertes Verfahren zur Bearbeitung einer Menge an Rechenaufgaben durch ein System, umfassend zumindest eine erste Rechenvorrichtung (ELC1-ELCN) und zumindest zwei
5 zweite Rechenvorrichtungen (EAC1-EACN), jeweils mit einem Prozessor und einem Speicher, wobei die zumindest eine erste Rechenvorrichtung (ELC1-ELCN) mit den zumindest zwei zweiten Rechenvorrichtungen (EAC1-EACN) durch jeweilige Kommunikationsverbindungen verbunden sind und folgende Schritte ausge-
10 führt werden:
- i) Bestimmen jeweils einer Gewichtsfunktion (w11-w13, wN1-wN3) für jede Rechenaufgabe aus der Menge an Rechenaufgaben zwischen der zumindest einen ersten Rechenvorrichtung (ELC1-ELCN) und den zumindest zwei
15 zweiten Rechenvorrichtungen (EAC1-EACN), wobei die Gewichtsfunktion aufweist:
- eine Priorität der Rechenaufgabe,
 - einen Kommunikations-Parameter bezüglich der Kommunikationsverbindung,
 - 20 • einen Reihungsfaktor für die Bearbeitung der Rechenaufgabe, vorzugsweise beginnend bei null,
 - einen Speicher-Parameter für den benötigten Speicherbedarf bei der Ausführung der Rechenaufgabe,
 - einen Ausführungs-Parameter für den benötigten Re-
25 chenbedarf bei der Ausführung der Rechenaufgabe,
- ii) Ermitteln der Ausführungskapazität hinsichtlich Kommunikations-, Speicher- und Ausführungs-Eigenschaften der zumindest zwei zweiten Rechenvorrichtungen (EAC1-EACN),
- 30 iii) Sortieren der zuvor bestimmten Gewichtsfunktionen (w11-w13, wN1-wN3), und Festlegen einer Teilmenge der sortierten Gewichtsfunktionen, welche innerhalb

eines vorbestimmen Wertebereichs der ermittelten Ausführungskapazität liegen,

iv) Verteilen der Teilmenge der Rechenaufgaben, welche der Teilmenge der sortierten Gewichtsfunktionen zugeordnet ist, an die zumindest zwei zweiten Rechenvorrichtungen (EAC1-EACN), Ausführen der Teilmenge an Rechenaufgaben und Löschen der zugehörigen Gewichtsfunktion,

v) Erhöhen des Reihungsfaktors für die Rechenaufgaben, die im vorhergehenden Schritt aufgrund der Ausführungskapazität nicht verteilt wurden,

vi) Solange Gewichtsfunktionen vorliegen, Fortsetzen mit Schritt ii), sonst Beenden.

2. Verfahren nach dem vorhergehenden Anspruch, wobei jeweils die Gewichtsfunktion bestimmt wird, indem die Priorität, der Kommunikations-Parameter und der Reihungsfaktor jeweils in einem Wertebereich zwischen 0 und 1 liegen und deren Summe einen Rechen-Parameter bilden, und der Speicher-Parameter und der Ausführungs-Parameter jeweils in einem Wertebereich zwischen 0 und 1 liegen und deren Summe einen Berechnungs-Parameter bilden, und die Differenz von Rechen-Parameter und Berechnungs-Parameter die Gewichtsfunktion bilden.

3. Verfahren nach dem vorhergehenden Anspruch, wobei nach Schritt i) geprüft wird, ob die jeweilige Gewichtsfunktion gleich null ist, und falls ja, die zugehörige Rechenaufgabe in Teil-Rechenaufgaben unterteilt wird und die Teil-Rechenaufgaben der Menge an Rechenaufgaben hinzugefügt wird und die aktuelle Rechenaufgabe aus der Menge an Rechenaufgaben entfernt wird und das Verfahren mit Schritt i) fortgesetzt wird.

4. Verwendung der Verfahrens nach einem der vorhergehenden Ansprüche zur Anomalie-Erkennung beim Betrieb eines techni-

schen Geräts (TD1), welches von einem Klienten-Server-System (CSS) mit einem Server (S) und einen Klienten (K1) mit der zumindest einen ersten Rechenvorrichtung (ELC1-ELCN) betrieben wird und zusätzlich folgende Verfahrens-Schritte ausgeführt werden:

- a) Erfassen aus dem Betrieb des technischen Geräts (TD1) und Bereitstellen von ersten Trainings-Daten (DT) vom Klienten (K1) an den Server (S), Erzeugen und Trainieren eines ersten Modells in Form eines Auto-Encoders durch den Server, und Bereitstellen des ersten Modells in Form von ersten Gewichten vom Server (S) an den Klienten (K1), sowie Laden und Speichern des ersten Modells in der zumindest einen ersten Rechenvorrichtung (ELC1-ELCN),
- b) Erfassen von Betriebs-Daten des technischen Geräts (TD1) und Bestimmen einer Anomalie-Kenngröße hinsichtlich der Übereinstimmung mit dem ersten Modell durch die zumindest eine erste Rechenvorrichtung (ELC1-ELCN),
- c) Prüfen, ob die Anomalie-Kenngröße innerhalb eines vorbestimmten Werte-Bereichs liegt, wenn ja, fortfahren mit Schritt b), sonst fortfahren mit Schritt d),
- d) Bereitstellen der ersten Modell-Gewichte des Auto-Encoders und Speichern mit der Anomalie-Kenngröße im Speicher durch die zumindest eine erste Rechenvorrichtung (ELC1-ELCN),
- e) Prüfen, ob eine vorbestimmte Anzahl an gespeicherten ersten Modell-Gewichte des Auto-Encoders für jeweilige Anomalie-Kenngrößen erreicht ist, wenn ja, dann Fortfahren mit Schritt b), sonst Fortfahren mit Schritt f),
- f) Übermitteln der gespeicherten ersten Modell-Gewichte für jeweilige Anomalie-Kenngrößen an die zumindest zwei zweiten Rechenvorrichtungen (EAC1-EACN), welche vom Klienten (K1) umfasst ist und Berechnen eines zweiten Modells in Form eines Auto-Encoders mit zweiten Modell-Gewichten

durch Bestimmen der Inferenz durch eine der zumindest zwei zweiten Rechenvorrichtungen (EAC1-EACN),

g) Übermitteln der zweiten Modell-Gewichte des zuvor berechneten zweiten Modells an die zumindest eine erste Rechen-
5 vorrichtung (ELC1-ELCN), und Löschen der im Schritt f) übermittelten ersten Modell-Gewichte aus dem Speicher der zumindest einen ersten Rechenvorrichtung (ELC1-ELCN), und Fortfahren mit Schritt b),

5. Verwendung nach dem vorhergehenden Anspruch, wobei die
10 Anomalie-Kenngröße der Rekonstruktions-Fehler des Auto-Encoders ist.

6. Verwendung nach einem der Ansprüche 4 oder 5, wobei die Größe der Speicherbelegung des Speichers der zumindest einen ersten Rechenvorrichtung (ELC1-ELCN) zum Speichern der im
15 Schritt e) vorbestimmte Anzahl an gespeicherten ersten Modell-Gewichte des Auto-Encoders proportional zur Bandbreite der Kommunikations-Verbindung (C) zwischen der zumindest einen ersten Rechenvorrichtung (ELC1-ELCN) und der zumindest zwei zweiten Rechenvorrichtungen (EAC1-EACN) vorgesehen ist.

20 7. Verwendung nach einem der Ansprüche 4 bis 6, wobei die im Schritt f) berechneten zweiten Modell-Gewichte vom Klienten (K1) dem Server (S) bereitgestellt werden, und das Klienten-Server-System vorzugsweise einen weiteren Klienten (K2) mit einem weiteren technischen Gerät (TD2) aufweist und der
25 Server (S) dazu eingerichtet ist, die zweiten Modell-Gewichte dem weiteren Klienten (K2) durch föderiertes Lernen bereitzustellen.

8. Rechen-System zur Bearbeitung einer Menge an Rechenaufgaben, umfassend zumindest eine erste Rechenvorrich-
30 tung (ELC1-ELCN) und zumindest zwei zweite Rechenvorrichtungen (EAC1-EACN), jeweils mit einem Prozessor und einem Speicher, wobei die zumindest eine erste Rechenvorrichtung (ELC1-

ELCN) jeweils mit den zumindest zwei zweiten Rechenvorrichtungen (EAC1-EACN) durch eine Kommunikationsverbindung verbunden ist und das Rechen-System dazu eingerichtet ist, das Verfahren nach einem der Ansprüche 1 bis 3 auszuführen.

5 9. Computerprogramm, umfassend Befehle, welche bei deren Ausführung durch einen Computer diesen veranlassen, das Verfahren nach einem der Ansprüche 1 bis 3 auszuführen.

10. Elektronisch lesbarer Datenträger mit darauf gespeicherten lesbaren Steuerinformationen, welche zumindest das Computerprogramm nach dem vorhergehenden Anspruch umfassen und
10 derart ausgestaltet sind, dass sie bei Verwendung des Datenträgers in einer Recheneinrichtung das Verfahren nach einem der Ansprüche 1 bis 3 durchführen.

11. Datenträgersignal, welches das Computerprogramm nach Anspruch 9 überträgt.
15

FIG 1

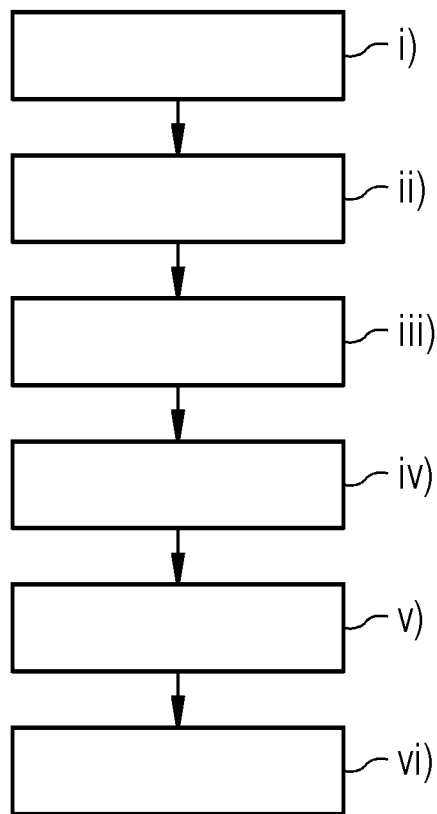
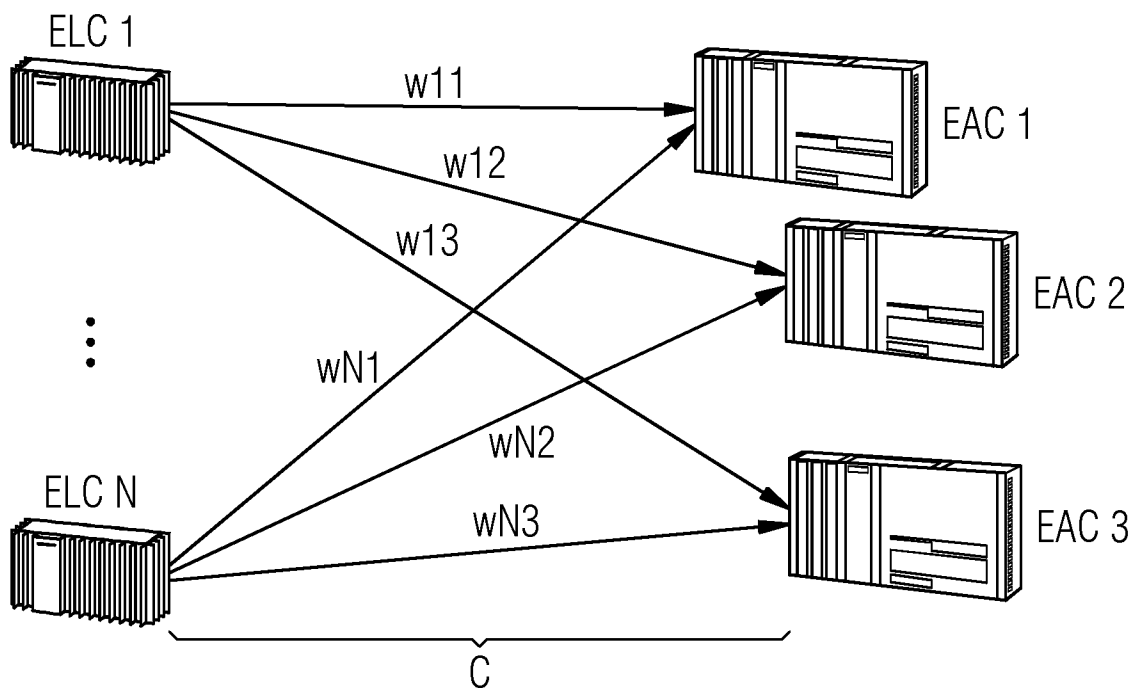


FIG 2



2/2

FIG 3

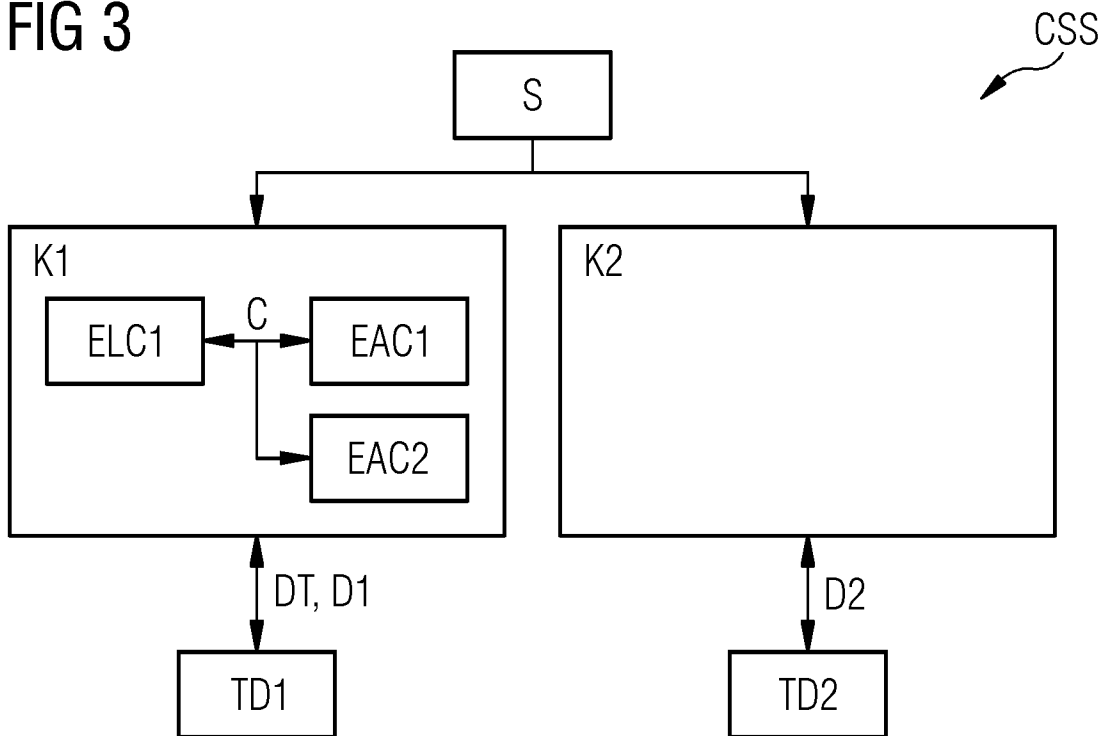
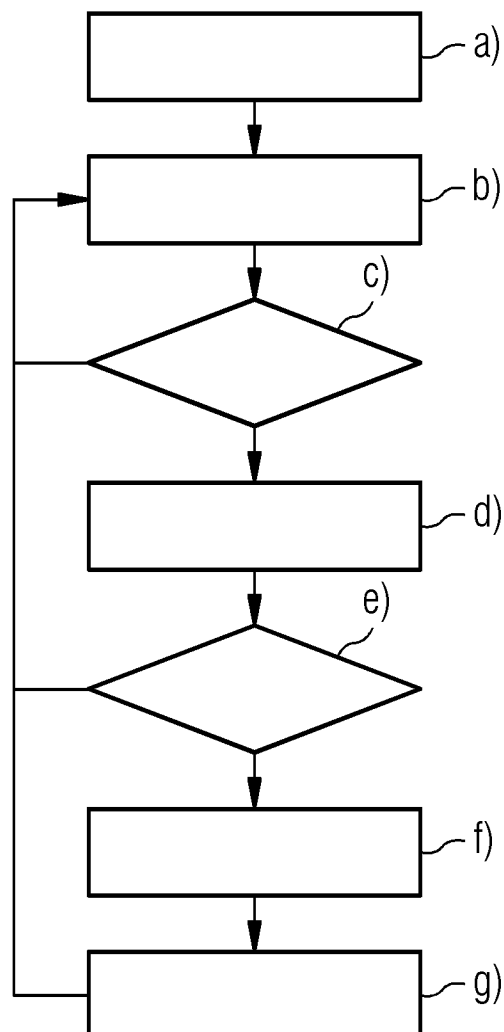


FIG 4



INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP2022/065756

A. CLASSIFICATION OF SUBJECT MATTER**G06F 9/50**(2006.01)i; **G06N 3/0464**(2023.01)i; **G06N 3/088**(2023.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F; G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US 2009254914 A1 (SUNDARESAN NAGARAJ [US] ET AL) 08 October 2009 (2009-10-08) paragraph [0007] - paragraphs [0014], [0035] figure 1	1, 2, 8-11 3
Y	EP 0652513 A1 (IBM [US]) 10 May 1995 (1995-05-10) page 2, line 39 - page 3, line 6	3
X	WO 2021050790 A1 (C3 AI INC [US]) 18 March 2021 (2021-03-18) paragraph [0010] - paragraph [0020] paragraph [0053] - paragraph [0056] figure 1	4-7
A	Jinwon An ET AL. "SNU Data Mining Center 2015-2 Special Lecture on IE Variational Autoencoder based Anomaly Detection using Reconstruction Probability" 27 December 2015 (2015-12-27), Retrieved from the Internet: http://dm.snu.ac.kr/static/docs/TR/SNUDM-TR-2015-03.pdf XP055688241 the whole document	4-7

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

06 March 2023

Date of mailing of the international search report

28 April 2023

Name and mailing address of the ISA/EP

European Patent Office
p.b. 5818, Patentlaan 2, 2280 HV Rijswijk
Netherlands

Telephone No. (+31-70)340-2040

Facsimile No. (+31-70)340-3016

Authorized officer

Mühlenbrock, Martin

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP2022/065756

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. Claims: 1-3, 8-11

A method for processing a set of computational tasks using a system

2. Claims: 4-7

A method for detecting anomalies during operation of a technical device

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☒ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/EP2022/065756

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US	2009254914	A1	08 October 2009	NONE	
EP	0652513	A1	10 May 1995	EP 0652513 A1	10 May 1995
				JP 3414520 B2	09 June 2003
				JP H07182185 A	21 July 1995
				US 5437032 A	25 July 1995
WO	2021050790	A1	18 March 2021	AU 2020345874 A1	31 March 2022
				CA 3154145 A1	18 March 2021
				CN 114730390 A	08 July 2022
				EP 4028963 A1	20 July 2022
				JP 2022548227 A	17 November 2022
				KR 20220062329 A	16 May 2022
				US 2022260988 A1	18 August 2022
				WO 2021050790 A1	18 March 2021

Feld Nr. II Bemerkungen zu den Ansprüchen, die sich als nicht recherchierbar erwiesen haben (Fortsetzung von Punkt 2 auf Blatt 1)

Gemäß Artikel 17(2)a) wurde aus folgenden Gründen für bestimmte Ansprüche kein internationaler Recherchenbericht erstellt:

1. ☐ Ansprüche Nr.
weil sie sich auf Gegenstände beziehen, zu deren Recherche diese Behörde nicht verpflichtet ist, nämlich

2. ☐ Ansprüche Nr.
weil sie sich auf Teile der internationalen Anmeldung beziehen, die den vorgeschriebenen Anforderungen so wenig entsprechen, dass eine sinnvolle internationale Recherche nicht durchgeführt werden kann, nämlich

3. ☐ Ansprüche Nr.
weil es sich dabei um abhängige Ansprüche handelt, die nicht entsprechend Satz 2 und 3 der Regel 6.4 a) abgefasst sind.

Feld Nr. III Bemerkungen bei mangelnder Einheitlichkeit der Erfindung (Fortsetzung von Punkt 3 auf Blatt 1)

Diese Internationale Recherchenbehörde hat festgestellt, dass diese internationale Anmeldung mehrere Erfindungen enthält:

siehe Zusatzblatt

1. ☒ Da der Anmelder alle erforderlichen zusätzlichen Recherchegebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht auf alle recherchierbaren Ansprüche.

2. ☐ Da für alle recherchierbaren Ansprüche die Recherche ohne einen Arbeitsaufwand durchgeführt werden konnte, der zusätzliche Recherchegebühr gerechtfertigt hätte, hat die Behörde nicht zur Zahlung solcher Gebühren aufgefordert.

3. ☐ Da der Anmelder nur einige der erforderlichen zusätzlichen Recherchegebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht nur auf die Ansprüche, für die Gebühren entrichtet worden sind, nämlich auf die Ansprüche Nr.

4. ☐ Der Anmelder hat die erforderlichen zusätzlichen Recherchegebühren nicht rechtzeitig entrichtet. Dieser internationale Recherchenbericht beschränkt sich daher auf die in den Ansprüchen zuerst erwähnte Erfindung;; diese ist in folgenden Ansprüchen erfasst:

Bemerkungen hinsichtlich eines Widerspruchs

- ☐ Der Anmelder hat die zusätzlichen Recherchegebühren unter Widerspruch entrichtet und die gegebenenfalls erforderliche Widerspruchsgebühr gezahlt.
- ☐ Die zusätzlichen Recherchegebühren wurden vom Anmelder unter Widerspruch gezahlt, jedoch wurde die entsprechende Widerspruchsgebühr nicht innerhalb der in der Aufforderung angegebenen Frist entrichtet.
- ☒ Die Zahlung der zusätzlichen Recherchegebühren erfolgte ohne Widerspruch.

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES INV. G06F9/50 G06N3/0464 G06N3/088 ADD.		
Nach der Internationalen Patentklassifikation (IPC) oder nach der nationalen Klassifikation und der IPC		
B. RECHERCHIERTE GEBIETE Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole) G06F G06N		
Recherchierte, aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen		
Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe) EPO-Internal, WPI Data		
C. ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	US 2009/254914 A1 (SUNDARESAN NAGARAJ [US] ET AL) 8. Oktober 2009 (2009-10-08)	1, 2, 8-11
Y	Absatz [0007] – Absätze [0014], [0035] Abbildung 1	3
Y	EP 0 652 513 A1 (IBM [US]) 10. Mai 1995 (1995-05-10) Seite 2, Zeile 39 – Seite 3, Zeile 6	3
X	WO 2021/050790 A1 (C3 AI INC [US]) 18. März 2021 (2021-03-18) Absatz [0010] – Absatz [0020] Absatz [0053] – Absatz [0056] Abbildung 1	4-7
	----- -/-	
☒ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen ☒ Siehe Anhang Patentfamilie		
* Besondere Kategorien von angegebenen Veröffentlichungen : "A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist "E" frühere Anmeldung oder Patent, die bzw. das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist "L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt) "O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht "P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist "T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist "X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden "Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist "&" Veröffentlichung, die Mitglied derselben Patentfamilie ist		
Datum des Abschlusses der internationalen Recherche		Absendedatum des internationalen Recherchenberichts
6. März 2023		28/04/2023
Name und Postanschrift der Internationalen Recherchenbehörde Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Bevollmächtigter Bediensteter Mühlenbrock, Martin

C. (Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	Jinwon An ET AL: "SNU Data Mining Center 2015-2 Special Lecture on IE Variational Autoencoder based Anomaly Detection using Reconstruction Probability", , 27. Dezember 2015 (2015-12-27), XP055688241, Gefunden im Internet: URL: http://dm.snu.ac.kr/static/docs/TR/SNU DM-TR-2015-03.pdf das ganze Dokument -----	4-7

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP2022/065756

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
US 2009254914 A1	08-10-2009	KEINE	
EP 0652513 A1	10-05-1995	EP 0652513 A1	10-05-1995
		JP 3414520 B2	09-06-2003
		JP H07182185 A	21-07-1995
		US 5437032 A	25-07-1995
WO 2021050790 A1	18-03-2021	AU 2020345874 A1	31-03-2022
		CA 3154145 A1	18-03-2021
		CN 114730390 A	08-07-2022
		EP 4028963 A1	20-07-2022
		JP 2022548227 A	17-11-2022
		KR 20220062329 A	16-05-2022
		US 2022260988 A1	18-08-2022
		WO 2021050790 A1	18-03-2021

WEITERE ANGABEN**PCT/ISA/ 210**

Die internationale Recherchenbehörde hat festgestellt, dass diese internationale Anmeldung mehrere (Gruppen von) Erfindungen enthält, nämlich:

1. Ansprüche: 1-3, 8-11

Verfahren zur Bearbeitung einer Menge an Rechenaufgaben
durch ein System

2. Ansprüche: 4-7

Verfahren zur Anomalie-Erkennung beim Betrieb eines
technischen Geräts
