(19) Deutsches Patent- und Markenamt

(10) **DE 10 2018 216 310 A1** 2020.03.26

(12) # Offenlegungsschrift

(21) Aktenzeichen: **10 2018 216 310.3**
(22) Anmeldetag: **25.09.2018**
(43) Offenlegungstag: **26.03.2020**

(51) Int Cl.: ***H04L 9/06*** (2006.01)
G06N 3/02 (2006.01)
G06F 21/55 (2013.01)

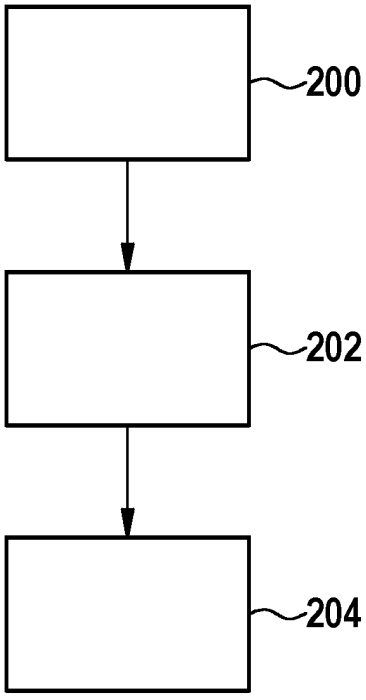
(71) Anmelder:
Robert Bosch GmbH, 70469 Stuttgart, DE

(72) Erfinder:
Schorn, Christoph, 71229 Leonberg, DE; Hettwer, Benjamin, 70180 Stuttgart, DE; Gehrer, Stefan, 99099 Erfurt, DE

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **Verfahren und Vorrichtung zum Verarbeiten von Daten eines technischen Systems**

(57) Zusammenfassung: Verfahren zum Verarbeiten von Daten eines technischen Systems, wobei das technische System wenigstens eine Implementierung eines kryptografischen Verfahrens aufweist, wobei das Verfahren zum Verarbeiten von Daten des technischen Systems die folgenden Schritte aufweist: Modellieren wenigstens eines Teils des technischen Systems mittels wenigstens eines, insbesondere künstlichen, neuronalen Netzes, Trainieren des neuronalen Netzes unter Verwendung bekannter erster Daten und wenigstens einer physikalischen Größe des technischen Systems, wodurch ein trainiertes neuronales Netz erhalten wird, Ermitteln von einen Betrieb der wenigstens einen Implementierung charakterisierenden Seitenkanalinformationen in Abhängigkeit des trainierten neuronalen Netzes.



Beschreibung

Stand der Technik

[0001] Die Offenbarung betrifft ein Verfahren zum Verarbeiten von Daten eines technischen Systems, wobei das technische System wenigstens eine Implementierung eines kryptografischen Verfahrens aufweist. Die Offenbarung betrifft ferner eine Vorrichtung zum Verarbeiten von Daten eines technischen Systems.

[0002] Seitenkanalangriffe gegen Implementierungen von kryptografischen Verfahren in einem technischen System wie z.B. einem Mikrocontroller sind bekannt. Seitenkanalgriffe beschreiben eine Klasse von Techniken mittels derer vertrauliche Parameter einer kryptografischen Implementierung extrahiert werden können (z.B. Schlüsselmaterial eines geheimen kryptografischen Schlüssels). Dabei werden bestimmte Seitenkanalinformationen wie z.B. der Stromverbrauch, ein Zeitverhalten usw. analysiert um das von dem kryptografischen Verfahren verwendete Geheimnis zu berechnen bzw. Informationen hierüber zu erlangen.

[0003] Um Seitenkanalangriffe zu erschweren werden bei konventionellen Systemen bestimmte Gegenmaßnahmen in kryptographische Implementierungen eingebaut. Diese können grob in zwei Klassen eingeteilt werden: 1. Masking, 2. Hiding. Mittels Masking werden Zwischenergebnisse, die während der Berechnung einer kryptografischen Operation anfallen, randomisiert (maskiert), um die Datenabhängigkeit z.B. des Stromverbrauchs vom geheimen Schlüssel zu unterbrechen. Beim Hiding wird versucht, die Charakteristika des Stromverbrauchs so zu verändern, dass alle Operation entweder dieselbe Menge an Strom verbrauchen oder dass der Stromverbrauch für alle Operationen zufällig ist (jedoch nicht datenabhängig). Um zu evaluieren, ob eine oder mehrere Gegenmaßnahmen gegen Seitenkanalattacken richtig implementiert wurden, werden neben praktischen Angriffen in der Regel statistische Methoden wie Test Vector Leakage Assessment (TVLA) durchgeführt. TVLA und verwandte Methoden liefern jedoch nur eine Aussage darüber, ob eine Abhängigkeit zwischen Stromverbrauch und verarbeiteten Daten messbar ist. Sie geben einem Entwickler jedoch keine Aussage, wie stark die Leckage (englisch: ‚leakage') von Seitenkanalinformationen ist und wodurch sie verursacht wird. Auch ist nicht sichergestellt, ob die Leckage auch für einen praktischen Angriff ausgenutzt werden kann.

Offenbarung der Erfindung

[0004] Es ist Aufgabe der vorliegenden Erfindung, ein Verfahren und eine Vorrichtung der eingangs genannten Art dahingehend zu verbessern, dass ein Gebrauchsnutzen gesteigert wird und die genannten Nachteile des Stands der Technik vermindert oder vermieden werden.

[0005] Bevorzugte Ausführungsformen beziehen sich auf ein Verfahren zum Verarbeiten von Daten eines technischen Systems, wobei das technische System wenigstens eine Implementierung eines kryptografischen Verfahrens aufweist, wobei das Verfahren zum Verarbeiten von Daten des technischen Systems die folgenden Schritte aufweist: Modellieren wenigstens eines Teils des technischen Systems mittels wenigstens eines, insbesondere künstlichen, neuronalen Netzes, Trainieren des neuronalen Netzes unter Verwendung bekannter erster Daten und wenigstens einer physikalischen Größe des technischen Systems, wodurch ein trainiertes neuronales Netz erhalten wird, Ermitteln von einen Betrieb der wenigstens einen Implementierung charakterisierenden Seitenkanalinformationen in Abhängigkeit des trainierten neuronalen Netzes. Das Prinzip gemäß den Ausführungsformen ermöglicht vorteilhaft, einen Betrieb der wenigstens eine Implementierung charakterisierende Seitenkanalinformationen **SCI** in Abhängigkeit des trainierten neuronalen Netzes **NN'** zu ermitteln.

[0006] Bei bevorzugten Ausführungsformen kann das künstliche neuronale Netz beispielsweise als CNN (convolutional neural network) ausgebildet sein.

[0007] Bei weiteren bevorzugten Ausführungsformen ist vorgesehen, dass das kryptografische Verfahren ein Verschlüsselungsverfahren ist, insbesondere vom AES-Typ. Für Details zu AES (advanced encryption standard) s. z.B. https://doi.org/10.6028/NIST.FIPS.197. Bei weiteren bevorzugten Ausführungsformen kann das kryptografische Verfahren auch ein anders Verfahren als das beispielhaft genannte Verschlüsselungsverfahren umfassen, z.B. eine Hashwertbildung oder dergleichen.

[0008] Bei weiteren bevorzugten Ausführungsformen ist vorgesehen, dass die bekannten ersten Daten Eingangsdaten für die Implementierung des kryptografischen Verfahrens sind, und/oder dass die wenigstens eine physikalische Größe des technischen Systems einen Stromverbrauch und/oder elektromagnetische Abstrahlungen umfasst.

[0009] Bei weiteren bevorzugten Ausführungsformen ist vorgesehen, dass der Schritt des Ermittelns ein Visualisieren von wenigstens einer Komponente des trainierten neuronalen Netzes und/oder aus dem trainierten neuronalen Netz ableitbaren Daten umfasst. Dadurch können kritische Operationen in Hinsicht auf Seitenkanalangriffe effizient evaluiert werden.

[0010] Bei weiteren bevorzugten Ausführungsformen ist vorgesehen, dass der Schritt des Ermittelns

die Ermittlung eines Zeitpunkts bei der Ausführung des kryptografischen Verfahrens durch die Implementierung umfasst, zu dem, insbesondere bezogen auf andere Zeitbereiche der Ausführung des kryptografischen Verfahrens, eine vergleichsweise große Leckage von Seitenkanalinformationen auftritt.

[0011] Bei weiteren bevorzugten Ausführungsformen ist vorgesehen, dass das trainierte neuronale Netz eine Ausgabefunktion S aufweist, wobei mittels Fehlerrückführung (englisch: backpropagation), eine Ableitung Sc in Bezug auf eine Seitenkanalmesswertreihe (englisch: side channel trace), ermittelt wird, wobei c eine Ausgabeklasse des trainierten neuronalen Netzes für einen korrekten Ausgabewert repräsentiert, wobei insbesondere die vorstehend genannten Schritte für eine Vorgebbare Anzahl von Eingangs-Seitenkanalmesswertreihen wiederholt werden.

[0012] Weitere bevorzugte Ausführungsformen beziehen sich auf eine Vorrichtung zum Verarbeiten von Daten eines technischen Systems, wobei das technische System wenigstens eine Implementierung eines kryptografischen Verfahrens aufweist, wobei die Vorrichtung zur Ausführung der folgenden Schritte ausgebildet ist: Modellieren wenigstens eines Teils des technischen Systems mittels wenigstens eines, insbesondere künstlichen, neuronalen Netzes, Trainieren des neuronalen Netzes unter Verwendung bekannter erster Daten und wenigstens einer physikalischen Größe des technischen Systems, wodurch ein trainiertes neuronales Netz erhalten wird, Ermitteln von einen Betrieb der wenigstens einen Implementierung charakterisierenden Seitenkanalinformationen in Abhängigkeit des trainierten neuronalen Netzes.

[0013] Weitere bevorzugte Ausführungsformen sehen vor, dass die Vorrichtung zur Ausführung des Verfahrens gemäß den Ausführungsformen ausgebildet ist.

[0014] Weitere bevorzugte Ausführungsformen beziehen sich auf eine Verwendung des Verfahrens gemäß den Ausführungsformen und/oder der Vorrichtung gemäß den Ausführungsformen zur Bewertung einer Implementierung eines kryptografischen Verfahrens in Bezug auf eine Leckage von Seitenkanalinformationen.

[0015] Weitere Merkmale, Anwendungsmöglichkeiten und Vorteile der Erfindung ergeben sich aus der nachfolgenden Beschreibung von Ausführungsbeispielen der Erfindung, die in den Figuren der Zeichnung dargestellt sind. Dabei bilden alle beschriebenen oder dargestellten Merkmale für sich oder in beliebiger Kombination den Gegenstand der Erfindung, unabhängig von ihrer Zusammenfassung in den Patentansprüchen oder deren Rückbeziehung sowie unabhängig von ihrer Formulierung bzw. Darstellung in der Beschreibung bzw. in der Zeichnung.

[0016] In der Zeichnung zeigt:

Fig. 1 schematisch ein Blockdiagramm eines technischen Systems gemäß einer Ausführungsform,

Fig. 2 schematisch ein vereinfachtes Flussdiagramm eines Verfahrens gemäß bevorzugter Ausführungsformen,

Fig. 3 schematisch ein neuronales Netz gemäß einer Ausführungsform,

Fig. 4 schematisch ein Blockdiagramm gemäß weiterer Ausführungsformen,

Fig. 5 schematisch eine Visualisierung gemäß weiterer Ausführungsformen, und

Fig. 6 schematisch ein Blockdiagramm einer Vorrichtung gemäß bevorzugter Ausführungsformen.

[0017] **Fig. 1** zeigt schematisch ein Blockdiagramm eines technischen Systems **100** gemäß einer Ausführungsform. Das technische System **100** weist wenigstens eine Implementierung **110** eines kryptografischen Verfahrens auf. Beispielsweise kann die Implementierung **110** eine Recheneinheit **112** und eine zugehörige Software **114**, die den Betrieb der Recheneinheit **112** steuert, umfassen.

[0018] Bei weiteren bevorzugten Ausführungsformen, die aus Gründen der Übersichtlichkeit nicht abgebildet sind, ist auch denkbar, dass die Implementierung **110** rein hardwarebasiert ausgebildet ist, beispielsweise mittels eines anwendungsspezifischen integrierten Schaltkreises (ASIC).

[0019] Beispielhaft wird für die weitere Beschreibung davon ausgegangen, dass es sich bei dem kryptografischen Verfahren um ein Verschlüsselungsverfahren basierend auf dem advanced encryption standard, AES, handelt. Für Details zu AES (advanced encryption standard) s. z.B. https://doi.org/10.6028/NIST.FIPS.197.

[0020] Dem Verschlüsselungsverfahren bzw. seiner Implementierung **110** in dem technischen System **100** werden bei bevorzugten Ausführungsformen Eingangsdaten **ED** zugeführt, bei denen es sich beispielsweise um einen zu verschlüsselten Klartext und wenigstens einen (geheimen) kryptografischen Schlüssel handelt. Das Verschlüsselungsverfahren ermittelt hieraus Ausgangsdaten **AD**, bei denen es sich beispielsweise um mittels des kryptografischen Schlüssels AES-verschlüsselte Eingangsdaten **ED** handelt. Während der Ausführung des Verschlüsselungsverfahrens mittels der Implementierung **110** sind im Bereich des technischen Systems **100** in der

Regel ein oder mehrere physikalische Größen **O** ermittelbar, deren zeitlicher Verlauf zumindest zeitweise von den während der Ausführung des Verschlüsselungsverfahrens verarbeiteten Daten abhängt, insbesondere von den Eingangsdaten **ED** und/oder dem kryptografischen Schlüssel und/oder den Ausgangsdaten **AD**. Beispielsweise kann es sich bei den physikalischen Größen **O** um eine elektrische Energieaufnahme bzw. einen zeitlichen Verlauf eines Stroms handeln, den das technische System **100** bzw. die Implementierung **110** aufnimmt. Alternativ oder ergänzend kann es sich bei den physikalischen Größen **O** auch um einen zeitlichen Verlauf von elektrischen und/oder magnetischen Feldern oder dergleichen handeln. Diese physikalischen Größen **O** sind ggf. durch einen Angreifer ermittelbar und ermöglichen die Ausführung von Seitenkanalattacken auf die Implementierung **110**.

[0021] Bei bevorzugten Ausführungsformen ist der nachstehend beispielhaft unter Bezugnahme auf das Flussdiagramm gemäß **Fig. 2** beschriebene Ablauf vorgeschlagen, der es ermöglicht, effizient Informationen über die konkrete Implementierung **110** des kryptografischen Verfahrens mittels des Systems **100** zu erlangen, insbesondere in Bezug auf eine mögliche Leckage (englisch: leakage) von Seitenkanalinformationen („SCA leakage“).

[0022] In einem ersten Schritt **200** wird wenigstens ein Teil des technischen Systems **100** mittels wenigstens eines, insbesondere künstlichen, neuronalen Netzes modelliert. Beispielhaft wird in Schritt **200** demnach ein neuronales Netz bereitgestellt, das zur Nachbildung des wenigstens einen Teils des technischen Systems **100**, insbesondere bezüglich der Implementierung **110** des kryptografischen Verfahrens, verwendbar ist. **Fig. 3** zeigt hierzu beispielhaft ein vereinfachtes Blockdiagramm. Bei dem neuronalen Netz **NN** kann es sich beispielsweise um ein tiefes („deep“, also vergleichsweise viele Schichten aufweisendes) neuronales Netz handeln, insbesondere vom CNN (convolutional neural network)-Typ.

[0023] In einem zweiten Schritt **202** (vgl. **Fig. 2**) wird das neuronale Netz **NN** unter Verwendung bekannter erster Daten **D1** (**Fig. 3**) und wenigstens einer physikalischen Größe **G1** des technischen Systems **100** trainiert, wodurch ein trainiertes neuronales Netz erhalten wird. Die bekannten ersten Daten **D1** umfassen beispielsweise den zu verschlüsselten Klartext. Die wenigstens eine physikalische Größe **G1** umfasst beispielsweise die vorstehend unter Bezugnahme auf **Fig. 1** beschriebene Größe **O** (beispielsweise zeitlicher Verlauf des von dem technischen System **100** bzw. der Implementierung **110** aufgenommenen Stroms). Bevorzugt kann dem neuronalen Netz **NN** für das Training **202** (**Fig. 2**) zusätzlich der für die Verschlüsselung verwendete kryptografische Schlüssel und/oder ein von dem Schlüssel abhängiger Zwischenwert, vergleiche das Bezugszeichen v aus **Fig. 3**, zugeführt werden. Bei besonders bevorzugten Ausführungsformen hat der Schritt **202** des Trainings zum Gegenstand, das neuronale Netz **NN** dahingehend zu trainieren, dass es die vergleichsweise komplexe Funktion der Implementierung **110**, also die Ausführung eines kryptografischen Verfahrens in dem spezifischen technischen System **100**, nachbildet. Hierfür können bei weiteren bevorzugten Ausführungsformen an sich bekannte Technologien des deep learning verwendet werden. Grundlegende Aspekte des deep learning bei neuronalen Netzwerken sind beispielsweise in http://www.deeplearningbook.org/ von Goodfellow et al. (@book{Goodfellow-et-al-2016, title={Deep Learning}, author={Ian Goodfellow and Yoshua Bengio and Aaron Courville}, publisher={MIT Press}, note={\url{http://www.deeplearningbook.org}}, year={2016} }) beschrieben.

[0024] **Fig. 4** zeigt schematisch ein vereinfachtes Blockdiagramm von Aspekten des vorstehend unter Bezugnahme auf **Fig. 2** beschriebenen Ablaufs. Das gemäß bevorzugten Ausführungsformen bereitgestellte neuronale Netz **NN** wird gemäß Schritt **202** aus **Fig. 2** trainiert, wodurch das trainierte neuronale Netz **NN'** erhalten wird, vergleiche den Blockpfeil **A1** aus **Fig. 4**. Das trainierte neuronale Netz **NN'** ist bei bevorzugten Ausführungsformen mit anderen Worten in der Lage, zugeführten Eingangsdaten wie beispielsweise einem zu verschlüsselnden Klartext und der physikalischen Größe **G1** (beispielsweise ein zeitlicher Verlauf eines Stroms, der durch die Implementierung **110** bei der Verarbeitung des zu verschlüsselten Klartext aufgenommen wird) einen kryptografischen Schlüssel zuzuordnen. Auf diese Weise kann das trainierte neuronale Netz **NN'** beispielsweise zur Ausführung von Seitenkanalattacken auf die Implementierung **110** genutzt werden.

[0025] Bei besonders bevorzugten Ausführungsformen ist zusätzlich zu den vorstehend bereits unter Bezugnahme auf **Fig. 2** beschriebenen Schritten **200**, **202** jedoch vorgesehen, in Abhängigkeit des trainierten neuronalen Netzes **NN'** einen Betrieb der wenigstens einen Implementierung **110** charakterisierende Seitenkanalinformationen **SCI** (**Fig. 4**) zu ermitteln, vergleiche Schritt **204** gemäß **Fig. 2**.

[0026] Dem Prinzip der Ausführungsformen folgend ist erkannt worden, dass das trainierte neuronale Netz **NN'** wichtige Informationen über eine (mögliche) Leckage von Seitenkanalinformationen, beispielsweise in Form der physikalischen Größe **G1**, der Implementierung **110** aufweist. Diese Informationen können vorteilhaft im Rahmen des Schritts **204** ausgewertet werden, beispielsweise um die Leckage von Seitenkanalinformationen qualitativ und/oder quantitativ zu bewerten. Damit wird vorteilhaft eine Untersuchung der Implementierung **110** des krypto-

grafischen Verfahrens durch das technische System **100** ermöglicht, die beispielsweise in einem Prozess einer Entwicklung der Implementierung **110** eines kryptografischen Verfahrens genutzt werden kann. Mit den Erkenntnissen, die gemäß dem Prinzip der Ausführungsformen unter Verwendung des trainierten neuronalen Netzes **NN'** erhalten werden können, können beispielsweise zukünftig zu implementierende kryptografische Verfahren bzw. die entsprechenden Implementierungen **110** dahingehend optimiert werden, dass sie eine geringere Leckage von Seitenkanalinformationen aufweisen bzw. die Leckage so beeinflusst werden kann, dass erfolgreiche Seitenkanalattacken rechnerisch aufwendiger, also erschwert, werden.

[0027] Beispielsweise kann bei weiteren bevorzugten Ausführungsformen eine geplante Implementierung **110** mittels des vorstehend unter Bezugnahme auf **Fig. 2**, **Fig. 4** beschriebenen Ablaufs untersucht werden. Gegebenenfalls kann, in Abhängigkeit von dem Ausgang des Schritts **204**, die geplante Implementierung **110** sodann ein oder mehrmalig modifiziert und der betreffende Ablauf wiederholt werden, bevorzugt so lange, bis die durch Schritt **204** ermittelbaren Seitenkanalinformationen **SCI** den Erwartungen an die geplante Implementierung **110** (beispielsweise einer Härtung gegen Seitenkanalattacken) entsprechen.

[0028] Bei weiteren bevorzugten Ausführungsformen ist vorgesehen, dass der Schritt des Ermittelns **204** (**Fig. 2**) ein Visualisieren von wenigstens einer Komponente des trainierten neuronalen Netzes **NN'** und/oder aus dem trainierten neuronalen Netz **NN'** ableitbaren Daten umfasst. Dadurch können kritische Operationen in Hinsicht auf Seitenkanalangriffe besonders effizient evaluiert werden.

[0029] Bei weiteren bevorzugten Ausführungsformen kann das Visualisieren die Anwendung wenigstens einer der folgenden Techniken umfassen:

a) Erstellung einer eindimensionalen (1-D) Heatmap basierend auf dem Prinzip der sogenannten „Saliency Maps“, beschrieben in: K. Simonyan, A. Vedaldi, and A. Zisserman. „Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps.“ ArXiv e-prints, December 2013, für Details vgl. https://arxiv.org/abs/1312.6034,

b) M. D Zeiler and R. Fergus. „Visualizing and Understanding Convolutional Networks“. ArXiv e-prints, November 2013, https://arxiv.org/abs/1311.2901,

c) G. Montavon, S. Bach, A. Binder, W. Samek, and K.-R. Müller. „Explaining NonLinear Classification Decisions with Deep Taylor Decomposition.“ ArXiv e-prints, December, 2015, https://arxiv.org/abs/1512.02479,

d) R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. „Grad-cam: Visual explanations from deep networks via gradient-based localization“, 2017 IEEE International Conference on Computer, Vision (ICCV), pages 618- 626, Oct 2017.

[0030] Bei weiteren bevorzugten Ausführungsformen ist vorgesehen, dass der Schritt **204** des Ermittelns die Ermittlung eines Zeitpunkts bei der Ausführung des kryptografischen Verfahrens durch die Implementierung **110** (**Fig. 1**) umfasst, zu dem, insbesondere bezogen auf andere Zeitbereiche der Ausführung des kryptografischen Verfahrens, eine vergleichsweise große Leckage von Seitenkanalinformationen auftritt. Daraus kann vorteilhaft abgeleitet werden, welche Daten bezüglich der physikalischen Größe **G1** für das trainierte neuronale Netz **NN'** ausschlaggebend sind, um auf ein entsprechendes Klassifikationsergebnis zu gelangen. Mit anderen Worten wird bei dieser Ausführungsform die Tatsache ausgenutzt, dass aus einer besonderen Aktivität bestimmter Teile des trainierten neuronalen Netzes **NN'** sowohl auf die Art der betrachteten Daten als auch beispielsweise auf einen Zeitpunkt innerhalb eines Zeitverlaufs der physikalischen Größe geschlossen werden kann, der für das betreffende Klassifikationsergebnis von Bedeutung ist.

[0031] Bei weiteren bevorzugten Ausführungsformen ist vorgesehen, dass das trainierte neuronale Netz **NN'** eine Ausgabefunktion S aufweist, wobei mittels Fehlerrückführung (englisch: backpropagation) eine Ableitung Sc in Bezug auf eine Seitenkanalmesswertreihe (englisch: side channel trace) ermittelt wird, wobei c eine Ausgabeklasse des trainierten neuronalen Netzes für einen korrekten Ausgabewert repräsentiert, wobei insbesondere die vorstehend genannten Schritte für eine Vorgebbare Anzahl von Eingangs-Seitenkanalmesswertreihen wiederholt werden. Auf diese Weise kann eine sog. „Saliency Trace“ ermittelt werden, die z.B. in Form der 1-D Heatmap HM gemäß **Fig. 5** grafisch darstellbar ist, die einen zeitlichen Verlauf (horizontale Achse t) eines Maßes für eine Seitenkanalleckage angibt. Dabei markiert der erste Bereich **B1** der Heatmap HM Zeitpunkte mit einem hohen Grad an Seitenkanal-Leckage (englisch: side channel leakage), wohingegen die weiteren Bereiche **B2**, **B3** einen geringeren Grad an Seitenkanal-Leckage angeben. Details zu dieser Ausführungsform sind weiter unten unter Bezugnahme des dort beschriebenen „Algorithmus 1“ aufgeführt.

[0032] Weitere bevorzugte Ausführungsformen beziehen sich auf eine Vorrichtung **1000**, vgl. das Blockdiagramm aus **Fig. 6**, zum Verarbeiten von Daten eines technischen Systems **100** (**Fig. 1**), wobei das technische System **100** wenigstens eine Implementierung **110** eines kryptografischen Verfahrens auf-

weist, wobei die Vorrichtung **1000** (**Fig. 6**) zur Ausführung der folgenden Schritte ausgebildet ist: Modellieren **200** (**Fig. 2**) wenigstens eines Teils des technischen Systems **100** mittels wenigstens eines, insbesondere künstlichen, neuronalen Netzes **NN** (**Fig. 3**), Trainieren **202** (**Fig. 2**) des neuronalen Netzes **NN** unter Verwendung bekannter erster Daten **D1** und wenigstens einer physikalischen Größe **G1** des technischen Systems **100**, wodurch ein trainiertes neuronales Netz **NN'** (**Fig. 4**) erhalten wird, Ermitteln **204** (**Fig. 2**) von einen Betrieb der wenigstens einen Implementierung **110** charakterisierenden Seitenkanalinformationen **SCI** (**Fig. 4**) in Abhängigkeit des trainierten neuronalen Netzes **NN'**.

[0033] Weitere bevorzugte Ausführungsformen sehen vor, dass die Vorrichtung **1000** (**Fig. 6**) zur Ausführung des Verfahrens gemäß den Ausführungsformen ausgebildet ist.

[0034] Weitere bevorzugte Ausführungsformen beziehen sich auf eine Verwendung des Verfahrens gemäß den Ausführungsformen und/oder der Vorrichtung **1000** gemäß den Ausführungsformen zur Bewertung einer Implementierung **110** eines kryptografischen Verfahrens in Bezug auf eine Leckage von Seitenkanalinformationen.

[0035] Bei bevorzugten Ausführungsformen weist die Vorrichtung **1000** eine Recheneinheit **1002** auf, die beispielsweise einen Mikroprozessor und/oder einen Mikrocontroller und/oder einen anwendungsspezifischen integrierten Schaltkreis (ASIC) und/oder einen programmierbaren Logikbaustein (beispielsweise FPGA, field programmable gate array) und/oder einen digitalen Signalprozessor (DSP) und/oder einen Tensorprozessor oder eine Kombination hieraus aufweist. Der Recheneinheit **1002** ist eine Speichereinheit **1004** zugeordnet, die bevorzugt einen flüchtigen Speicher **1004a**, insbesondere Arbeitsspeicher (RAM), und/oder einen nichtflüchtigen Speicher **1004b**, insbesondere Flash-EEPROM, aufweist, und die zur zumindest zeitweisen Speicherung von wenigstens einem Computerprogramm **PRG1**, **PRG2** (und/oder von Daten) für die Recheneinheit **1002** ausgebildet ist. Beispielsweise kann das erste Computerprogramm **PRG1** die Ausführung des vorstehend beispielhaft unter Bezugnahme auf **Fig. 2** beschriebenen Verfahrens durch die Vorrichtung **1000** bzw. die Recheneinheit **1002** steuern. Beispielsweise kann das zweite Computerprogramm **PRG2** das neuronale Netz **NN**, **NN'** bzw. seinen Betrieb steuern. Die Parameter des neuronalen Netzes **NN**, welche beispielsweise während des Trainings **202** gelernt bzw. modifiziert werden, können zumindest zeitweise in der Speichereinheit **1004** gespeichert werden. Bei weiteren bevorzugten Ausführungsformen kann die Vorrichtung **1000** über eine Schnittstelle **1006** verfügen, die insbesondere zur Eingabe der wenigstens einen physikalischen Größe **G1**, **O**, also z.B. zur Ermittlung von Seitenkanalinformationen in Form von elektromagnetischen Abstrahlungen und/oder eines zeitlichen Verlauf eines durch die Implementierung **110** aufgenommenen Stroms, ausgebildet ist.

[0036] Das Prinzip gemäß den Ausführungsformen ermöglicht vorteilhaft, einen Betrieb der wenigstens eine Implementierung **110** (**Fig. 1**) charakterisierende Seitenkanalinformationen **SCI** in Abhängigkeit des trainierten neuronalen Netzes **NN'** zu ermitteln. Bei besonders bevorzugten Ausführungsformen werden Visualisierungstechniken für neuronale Netze auf das trainierte neuronale Netz **NN'** angewandt, um beispielsweise Zeitpunkte einer größten Seitenkanal-Leckage (bezogen auf die Ausführung des kryptografischen Verfahrens mittels der Implementierung **110**) einfach identifizierbar zu machen.

[0037] Bei weiteren bevorzugten Ausführungsformen kann der in **Fig. 2** beschriebene Ablauf auch zur Ausführung von Seitenkanalattacken auf Implementierungen **110** von kryptografischen Verfahren genutzt werden.

[0038] Nachstehend sind weitere Aspekte und Ausführungsformen beschrieben.

[0039] Bei weiteren Ausführungsformen ist erkannt worden, dass Seitenkanalangriffe mit (tiefen) neuronalen Netzen **NN** besonders mächtig sind. Durch ihre Fähigkeit, hochkomplexe Funktionen zu approximieren, können auch geschützte Implementierungen **110** effektiv und effizient angegriffen werden. Im Gegensatz zu konventionellen Angriffstechniken sind sie des Weiteren in der Lage die für die Klassifikation relevanten Datenpunkte (beispielsweise ausgewählte Abtastwerte des zeitlichen Verlaufs des Stroms) selbständig zu extrahieren. Bei weiteren bevorzugten Ausführungsformen ist insbesondere keine aufwendige Vorverarbeitung nötig.

[0040] Besonders bevorzugte Ausführungsformen beziehen sich auf die Nutzung eines wie vorstehend unter Bezugnahme auf **Fig. 2** trainierten neuronalen Netzes **NN'** als Werkzeug, um die mögliche Effektivität von etwaigen Gegenmaßnahmen (z.B. Härtung der Implementierung **110** gegen mögliche Seitenkanalattacken, z.B. mittels Masking und/oder Hiding) zu evaluieren. Durch Visualisierung, insbesondere Generierung von 1-D Heatmaps HM (**Fig. 5**, z.B. während oder basierend auf der Ermittlung **204**) gemäß weiteren bevorzugten Ausführungsformen wird es möglich, genau solche Zeitpunkte **B1** in Seitenkanalmessungen **O**, **G1** zu identifizieren, die verantwortlich für ein mögliches Gelingen eines Seitenkanal-Angriffs waren ("Wohin „schaut“ das trainierte neuronale Netz **NN'** im Trace (Messwertreihe von wenigstens einer physikalischen Größe **O**, **G1**), um auf das Klassifikationsergebnis zu kommen"). Da dem Entwickler der Implementierung **110** (**Fig. 1**) bekannt ist, welche

Operationen mit entsprechenden Daten bei der Ausführung des kryptografischen Verfahrens zu welchem Zeitpunkt ausgeführt werden, ist er nun in der Lage Korrekturen bzw. Verbesserungen an der Implementierung **110** vorzunehmen, um die SCA leakage (Seitenkanal-Leckage) zu verringern.

[0041] Nachfolgend ist ein Ablauf eines Seitenkanalangriffes mit Hilfe von Deep Learning (**DL**) gemäß weiteren Ausführungsformen dargestellt. Der kryptografische Algorithmus wird auf einem Gerät (z.B. System **100** aus **Fig. 1**) ausgeführt, dessen Seitenkanalinformationen **O** genutzt werden sollen, um die Implementierung **110** zu brechen. Bei weiteren Ausführungsformen läuft der Angriff in zwei Phasen ab:

1. „Profiling“ Phase: Trainieren des neuronalen Netzes **NN** (**Fig. 3**). Hierfür wird bei weiteren Ausführungsformen die Implementierung **110** auf einem Profiling Device, z.B. realisiert in Form des technischen Systems **100** nach **Fig. 1**, mit einer Serie von bekannten Klartexten m und i.d.R. (geheimen) kryptographischen Schlüsseln k ausgeführt. Parallel dazu wird bei weiteren Ausführungsformen die wenigstens eine physikalische Größe **O** bzw. **G1**, bspw. Stromverbrauch oder elektromagnetische Abstrahlung, gemessen. Das neuronale Netz **NN** wird nun trainiert mit den physikalischen Seitenkanalmessungen („Traces“, abgeleitet aus den Größen **O**, **G1**) und eventuellen Zusatzinformationen (wie z.B. dem Klartext m) als Eingangsgröße und dem verwendeten Schlüssel v oder einem Zwischenwert der vom Schlüssel abhängt, als „Ground Truth“. Dadurch wird das trainierte neuronale Netz **NN'** erhalten.

2. „Attack“ Phase: Angreifen eines Gerätes mit derselben Implementierung **110** („Target Device“) mit dem Ziel des Herausfindens des (geheimen) kryptographischen Schlüssels. Gegeben sind bei weiteren Ausführungsformen bspw. der Klartext und zugehörige gemessene physikalische Größen **G1**, **O** („Traces“). Das trainierte neuronale Netz **NN'** liefert bei weiteren Ausführungsformen eine Wahrscheinlichkeitsverteilung v* über alle möglichen Werte der angegriffenen Operation. Mithilfe z.B. der Maximum-Likelihood Methode kann bei weiteren Ausführungsformen daraus der geheime Schlüssel k bestimmt werden. Alternativ oder ergänzend zu dem eigentlichen Angriff auf das Target Device kann das in der Profiling Phase erstellte Device Model (Modellierung der Implementierung **110** mit dem neuronalen Netz **NN**) bei weiteren Ausführungsformen auch genutzt werden, um die Zeitpunkte mit der größten Seitenkanal-Leckage zu bestimmen. Dazu können bei weiteren Ausführungsformen die vorstehend bereits genannten Methoden zur Visualisierung von neuronalen Netzen (oder von Daten hiervon) genutzt werden, s. die Punkte a) bis d) weiter oben.

[0042] Bei weiteren Ausführungsformen werden in 1-D Seitenkanal Traces (Zeitverlauf einer im Rahmen einer Seitenkanalattacke ermittelbaren physikalischen Größe **G1**, **O**) solche Zeitpunkte visualisiert, welche entscheidend sind um im Rahmen der betreffenden Seitenkanalattacke den korrekten geheimen Schlüssel zu bestimmen. Eine Methode hierfür gemäß weiterer Ausführungsformen ist nachfolgend als „Algorithmus 1“ dargestellt. Der Algorithmus **1** beschreibt im Wesentlichen folgenden Ablauf: Für ein trainiertes neuronales Netz **NN'** mit einer Ausgabefunktion S wird mithilfe von Backpropagation (Fehlerrückführung) die Ableitung Sc in Bezug auf einen bestimmten Seitenkanal-Trace X berechnet, wobei c die Ausgabeklasse des trainierten neuronalen Netzes **NN'** für den korrekten Schlüssel repräsentiert. Dies wird bevorzugt für ein Sammlung von n vielen Traces wiederholt und im Anschluss wird der Mittelwert gebildet. Das Ergebnis ist ein sogenannter „Saliency Trace“, welcher z.B. als 1-D Heatmap wie vorstehend bereits unter Bezugnahme auf **Fig. 5** beschrieben dargestellt werden kann. In diesem Beispiel befindet sich fast die komplette Information, die für einen erfolgreichen Angriff durch das trainierte neuronale Netz **NN'** entscheidend ist, in einem kleinen Bereich **B1** um den Zeitpunkt **1500**. Bei weiteren Ausführungsformen können Entwickler dieses Wissen aus der Heatmap HM nutzen, um bei dem Entwurf der Implementierung **110** gezielt Gegenmaßnahmen zu implementieren, da ihnen bekannt ist welche (kryptografischen) Operationen zu diesem Zeitpunkt auf dem angegriffenen Gerät durchgeführt werden. Hierzu kann bei weiteren Ausführungsformen z. B. eine auf dem Prinzip des Higher-Order Masking beruhende Technik verwendet werden. Zu Details bezüglich Higher-Order Masking s. z.B. Matthieu Rivain and Emmanuel Prouff, Provably secure higher-order masking of aes. Cryptology ePrint Archive, Report 2010/441, 2010, https://eprint.iacr.org/2010/441.

[0043] Nachstehend ist der vorstehend bereits angesprochene Algorithmus 1 zum Erstellen einer sog. Saliency Trace gemäß weiterer bevorzugter Ausführungsformen beschrieben, wobei einzelnen Zeilen jeweils eine Zeilennummer mit Doppelpunkt vorangestellt ist. Eingangsgrößen sind: $X = \{X\}_{i=1}^{n}$, also n viele Eingangs-Traces (z.B. Zeitverlauf der phys. Größe **G1**), $C = \{c\}_{i=1}^{n}$, also n viele Ausgabeklassen des trainierten neuronalen Netzwerks **NN'** für den jeweiligen korrekten kryptographischen Schlüssel, S (X), also eine „class score function“ des trainierten

neuronalen Netzes **NN'**, und als Ausgangsgröße wird X_{SAL} erhalten, also die mittlere „salience trace“.

1: $\omega \leftarrow 0$

2: $i \leftarrow n$

3: while $i \neq 0$ do

4: $\omega \leftarrow \omega + \left.\frac{\partial S_c}{\partial X}\right|_{x_i}$

5: $i \leftarrow i - 1$

6: end while

7: $X_{SAL} \leftarrow \frac{\omega}{n}$

[0044] In Zeile 1 wird der Variable ω der Wert „0“ zugewiesen, in Zeile 2 wird der Zählervariable i der Wert n zugewiesen, die Zeilen 3 bis 6 definieren eine Schleife mit n vielen Durchläufen, in der die (Ergebnis-)Variable ω sukkzessive in Abhängigkeit des Terms $\left.\frac{\partial S_c}{\partial X}\right|_{x_i}$ verändert wird, wobei schließlich in Zeile 7 die mittlere „salience trace“ erhalten wird.

[0045] Das Prinzip gemäß den Ausführungsformen ermöglicht vorteilhaft eine detaillierte Analyse und Bewertung von Implementierungen **110** insbesondere kryptografischer (aber auch sonstiger) Verfahren bezüglich einer Seitenkanal-Leckage. Weiter ermöglicht das Prinzip die Ausführung von Seitenkanalattacken unter Verwendung des trainierten neuronalen Netzes **NN'**.

Patentansprüche

1. Verfahren zum Verarbeiten von Daten eines technischen Systems (100), wobei das technische System (100) wenigstens eine Implementierung (110) eines kryptografischen Verfahrens aufweist, wobei das Verfahren zum Verarbeiten von Daten des technischen Systems (100) die folgenden Schritte aufweist: Modellieren (200) wenigstens eines Teils des technischen Systems (100) mittels wenigstens eines, insbesondere künstlichen, neuronalen Netzes (NN), Trainieren (202) des neuronalen Netzes (NN) unter Verwendung bekannter erster Daten (D1) und wenigstens einer physikalischen Größe (G1) des technischen Systems (100), wodurch ein trainiertes neuronales Netz (NN') erhalten wird, Ermitteln (204) von einen Betrieb der wenigstens einen Implementierung (110) charakterisierenden Seitenkanalinformationen (SCI) in Abhängigkeit des trainierten neuronalen Netzes (NN').

2. Verfahren nach Anspruch 1, wobei das kryptografische Verfahren ein Verschlüsselungsverfahren ist, insbesondere vom AES-Typ.

3. Verfahren nach wenigstens einem der vorstehenden Ansprüche, wobei die bekannten ersten Daten (D1) Eingangsdaten (ED) für die Implementierung (110) des kryptografischen Verfahrens sind, und/oder wobei die wenigstens eine physikalische Größe (G1) des technischen Systems (100) einen Stromverbrauch und/oder elektromagnetische Abstrahlungen umfasst.

4. Verfahren nach wenigstens einem der vorstehenden Ansprüche, wobei der Schritt des Ermittelns (204) ein Visualisieren von wenigstens einer Komponente des trainierten neuronalen Netzes (NN') und/oder aus dem trainierten neuronalen Netz (NN') ableitbaren Daten umfasst.

5. Verfahren nach wenigstens einem der vorstehenden Ansprüche, wobei der Schritt des Ermittelns (204) die Ermittlung eines Zeitpunkts (B1) bei der Ausführung des kryptografischen Verfahrens durch die Implementierung (110) umfasst, zu dem, insbesondere bezogen auf andere Zeitbereiche (B2, B3) der Ausführung des kryptografischen Verfahrens, eine vergleichsweise große Leckage von Seitenkanalinformationen (SCI) auftritt.

6. Verfahren nach wenigstens einem der vorstehenden Ansprüche, wobei das trainierte neuronale Netz (NN') eine Ausgabefunktion S aufweist, wobei mittels Fehlerrückführung, insbesondere backpropagation, eine Ableitung S_c in Bezug auf eine Seitenkanalmesswertreihe, insbesondere side channel trace, ermittelt wird, wobei c eine Ausgabeklasse des trainierten neuronalen Netzes (NN') für einen korrekten Ausgabewert repräsentiert, wobei insbesondere die vorstehend genannten Schritte für eine vorgebbare Anzahl von Eingangs-Seitenkanalmesswertreihen wiederholt werden.

7. Vorrichtung (1000) zum Verarbeiten von Daten eines technischen Systems (100), wobei das technische System (100) wenigstens eine Implementierung (110) eines kryptografischen Verfahrens aufweist, wobei die Vorrichtung (1000) zur Ausführung der folgenden Schritte ausgebildet ist: Modellieren (200) wenigstens eines Teils des technischen Systems (100) mittels wenigstens eines, insbesondere künstlichen, neuronalen Netzes (NN), Trainieren (202) des neuronalen Netzes (NN) unter Verwendung bekannter erster Daten (D1) und wenigstens einer physikalischen Größe (G1) des technischen Systems (100), wodurch ein trainiertes neuronales Netz (NN') erhalten wird, Ermitteln (204) von einen Betrieb der wenigstens einen Implementierung (110) charakterisierenden Seitenkanalinformationen (SCI) in Abhängigkeit des trainierten neuronalen Netzes (NN').

8. Vorrichtung (1000) nach Anspruch 7, wobei die Vorrichtung (1000) zur Ausführung des Verfahrens

nach wenigstens einem der Ansprüche 2 bis 6 ausgebildet ist.

9. Verwendung des Verfahrens nach wenigstens einem der Ansprüche 1 bis 6 und/oder der Vorrichtung (1000) nach wenigstens einem der Ansprüche 7, 8 zur Bewertung einer Implementierung (110) eines kryptografischen Verfahrens in Bezug auf eine Leckage von Seitenkanalinformationen (SCI).

Es folgen 3 Seiten Zeichnungen

Anhängende Zeichnungen

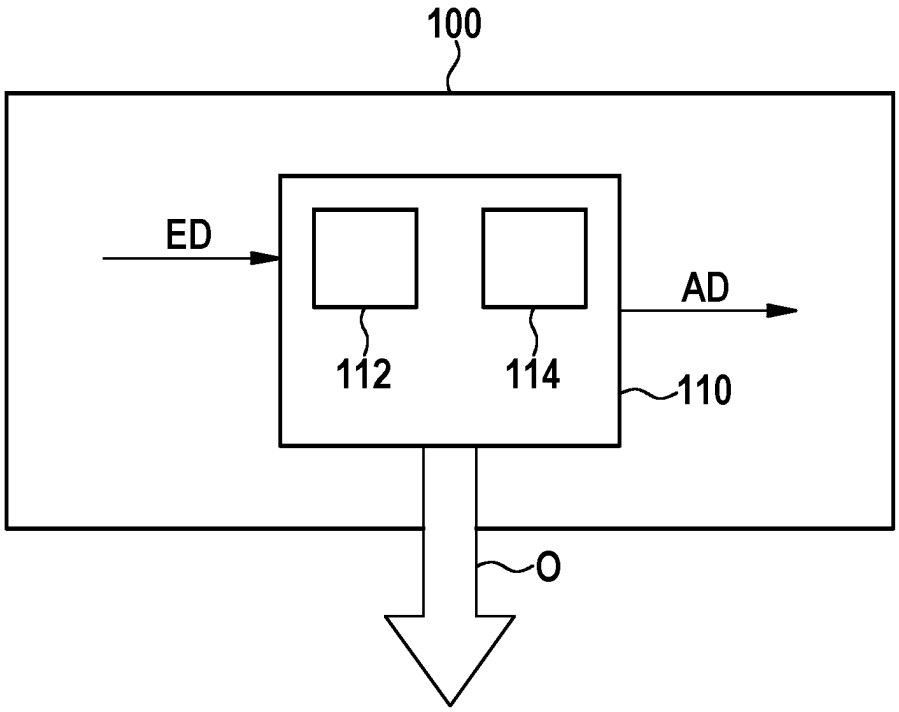


Fig. 1

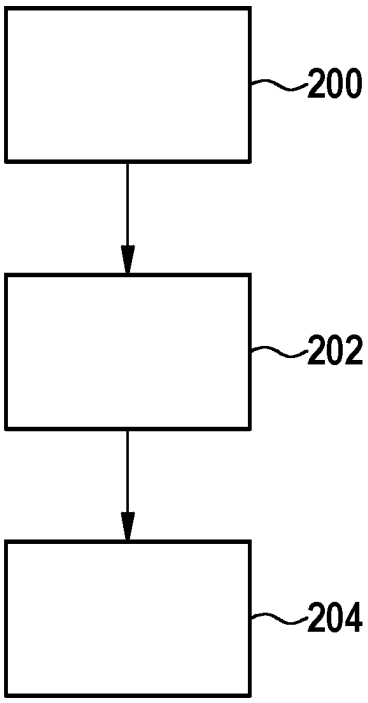


Fig. 2

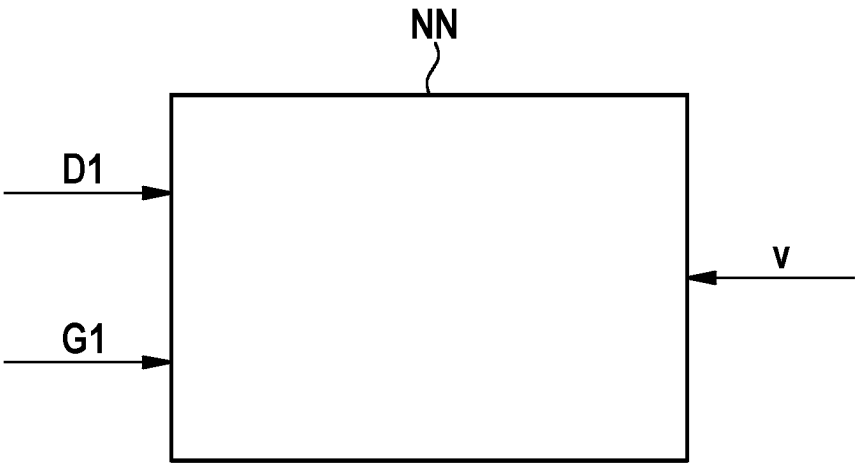


Fig. 3

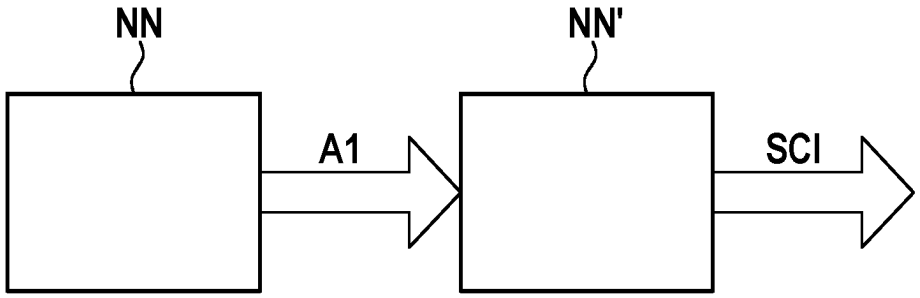


Fig. 4

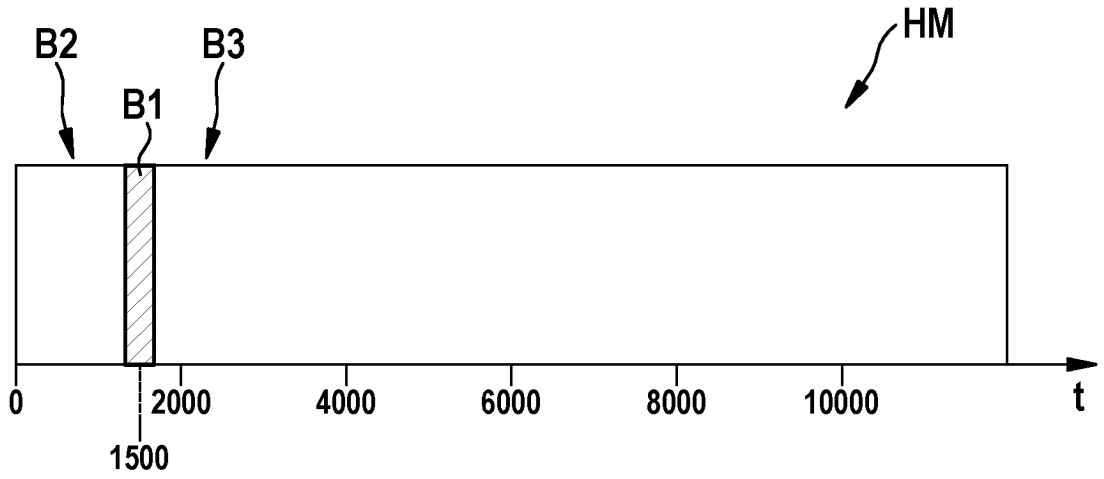

B2
B3
B1
HM
0
2000
4000
6000
8000
10000
t
1500

Fig. 5

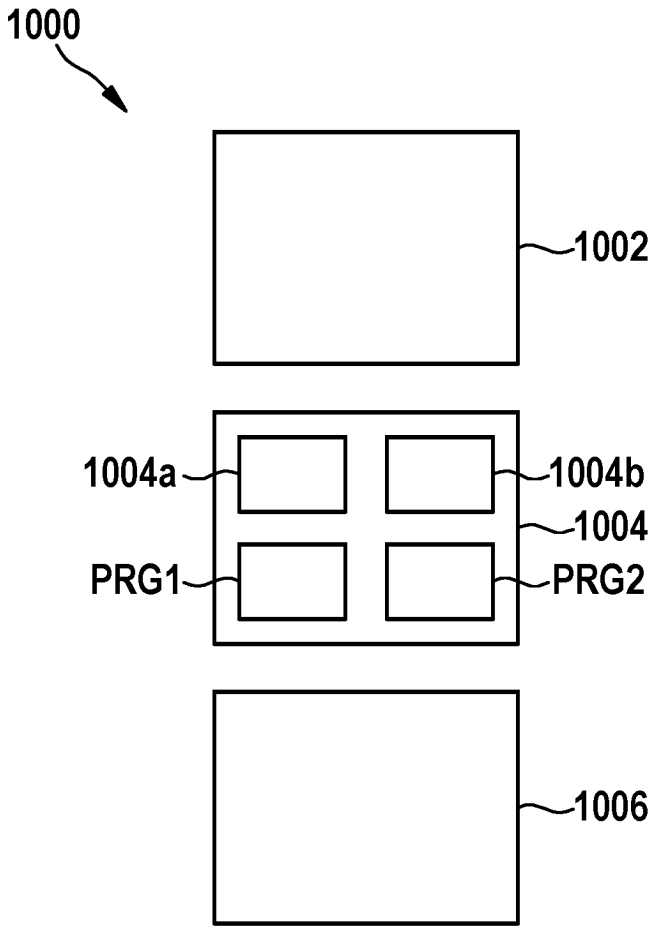

1000
1002
1004a
1004b
1004
PRG1
PRG2
1006

Fig. 6