

Wir wollen den 22. Januar 2021 zum Anlass nehmen, den "Vertrag über das Verbot von Kernwaffen", TPNW, mit diesem Artikel zu würdigen, da der Vertrag an diesem Tag in Kraft getreten ist.

Dies ist die deutschsprachige Version des Artikels „Unintended Nuclear War“ in der Zeitschrift „Künstliche Intelligenz“, Heft 1, 2021, der Gesellschaft für Informatik.

Online-Version: <https://doi.org/10.1007/s13218-021-00710-0>

Link zu dieser Version: www.fwes.info/bs-akav-ki-2021-1.pdf

Atomkrieg aus Versehen

Karl-Hans Bläsius

Hochschule Trier

Fachbereich Informatik

blaesius@hochschule-trier.de

Jörg Siekmann

Universität des Saarlandes

Fachbereich Informatik/DFKI

Joerg.Siekmann@dfki.de

Mit der Resolution 71/258 hat die Generalversammlung der Vereinten Nationen beschlossen, im Jahr 2017 eine UN-Konferenz einzuberufen, um ein rechtsverbindliches Instrument zum Verbot von Atomwaffen auszuhandeln, das zu ihrer vollständigen Abschaffung führen soll. Am 7. Juli 2017 stimmten 122 Staaten für diesen Antrag (bei einer Gegenstimme und einer Enthaltung). Der Vertrag verbietet die Herstellung, den Besitz und den Einsatz von Atomwaffen. Am 24. Oktober 2020 ratifizierte der 50. Staat den Vertrag und der TPNW ist am 22. Januar 2021 in Kraft getreten. Da die Atommächte die Konferenz jedoch boykottiert und nicht unterschrieben haben, hat der Vertrag derzeit nur einen politischen und symbolischen Wert. Das Wettrüsten inklusive Atomwaffen geht weiter.

Im Jahr 2017 erhielt die International Campaign to Abolish Nuclear Weapons (ICAN) den Friedensnobelpreis für "ihre Arbeit, um auf die katastrophalen humanitären Folgen eines jeden Einsatzes von Atomwaffen aufmerksam zu machen, und für ihre bahnbrechenden Bemühungen um ein vertragliches Verbot solcher Waffen" und forderte die Atommächte zu ernsthaften Abrüstungsverhandlungen auf.

Dieser Artikel befasst sich mit der Problematik, ob auf KI-Techniken basierende Entscheidungen in Frühwarnsystemen sinnvoll sein können und ob diese Systeme dann im Hinblick auf mögliche Fehlalarme sicherer sind.

KI in militärischen Frühwarn- und Entscheidungssystemen

Die Sicherung der nuklearen Zweitschlagsfähigkeit ist die Grundlage der Abschreckungsstrategie, die bisher jeden potenziellen Angreifer davon abgehalten hat, einen nuklearen Angriff zu starten: "Wer zuerst schießt, stirbt als Zweiter".

Um bei einer Bedrohung der Zweitschlagsfähigkeit reagieren zu können, haben die Atommächte komplexe Frühwarn- und Entscheidungsunterstützungssysteme entwickelt und installiert mit dem Ziel, einen Angriff rechtzeitig zu erkennen, um die eigenen nuklearen Trägersysteme vor dem zerstörerischen Einschlag aktivieren zu können.

Eine solche Strategie wird als "launch-on-warning" bezeichnet.

Diese Systeme erkennen einen Angriff auf Basis von Sensordaten, wie Radar, Licht oder Ultraschall, und basieren auf einer computergestützten Interpretation dieser Daten.

Aufbau und Funktionsweise von Frühwarnsystemen wurden vor allem aus den USA durch verschiedene Untersuchungsberichte und Veröffentlichungen in den 1980er Jahren bekannt. Eine frühe Kommandozentrale der USA war NORAD (North American Aerospace Defence Command), die bereits 1957 ihren Betrieb aufnahm und ihre Computersysteme enthielten 1983 bereits etwa zehn Millionen Codezeilen. Seitdem sind diese Systeme weitgehend modernisiert worden, um dem neuen Waffenarsenal Rechnung zu tragen, und sie sind jetzt wesentlich komplexer.

Wie jedes große kombinierte Soft- und Hardwaresystem sind diese riesigen Anlagen anfällig für Fehler, was zu einem versehentlichen Atomkrieg führen könnte.

Obwohl die Zeit für die Entscheidung zwischen einem gemeldeten Angriff und dem Start von Raketen für den Gegenangriff in den letzten Jahren auf wenige Minuten gesunken ist, liegt die endgültige Entscheidung - nicht zuletzt wegen der Fehleranfälligkeit solcher Systeme - immer noch beim Oberbefehlshaber, also z.B. dem Präsidenten der Vereinigten Staaten.

Durch die zunehmende Anzahl unterschiedlicher Sensortypen, Satelliten im Weltraum und Überwachungssysteme steigt die Menge der verfügbaren Daten und Informationen in einer konkreten Situation überproportional an. So werden für die Klassifizierung von Sensordaten und die Bewertung einer Alarmsituation immer mehr Systeme der Künstlichen Intelligenz für Teilaufgaben eingesetzt, die die endgültige menschliche Entscheidung vorbereiten.

Das Ende des INF-Vertrags (Intermediate Range Nuclear Forces) hat zu einem neuen Wettrüsten geführt, auch mit Hyperschallraketen, die diese Zeitspanne noch weiter verkürzen. Daher haben Politiker und Militärs die Erwartung, dass KI-Systeme in der Lage sein werden, innerhalb dieser sehr kurzen verfügbaren Zeit bessere Entscheidungen zu treffen als militärisches Personal.

Falsche Alarme und der politische Kontext

Wegen der Unsicherheit der Daten stützt das militärische Personal seine Entscheidungen auch auf kontextuelles Wissen über die politische Lage und die Einschätzung des Gegners. So entschied z. B. die Bedienmannschaft des amerikanischen Frühwarnsystems, dass der durch mehrere auf die USA zusteuernde Raketen ausgelöste Alarm am 5. Oktober 1960 falsch sein müsse und ein Vergeltungsschlag keinen Sinn mache, weil sich der sowjetische Staatschef zu diesem Zeitpunkt auf einem Staatsbesuch in New York befand.

Das heißt, auch bei einer maschinenbasierten Entscheidung muss kontextuelles Wissen über die politische Weltlage in die Bewertung von Alarmmeldungen einfließen, und auch dieses Wissen ist unsicher, vage und unvollständig. Das Ergebnis der Analyse durch ein KI-System ist daher nur im Rahmen einer statistischen Wahrscheinlichkeit richtig.

Zwei Beispiele Pro und Contra einer maschinenbasierten Entscheidung

Beispiel 1:

Im Januar 2020 töteten die USA den iranischen General Soleimani mit einem Drohnenangriff und als Vergeltung griff der Iran einige Tage später amerikanische Stellungen im Irak an. Kurz darauf wurde ein ukrainisches Verkehrsflugzeug versehentlich im Iran abgeschossen, da die Kontrollmannschaft zu dem Schluss kam, dass es sich bei dem Flugobjekt um einen angreifenden Marschflugkörper handeln könnte.

In dieser Situation hätte ein Computer vermutlich eine bessere Entscheidung getroffen, weil die reinen Fakten, wie z. B. die Größe des Radarsignals, wahrscheinlich besser interpretiert worden wären und dann gegen einen Marschflugkörperangriff gesprochen hätten. Außerdem hätte eine Maschine auch in der kurzen Zeit mehr Informationen, wie z.B. zivile Flugpläne, berücksichtigen können.

Die Fehlentscheidung kam vor allem deshalb zustande, weil die Besatzung mit Krieg und einem Angriff der USA gerechnet und den politischen Kontext offensichtlich überbewertet hatte.

Beispiel 2:

Ein Satellit des russischen Frühwarnsystems meldet am 26. September 1983 fünf anfliegende Interkontinentalraketen (ICBM). Da die korrekte Funktion des Satelliten überprüft und festgestellt wurde, hätte der diensthabende russische Offizier, Stanislaw Petrow, diese Information vorschriftsmäßig an seinen Vorgesetzten und an den Staatschef der Sowjetunion, J.W. Andropow, weiterleiten müssen. Er hielt jedoch einen Angriff der Amerikaner mit nur fünf Raketen für unwahrscheinlich und entschied trotz der vorliegenden Daten, dass es sich wahrscheinlich um einen Fehlarbeit handelt und verhinderte so eine Katastrophe mit Atomschlag und Gegenangriff. Der Vorfall ereignete sich in einer instabilen politischen Situation: Die Modernisierung mit Mittelstreckenraketen stand an und wenige Wochen zuvor hatten die Sowjets versehentlich ein koreanisches Passagierflugzeug über internationalen Gewässern

abgeschossen. Ein KI-System hätte den Angriff vermutlich eher als real eingeschätzt und aufgrund dieser Fakten einen Gegenangriff eingeleitet.

Petrov jedoch hatte emotional auf einen Fehlalarm gehofft und wollte nicht für den millionenfachen Tod von Menschen verantwortlich sein und entschied sich gegen die Vorschriften.

Neue Technologien und Tests

Bei neuen Technologien gibt es immer Ängste, ob berechtigt oder nicht. So galten beispielsweise schnelle Zugfahrten im 19. Jahrhundert als gefährlich und aktuell wird in den Medien über die Gefahr von autonomen Fahrzeugen diskutiert. Allerdings hat die Sicherheit mit neuen Technologien in der Regel zugenommen, und viele Experten glauben, dass autonomes Fahren das Unfallrisiko deutlich reduzieren wird. Doch gerade der wiederholte Zyklus von "Versuch und Irrtum" ist wichtig, um aus Innovationen zuverlässige Technologien zu entwickeln. Die Fehlklassifikation einer Lkw-Plane als "freie Straße" beim früheren Einsatz eines Tesla, die zu einem schweren Unfall führte, wurde korrigiert und hat zu besseren und zuverlässigeren Klassifizierungsmethoden geführt. In solchen Fällen halten sich die Schäden in Grenzen, wohingegen ein größerer nuklearer Schlagabtausch nicht auf einen lokalen Bereich der Welt beschränkt werden kann, da er in Kombination mit dem anschließenden nuklearen Winter Milliarden von Toten bedeuten und möglicherweise das menschliche Leben auf der Erde auslöschen würde. Außerdem können die Frühwarnsysteme nur mit Hilfe von Simulationssoftware getestet werden, da sie offensichtlich nicht in der Realität erprobt werden können.

Die Atomkriegsuhr

1947 wurde die "Atomic Doomsday Clock" eingerichtet, um vor der Gefahr eines drohenden Atomkrieges zu warnen. Die Uhr wird einmal im Jahr von Atomwissenschaftlern und Nobelpreisträgern neu eingestellt und die Gründe für die Einstellung werden im Bulletin of the Atomic Scientists veröffentlicht. Die erste Einstellung 1947 erfolgte auf sieben Minuten vor zwölf, 1984 wurde sie wegen des damals beschleunigten Wettrüstens auf drei Minuten vor zwölf gestellt und 1991 - wegen des Abrüstungsabkommens zwischen Michail Gorbatschow und Ronald Reagan - auf siebzehn Minuten vor Zwölf zurückgestellt. Aufgrund der sich zuletzt verschlechternden politischen und militärischen Lage wurde sie 2020 auf den historischen Tiefstand von 100 Sekunden gesetzt. Bei der neuen Festlegung am 27. Januar 2021 wurde diese Einstellung beibehalten.

Gründe für diese dramatische Warnung sind, dass die Atommächte derzeit ihre Atomwaffen modernisieren und sogar ausbauen, dass die meisten wichtigen Verträge zur Begrenzung und zum gegenseitigen Vertrauen außer Kraft gesetzt wurden und dass der Klimawandel und die damit verbundene Verschlechterung der Lebensbedingungen zu Konflikten führen könnte.

Zusammenfassung:

Während KI-Systeme für die politische Bewertung und militärische Aufklärung durchaus nützlich sein können, ist die endgültige Entscheidung für einen Gegenangriff durch einen Computer nicht akzeptabel. Das Überleben der Menschheit als Ganzes sollte niemals von der Entscheidung eines einzelnen Menschen oder einer Maschine abhängen. Maschinen können aufgrund der unsicheren und unvollständigen Datenlage eingehende Alarmmeldungen nicht mit hundertprozentiger Sicherheit auswerten und wenn der versehentliche Gegenschlag erst einmal eingeleitet ist, gibt es keine Chance zur Korrektur.

Mit mehreren Kollegen haben wir eine Website eingerichtet und pflegen sie

www.atomkrieg-aus-versehen.de

www.unintended-nuclear-war.eu

Weitere Hinweise:

Dieser Artikel basiert auf dem detaillierteren Beitrag:

K.H. Bläsius, J. Siekmann; Computergestützte Frühwarn- und Entscheidungssysteme, 2021, <https://www.fwes.info/fwes-21-1.pdf> (auch eine englische Version ist wählbar).

K.H. Bläsius, J. Siekmann: Computergestützte Frühwarn- und Entscheidungssysteme, Informatik-Spektrum, Band 10, Heft 1, 24-39, 1987

K.H. Bläsius, J. Siekmann: Frühwarnsysteme und Cyberangriffe – gefährliche Wechselwirkungen möglich, Behördenspiegel August 2019, S. 44, in: https://issuu.com/behoerden_spiegel/docs/2019_august

K.H. Bläsius, J. Siekmann: Künstliche Intelligenz in Frühwarnsystemen. in: Newsletter des Behördenspiegels Verteidigung, Streitkräfte, Wehrtechnik, am 19.12.2019, https://www.fwes.info/nl_defence_253_02.jpg