Eberhard von Faber, Arndt Kohler

Die Lücke: Informationssicherheit in Systemen mit künstlicher Intelligenz

Wie Algorithmen und künstliche Intelligenz zur Gefahr für die IT-Sicherheit werden

Alle bekannten Maßnahmen und Methoden für die IT-Sicherheit beruhen auf der Definition eines Soll-Zustandes, der Kontrolle des Ist-Zustandes und dem gezielten, korrigierenden Eingriff. Aber dies sind nicht die einzigen Annahmen für eine funktionierende IT-Sicherheit, die der Beitrag ebenso analysiert wie die Probleme, die sich durch den Einsatz von maschinellem Lernen und künstlicher Intelligenz ergeben. Denn letztere können aus unseren herkömmlichen Sicherheitsmaßnahmen stumpfe Schwerter machen und selbst zur Waffe werden.

1 Eingrenzung des Themas

Die Welt ist voller Bedrohungen und die „Security" nimmt sich der meisten davon an. Abbildung 1 zeigt einige der Bedrohungen. Die linke Spalte enthält die Quelle, die oberste Zeile das Opfer.

IT-Experten interessieren sich grundsätzlich für alle Fälle, die von den beiden Zeilen und den beiden Spalten mit „IT" eingeschlossen werden, sie bilden ein Kreuz. Lassen wir zweimal „Hollywood"=Fiktion und einmal „Unfall" beiseite, verbleiben acht Fälle. Die sechs mit „Qualität" bezeichneten Fälle sind besonders wichtig. Leider ist das Interesse daran bisher nicht besonders hoch, obwohl der Fall „IT-Systeme gefährden Menschen" zunehmend in den Fokus rücken sollte. Er ist eng verbunden mit dem Fall „Menschen bedrohen IT-Systeme", denn eigentlich können die Fälle aufeinander aufbauen, was die Matrix dreidimensional machen würde. Der Verein „Zero Outage Industry Standard" [1] nimmt sich des Themas „Qualität in der IT einschließlich IT-Sicherheit" an. In diesem Beitrag geht es um die verbliebenen zwei Fälle „IT-Security" und speziell um den **fett** hervorgehobenen. Doch wir werden zwei Schritte zulassen und die „IT mit künstlicher Intelligenz" als Mittel des Angreifers und als Mittel zur Verteidigung ansehen und getrennt betrachten.

Doch auch dieses **eine** verbliebene Kästchen hat es in sich. In diesem Beitrag geht es aber weder um die rechtlichen Dimensionen von künstlicher Intelligenz (siehe hierzu [2]), noch um den rechtskonformen und verantwortungsvollen Einsatz von Algorithmen (siehe [3]). Im Folgenden geht es um Technologie; es geht um IT-Sicherheit im engeren Sinne.

Prof. Dr. Eberhard von Faber

T-Systems, Chief Security Advisor, IT Division; Arbeitsgebiete: Sicherheitsarchitektur, Entwickler von ESARIS, sichere IT-Produktion, sicheres IT-Outsourcing, Prozess- und ITIL-Integration, Standardisierung, Cloud, IAM
E-Mail: Eberhard.vonFaber@th-brandenburg.de

Arndt Kohler

IBM, Head of IoT Security, Security Division;
Arbeitsgebiete: Internet of Things Security, Operational Technologies Security, Security Consulting & Architecture, Security Operation
E-Mail: Arndt.Kohler@de.ibm.com

Abbildung 1 | Bedrohungen und Arbeitsgebiete

bedrohen → ↑	Menschen	IT-Systeme/ Daten	IT mit Künstlicher Intelligenz	Maschinen
Menschen	Krieg/Terror	IT-Security	**IT-Security**	Sabotage/ Terror
IT-Systeme	Qualität/ IT-Security (Safety)	Qualität	Qualität	Qualität
IT mit Künstlicher Intelligenz	(Hollywood)	Qualität	(Hollywood)	Qualität
Maschinen	Safety	Unfall	- -	Cyber-Krieg

Der Beitrag analysiert die Basis bisheriger IT-Sicherheitslösungen und skizziert die Probleme, die sich durch den Einsatz von Algo-

rithmen bzw. künstlicher Intelligenz in den geschäftlichen Anwendungen sowie deren Nutzung durch Angreifer ergeben können. Werden wir den von „künstlicher Intelligenz" kontrollierten Bereich als Black-Box betrachten und zum Perimeter-Schutz zurückkehren müssen? Der Beitrag kann keine endgültigen Lösungen bieten, sondern sieht seine primäre Aufgabe darin, auf mögliche kritische und grundlegende Schwierigkeiten hinzuweisen. Die Autoren verbinden damit die Hoffnung, dass die IT-Sicherheit wiederum nachzieht und Lösungen zur Absicherung „intelligenter, algorithmen-gesteuerter IT" entwickelt, wie sie es in den zurückliegenden 50 Jahren immer bewiesen hat.

2 Ausgangslage

2.1 Was sind Maschinelles Lernen und Künstliche Intelligenz?

Im engeren Sinne bezieht sich der Begriff künstliche Intelligenz (KI) auf den Versuch, menschenähnliche Entscheidungsstrukturen in einem nicht-eindeutigen Umfeld nachzubilden. Computer werden so programmiert, dass sie eigenständig Probleme bearbeiten und lösen sowie Entscheidungen fällen können.

Die Verfahren, die die Autoren beim Schreiben dieses Beitrages hauptsächlich im Blick haben, basieren auf Maschinellem Lernen (ML). Darunter versteht man die maschinelle Generierung von Wissen aus Erfahrung. Ein System lernt aus Beispielen (Lerndaten) und kann diese nach Beendigung der Lernphase verallgemeinern, also auf andere neue Konstellationen (Nutzdaten) anwenden. Das System „erkennt" Muster und Gesetzmäßigkeiten in den Lerndaten und wendet diese bei der Verarbeitung von Nutzdaten an.

Das Anlernen erfolgt durch Vorzeigen des Ergebnisses, nicht aber durch Vorgabe der Kriterien (Muster, Gesetzmäßigkeiten). Diese extrahiert das System selbst, was das Lernen und die spätere „Intelligenz" ausmacht. Eine Realisierungsmöglichkeit bietet die Nachahmung neuronaler Netze. Dabei werden aus vielen Eingangsparametern einige Ausgangssignale erzeugt. Dies erfolgt durch Gewichtung der Eingangssignale (mittels Parametern) und die Verbindung zwischen allen Eingangssignalen und allen Ausgangssignalen über eine einfache Rechenvorschrift.[1]

Beim Anlernen werden die Werte aller Parameter bestimmt. Deren Anzahl kann in nicht allzu komplexen Fällen mehr als 100 Millionen betragen. D.h., wir haben es mit einem riesigen Gleichungssystem zu tun, dass in seiner Gänze kaum zu überblicken ist. Das bedeutet aber auch, dass man nicht wissen kann, nach welchen Kriterien, Mustern, Gesetzmäßigkeiten das System das Ergebnis bestimmt hat! Nicht einmal, welche Eingangsparameter wirklich relevant sind. Man sieht nur das Resultat.

Nur manchmal kann es im Nachhinein auffallen, dass etwas schief gegangen ist. So experimentiert Amazon mit künstlicher Intelligenz im Einstellungsprozess zur Bewerberauswahl. Es zeigte sich, dass Frauen systematisch benachteiligt wurden, was an den zum Anlernen verwendeten Daten lag. In ihnen waren Frauen schlicht unterrepräsentiert. Die künstliche Intelligenz hielt dies für ein Kriterium.

1 Die jeweilige Rechenvorschrift entspricht einem „Neuron mit Parametern". Diese bilden eine Schicht. Beim „Deep Learning" (DL) erfolgt die Verbindung zwischen Eingang und Ausgang nicht über eine Ebene oder Schicht von Neuronen, sondern über mehrere hintereinanderliegende Schichten.

Grundsätzlich handelt es sich bei künstlicher Intelligenz auch um Entscheidungen, die im Mittel richtig sind, im Einzelfall aber komplett falsch sein können. Während Menschen in vielen Fällen „verstehen", also Zusammenhänge zwischen Ursache und Wirkung erkennen und ihren Entscheidungen zu Grunde legen, ist dies bei künstlicher Intelligenz nicht der Fall. Das System filtert statistische Korrelationen, die nicht zu verwechseln sind mit Wirkungszusammenhängen. Von letzteren hat das System „keine Ahnung". Aus den aufgeführten Gründen kann man an der „Intelligenz" solcher Systeme zweifeln. Im Folgenden verwenden wir den Begriff „Künstliche Intelligenz" ganz wertfrei als Terminus Technicus und schreiben ihn entsprechend groß. In Zusammenhängen verwenden wir oft die Abkürzung „KI".

Die grundlegenden Methoden sind schon länger bekannt. Die Technologien erleben aber heute einen Boom, weil 1.) sich die Rechen- und Speichertechnik hinsichtlich ihrer Leistungsfähigkeit und Verfügbarkeit dramatisch weiterentwickelt haben und 2.) in vielen Bereichen mit „Big Data" riesige Datenmengen zum Anlernen der Künstlichen Intelligenz zur Verfügung stehen.

2.2 Auswahl und Implementierung von IT-Sicherheitsmaßnahmen heute

Im Folgenden wird einmal dargestellt, wie die Auswahl geeigneter Sicherheitsmaßnahmen ohne Berücksichtigung von KI erfolgt. Das wird es erlauben, im Abschnitt 2.3 herausarbeiten zu können, auf welchen Fundamenten die IT-Sicherheit heute beruht. Das wird es ermöglichen, im Verlaufe dieses Beitrages die Schwierigkeiten bei der Absicherung von Systemen mit KI zu beschreiben.

Soll ein IT-System adäquat geschützt werden, so geht es um die Frage, welche Sicherheitsmaßnahmen bzw. Sicherheitslösungen an welcher Stelle integriert werden müssen. Dies ist keine einfache Aufgabe. Der Verweis auf einen risikobasierten Ansatz und auf ein Kosten-Nutzung-Verhältnis greift zu kurz, denn die eigentliche Aufgabe besteht darin zu analysieren, wo und wodurch Risiken entstehen und für welche Sicherheitsmaßnahmen Kosten zu veranschlagen sind. Der Begriff Sicherheitskonzept zeigt schon eher die prinzipielle Vorgehensweise auf, aber oft ist es dem Wissen und der Intuition des Bearbeiters überlassen, den Bezug zur IT herzustellen.

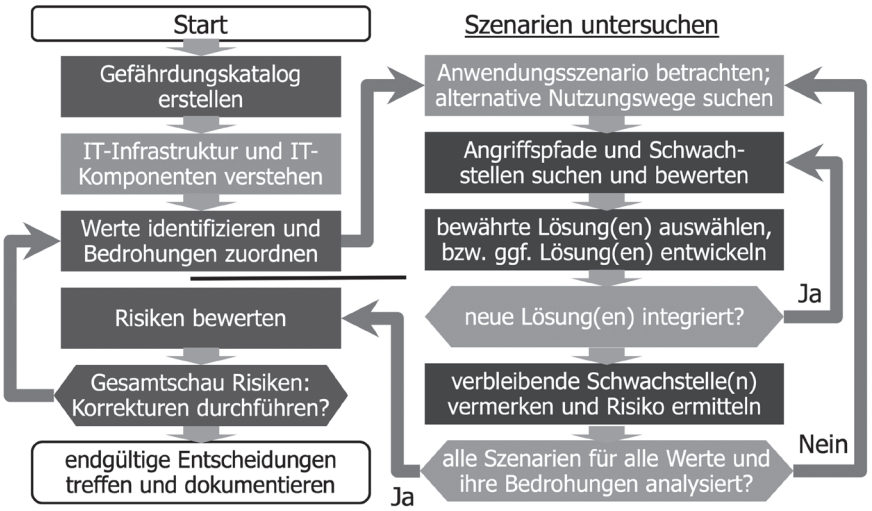
Abbildung 2 zeigt ein mögliches Vorgehen, das nun kurz skizziert wird, wobei bereits auf mögliche Probleme mit Künstlicher Intelligenz hingewiesen wird. Details zum Verfahren findet man im Buch „Joint Security Management (JSM)" [4].

Die linke Seite enthält allzu bekannte Schritte wie Gefährdungskatalog erstellen, Werte identifizieren und Bedrohungen zuordnen, Risiken bewerten sowie Gesamtschau und endgültige Entscheidung treffen. Der Punkt „IT-Infrastruktur und IT-Komponenten verstehen" ist eine ganz wesentliche Aufgabe, enthält aber bereits ein erstes mögliches Problem. Die Unternehmenswerte stecken in der IT (in Form von Daten und IT-Services), und mögliche Schwachstellen hinsichtlich der IT-Sicherheit ergeben sich aus ihrer Konstruktion bzw. Mängeln in ihrer Absicherung. Wird Künstliche Intelligenz eingesetzt, so sind, wie weiter unten noch eingehender ausgeführt wird, dem Verständnis der IT-Infrastruktur und ihrer Komponenten Grenzen gesetzt.

Die rechte Seite in Abbildung 2 umfasst Schritte zur Analyse, ob die Werte realen Risiken ausgesetzt sind und welche dies sind. Denn aus einer Bedrohung wird erst dann ein Risiko, wenn sie

(mit einer bestimmten Wahrscheinlichkeit) eine Schwachstelle ausnutzen kann. Sucht man nach Schwachstellen (oder Angriffsmöglichkeiten), so müssen die möglichen Informationsflüsse untersucht werden. Diese werden nicht allein durch die IT-Architektur sondern durch die Konfiguration von IT-Komponenten bestimmt. Auf dieser Basis versteht man, was ein Angreifer an Funktionalität vorfindet. Die Suche nach weiteren Nutzungsszenarien bzw. -wegen ist wichtig, da der Angreifer natürlich Funktionalitäten in einer Weise nutzen wird, wie dies nicht vorgesehen war. In den allermeisten Fällen setzen IT-Experten bereits hier an und integrieren eine Sicherheitslösung. In anderen Fällen ist eine systematische Suche nach weiteren Schwachstellen nötig. Auf dieser Basis können bewährte IT-Sicherheitslösungen (best practices) ausgewählt werden, die die Schwachstelle schließen. Danach wird die Suche nach Angriffspfaden und Schwachstellen fortgesetzt.

Abbildung 2 | Risikogesteuertes Verfahren zur Auswahl von Sicherheitslösungen



Die Abbildung zeigt, wie die Suche erfolgt und wie schrittweise Schwachstellen durch die Integration von IT-Sicherheitslösungen beseitigt werden. Dies erfordert

- das Verständnis von IT-Infrastruktur und IT-Komponenten,
- das Verständnis möglicher Informationsflüsse (einschließlich aller Angriffspfade) und
- die Möglichkeit, gezielt einzugreifen und gefundene Schwachstellen durch die Integration von IT-Sicherheitslösungen zu beseitigen.

Alle drei Voraussetzungen sind beim Einsatz von Künstlicher Intelligenz nur bedingt oder gar nicht gegeben. Stellen wir uns dazu vor, dass Algorithmen bzw. Künstliche Intelligenz eingesetzt werden. Sie sind Teil der IT, die damit Komponenten bekommt, deren innere Wirkungsweise per definitionem nicht oder nur teilweise verstehbar ist. Sie steuern zudem die Anwendungsszenarien bzw. primär die gewünschten Informationsflüsse selbständig in einer Weise, dass dies für den Sicherheitsanalytiker weder vorhersehbar noch im Nachhinein verstehbar ist; es handelt sich schließlich um eine unabhängig arbeitende „Intelligenz". Auch kann er, denn das liegt ebenfalls in der Natur der Sache, die Informationsströme nicht beeinflussen, denn das soll ja durch die Algorithmen bzw. die Künstliche Intelligenz geleistet werden.

2.3 Basis der IT-Sicherheit heute

Alle bekannten Maßnahmen und Methoden für die IT-Sicherheit beruhen auf der Definition eines Soll-Zustandes, der Kontrolle des Ist-Zustandes und dem gezielten, korrigierenden Eingriff. Doch es gibt weitere Voraussetzungen bzw. wichtige Details zu den bereits genannten. Drei werden im Folgenden skizziert.

1. Bekannte Informationsflüsse (im gesamten System): Die IT-Infrastruktur und die durch sie unterstützten Anwendungsszenarien (gewünschte Informationsflüsse usw.) müssen bekannt sein. Dann kann der IT-Sicherheitsanalytiker beginnen, nach Angriffspfaden und Schwachstellen zu suchen und diese hinsichtlich der Wahrscheinlichkeit ihrer Ausnutzung bewerten. Auf Basis dieser Analyse werden ggf. Sicherheitslösungen für die Implementierung ausgewählt, die die entsprechende Schwachstelle schließen und damit Risiken mindern oder ganz beseitigen. – Bestimmt die Künstliche Intelligenz den Informationsfluss selbst, so ist es schwierig, gewollten bzw. funktionsbedingten Informationsfluss von verdächtigem bzw. feindseligen Informationsfluss zu unterscheiden. Es ist auch schwierig, bestimmte Informationsflüsse zu unterbinden, um zu verhindern, dass diese zu Sicherheitsschutzverletzungen führen können. Es ist also, um es kurz zu sagen, nur schwer möglich, den Soll-Zustand bezüglich der Informationsflüsse festzulegen.
2. Verständnis der IT-Funktionalität (eines IT-Objekts): Die Notwendigkeit, den Soll-Zustand zu kennen, betrifft auch die IT, speziell ihre Software-Komponenten wie Applikationen selbst. Hier geht es aber um die Funktion (also die Informationsflüsse zu und vom IT-Objekt sowie innerhalb dessen). Zum System gehörende Software unterscheidet sich von feindseliger durch ihre Funktionalität. Oft schließt man von der Herkunft der Software darauf, ob sie verdächtig bzw. feindselig ist oder nicht. Ist dies nicht einfach möglich, basiert die Entscheidung zwischen Gut und Böse auf Erfahrungen oder Analysen der Funktionalität der Software. Weitere Schwachstellen werden beseitigt, indem Zugriffe mit Hilfe zentral verwalteter digitaler Identitäten gesteuert und beschränkt werden. Das gleiche wird auch durch Verschlüsselung von Informationen erreicht bzw. durch die Verwendung von integritätssichernder Maßnahmen wie Signaturen. Schwachstellen werden auch durch das Filtern und Ersetzen von Daten und Befehlen beseitigt. Zugriffskontrolle, Verschlüsselung, Integritätssicherung und (manipulierende) Filter können nur dann eingesetzt werden, wenn die gewollte IT-Funktionalität (eines IT-Objekts) mehr oder minder vollständig verstanden wird. – Künstliche Intelligenz ist gerade dadurch gekennzeichnet, dass ihre Funktionsweise nicht vollständig durchschaubar bzw. einfach verstehbar ist. Dadurch werden einige der heute üblichen IT-Sicherheitslösungen an Wirksamkeit verlieren oder Einschränkungen hinsichtlich ihrer Einsatzmöglichkeiten erfahren.
3. Konfrontation bzw. Duellsituation: Angreifer und Verteidiger stehen sich direkt gegenüber (Duell). Der Angreifer versucht, die Sicherheitsmaßnahmen, die der Verteidiger aufgebaut hat und pflegt, direkt zu überwinden oder Lücken zu

finden. Das Ziel besteht darin, die Vertraulichkeit, Integrität und/oder Verfügbarkeit von Werten zu verletzen oder weitere Angriffsschritte vorzubereiten. – Im Falle von Systemen mit Künstlicher Intelligenz stehen sich Angreifer und Verteidiger nicht unbedingt direkt gegenüber. Sie interagieren auch nicht unbedingt. Der Angreifer kann z.B. die Nutzer bzw. die von ihnen erzeugten Nutzdaten beeinflussen oder manipulieren. Diese werden von einem KI-basierten Sicherheitssystem des Verteidigers aber z.B. zum Anlernen verwendet. Dadurch dass der Angreifer diese manipuliert, kann er seinen späteren Angriff verschleiern. Der Angreifer „spielt also über die Bande".

3 Situation mit künstlicher Intelligenz

3.1 Basis für den Einsatz von Künstlicher Intelligenz

Allgemein funktioniert KI nur dann richtig, wenn die Trainingsdaten und die Nutzdaten, die im Betrieb verwendet werden, qualitativ, also hinsichtlich ihrer Verteilung, übereinstimmen. Außerdem müssen Trainingsdaten und die Nutzdaten beide über das zu filternde Merkmal verfügen. Daraus ergeben sich die nachfolgend beschriebenen Voraussetzungen für einen erfolgreichen Einsatz, die gleichzeitig auf Probleme bzw. Angriffsmöglichkeiten hinweisen:

1. Stabilität: Das bedeutet, dass sich die Situation nicht schnell verändern darf. Ist dies der Fall, muss das System mit neuen Daten wieder neu angelernt werden.
2. Integrität der Lerndaten: Auch müssen die Lerndaten in dem Sinne integer sein, dass sie durch einen Angreifer nicht manipuliert wurden. Eine Manipulation kann zum Beispiel durch Beeinflussung der Nutzer geschehen aber auch dadurch, dass der Angreifer in der Lernphase bereits aktiv ist, so dass er später nicht erkannt wird.
3. Integrität des Lernvorgangs: Es muss auch ausgeschlossen sein, dass ein Angreifer den Lernvorgang gezielt manipuliert, indem feindselige Beispiele eingeschleust werden. Dabei muss bedacht werden, dass sich die „Wahrnehmung" von KI-Systemen von der menschlichen ganz wesentlich unterscheidet. Die Änderung weniger Pixel in einem Bild kann für die Künstliche Intelligenz aus einem Gesicht ein Auto machen [5], während ein Mensch den Unterschied sofort erkennt.
4. Kennzeichnung: Oft mangelt es aber auch an für das Anlernen benötigten Informationen. Bei komplexen IT-Systemen sind diese schwer zu identifizieren und parametrisieren. Dabei sind viele Fragen zu klären: Wie werden Beispiele für den „Gut-Fall" definiert? Welche Möglichkeiten bestehen, diese „Gut-Fälle" im Sinne der Ergebnismaximierung über die Zeit herzustellen und für das Lernen zur Verfügung zu stellen? Und welche Rahmenbedingungen, harte wie weiche, sollen einem solchen KI-System vorgegeben werden?

3.2 Probleme für die IT-Sicherheit aufgrund von KI

Stellen wir uns ein komplexes KI-System in der IT vor, das im Rahmen definierter Parameter selbständig lernen kann und Entscheidungen trifft. Über verschiedene Layer des OSI-Modells hinweg erfolgt über die Zeit eine sich verändernde Nutzung von IT-Ressourcen. Die Interaktion von IT-Elementen untereinander passt sich entsprechend an. Eine nachvollziehbare Begründung für diese Veränderung, sowie der Nachweis über die Integrität der Lerndaten und des Lernvorgangs wird dabei nicht zwingend erhoben.

In den Veränderungen einen Advanced Persistent Threat (APT)[2] zu erkennen wäre nur möglich, wenn es dieses geänderte Verhaltensmuster in einer vergleichbaren KI-Umgebung in der Vergangenheit bereits gab und dieses erfolgreich analysiert und dokumentiert wurde. Die APT-Erkennung durch KI funktioniert entgegen vielen Marketingbehauptungen deshalb in der Regel nicht [6]. Das KI-basierte Sicherheitssystem wurde mit Hilfe von Daten aus der Vergangenheit trainiert. Es kann einen APT also nur erkennen, wenn das Verhaltensmuster bereits auftrat und als feindselig erkannt wurde. Ändert der Angreifer sein Verhaltensmuster aber oder agiert er völlig neu, so wird das KI-basierte Sicherheitssystem ihn nicht erkennen können.

Aktive Sicherheitssysteme innerhalb der von KI kontrollierten IT wie z.B. Firewalls und Identitäts- und Zugriffsmanagement können rasch zu einem Hemmschuh für den Betrieb der KI und damit der IT werden. Die Administratoren, die die Sicherheitssysteme verwalten, können unter Umständen nicht schnell genug den veränderten Bedarf antizipieren und die Regeln anpassen.

Auch klassische SIEM-Systeme[3] stoßen in einem komplexen KI System rasch an ihre Grenzen. Anomalien definieren sich nicht zuletzt durch ein sich änderndes Nutzungsverhalten der IT – was letztlich jeder Entscheidung des KI Systems entspricht. Das gleiche gilt für IDS/IPS.[4] Sowohl SIEM-Lösungen als auch IDS/IPS könnten ihre Wirkung also zunehmend verfehlen. Schon heute ist es so, dass deren Wirksamkeit davon abhängt, dass sie an die von Menschen veränderten Anwendungsszenarien und Datenströme z.B. bei der Installation neuer Applikationen angepasst werden müssen. Werden Anwendungsszenarien und Datenströme jedoch zunehmend automatisch und ohne menschliche Eingriffe von Algorithmen bzw. Künstlicher Intelligenz gesteuert, so wird der Anpassung der genannten Lösungen für die IT-Sicherheit der Boden entzogen.

Am praktikabelsten wird es sein, die IT-Sicherheit auf den Schutz der Schnittstellen des KI-Systems mit seiner Umwelt zu beschränken – unter der Preisgabe der Überwachung jeglicher Aktivitäten innerhalb des KI-Systems. Werden wir also den von Künstlicher Intelligenz kontrollierten Bereich als Black-Box betrachten und zum Perimeter-Schutz zurückkehren müssen? Keine schöne Aussicht, denn das kann nicht der Ansatz zum Schutz komplexer KI-Systeme sein.

3.3 Fallbeispiel: Produktionsplanung und steuerung

Ein konkretes Beispiel soll die Überlegungen verdeutlichen. Produktionsplanung und Steuerung sind in vielen Industrien eine Herausforderung. Die nächste industrielle Revolution (Industrie 4.0) verspricht Einzelfertigung im Rahmen einer Serienprodukt-

2 APT: eine fortgeschrittene, andauernde Bedrohung bzw. eigentlich ein Angriff, der durch seine hohe Komplexität und lange Dauer gekennzeichnet ist

3 SIEM: Security Information and Event Management; Security Information Management (SIM) umfasst das Sammeln und Analysieren historischer Daten wie Log-Daten inklusive Log-Management und die (automatische) Überwachung der Einhaltung von Compliance-Vorschriften (wie Richtlinien zu Härtung und Aktualisierung mit Patches). Security Event Management (SEM) umfasst die Echtzeitüberwachung von ICT-Systemen und die Analyse von Ereignissen einschließlich Alarmen usw. SIEM ist die Kombination aus beiden und unterstützt das Incident-Management bezüglich Cyber Defense-Aktivitäten.

4 IDS/IPS: Intrusion Detection /Prevention Systems

tion. Beispielsweise in der Automobilindustrie wird bereits heute in einem Produktionsjahr selten ein und dasselbe Fahrzeug zweimal produziert. Zu groß sind die Anzahl des wählbaren Zubehörs, die Anzahl ihrer Varianten und Kombinationen. Hinter jedem Zubehör verbergen sich Lieferanten mit entsprechender Logistik. All dies muss in definierter Reihenfolge analog eines straffen Zeitplans am Montageband ankommen, ansonsten kann das Fahrzeug nicht produziert werden.

Dies ist eine Mammutaufgabe, die rasch den Ruf nach KI bei der Planung und Steuerung weckt. Ein komplexes System, welches aus Warenbewegungen, Produktionseinheiten und dem Faktor Mensch besteht, ist durch wenig Stabilität gekennzeichnet. Es gibt zudem diverse Schnittstellen in kaufmännische Systeme wie Einkauf, Controlling usw. Aus der technischer Sicht der IT handelt es sich um eine Mischung vielfältiger Komponenten, angefangen von physischen Maschinen und Transporteinheiten über deren Steuerungssysteme bis zu den IT-Anwendungen.

Ein KI-System mit einem derartig breiten Betätigungsfeld stellt ein hochkritisches Angriffsziel dar. Ein Angriff kann nicht „nur" zu Produktionsausfall oder fehlerhaften Produkten führen. Es gilt, auch die Sicherheit der Angestellten und Nutzer zu gewährleisten. Ausschließlich die Schnittstellen des KI-Systems nach außen abzusichern, ist keine Lösung. Das KI-System selbst muss gesichert werden. Dazu gehört es unter anderem, die Lerndaten und den Lernprozess zu überwachen. Der Wirkungsbereich des KI-Systems darf kein sicherheitsfreier Raum sein. Wenn sich einige der althergebrachten Sicherheitskomponenten als wenig zweckmäßig erweisen, ist zu überlegen, andere auszubauen und auf das gesamte KI-System auszuweiten. Das bedeutet, dass Bereiche, die aktuell keine oder wenige Sicherheitskomponenten aufweisen, nachgerüstet werden. So könnten z.B. die physischen Produktions- und Transporteinheiten so nachgerüstet werden, dass sie robuster gegen mögliche Fehlsteuerungen sind.

Die grundlegende Aussage – ein Cyberangriff ist keine Frage des „ob", sondern des „wann" – gilt für alle Komponenten des KI-Systems und auch die KI selbst. Daher müssen wir uns Gedanken machen, wie wir die Folgen eines Angriffs minimieren oder gar beseitigen könnten. Dies sollte im Idealfall natürlich so erfolgen, ohne dass das System auf null zurückgesetzt wird.

4 Verbesserung der IT-Sicherheit durch KI

Könnte denn die IT-Sicherheit durch den Einsatz von Künstlicher Intelligenz wenigstens verbessert werden?

Künstliche Intelligenz bzw. „smarte Algorithmen" werden schon jetzt an vielen Stellen eingesetzt, um die IT-Sicherheit zu verbessern. Ein Beispiel ist die Erkennung von Schadcode (Anti-Malware). Allerdings haben die Lösungen zumindest anfangs z.B. WannaCry nicht erkannt. Eine gute Implementierung kann aber zumindest den Großteil der Schadsoftware automatisiert und sicher erkennen [5]. Trotzdem wird man mit Fehlalarmen („fault positives") zu kämpfen haben. Insbesondere firmenspezifische, selten genutzte Software wird als solche erkannt werden, da ihr Profil nicht Teil der Lerndaten war. Die Dynamik der Schadsoftware zeigt außerdem, dass die Anti-Malware-Lösung regelmäßig neu angelernt werden muss.

Eine Hauptanwendung sind SIEM-Lösungen, die die Grundlage bilden bzw. die Infrastruktur stellen für Security-Operations-Center (SOC) bzw. Cyber-Defense-Center (CDC). SIEM-Lösungen sammeln, filtern, normalisieren, korrelieren und analysieren Vorfalldaten (events) aus vielen Quellen wie Firewalls, IDS/IPS-Systemen, Network-Admission-Control (NAC), Anti-Malware, Data-Leakage-Protection (DLP) und Authentisierungsdiensten. Künstliche Intelligenz bzw. „big data"-Lösungen helfen hier enorm. Allerdings müssen auch sie, wie man dies von IDS-IPS-Systemen kennt, regelmäßig der veränderten IT angepasst bzw. neu angelernt werden.

Entsprechende Daten vorausgesetzt, können Künstliche Intelligenz bzw. „smarte Algorithmen" auch für die Verbesserung des Zugriffsmanagements eingesetzt werden. Dabei geht es darum, dass die Zugriffsrechte dynamisch um Informationen ergänzt werden, die z.B. „unlogische" Zugriffe ausschließen. Damit ist man sehr nahe an einem vierten Anwendungsbereich. Er berührt nicht direkt die IT-Sicherheit, ist aber damit verwandt. Künstliche Intelligenz bzw. „smarte Algorithmen" können ganz gut zur Betrugserkennung („fraud protection") und zum Auffinden „missbräuchlichen" Verhaltens etwa in sozialen Medien eingesetzt werden. Bei all diesen Anwendungen muss aber bedacht werden, dass wir voraussetzen, Lerndaten zu besitzen, die unserem Soll-Zustand der IT-Sicherheit entsprechen. Können Daten, die wir real beobachten, überhaupt dem Soll-Zustand entsprechen?

5 Diskussion und Ausblick

Der Einsatz von Algorithmen bzw. künstlicher Intelligenz in geschäftlichen Anwendungen schafft neue Probleme. Schon die etablierten Verfahren zur Auswahl und Implementierung von IT-Sicherheitslösungen funktionieren nur eingeschränkt. Unsere heutige IT-Sicherheit basiert darauf, dass die Informationsflüsse bekannt und relativ stabil sind. Es ist weiterhin erforderlich, die IT-Funktionalität aller IT-Objekte verstehen zu können, denn viele Sicherheitslösungen verändern diese oder schränken sie ein. Zudem ist es in der heutigen IT-Sicherheit in der Regel so, dass der Angreifer den Verteidigern in dem Sinne gegenüber steht, dass er die Befestigungen zu überwinden versucht. Insgesamt ergab die Analyse, dass die genannten Voraussetzungen beim Einsatz von Künstlicher Intelligenz nicht mehr unbedingt gegeben sind. Die Informationsflüsse steuert die Künstliche Intelligenz selbst und nicht vorhersehbar; deren Funktionsprinzipien sind nicht durchschaubar, und Angreifer beeinträchtigen die IT-Sicherheit eventuell indirekt.

Betrachtet man die Basis für den Einsatz Künstlicher Intelligenz näher, so werden spezifische Risiken bzw. Angriffsmöglichkeiten sichtbar. Sie wurden sowohl in Form grundsätzlicher Probleme als auch anhand eines Fallbeispiels diskutiert. Dies führte zu dem wenig ermutigenden Vorschlag, die IT-Sicherheit auf den Schutz der Schnittstellen des KI-Systems mit seiner Umwelt zu beschränken bzw. die Auswirkungen durch die Verstärkung anderer peripherer Maßnahmen zu verringern. Weder der Perimeter-Schutz noch die Verminderung der Auswirkungen (typisch für das Vorfallmanagement) sind besonders phantasievolle Schöpfungen oder gar neu.

Man könnte daher überlegen, ob die IDS/IPS- und SIEM-Lösungen als Reaktion darauf mit Algorithmen und Künstlicher Intelligenz ausgestattet werden sollten. Dies ist ja heute schon begrenzt der Fall, z.B. dann, wenn ihre Funktionsweise auf Big-Data-Analysen beruht. Es ist jedoch schwer vorstellbar, dass die „Intelligenz" der Sicherheitslösung der geschäftssteuernden

„Intelligenz“ so weit überlegen ist, dass sie die Schritte der anderen „verstehen“ oder „kontrollieren“ kann.

Das gleiche gilt übrigens auch dann, wenn Angreifer beginnen sollten, ihre Angriffe mit Hilfe „künstlicher Intelligenz“ durchzuführen. Sie könnten dadurch in die Lage versetzt werden, ihre Angriffe so zu verschleiern und zu verbergen, dass sie weder von IDS/IPS- noch von modernsten SIEM-Lösungen entdeckt werden können. Alle Strategien modernster Cyber Security Defense, die ja alle im Grunde auf SIEM (plus Threat Intelligence) basieren, würden ihre Übermacht verlieren.

Und wenn alles nicht hilft? Werden wir dann Künstliche Intelligenz einsetzen, um falsche Dokumente und Unternehmenswerte zu schaffen, um sie unter unsere echten zu mischen in der Hoffnung, letztere vor dem Zugriff der angreifenden Künstlichen Intelligenz zu schützen? Die angreifende „Intelligenz“ müsste dann gegen zwei „Intelligenzen“ kämpfen: gegen die, die sie entdecken will und die, die die Werte im Sinne von „security by obscurity“ versteckt. Vielleicht werden die falschen Dokumente und Unternehmenswerte noch als Lockmittel (wie „Honey Pots“) angesehen und mit zusätzlichen Sensoren durch eine dritte „Intelligenz“ überwacht, um den Angriff aufzuspüren. Der Fantasie sind keine Grenzen gesetzt. Hoffentlich bleibt dabei der Mensch nicht auf der Strecke, sonst wären wir beim Hollywood-Szenario aus Abbildung 1, das dann keine Fiktion mehr darstellen würde.

Für diesen Beitrag sehen die Autoren ihre primäre Aufgabe darin, auf mögliche kritische und grundlegende Schwierigkeiten hinzuweisen. Wirkliche Lösungen können nicht angeboten werden. Doch eine gute Problemanalyse wird die Grundlage für die Entwicklung von Lösungen zur Absicherung von zunehmend „intelligenter, algorithmen-gesteuerter IT“ sein.

Literatur

[1] https://www.zero-outage.com mit den Themen People, Platform, Processes, Security.

[2] Thomas Burri: Künstliche Intelligenz und internationales Recht, Mögliche Entwicklungen und Hindernisse; Datenschutz und Datensicherheit (DuD), Heft 10, 2018, Seiten 603-607

[3] Felix Bieker, Benjamin Bremert, Marit Hansen: Verantwortlichkeit und Einsatz von Algorithmen bei öffentlichen Stellen; Datenschutz und Datensicherheit (DuD), Heft 10, 2018, Seiten 608-612

[4] Eberhard von Faber and Wolfgang Behnsen: Joint Security Management: organisationsübergreifend handeln – Mehr Sicherheit im Zeitalter von Cloud-Computing, IT-Dienstleistungen und industrialisierter IT-Produktion; Springer-Vieweg, 2018, X+234 Seiten, ISBN 978-3-658-20833-2

[5] Thomas Hemker: Machen Maschinen die Welt sicherer? – Ein Kurzüberblick und Werkstattbericht zum Einsatz von künstlicher Intelligenz und maschinellem Lernen in der Sicherheitstechnologie; Datenschutz und Datensicherheit (DuD), Heft 10, 2018, Seiten 629-633

[6] Thomas Dullien: Maschinelles Lernen und künstliche Intelligenz in der Informationssicherheit, Fortschritte, Anwendungen und Einschränkungen; Datenschutz und Datensicherheit (DuD), Heft 10, 2018, Seiten 618-622