Πανεπιστήμιο Δυτικής Αττικής

Σχολή Μηχανικών

Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών

Πρόγραμμα Μεταπτυχιακών Σπουδών (ΠΜΣ) Κυβερνοασφάλεια (Cybersecurity)

Ευφυείς Μηχανισμοί Ασφάλειας σε Επίπεδο Υλικού με Χρήση Μεθόδων Μηχανικής Μάθησης

Παππά Αγλαΐα

AM cscyb19021

Επιβλέπων: Δρ. Εμμανουήλ Θ. Μιχαηλίδης, Διδάσκων ΠΜΣ

Αιγάλεω, Ιούνιος 2021

Εξεταστική Επιτροπή

Α/Α	ΟΝΟΜΑ ΕΠΏΝΥΜΟ	ΒΑΘΜΊΔΑ/ΙΔΙΟΤΗΤΑ/ΤΜΗΜΑ	ΨΗΦΙΑΚΗ ΥΠΟΓΡΑΦΗ
1	Στέφανος Γκρίτζαλης	Καθηγητής Τμήμα Ψηφιακών Συστημάτων του Πανεπιστημίου Πειραιώς	STEFANOS GKRITZALIS 27/07/2023 10:18
2	Παναγιώτης Γιαννακόπουλος	Καθηγητής Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών, ΠΑΔΑ	Panagiotis Giannakopoulos Digitally signed by Panagiotis Giannakopoulos Date: 2023.07.27 00:04:48 +03'00'
3	Εμμανουήλ Μιχαηλίδης	Ακαδημαϊκός Υπότροφος Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών, ΠΑΔΑ	Emmanouil Michailidis Digitally signed by Emmanouil Michailidis Date: 2023.07.27 09:20:29 +03'00'

Δήλωση συγγραφέα μεταπτυχιακής εργασίας

Η κάτωθι υπογεγραμμένη Αγλαΐα Παππά φοιτήτρια του προγράμματος μεταπτυχιακών σπουδών «κυβερνοασφάλεια» του Τμήματος μηχανικών πληροφορικής και υπολογιστών της σχολής μηχανικών του πανεπιστημίου Δυτικής Αττικής δηλώνω ότι:

«Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποίαν είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης οι οποίες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών η λέξεων είτε ακριβώς είτε παραφρασμένες αναφέρονται στο σύνολό τους με πλήρη αναφορά στους συγγραφείς, στον εκδοτικό οίκο ή το περιοδικό συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο.

Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από εμένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου όσο και του Ιδρύματος. Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου.»

Η δηλούσα

Αγλαΐα Παππά

Περιεχόμενα

Περίληψη

Η απειλή υλικού (hardware threat) είναι οτιδήποτε μπορεί να οδηγήσει σε απώλεια ή καταστροφή δεδομένων ή φυσική ζημιά στο υλικό**.** Όταν δε οι διάφορες συσκευές (κάθε μία με τις δικές της ευπάθειες) συνδέονται μεταξύ τους μέσω internet τότε το πρόβλημα της ασφάλειας υλικού γίνεται χαοτικό. Η αντιμετώπιση με ανθρώπινες δυνάμεις είναι αδύνατη. Η μηχανική μάθηση είναι αυτή που με τους έξυπνους αλγορίθμους της αναλαμβάνει τη διαχείριση της ασφάλειας εκατομμυρίων συσκευών ταυτόχρονα. Μηχανική Μάθηση σημαίνει α)μάθηση χωρίς ανθρώπινη παρέμβαση και β)μάθηση για πάντα. Μαθαίνει, συγκρίνει, αξιολογεί, προφυλάσσει, προβλέπει. Αν η γνώση είναι δύναμη τότε η εφαρμοσμένη γνώση είναι υπερδύναμη. Η παρούσα διπλωματική εστιάζει στην αντιμετώπιση των κυριότερων ευπαθειών και επιθέσεων υλικού με τα εργαλεία της Μηχανικής Μάθησης. Τέλος ειδική μνεία γίνεται στα προβλήματα που μπορούν να προκύψουν από την κακόβουλη χρήση της καθώς και και νέα πεδία προέκτασης αυτής.

Λέξεις-κλειδιά: Ασφάλεια υλικού, Μηχανική Μάθηση, Διαδίκτυο των Πραγμάτων, επιθέσεις, αντίμετρα

Abstract

Hardware threat is anything that can lead to data loss or destruction or physical damage to the hardware. When the various devices (each with its own vulnerabilities) are connected to each other via the internet then the problem of hardware security becomes chaotic. Tackling the problem with human efforts is impossible. Machine learning is the one that with its smart algorithms manages the security of millions of devices at the same time. Machine learning means a) learning without human intervention and b) learning forever. It Learns, it compares, it evaluates, it protects, it predicts. If knowledge is power then applied knowledge is superpower. The following dissertation focuses on dealing with the main vulnerabilities and hardware attacks with the tools of Machine Learning. Special mention is made on the problems that may arise from its malicious use as well as new areas of its extension.

Keywords: Hardware security, Machine Learning, Internet of Things, attacks, countermeasures

1. Υλικό

1.1 Τι είναι το υλικό

Το υλικό είναι οποιοδήποτε φυσικό συστατικό ενός συστήματος που περιέχει πλακέτα κυκλώματος, IC ή άλλα ηλεκτρονικά. Υλικό δεν είναι μόνο ο Η/Υ. Είτε πρόκειται για οθόνη, tablet ή smartphone, όλα αυτά είναι υλικό. Χωρίς υλικό, οποιαδήποτε συσκευή δεν θα υπήρχε και το λογισμικό δεν θα μπορούσε να χρησιμοποιηθεί.

Εικόνα 1: Η εικόνα είναι μια κάμερα Logitech, ένα παράδειγμα εξωτερικού περιφερειακού υλικού . Αυτή η συσκευή υλικού επιτρέπει στους χρήστες να τραβούν βίντεο ή φωτογραφίες και να τα μεταδίδουν μέσω του Διαδικτύου.

Παρακάτω είναι μια λίστα με εξωτερικό υλικό ή υλικό που βρίσκεται εκτός υπολογιστή:

- Flat-panel, monitor, και LCD
- Gamepad
- Χειριστήριο (Joystick)
- Πληκτρολόγιο
- Μικρόφωνο
- Ποντίκι
- Εκτυπωτής
- Προβολέας
- Scanner

- Ηχεία

- USB thumb drive

Παρακάτω είναι μια λίστα με εσωτερικό υλικό ή υλικό που βρίσκεται μέσα σε έναν υπολογιστή:

- CPU (κεντρική μονάδα επεξεργασίας) .
- Μονάδα δίσκου (π.χ. Blu-ray, CD-ROM, DVD, δισκέτα, σκληρός δίσκος και SSD).
- Ανεμιστήρας (ψύκτρα)
- Μόντεμ
- Μητρική πλακέτα
- Κάρτα δικτύου
- Παροχή ηλεκτρικού ρεύματος
- RAM
- Κάρτα ήχου
- Κάρτα βίντεο

1.2 Τι είναι το Internet of Things;

Το Διαδίκτυο των πραγμάτων ή Ίντερνετ των πραγμάτων (αγγλικά: Internet of things) αποτελεί το δίκτυο επικοινωνίας πληθώρας συσκευών, οικιακών συσκευών, αυτοκινήτων καθώς και κάθε αντικειμένου που ενσωματώνει ηλεκτρονικά μέσα, λογισμικό, αισθητήρες και συνδεσιμότητα σε δίκτυο ώστε να επιτρέπεται η σύνδεση και η ανταλλαγή δεδομένων. Απλούστερα, η φιλοσοφία του IoT είναι η σύνδεση όλων των ηλεκτρονικών συσκευών μεταξύ τους (τοπικό δίκτυο) ή με δυνατότητα σύνδεσης στο διαδίκτυο (παγκόσμιο ιστό).

Η έννοια "Things" (πράγματα) δεν είναι αυστηρά συνδεδεμένη με ορισμένα προϊόντα. Αναφέρεται σε μία ευρεία ποικιλία συσκευών εντελώς διαφορετικά μεταξύ τους, όπως για παράδειγμα αυτοκίνητα με ενσωματωμένους αισθητήρες, κάμερες, κλιματιστικά, φώτα, συστήματα ασφαλείας, smartwatches ακόμα και αυτοκίνητα των οποίων οι περίπλοκοι αισθητήρες εντοπίζουν αντικείμενα στην πορεία τους. Είναι μερικά από τα πολλά προϊόντα τεχνολογίας. Βασικό χαρακτηριστικό όλων είναι η σύνδεση μεταξύ τους με απώτερο σκοπό την δυνατότητα του χρήστη να τα ελέγχει από έναν υπολογιστή ή κινητό. Ο όρος Internet of Things επινοήθηκε την δεκαετία του 1990 από τον Kevin Ashton. ο οποίος είναι ένας από τους ιδρυτές του Auto-ID center στο MIT, ήταν μέρος μιας ομάδας που ανακάλυψε τον τρόπο να συνδέσει τα αντικείμενα με το Διαδίκτυο μέσω μιας ετικέτας RFID.

Μερικά κορυφαία παραδείγματα Internet-of-Things (IoT) είναι τα εξής:

- Συνδεδεμένες συσκευές
- Έξυπνα συστήματα ασφαλείας στο σπίτι
- Αυτόνομος γεωργικός εξοπλισμός
- Wearable health monitors
- Έξυπνος εργοστασιακός εξοπλισμός
- Wireless inventory trackers
- Ασύρματο internet εξαιρετικά υψηλής ταχύτητας
- Βιομετρικοί σαρωτές στον κυβερνοχώρο (Biometric cybersecurity scanners)
- Shipping container and logistics tracking

Οι συσκευές IoT περιλαμβάνουν συνήθως τέσσερα στοιχεία - αισθητήρες, συνδεσιμότητα, επεξεργασία δεδομένων και διεπαφή χρήστη. Οι αισθητήρες είναι αντικείμενα που συλλέγουν δεδομένα και τα στέλνουν μέσω του Διαδικτύου. Τα δεδομένα αποστέλλονται για αποθήκευση, επεξεργασία ή περαιτέρω διάδοση πληροφοριών. Η επεξεργασία γίνεται μέσω αλγορίθμων μηχανικής μάθησης. Η μηχανική εκμάθηση είναι όταν οι υπολογιστές μαθαίνουν με παρόμοιο τρόπο με τον άνθρωπο - συλλέγοντας δεδομένα από το περιβάλλον τους - και αυτό είναι που κάνει τις συσκευές IoT έξυπνες.

1.3 Ποιες είναι οι μεγαλύτερες απειλές ασφάλειας υλικού;

Η ανάπτυξη IoT οδηγεί σε αύξηση των απειλών για την ασφάλεια υλικού, με τους εισβολείς να επιτίθενται σε ευπάθειες στο υλικό, στο firmware και στο Unified Extensible Firmware Interface / BIOS λογισμικό που διασυνδέεται με το υλικό.

Η αναζήτηση των οργανισμών για αυτοματοποίηση μη αυτόματων εργασιών οδήγησε σε μια αναταραχή περιστατικών ασφάλειας υλικού. Σήμερα, τα πάντα, από τις κάμερες παρακολούθησης και τα συστήματα HVAC έως τις πλατφόρμες φυσικού ελεγκτή θυρών (physical door controller platforms) γίνονται IP-connected. Ως αποτέλεσμα, ορισμένες επιχειρήσεις προσπαθούν να μειώσουν το κόστος αγοράζοντας και εγκαθιστώντας IoT χαμηλού κόστους και υλικό «έξυπνου» κτηρίου. Ωστόσο, το υλικό αυτών των συσκευών συχνά βρίσκεται σε κίνδυνο παραβίασης - απειλώντας τη συνολική ασφάλεια ολόκληρου του εταιρικού δικτύου.

Κοινές απειλές για την ασφάλεια υλικού

Οι συσκευές IoT είναι ιδιαίτερα επικίνδυνες επειδή λειτουργούν ανεξάρτητα. Ο εντοπισμός του χρόνου που έγινε μια επίθεση σε ένα στοιχείο IoT είναι πιο δύσκολο από το να κάνουμε

τον ίδιο προσδιορισμό για διακομιστές, επιτραπέζιους υπολογιστές / φορητούς υπολογιστές ή έξυπνες συσκευές. Αυτό δεν σημαίνει ότι δεν υπάρχουν απειλές ασφάλειας υλικού και για αυτές τις συσκευές. Τα κοινά ελαττώματα ασφάλειας υλικού περιλαμβάνουν τα ακόλουθα:

Προεπιλεγμένοι κωδικοί πρόσβασης. Αυτό είναι πρωτίστως ένα ζήτημα για συσκευές και υλικό IoT χαμηλού κόστους που χρησιμοποιούν προεπιλεγμένους κωδικούς πρόσβασης (default passwords). Αυτοί οι κωδικοί πρόσβασης προστίθενται συνήθως στα επιχειρηματικά δίκτυα.

Μη προστατευμένη τοπική πρόσβαση. Σε πολλές περιπτώσεις, το IoT, το IIoT και το υλικό έξυπνου κτηρίου είναι προσβάσιμα τοπικά μέσω διαχειριζόμενου Ethernet ή σειριακής διεπαφής. Εάν αυτές οι συνδέσεις δεν είναι κλειδωμένες - είτε από διαμόρφωση/παραμετροποίηση δική μας είτε και από φυσική άποψη - ένας προσποιητής μπορεί να είναι σε θέση να θέσει σε κίνδυνο την υποδομή μιας εταιρείας παραβιάζοντας αυτές τις συσκευές ενώ επισκέπτεται το γραφείο, την αποθήκη ή το εργοστάσιο παραγωγής.

Παρωχημένο firmware/BIOS/Unified Extensible Firmware Interface συσκευής. Οι εταιρείες που κατασκευάζουν και πωλούν έξυπνα συστήματα HVAC(=Heating, ventilation, and air conditioning), ρομποτική εργοστασίων παραγωγής και άλλα εξαρτήματα IoT / IIoT που συνδέονται με IP δεν είναι απαραίτητα ειδικοί σε θέματα ασφάλειας πληροφορικής. Το firmware είναι συχνά γεμάτο σφάλματα και ελαττώματα ασφαλείας. Αυτή η ευπάθεια επιδεινώνεται από την ατημέλητη διαχείριση κώδικα, καθώς πολλά τμήματα πληροφορικής δεν ενημερώνουν τακτικά firmware σε αυτές τις συσκευές όταν κυκλοφορούν ενημερώσεις κώδικα ασφαλείας.

Purpose-built/custom chipsets. Τα προσαρμοσμένα chipset συνεχίζουν να αποτελούν μεγάλο μέρος του υλικού σε εταιρικά κέντρα δεδομένων ή σε επιτραπέζιους υπολογιστές προηγμένης τεχνολογίας. Επειδή αυτά τα ειδικά σχεδιασμένα τσιπ είναι προσαρμοσμένα για εξειδικευμένους σκοπούς, οι κριτικές ασφαλείας του κατασκευαστή δεν είναι τόσο έντονες όσο αυτές που πραγματοποιούνται για μάρκες που πρόκειται να εγκατασταθούν σε πολύ μεγαλύτερες ομάδες συσκευών. Με την πάροδο του χρόνου, οι χάκερ βρίσκουν ευπάθειες σε αυτές τις μάρκες, αναγκάζοντας τον κατασκευαστή να εκδώσει μια νέα ενημέρωση κώδικα.

Έλλειψη κρυπτογράφησης. Η κρυπτογράφηση συχνά λείπει σε λειτουργικές συσκευές τεχνολογίας που συνδέονται γρήγορα με IP. Μη κρυπτογραφημένα δεδομένα μπορούν είτε να συλλεχθούν μέσω του δικτύου είτε από κλεμμένες συσκευές που περιέχουν μη κρυπτογραφημένα δεδομένα που αποθηκεύονται απευθείας σε αυτά.

Παραδείγματα ευπαθειών υλικού από τον πραγματικό κόσμο

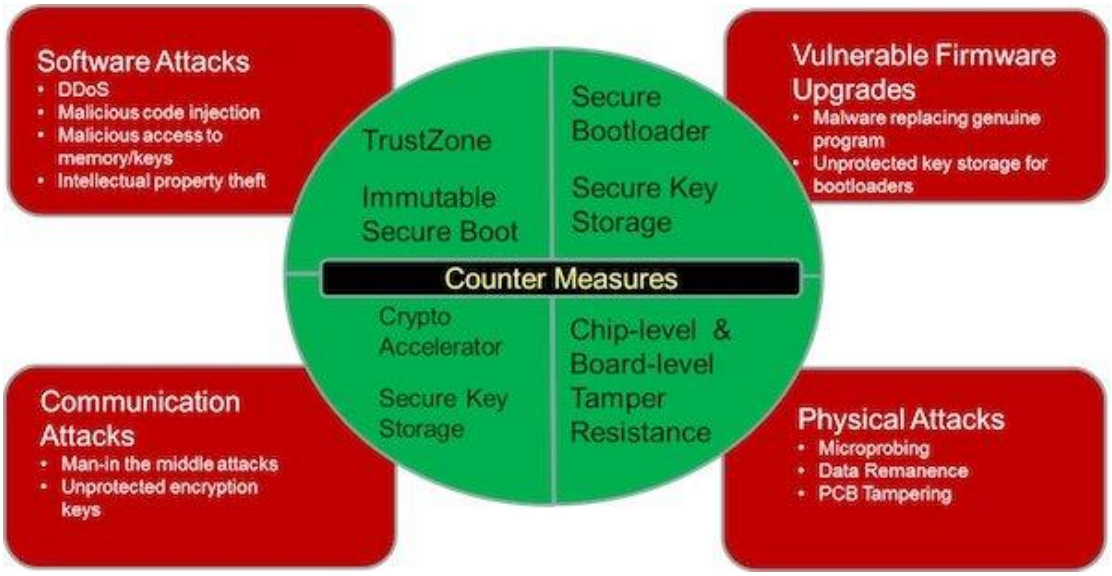
Τα νέα είναι γεμάτα λεπτομέρειες σχετικά με απειλές και ευπάθειες ασφάλειας υλικού. Στις αρχές του 2020, οι ερευνητές ασφαλείας προειδοποίησαν για ένα ελάττωμα ασφαλείας που εντοπίστηκε σε ορισμένους επεξεργαστές Intel που επέτρεψαν στους χάκερ να εγκαταστήσουν κακόβουλο λογισμικό σε επίπεδο υλικού, καθιστώντας έτσι την προστασία κακόβουλου λογισμικού που βασίζεται σε λειτουργικό σύστημα.

Πιο πρόσφατα, η Nvidia κυκλοφόρησε μια ενημερωμένη έκδοση κώδικα για να συνδέσει μια ευπάθεια που θα μπορούσε να επιτρέψει στους χάκερ να πάρουν τον έλεγχο της γραμμής διακομιστών DGX υψηλής τεχνολογίας της εταιρείας. Αυτοί οι τύποι διακομιστών είναι συνηθισμένοι σε επιχειρήσεις που εκτελούν προηγμένη AI και μηχανική μάθηση, θέτοντας σε κίνδυνο ευαίσθητα δεδομένα.

Ένα τελευταίο παράδειγμα η ευπάθεια υλικού που εντοπίστηκε πρόσφατα στο έξυπνο τηλεχειριστήριο της XR11 που ελέγχεται με φωνή της Comcast. Εάν ένας χρήστης ενημέρωνε το τηλεχειριστήριο με παραβιασμένη έκδοση firmware, θα μπορούσε αποτελεσματικά να μετατραπεί σε συσκευή ακρόασης. Ευτυχώς, οι ερευνητές ασφαλείας βρήκαν το ελάττωμα και ενημέρωσαν την Comcast, η οποία γρήγορα αντέδρασε και έβγαλε μια ενημέρωση ασφαλείας.

Συμπερασματικά

Τα ζητήματα ασφάλειας υλικού διαφέρουν από τα κενά ασφαλείας που επικεντρώνονται στο λογισμικό καθώς τα πρώτα επηρεάζουν συνήθως εξειδικευμένα προϊόντα με τα οποία το προσωπικό ασφαλείας πληροφορικής δεν είναι υπερβολικά εξοικειωμένο. Εργαλεία όπως τα Intrusion Detection Systems – IDS χρησιμοποιούν AI για να βασίσουν την κανονική συμπεριφορά και να ενεργοποιήσουν μια ειδοποίηση όταν αυτή η συμπεριφορά υπερβαίνει ένα καθορισμένο όριο. Το υλικό που αλλάζει απότομα από τον «κανόνα» είναι ένα ενδεικτικό σημάδι ότι κάτι συμβαίνει/υπάρχει κίνδυνος.



Εικόνα 2: Μια άποψη των φυσικών και απομακρυσμένων απειλών ασφαλείας σε έναν κόμβο IoT και οι αντίστοιχες μετρικές τους, οι οποίες είναι ενσωματωμένες σε συστήματα για προστασία από τις επιθέσεις.

1.4 Σκοπός αυτής της εργασίας

Ο πρωταρχικός σκοπός αυτής της εργασίας είναι να δείξει τις τελευταίες εξελίξεις στην εφαρμογή τεχνικών που βασίζονται στη μηχανική μάθηση σε τομείς ασφάλειας υλικού και να παρέχει μια γενική κατανόηση και ένα οδηγό σε όσους θέλουν να συσχετίσουν τη μηχανική μάθηση με την ασφάλεια υλικού.

Το κεφάλαιο 2 πραγματεύεται βασικές έννοιες και εφαρμογές της Μηχανικής Μάθησης και προσπαθεί να εξηγήσει τη σπουδαιότητά της στον τομέα της ασφάλειας υλικού.

Το κεφάλαιο 3 περιγράφει τους βασικούς αλγορίθμους και τις ιδιότητες τους. Οι αλγόριθμοι και οι μέθοδοι μηχανικής μάθησης που αναφέρονται συγκρίνονται, λαμβάνοντας υπόψη την πολυπλοκότητα που επιφέρουν στο σύστημα, καθώς και την αποτελεσματικότητά τους.

Στο κεφάλαιο 4 αναλύονται τα σημαντικότερα αντίμετρα απέναντι σε επιθέσεις υλικού, τα οποία είναι βασισμένα σε μεθόδους μηχανικής μάθησης. Γίνεται αναφορά σε τεχνικές μηχανικής μάθησης για ανίχνευση εισβολών, σε αλγορίθμους προστασίας υλικού και σε τεχνικές για το υλικό των συσκευών του IoT. Στο κεφάλαιο αυτό, αναφέρονται και παραδείγματα από εταιρείες που έκαναν χρήση από τέτοιες προσεγγίσεις, όπως η Exabeam η οποία χρησιμοποίησε τη Μηχανική Μάθηση για την προστασία στο υλικό των IoT συσκευών.

Επιπρόσθετα, είναι αποδεδειγμένο ότι πρέπει να ακολουθούνται συγκεκριμένες οδηγίες κατά την εφαρμογή της μηχανικής εκμάθησης σε οποιοδήποτε πρόβλημα που σχετίζεται με την ασφάλεια υλικού, όπως η επιλογή της κατάλληλης προεπεξεργασίας δεδομένων, η χρήση ενός

ισορροπημένου εκπαιδευτικού συνόλου δεδομένων, η χρήση διαφορετικών μετρήσεων αξιολόγησης απόδοσης και επίσης η κατάρτιση μοντέλων μηχανικής μάθησης πρέπει να γίνεται επαναληπτικά. Παρουσιάζεται, ότι η απόδοση των μοντέλων μηχανικής μάθησης εξαρτάται από συγκεκριμένους παράγοντες, όπως το μέγεθος και η αξία των επιλεγμένων χαρακτηριστικών δεδομένων, το επίπεδο θορύβου στα δεδομένα και επίσης το μέγεθος των εκπαιδευτικών δεδομένων.

Στο κεφάλαιο 5 αναφέρονται οι προκλήσεις και οι κίνδυνοι από την (κακόβουλη) χρήση της Μηχανικής Μάθησης και που θα πρέπει να εστιάσουμε την προσοχή μας για να τους αποφύγουμε.

Στο κεφάλαιο 6 οι μελέτες συνεχίζονται μέχρι σήμερα σε αυτόν τον τομέα, στο πλαίσιο της συνεργασίας της τεχνητής νοημοσύνης και των μεγάλων δεδομένων, προσπαθώντας να ξεπεράσουν τυχόν εμπόδια ή προβλήματα.

2. Μηχανική Μάθηση (Machine Learning)

2.1 Βασικές Έννοιες

Η μηχανική εκμάθηση είναι μια εφαρμογή τεχνητής νοημοσύνης (AI) που παρέχει στα συστήματα τη δυνατότητα αυτόματης μάθησης και βελτίωσης από την εμπειρία χωρίς να προγραμματίζονται ρητά. Η μηχανική μάθηση επικεντρώνεται στην ανάπτυξη προγραμμάτων υπολογιστών που μπορούν να έχουν πρόσβαση σε δεδομένα και να τα χρησιμοποιούν για να μάθουν μόνοι τους.

Η διαδικασία της μάθησης ξεκινά με παρατηρήσεις ή δεδομένα, όπως παραδείγματα, άμεση εμπειρία ή οδηγίες, προκειμένου να αναζητηθούν μοτίβα στα δεδομένα και να ληφθούν καλύτερες αποφάσεις στο μέλλον με βάση τα παραδείγματα που παρέχουμε. Ο πρωταρχικός στόχος είναι να επιτρέπεται στους υπολογιστές να μαθαίνουν αυτόματα χωρίς ανθρώπινη παρέμβαση ή βοήθεια και να προσαρμόζουν ανάλογα τις ενέργειες.

Διαφορά μεταξύ AI και Machine Learning

Πίνακας 1: Διαφορές Τεχνητής Νοημοσύνης και Μηχανικής Μάθησης

Τεχνητή νοημοσύνη	**Μηχανική εκμάθηση**
Η τεχνητή νοημοσύνη είναι μια τεχνολογία που επιτρέπει σε μια μηχανή να προσομοιώνει την ανθρώπινη συμπεριφορά.	Η μηχανική εκμάθηση είναι ένα υποσύνολο της τεχνητής νοημοσύνης που επιτρέπει σε μια μηχανή να μαθαίνει αυτόματα από προηγούμενα δεδομένα χωρίς προγραμματισμό ρητά.
Ο στόχος της τεχνητής νοημοσύνης είναι να κάνει ένα έξυπνο σύστημα υπολογιστών σαν τους ανθρώπους για να λύσει πολύπλοκα προβλήματα.	Ο στόχος του ML είναι να επιτρέψει στις μηχανές να μαθαίνουν από δεδομένα, ώστε να μπορούν να παρέχουν ακριβή έξοδο.
Στο AI, φτιάχνουμε έξυπνα συστήματα για να εκτελέσουμε οποιαδήποτε εργασία σαν τον άνθρωπο.	Στο ML, διδάσκουμε μηχανές με δεδομένα να εκτελούν μια συγκεκριμένη εργασία και να δίνουν ένα ακριβές αποτέλεσμα.
Η μηχανική μάθηση και η βαθιά μάθηση είναι τα δύο κύρια υποσύνολα του AI.	Η βαθιά μάθηση είναι ένα κύριο υποσύνολο της μηχανικής μάθησης.
Η τεχνητή νοημοσύνη έχει ένα πολύ ευρύ φάσμα πεδίων.	Η μηχανική μάθηση έχει περιορισμένο πεδίο εφαρμογής.
Η AI εργάζεται για τη δημιουργία ενός ευφυούς συστήματος που μπορεί να εκτελεί διάφορες πολύπλοκες εργασίες.	Η μηχανική εκμάθηση εργάζεται για τη δημιουργία μηχανών που μπορούν να εκτελέσουν μόνο εκείνες τις συγκεκριμένες εργασίες για τις οποίες έχουν εκπαιδευτεί.
Το σύστημα AI ανησυχεί για τη μεγιστοποίηση των πιθανοτήτων επιτυχίας.	Η μηχανική μάθηση ασχολείται κυρίως με την ακρίβεια και τα πρότυπα.

Οι κύριες εφαρμογές του AI είναι το **Siri, η υποστήριξη πελατών που χρησιμοποιεί catboats**, το Expert System, το Online παιχνίδι, έξυπνο ανθρωποειδές ρομπότ κ.λπ.	Οι κύριες εφαρμογές της μηχανικής μάθησης είναι το **σύστημα σύστασης σε απευθείας σύνδεση**, **οι αλγόριθμοι αναζήτησης Google**, **οι προτάσεις αυτόματων ετικετών φίλων Facebook** κ.λπ.
Με βάση τις δυνατότητες, το AI μπορεί να χωριστεί σε τρεις τύπους, που είναι: **Weak AI**, **General AI** και **Strong AI** .	Η μηχανική μάθηση μπορεί επίσης να διαιρεθεί σε τρία κυρίως είδη που είναι **Εποπτευόμενοι μάθησης**, **Μάθηση χωρίς επίβλεψη** και **ενίσχυση της μάθησης** .
Περιλαμβάνει μάθηση, συλλογιστική και αυτο-διόρθωση.	Περιλαμβάνει μάθηση και αυτο-διόρθωση όταν εισάγονται με νέα δεδομένα.
Το AI ασχολείται πλήρως με τα δομημένα, ημι-δομημένα και μη δομημένα δεδομένα.	Η μηχανική εκμάθηση ασχολείται με δομημένα και ημι-δομημένα δεδομένα.

2.2 Εφαρμογές

Η μηχανική εκμάθηση είναι μια λέξη-κλειδί για τη σημερινή τεχνολογία και αναπτύσσεται πολύ γρήγορα μέρα με τη μέρα. Χρησιμοποιούμε μηχανική εκμάθηση στην καθημερινή μας ζωή ακόμη και χωρίς να το γνωρίζουμε, όπως οι Χάρτες Google, ο βοηθός της Google, η Alexa κ.λπ. Ακολουθούν μερικές πιο δημοφιλείς πραγματικές εφαρμογές της Μηχανικής Μάθησης:

1. Αναγνώριση εικόνας

Η αναγνώριση εικόνας είναι μία από τις πιο κοινές εφαρμογές της μηχανικής μάθησης. Χρησιμοποιείται για τον εντοπισμό αντικειμένων, ατόμων, τοποθεσιών, ψηφιακών εικόνων κ.λπ. Η δημοφιλής περίπτωση χρήσης της αναγνώρισης εικόνας και της ανίχνευσης προσώπου είναι η «Automatic friend tagging suggestion»: Το Facebook μας παρέχει μια δυνατότητα προτάσεων αυτόματης προσθήκης ετικετών φίλων. Κάθε φορά που ανεβάζουμε μια φωτογραφία με τους φίλους μας στο Facebook, τότε λαμβάνουμε αυτόματα μια πρόταση με ετικέτες με όνομα και η τεχνολογία πίσω από αυτό είναι ο αλγόριθμος αναγνώρισης και αναγνώρισης προσώπου της μηχανικής μάθησης. Βασίζεται στο project με τίτλο " Deep Face ", το οποίο είναι υπεύθυνο για την αναγνώριση προσώπου και την αναγνώριση προσώπων στην εικόνα.

2. Αναγνώριση ομιλίας

Ενώ χρησιμοποιούμε το Google, έχουμε την επιλογή " Αναζήτηση με φωνή "· υπόκειται στην αναγνώριση ομιλίας και είναι μια δημοφιλής εφαρμογή μηχανικής μάθησης.

Η αναγνώριση ομιλίας είναι μια διαδικασία μετατροπής φωνητικών οδηγιών σε κείμενο και είναι επίσης γνωστή ως " Ομιλία σε κείμενο " ή " Αναγνώριση ομιλίας υπολογιστή ". Προς το παρόν, οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται ευρέως από διάφορες εφαρμογές αναγνώρισης ομιλίας. Οι βοηθοί Google, Siri, Cortana και Alexa χρησιμοποιούν τεχνολογία αναγνώρισης ομιλίας για να ακολουθήσουν τις φωνητικές οδηγίες.

3. Πρόβλεψη κυκλοφορίας

Αν θέλουμε να επισκεφθούμε ένα νέο μέρος, παίρνουμε τη βοήθεια των Χαρτών Google, που μας δείχνει τη σωστή διαδρομή με τη συντομότερη διαδρομή και προβλέπει τις συνθήκες κυκλοφορίας. Προβλέπει τις συνθήκες κυκλοφορίας, όπως εάν η αραιή κυκλοφορία, αργή κίνηση ή έντονη συμφόρηση με τη βοήθεια δύο τρόπων: α) Τοποθεσία σε πραγματικό χρόνο του οχήματος από την εφαρμογή Google Maps και αισθητήρες και β) μέσος χρόνος που χρειάστηκε τις τελευταίες ημέρες. Όλοι όσοι χρησιμοποιούν το Google Map βοηθούν αυτήν την εφαρμογή να βελτιωθεί. Παίρνει πληροφορίες από τον χρήστη και στέλνει πίσω στη βάση δεδομένων του για να βελτιώσει την απόδοση.

4. Συστάσεις προϊόντων (Product recommendations)

Η μηχανική εκμάθηση χρησιμοποιείται ευρέως από διάφορες εταιρείες ηλεκτρονικού εμπορίου και ψυχαγωγίας όπως το Amazon, το Netflix κ.λπ., για σύσταση προϊόντος στον χρήστη. Κάθε φορά που ψάχνουμε για κάποιο προϊόν στο Amazon, τότε αρχίζουμε να λαμβάνουμε μια διαφήμιση για το ίδιο προϊόν ενώ σερφάρουμε στο Διαδίκτυο στο ίδιο πρόγραμμα περιήγησης και αυτό οφείλεται στη μηχανική εκμάθηση. Η Google κατανοεί το ενδιαφέρον των χρηστών χρησιμοποιώντας διάφορους αλγόριθμους μηχανικής εκμάθησης και προτείνει το προϊόν σύμφωνα με το ενδιαφέρον των πελατών. Ομοίως, όταν χρησιμοποιούμε το Netflix, βρίσκουμε κάποιες προτάσεις για ψυχαγωγικές σειρές, ταινίες κ.λπ.

5. Αυτοκινούμενα αυτοκίνητα

Μία από τις πιο συναρπαστικές εφαρμογές της μηχανικής μάθησης είναι τα αυτοκίνητα αυτο-οδήγησης. Η μηχανική μάθηση διαδραματίζει σημαντικό ρόλο στα αυτοκίνητα αυτο-οδήγησης. Η Tesla, η πιο δημοφιλής εταιρεία κατασκευής αυτοκινήτων εργάζεται για αυτο-οδήγηση αυτοκινήτου. Χρησιμοποιεί μη επιτηρούμενη μέθοδο εκμάθησης για να εκπαιδεύσει τα μοντέλα αυτοκινήτων για να ανιχνεύει άτομα και αντικείμενα κατά την οδήγηση.

6. Φιλτράρισμα ανεπιθύμητων μηνυμάτων ηλεκτρονικού ταχυδρομείου και κακόβουλου προγράμματος

Κάθε φορά που λαμβάνουμε ένα νέο email, φιλτράρεται αυτόματα ως σημαντικό, κανονικό και ανεπιθύμητο. Πάντα λαμβάνουμε ένα σημαντικό μήνυμα στα εισερχόμενά μας με τα σημαντικά σύμβολα και τα ανεπιθύμητα μηνύματα στο πλαίσιο ανεπιθύμητων μηνυμάτων μας, και η τεχνολογία πίσω από αυτό είναι η Μηχανική εκμάθηση. Ακολουθούν ορισμένα φίλτρα ανεπιθύμητης αλληλογραφίας που χρησιμοποιούν το Gmail:

Φίλτρο περιεχομένου

Φίλτρο κεφαλίδας

Γενικό φίλτρο black list

Φίλτρα βάσει κανόνων

Permission filters

Ορισμένοι αλγόριθμοι μηχανικής εκμάθησης, όπως το Multi-Layer Perceptron, το tree Decision και ο ταξινομητής Naïve Bayes, χρησιμοποιούνται για φιλτράρισμα ανεπιθύμητων μηνυμάτων email και εντοπισμό κακόβουλου λογισμικού.

7. Εικονικός προσωπικός βοηθός

Έχουμε διάφορους εικονικούς προσωπικούς βοηθούς όπως ο Βοηθός Google, Alexa, Cortana, Siri . Όπως υποδηλώνει το όνομα, μας βοηθούν στην εύρεση των πληροφοριών χρησιμοποιώντας τη φωνητική μας οδηγία. Αυτοί οι βοηθοί μπορούν να μας βοηθήσουν με διάφορους τρόπους μόνο με τις φωνητικές μας οδηγίες, όπως Αναπαραγωγή μουσικής, κλήση σε κάποιον, Άνοιγμα email, Προγραμματισμός ραντεβού κ.λπ. Αυτοί οι βοηθοί καταγράφουν τις φωνητικές μας οδηγίες, τις στέλνουν μέσω του διακομιστή σε ένα cloud και την αποκωδικοποιούν χρησιμοποιώντας αλγόριθμους ML και ενεργούν ανάλογα.

8. Διαδικτυακή ανίχνευση απάτης

Η μηχανική εκμάθηση καθιστά την ηλεκτρονική μας συναλλαγή ασφαλή και ασφαλή εντοπίζοντας συναλλαγή απάτης. Κάθε φορά που πραγματοποιούμε κάποια διαδικτυακή συναλλαγή, μπορεί να υπάρχουν διάφοροι τρόποι με τους οποίους μπορεί να πραγματοποιηθεί μια δόλια συναλλαγή, όπως ψεύτικοι λογαριασμοί, ψεύτικες ταυτότητες και κλοπή χρημάτων

στη μέση μιας συναλλαγής. Έτσι, για να το εντοπίσουμε αυτό, το δίκτυο Feed Forward Neural μας βοηθά ελέγχοντας αν πρόκειται για πραγματική συναλλαγή ή για συναλλαγή απάτης. Για κάθε γνήσια συναλλαγή, η έξοδος μετατρέπεται σε ορισμένες τιμές κατακερματισμού και αυτές οι τιμές γίνονται η είσοδος για τον επόμενο γύρο. Για κάθε γνήσια συναλλαγή, υπάρχει ένα συγκεκριμένο μοτίβο που αλλάζει για τη συναλλαγή απάτης, το εντοπίζει και κάνει τις διαδικτυακές μας συναλλαγές πιο ασφαλείς.

9. Συναλλαγές χρηματιστηρίου (Stock Market trading)

Η μηχανική μάθηση χρησιμοποιείται ευρέως στο χρηματιστήριο. Στο χρηματιστήριο, υπάρχει πάντα ο κίνδυνος αύξησης των μετοχών, επομένως για το βραχυπρόθεσμο νευρωνικό δίκτυο μνήμης αυτής της μηχανικής μάθησης χρησιμοποιείται για την πρόβλεψη των τάσεων του χρηματιστηρίου.

10. Ιατρική διάγνωση

Στην ιατρική επιστήμη, η μηχανική μάθηση χρησιμοποιείται για τη διάγνωση ασθενειών. Με αυτό, η ιατρική τεχνολογία αναπτύσσεται πολύ γρήγορα και μπορεί να δημιουργήσει μοντέλα 3D που μπορούν να προβλέψουν την ακριβή θέση των βλαβών στον εγκέφαλο. Βοηθά στην εύρεση όγκων του εγκεφάλου και άλλων σχετικών με τον εγκέφαλο ασθενειών.

11. Αυτόματη μετάφραση γλώσσας

Σήμερα, αν επισκεφθούμε ένα νέο μέρος και δεν γνωρίζουμε τη γλώσσα, τότε δεν είναι καθόλου πρόβλημα, καθώς και αυτή η μηχανική μάθηση μας βοηθά μετατρέποντας το κείμενο σε γνωστές γλώσσες μας. Το Google GNMT (Google Neural Machine Μετάφραση) παρέχει αυτήν τη δυνατότητα, η οποία είναι μια Νευρωνική Μηχανική Εκμάθηση που μεταφράζει το κείμενο στη γνωστή μας γλώσσα, και το ονομάστηκε ως αυτόματη μετάφραση. Η τεχνολογία πίσω από την αυτόματη μετάφραση είναι ένας sequence to sequence αλγόριθμος εκμάθησης, ο οποίος χρησιμοποιείται με την αναγνώριση εικόνας και μεταφράζει το κείμενο από τη μία γλώσσα στην άλλη.

2.3 Αξιοποίηση στην ασφάλεια υλικού (hardware security)

Η αξιοσημείωτη επιτυχία της μηχανικής μάθησης (ML) σε πολλούς ερευνητικούς τομείς έχει εμπνεύσει ακαδημαϊκές και βιομηχανικές κοινότητες να διερευνήσουν τις δυνατότητές της για την αντιμετώπιση επιθέσεων trojan υλικού. Ενώ πολυάριθμα έργα έχουν δημοσιευτεί την τελευταία δεκαετία, λίγα δημοσιευμένα έργα, από όσο γνωρίζουμε, έχουν επανεξετάσει

συστηματικά τα επιτεύγματα και ανέλυσαν τις υπόλοιπες προκλήσεις σε αυτόν τον τομέα. Για να καλυφθεί αυτό το κενό, το άρθρο στο [5] ερευνά προσεγγίσεις που βασίζονται στη ML ενάντια σε επιθέσεις HTs που είναι διαθέσιμες στη βιβλιογραφία. Συγκεκριμένα, παρέχεται πρώτα μια ταξινόμηση όλων των πιθανών επιθέσεων HTs και στη συνέχεια εξετάζονται οι πρόσφατες εξελίξεις από τέσσερις προοπτικές, δηλαδή από πλευράς ανίχνευσης HTs, σχεδιασμού ασφάλειας (Design-For-Security - DFS), ασφάλειας διαύλου και ασφάλειας αρχιτεκτονικής. Με βάση την ανασκόπηση αυτή αναφέρονται περαιτέρω συμπεράσματα και προκλήσεις που προέκυψαν από προηγούμενες μελέτες. Είναι αξιοσημείωτο ότι οι νέες απειλές HTs εμφανίζονται συνεχώς και έχουν εξελιχθεί πέρα από τα κυκλώματα και σε επίπεδο στοιχείων, συσκευών και ακόμη και συμπεριφορών, γεγονός που διακυβεύει την ασφάλεια και την αξιοπιστία του συνολικού οικοσυστήματος υλικού. Επομένως, οι απειλές HTs διαιρούνται σε τέσσερα επίπεδα και προτείνεται ένα μοντέλο αναφοράς άμυνα σε trojan υλικού (Hardware Trojan Defense - HTD) από την οπτική του συνολικού οικοσυστήματος υλικού, κατηγοριοποιώντας εκεί τις απειλές και τις απαιτήσεις ασφαλείας σε κάθε επίπεδο για να παραχθεί μια κατευθυντήρια γραμμή για μελλοντική έρευνα προς αυτήν την κατεύθυνση.

2.4 Γιατί τώρα;

Τα συστήματα AI / ML / DL βρίσκονται ακόμη στην απαρχή τους, παρά το γεγονός ότι έχουν ερευνηθεί με τον έναν ή τον άλλον τρόπο από τη δεκαετία του 1950. Αλλά μόλις αυτή τη δεκαετία η αγορά για συστήματα AI / ML / DL ξεκίνησε λόγω συνδυασμού πολλών παραγόντων:

- Υπάρχει αρκετή ισχύς επεξεργασίας και μνήμη για την επεξεργασία αλγορίθμων AI / ML / DL, τόσο σε κέντρα δεδομένων για εκπαίδευση όσο και σε όλο το δίκτυο για εξαγωγή συμπερασμάτων.

- Υπάρχουν πραγματικές εφαρμογές για αυτήν την τεχνολογία, επομένως υπάρχουν χρήματα για την ανάπτυξη καλύτερων αλγορίθμων και αποτελεσματικότερων αρχιτεκτονικών υλικού.

- Η τεχνολογία επέτρεψε την ανάπτυξη αλγορίθμων σε υπολογιστές και όχι στο χέρι, επιτρέποντας στις εταιρείες να ξεκινήσουν με αλγόριθμους από κάποια βάση αντί να προσπαθούν να αναπτύξουν τους δικούς τους.

2.5 Ήρθε για να μείνει

Όλα αυτά επέτρεψαν στα AI / ML / DL να συνεχίσουν την έρευνα που είχαν ξεκινήσει μεγάλες εταιρείες υπολογιστών όπως η IBM και η Digital Equipment στις αρχές της δεκαετίας του 1990. Έκτοτε, η IBM συνέχισε τις προσπάθειές της, συνεργάστηκε με cloud providers όπως την Amazon, την Microsoft, την Google, καθώς και την Alibaba, το Facebook και πολλές μικρότερες εταιρείες. Πλέον, ξοδεύονται δισεκατομμύρια δολάρια στο κομμάτι της έρευνας από κυβερνήσεις σε όλο τον κόσμο.

Σύμφωνα με μια νέα έκθεση του Brookings Institution, οι επενδύσεις σε μηχανική μάθηση αυξάνονται σε ένα ευρύ φάσμα αγορών, συμπεριλαμβανομένης της εθνικής ασφάλειας, του οικονομικού τομέα, της υγειονομικής περίθαλψης, της ποινικής δικαιοσύνης, των μεταφορών και των έξυπνων πόλεων. Η ανταμοιβή, σύμφωνα με την PricewaterhouseCoopers, ανέρχεται σε 15,7 τρισεκατομμύρια δολάρια σε δυνητική συνεισφορά στην παγκόσμια οικονομία έως το 2030.

"Αν πιστοποιήσω τη λειτουργία ενός προϊόντος και το πουλήσω και στην συνέχεια αλλάξει συμπεριφορά, τότε χάνει την αξία του και είμαι υπόλογος", δήλωσε ο Wally Rhines, πρόεδρος και διευθύνων σύμβουλος της Mentor, της Siemens Business. "Τι κάνουμε γι' αυτό; Τι κάνει ένας κατασκευαστής αυτοκινήτων; Με τη βοήθεια της μηχανικής μάθησης όταν ένα κύκλωμά δεν λειτουργεί του δίνει τη δυνατότητα να αυτο-τεστάρεται με ένα σύνολο κριτηρίων που έχει ορίσει ο κατασκευαστής του συστήματος και να έχει δυναμικό αυτοέλεγχο. Το σύστημα εξελίσσεται με την πάροδο του χρόνου όσο λαμβάνει όλο και περισσότερα δεδομένα και εφαρμόζει μηχανική μάθηση ώστε να επαληθεύει ότι δεν έχει τροποποιηθεί με τέτοιο τρόπο που θα μπορούσε να καταστεί επικίνδυνο ή μη λειτουργικό".

Εν κατακλείδι, η επίλυση αυτού του προβλήματος έχει ανοδική πορεία για εταιρείες που μπορούν να το αυτοματοποιήσουν και η βιομηχανία εξόρυξης δεδομένων γνωρίζει καλά την ευκαιρία όχι μόνο για τη χρήση των AI/ ML /DL εσωτερικά, αλλά και για την ανάπτυξη εργαλείων που μπορούν να ενισχύσουν την ανάπτυξη και την ασφάλεια των αλγορίθμων.

3. Αλγόριθμοι μηχανικής μάθησης

3.1 Βασικότερες κατηγορίες

Υπάρχουν τρεις τύποι των πιο δημοφιλών αλγορίθμων Μηχανικής Μάθησης:

1. Εποπτευόμενη μάθηση:

Αυτός ο αλγόριθμος αποτελείται από μια μεταβλητή στόχου / αποτελέσματος (ή εξαρτώμενη μεταβλητή) η οποία πρέπει να προβλεφθεί από ένα δεδομένο σύνολο προγνωστικών (ανεξάρτητες μεταβλητές). Χρησιμοποιώντας αυτά τα σύνολα μεταβλητών, δημιουργούμε μια συνάρτηση που αντιστοιχίζει τις εισόδους στις επιθυμητές εξόδους. Η διαδικασία εκπαίδευσης συνεχίζεται έως ότου το μοντέλο επιτύχει ένα επιθυμητό επίπεδο ακρίβειας στα δεδομένα εκπαίδευσης. Παραδείγματα εποπτευόμενης μάθησης: παλινδρόμηση, δέντρο απόφασης, τυχαίο δάσος, KNN, λογιστική παλινδρόμηση κ.λπ.

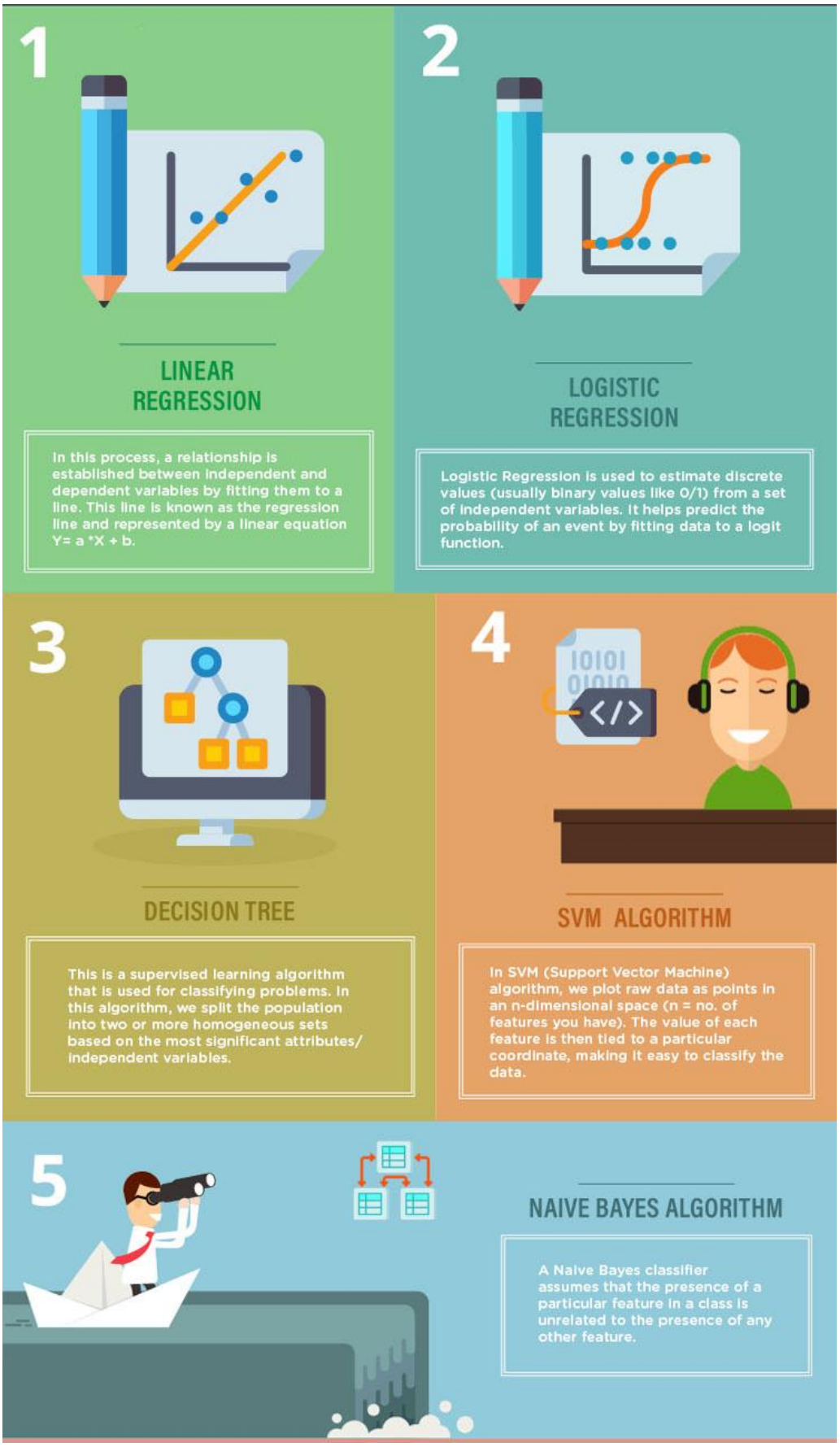
2. Μη εποπτευόμενη μάθηση:

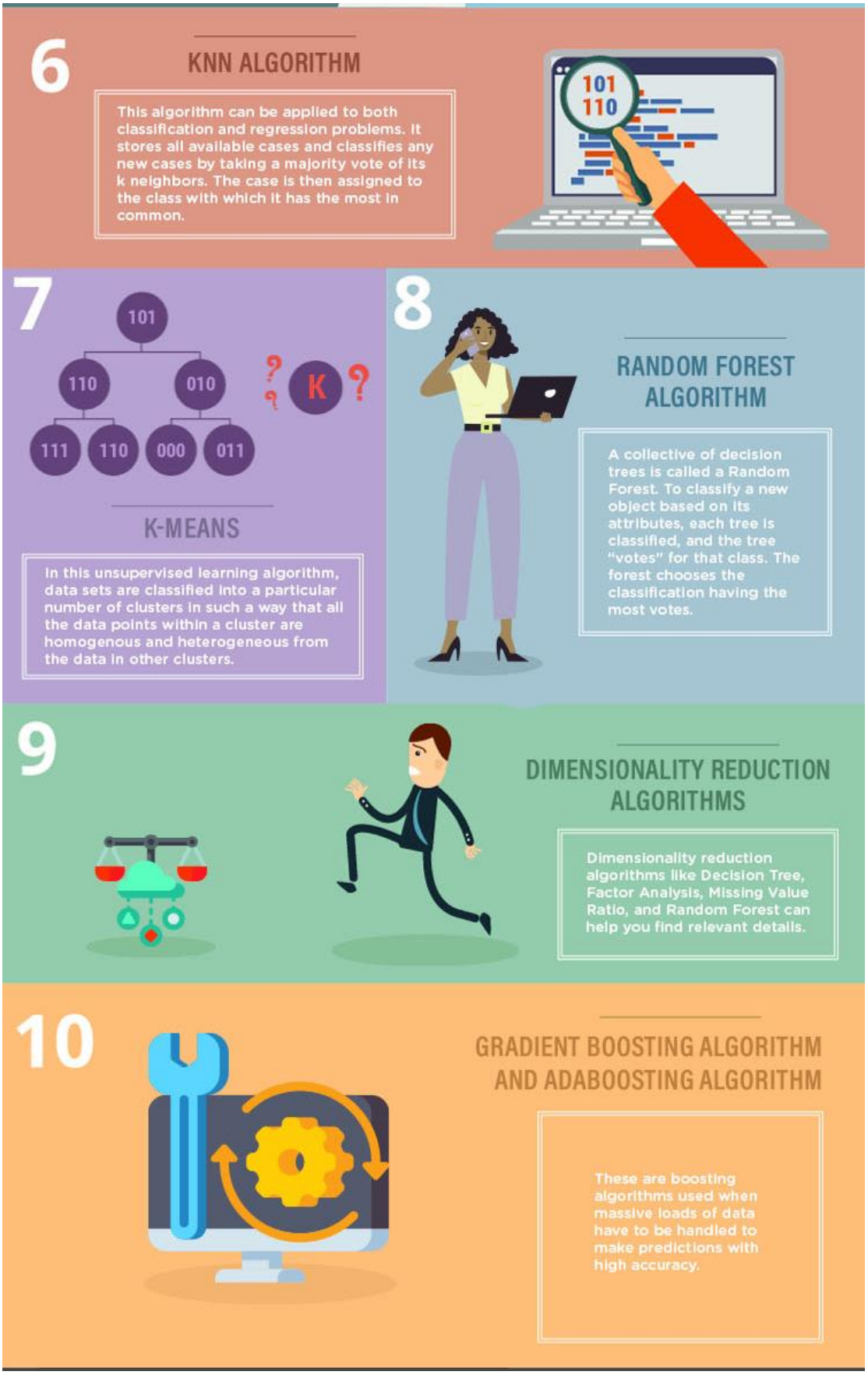
Σε αυτόν τον αλγόριθμο, δεν έχουμε καμία μεταβλητή στόχου ή αποτελέσματος για την πρόβλεψη/εκτίμηση. Χρησιμοποιείται για τη συγκέντρωση πληθυσμού σε διαφορετικές ομάδες, το οποίο χρησιμοποιείται ευρέως για την τμηματοποίηση πελατών σε διαφορετικές ομάδες για συγκεκριμένη παρέμβαση. Παραδείγματα μη εποπτευόμενης μάθησης: Αλγόριθμος Apriori, K-mean.

3. Μάθηση Ενίσχυσης:

Χρησιμοποιώντας αυτόν τον αλγόριθμο, το μηχάνημα είναι εκπαιδευμένο να λαμβάνει συγκεκριμένες αποφάσεις. Λειτουργεί με αυτόν τον τρόπο: το μηχάνημα εκτίθεται σε περιβάλλον όπου εκπαιδεύεται συνεχώς χρησιμοποιώντας δοκιμές και σφάλματα. Αυτό το μηχάνημα μαθαίνει από την εμπειρία του παρελθόντος και προσπαθεί να αντλήσει την καλύτερη δυνατή γνώση για να λάβει ακριβείς επιχειρηματικές αποφάσεις. Παράδειγμα εκμάθησης ενίσχυσης: Διαδικασία απόφασης Markov.

Αλγόριθμοι μηχανικής εκμάθησης





Εικόνα 3: Κορυφαίοι αλγόριθμοι μηχανικής μάθησης

Γραμμική παλινδρόμηση

Ένα παράδειγμα για τη λειτουργικότητα αυτού του αλγορίθμου είναι η τακτοποίηση τυχαίων κορμών ξύλου αυξάνοντας τη σειρά του βάρους τους. Υπάρχει μια παγίδα: δεν μπορούμε να σταθμίσουμε κάθε κούτσουρο. Πρέπει να μαντέψουμε το βάρος του απλώς κοιτάζοντας το ύψος και το πλάτος του κούτσουρου (οπτική ανάλυση) και να τα τακτοποιήσουμε χρησιμοποιώντας έναν συνδυασμό αυτών των ορατών παραμέτρων. Έτσι είναι η γραμμική παλινδρόμηση στη μηχανική μάθηση. Σε αυτήν τη διαδικασία, δημιουργείται μια σχέση μεταξύ ανεξάρτητων και εξαρτημένων μεταβλητών προσαρμόζοντάς τις σε μια γραμμή. Αυτή η γραμμή είναι γνωστή ως γραμμή παλινδρόμησης και αντιπροσωπεύεται από μια γραμμική εξίσωση Y= a*X+ b.

Σε αυτήν την εξίσωση:

Y - Εξαρτώμενη μεταβλητή

α - Κλίση

X - Ανεξάρτητη μεταβλητή

β - Παρεμβολή

Οι συντελεστές a & b προκύπτουν ελαχιστοποιώντας το άθροισμα της τετραγωνικής διαφοράς απόστασης μεταξύ σημείων δεδομένων και της γραμμής παλινδρόμησης.

Λογιστική παλινδρόμηση

Ο Logistic Regression χρησιμοποιείται για την εκτίμηση διακριτών τιμών (συνήθως δυαδικών τιμών όπως 0/1) από ένα σύνολο ανεξάρτητων μεταβλητών. Βοηθά στην πρόβλεψη της πιθανότητας ενός συμβάντος προσαρμόζοντας δεδομένα σε μια συνάρτηση logit.

Δέντρο απόφασης

Ο αλγόριθμος Decision Tree στη μηχανική μάθηση είναι ένας από τους πιο δημοφιλείς αλγόριθμους που χρησιμοποιούνται σήμερα. Αυτός είναι ένας εποπτευόμενος μαθησιακός αλγόριθμος που χρησιμοποιείται για την ταξινόμηση προβλημάτων. Λειτουργεί καλά ταξινομώντας τόσο κατηγορίες όσο και συνεχείς εξαρτώμενες μεταβλητές. Σε αυτόν τον αλγόριθμο, χωρίζουμε τον πληθυσμό σε δύο ή περισσότερα ομοιογενή σύνολα με βάση τα πιο σημαντικά χαρακτηριστικά/ανεξάρτητες μεταβλητές.

Αλγόριθμος SVM (Support Vector Machine)

Ο αλγόριθμος SVM είναι ένας αλγόριθμος ταξινόμησης στον οποίο βάζουμε ακατέργαστα δεδομένα ως σημεία σε έναν διαστατικό χώρο (όπου n είναι ο αριθμός των χαρακτηριστικών που έχουμε). Η τιμή κάθε δυνατότητας συνδέεται έπειτα με μια συγκεκριμένη συντεταγμένη, καθιστώντας εύκολη την ταξινόμηση των δεδομένων. Οι γραμμές που ονομάζονται ταξινομητές μπορούν να χρησιμοποιηθούν για να χωρίσουν τα δεδομένα και να τα σχεδιάσουν σε ένα γράφημα.

Αλγόριθμος Naive Bayes

Ένας ταξινομητής Naive Bayes υποθέτει ότι η παρουσία ενός συγκεκριμένου χαρακτηριστικού σε μια τάξη δεν σχετίζεται με την παρουσία οποιουδήποτε άλλου χαρακτηριστικού. Ακόμα κι αν αυτά τα χαρακτηριστικά σχετίζονται μεταξύ τους, ένας ταξινομητής Naive Bayes θα εξετάσει όλες αυτές τις ιδιότητες ανεξάρτητα κατά τον υπολογισμό της πιθανότητας ενός συγκεκριμένου αποτελέσματος. Ένα μοντέλο Naive Bayesian είναι εύκολο στη δημιουργία και χρήσιμο για τεράστια σύνολα δεδομένων. Είναι απλό και είναι γνωστό ότι ξεπερνά ακόμη και εξαιρετικά εξελιγμένες μεθόδους ταξινόμησης.

Αλγόριθμος KNN (K- Κοντινότεροι γείτονες)

Αυτός ο αλγόριθμος μπορεί να εφαρμοστεί σε προβλήματα ταξινόμησης και παλινδρόμησης. Προφανώς, στον κλάδο της Επιστήμης δεδομένων, χρησιμοποιείται ευρύτερα για την επίλυση προβλημάτων ταξινόμησης. Είναι ένας απλός αλγόριθμος που αποθηκεύει όλες τις διαθέσιμες περιπτώσεις και ταξινομεί τυχόν νέες περιπτώσεις, λαμβάνοντας την πλειοψηφία των k γειτόνων του. Στη συνέχεια, η υπόθεση ανατίθεται στην τάξη με την οποία έχει τα περισσότερα κοινά. Μια συνάρτηση απόστασης εκτελεί αυτήν τη μέτρηση. Το KNN μπορεί εύκολα να γίνει κατανοητό συγκρίνοντάς το με την πραγματική ζωή. Για παράδειγμα, εάν θέλουμε πληροφορίες για ένα άτομο, είναι λογικό να μιλήσουμε με τους φίλους και τους συναδέλφους του! Πράγματα που πρέπει να λάβουμε υπόψη πριν επιλέξουμε τον Αλγόριθμο K Nearest Neighbours:

- Το KNN είναι υπολογιστικά ακριβό
- Οι μεταβλητές πρέπει να κανονικοποιηθούν, διαφορετικά οι μεταβλητές υψηλότερου εύρους μπορούν να προκαλέσουν προκατάληψη στον αλγόριθμο
- Τα δεδομένα πρέπει ακόμη να υποστούν επεξεργασία.

K-Means

Είναι ένας αλγόριθμος μάθησης χωρίς επίβλεψη που επιλύει προβλήματα ομαδοποίησης. Τα σύνολα δεδομένων ταξινομούνται σε έναν συγκεκριμένο αριθμό συστάδων (ας ονομάσουμε αυτόν τον αριθμό K) με τέτοιο τρόπο ώστε όλα τα σημεία δεδομένων εντός ενός συμπλέγματος να είναι ομοιογενή και ετερογενή από τα δεδομένα σε άλλες συστάδες. Πώς το K-σημαίνει σχηματίζει συστάδες:

- Ο αλγόριθμος K-σημαίνει επιλέγει k αριθμό σημείων, που ονομάζονται κεντροειδή, για κάθε σύμπλεγμα
- Κάθε σημείο δεδομένων σχηματίζει ένα σύμπλεγμα με τα πλησιέστερα κεντροειδή, δηλαδή, συστάδες K.
- Δημιουργεί τώρα νέα κεντροειδή με βάση τα υπάρχοντα μέλη συμπλέγματος.
- Με αυτά τα νέα κεντροειδή, καθορίζεται η πλησιέστερη απόσταση για κάθε σημείο δεδομένων. Αυτή η διαδικασία επαναλαμβάνεται έως ότου τα κεντροειδή δεν αλλάξουν.

Random Forest Algorithm

Μια ομάδα δέντρων αποφάσεων ονομάζεται Random Forest . Για να ταξινομήσουμε ένα νέο αντικείμενο με βάση τις ιδιότητές του, κάθε δέντρο ταξινομείται και το δέντρο "ψηφίζει" για αυτήν την κλάση. Το δάσος επιλέγει την ταξινόμηση με τις περισσότερες ψήφους (πάνω από όλα τα δέντρα στο δάσος). Κάθε δέντρο φυτεύεται και καλλιεργείται ως εξής:

- Εάν ο αριθμός των περιπτώσεων στο training set είναι N, τότε λαμβάνονται τυχαία ένα δείγμα περιπτώσεων N. Αυτό το δείγμα θα είναι το εκπαιδευτικό σύνολο για την καλλιέργεια του δέντρου.
- Εάν υπάρχουν μεταβλητές εισόδου M, ο αριθμός m << M καθορίζεται έτσι ώστε σε κάθε κόμβο, οι μεταβλητές m επιλέγονται τυχαία από το M και η καλύτερη διαίρεση (split) σε αυτό το m χρησιμοποιείται για το διαχωρισμό του κόμβου. Η τιμή του m διατηρείται σταθερή κατά τη διάρκεια αυτής της διαδικασίας.
- Κάθε δέντρο καλλιεργείται στο μεγαλύτερο δυνατό βαθμό. Δεν υπάρχει κλάδεμα.

Αλγόριθμοι μείωσης διαστάσεων

Στον σημερινό κόσμο, τεράστιες ποσότητες δεδομένων αποθηκεύονται και αναλύονται από εταιρείες, κυβερνητικές υπηρεσίες και ερευνητικούς οργανισμούς. Ως επιστήμονας δεδομένων, γνωρίζουμε ότι αυτά τα ανεπεξέργαστα δεδομένα περιέχουν πολλές πληροφορίες - η πρόκληση είναι ο εντοπισμός σημαντικών προτύπων και μεταβλητών. Οι αλγόριθμοι

μείωσης διαστάσεων, όπως Decision Tree, Factor Analysis, Missing Value Ratio, and Random Forest μπορούν να μας βοηθήσουν να βρούμε σχετικές λεπτομέρειες.

Gradient Boosting Algorithm and AdaBoosting Algorithm

Αυτοί είναι οι αλγόριθμοι ενίσχυσης που χρησιμοποιούνται όταν πρέπει να χειριστούν τεράστια φορτία δεδομένων για να κάνουν προβλέψεις με υψηλή ακρίβεια. Το Boosting είναι ένας αλγόριθμος εκμάθησης συνόλου που συνδυάζει την προγνωστική ισχύ πολλών βασικών εκτιμητών για να βελτιώσει την ευρωστία. Με λίγα λόγια, συνδυάζει πολλαπλούς αδύναμους ή μέσους προγνωστικούς για να δημιουργήσει έναν ισχυρό προγνωστικό παράγοντα. Αυτοί οι αλγόριθμοι ενίσχυσης λειτουργούν πάντα καλά σε διαγωνισμούς επιστήμης δεδομένων όπως οι Kaggle, AV Hackathon, CrowdAnalytix. Αυτοί είναι οι πλέον προτιμώμενοι αλγόριθμοι μηχανικής μάθησης σήμερα.

3.2 Βασικές ιδιότητες αυτών-σύγκριση

Τα SVM και νευρωνικά δίκτυα λειτουργούν πολύ καλύτερα όταν δουλεύουν με πολλαπλές διαστάσεις και συνεχή χαρακτηριστικά. Από την άλλη πλευρά, τα λογικά συστήματα λειτουργούν καλύτερα για την αντιμετώπιση διακριτών/ιεραρχικών χαρακτηριστικών. Για τα μοντέλα νευρωνικών δικτύων και τα SVM, απαιτείται πλήθος δειγμάτων για την επίτευξη της μέγιστης ακρίβειας πρόβλεψης, ενώ τα συστήματα βασισμένα σε Naive Bayes (NB) ενδέχεται να απαιτούν νέα σύνολα δεδομένων.

Υπάρχει γενική συμφωνία ότι ο KNN είναι πολύ ευαίσθητος σε ασυσχέτιστα χαρακτηριστικά: αυτό το σημείο μπορεί να εξηγηθεί με βάση τον τρόπο που λειτουργεί ο αλγόριθμος. Οι περισσότεροι αλγόριθμοι δέντρων αποφάσεων ενδέχεται να μη λειτουργούν καλά με προβλήματα που απαιτούν διαγώνιο κατακερματισμό. Υπάρχει σχέση πολυγραμμικότητας και μια μη γραμμική σχέση μεταξύ των χαρακτηριστικών εισόδου και εξόδου όταν τα ANN και SVM λειτουργούν σωστά.

Ο αλγόριθμος Naive Bayes (NB) απαιτεί λιγότερο χώρο αποθήκευσης κατά τη διάρκεια των σταδίων εκπαίδευσης και ταξινόμησης: απαιτείται μνήμη για την αποθήκευση των a priori πιθανοτήτων και των δεσμευμένων. Ο βασικός αλγόριθμος KNN καταναλώνει μεγάλο χώρο αποθήκευσης για τη φάση εκπαίδευσης και ο χώρος εφαρμογής του είναι τουλάχιστον μεγαλύτερος από τον χώρο εκπαίδευσης. Επιπλέον, οι Naive Bayes και KNN μπορούν εύκολα να χρησιμοποιηθούν ως αυξανόμενοι μαθητές, ενώ οι κανονικοποιημένοι αλγόριθμοι δεν

μπορούν. Ο Naive Bayes είναι από τη φύση του ισχυρός για τις τιμές που λείπουν, επειδή αγνοούνται στο δυναμικό των υπολογισμών του και επομένως δεν επηρεάζουν την τελική απόφαση. Αντιθέτως, ο KNN και τα νευρωνικά δίκτυα απαιτούν ολόκληρες τις εγγραφές για να ολοκληρώσουν την εργασία τους. Τα SVM και ANN έχουν παρόμοιο λειτουργικό προφίλ. Ένας μεμονωμένος αλγόριθμος δεδομένων δεν αντισταθμίζει όλους τους άλλους αλγόριθμους σε όλα τα σύνολα δεδομένων.

Μπορούν να οριστούν διαφορετικά δεδομένα με διαφορετικούς τύπους μεταβλητών και ο αριθμός των στιγμιότυπων καθορίζει τον τύπο του αλγορίθμου που λειτουργεί καλύτερα. Ο παρακάτω πίνακας παρουσιάζει μια συγκριτική ανάλυση διαφόρων αλγορίθμων μάθησης, [20].

Πίνακας 2: Σύγκριση αλγορίθμων εκμάθησης (** αστέρια είναι η καλύτερη απόδοση και * αστέρι είναι η χειρότερη απόδοση**

	Δέντρα Απόφασης	**Νευρωνικά Δίκτυα**	**Naive Bayes**	**KNN**	**SVM**
Γενικότερη ακρίβεια	**	***	*	**	****
Ταχύτητα εκμάθησης ανάλογα με το πλήθος των χαρακτηριστικών και των στιγμιότυπων	***	*	****	****	*
Ταχύτητα ταξινόμησης	****	****	****	*	****
Ανοχή σε τιμές που δεν υπάρχουν	***	*	****	*	**
Ανοχή σε άσχετα χαρακτηριστικά	***	*	**	**	****
Ανοχή σε πλεονάζοντα χαρακτηριστικά	**	**	*	**	***
Ανοχή σε πολύ ανεξάρτητα χαρακτηριστικά	**	***	*	*	***
Αντιμετώπιση διακριτών/ δυαδικών/ συνεχών χαρακτηριστικών	****	*** (όχι τα διακριτά)	*** (όχι τα συνεχή)	*** (όχι τα άμεσα διακριτά)	** (όχι τα διακριτά)
Ανοχή στο θόρυβο	**	**	***	*	**
Αντιμετώπιση κινδύνου overfitting	**	*	***	***	**
Προσπάθειες για αυξανόμενη εκμάθηση	**	***	****	****	**
Επεξηγηματική ικανότητα/ σαφήνεια των γνώσεων. Ταξινομήσεις	****	*	****	**	*
Αντιμετώπιση παραμέτρων μοντέλου	***	*	****	***	*

3.3 Βασικές ιδιότητες αυτών σε περιβάλλον IoT - σύγκριση

Τα δεδομένα IoT διαφέρουν από τα άλλα πεδία σε σχέση με την ευρεία διακύμανση του όγκου, της ταχύτητας, της ποικιλίας και της ακρίβειας. Στην [28], πέντε συνηθισμένοι αλγόριθμοι μηχανικής μάθησης (KNN, Naive Bayes (NB), Δέντρα Απόφασης (Decision Trees - DTs), τυχαίο δάσος (Random Forest - RF) και λογιστική παλινδρόμηση (Logistic Regression – LR)) εφαρμόστηκαν σε πέντε διαφορετικά σύνολα δεδομένων IoT. Αυτοί οι αλγόριθμοι συγκρίθηκαν με βάση τα χαρακτηριστικά των συνόλων δεδομένων IoT όπως το μέγεθος, τον αριθμό των χαρακτηριστικών, τον αριθμό των κλάσεων, τις ανισορροπίες των κλάσεων, τις τιμές που λείπουν και τους χρόνους εκτέλεσης των αλγορίθμων, [28].

Η μελέτη στο [28], πραγματοποίησε ανάλυση απόδοσης βάσει των εξής παραμέτρων: του πίνακα λάθους, της ακρίβειας, της ανάκλησης, της μετρικής F, του χρόνου εκτέλεσης, της μετρικής kappa και της ακρίβειας. Επίσης, να σημειωθεί ότι τα σύνολα δεδομένων προεπεξεργάστηκαν και αντικαταστάθηκαν οι τιμές που λείπουν από την ενδιάμεση τιμή με τη χρήση του PCA για τη μείωση των διαστάσεων στα σύνολα δεδομένων. Παρατηρήθηκε ότι καθώς το μέγεθος του συνόλου δεδομένων αυξάνεται, η απόδοση των Naive Bayes και λογιστικής παλινδρόμησης μειώνεται ενώ η απόδοση των δέντρων απόφασης, του τυχαίου δάσους και του KNN δεν επηρεάζεται. Ο χρόνος εκτέλεσης όλων των αλγορίθμων αυξάνεται με την αύξηση του μεγέθους του συνόλου δεδομένων. Καθώς ο αριθμός των χαρακτηριστικών στο σύνολο δεδομένων αυξάνεται, η απόδοση των Naive Bayes και λογιστικής παλινδρόμησης μειώνεται ενώ η απόδοση των δέντρων απόφασης, του τυχαίου δάσους και του KNN αυξάνεται.

3.4 Κριτήρια επιλογής αλγορίθμου

Από τη βιβλιογραφία [22], [23] έχουμε τα εξής:

Τεχνική μείωσης διαστάσεων: Χρησιμοποιείται για τον εντοπισμό μοτίβων στα δεδομένα και την αντιμετώπιση υψηλού υπολογιστικού κόστους προβλημάτων. Χρήσιμοι για οπτικά και ακουστικά δεδομένα που περιλαμβάνουν ομιλία, βίντεο, εικόνες ή κείμενο, καθώς και κατά την απλοποίηση συνόλων δεδομένων, ώστε να ταιριάζουν καλύτερα σε ένα μοντέλο πρόβλεψης. Αλγόριθμοι: PCA, Μη – Αρνητική Παραγοντοποίηση Πινάκων (Non-negative Matrix Factorization - NMF), Kernel PCA, Graph-based kernel PCA, LDA, Γενικευμένη Ανάλυση Διακρίσεων (Generalized Discriminant Analysis – GDA), T Κατανεμημένη Στοχαστική Ενσωμάτωση Γείτονα (T-distributed Stochastic Neighbor Embedding - t-SNE), Autoencoder, Ομοιόμορφη Κατανομή και Προσέγγιση Πολλαπλών Μεταβλητών (Uniform

Manifold Approximation and Projection – UMAP) κ.ά. Για σύνολα δεδομένων πολλών διαστάσεων (δηλ. άνω των 10), η μείωση των διαστάσεων πραγματοποιείται συνήθως πριν από την εφαρμογή ενός αλγορίθμου KNN προκειμένου να αποφευχθούν οι συνέπειες των πολλαπλών διαστάσεων. Η εξαγωγή χαρακτηριστικών και η μείωση διαστάσεων μπορούν να συνδυαστούν σε ένα βήμα χρησιμοποιώντας την ανάλυση PCA, την ανάλυση LDA, ή τις τεχνικές NMF ως ένα στάδιο προ-επεξεργασίας μετά την ομαδοποίηση KNN σε διανύσματα χαρακτηριστικών στο χώρο μειωμένων διαστάσεων. Επικρατών αλγόριθμος και χαρακτηριστικά: Από την [22], για την εύρεση υποκείμενης δομής δεδομένων, επιλέχτηκε ο UMAP αλγόριθμος ο οποίος υπερισχύει στην κλιμάκωση των διαστάσεων και του μεγέθους ενός συνόλου δεδομένων αλλά και της γρήγορης προβολής.

Τεχνική συσταδοποίησης: ο στόχος της συσταδοποίησης είναι η ανίχνευση διακριτών ομάδων σε ένα σύνολο δεδομένων χωρίς ετικέτα, όπου οι χρήστες αναμένεται να καθορίσουν τα κριτήρια του τι είναι "σωστή" ομάδα, έτσι ώστε τα αποτελέσματα ομαδοποίησης να ικανοποιούν τις προσδοκίες τους. Αυτή η προσέγγιση χρησιμοποιείται για την τμηματοποίηση της αγοράς και των πελατών, σε μηχανές που κάνουν συστάσεις, για ταξινόμηση εγγράφων, ανίχνευση απάτης κ.ά. Αλγόριθμοι: KNN, K-means, Gaussian Mixture Model, Χωρική Ομαδοποίηση με βάση την Πυκνότητα σε Εφαρμογές Θορύβου (Density-Based Spatial Clustering of Applications with Noise – DBSCAN), Χωρική Ομαδοποίηση με βάση την Ιεραρχική Πυκνότητα σε Εφαρμογές Θορύβου (Hierarchical Density-Based Spatial Clustering of Applications with Noise – DBSCAN) κ.ά. Σύγκριση: Από την [24], η ασυμπτωτική πολυπλοκότητα με βάση τους πειραματισμούς δεν μπορεί να βελτιωθεί πολύ. Οι HDBSCAN και DBSCAN, ενώ έχουν πολυπλοκότητα μικρότερη από O(n^2), δεν μπορούν να επιτύχουν O(nlogn) στη δεδομένη διάσταση του συνόλου δεδομένων. Δεδομένου ότι η ομαδοποίηση HDBSCAN είναι πολύ καλύτερη από την K-Means και η κλιμάκωση είναι ακόμα αρκετά καλή, συστήνεται, αν δεν υπάρχει τεράστιος όγκος δεδομένων, η τεχνική HDBSCAN ως μια καλή επιλογή. Από πλευράς συνολικής απόδοσης, ο HDBSCAN είναι ο καλύτερος αλγόριθμος για την ομαδοποίηση. Αν πρέπει να συγκεντρωθούν δεδομένα πέρα από το εύρος που μπορεί να χειριστεί εύλογα ο HDBSCAN, τότε η επόμενη επιλογή είναι οι DBSCAN και K-Means· ο DBSCAN είναι ο πιο αργός από τα δύο, ειδικά για μεγάλα δεδομένα, αλλά η συσταδοποίηση K-Means μπορεί να είναι εξαιρετικά κακή – συνεπώς είναι ένα tradeoff. Επικρατών αλγόριθμος και χαρακτηριστικά: Από την [22], για την εύρεση υποκείμενης δομής δεδομένων, επιλέχτηκαν δύο αλγόριθμοι: ο DBSCAN αλγόριθμος και ο HDBSCAN. Ο DB-SCAN είναι καλύτερος στην εύρεση αυτών των υποκείμενων δομών δεδομένων και συσχετίσεων που

μπορούν να είναι χρήσιμες για τον καθορισμό προτύπων και την πρόβλεψη τάσεων. Για παράδειγμα, ο αλγόριθμος DBSCAN μπορεί να εφαρμοστεί σε βάσεις δεδομένων πελατών προκειμένου να βρει τα πιο συνήθη προϊόντα συγκεκριμένων χρηστών. Ο αλγόριθμος HDBSCAN είναι ακόμα καλύτερος και με βάση το κριτήριο ελαχίστου μεγέθους ομάδας, ενσωματώνει την ιεραρχία των ομάδων και, τέλος, εξάγει τις σταθερές ομάδες από το αντίστοιχο δέντρο.

Αλγόριθμος ανίχνευσης ανωμαλιών: χρησιμοποιείται σε συνδυασμό με τεχνικές μείωσης διαστάσεων και με βάση την πλειοψηφία των δειγμάτων μπορούν να εντοπιστούν οι κανονικές και μη τιμές. Επίσης χρησιμοποιείται και στη τη βελτίωση της ανάλυσης των ανωμαλιών. Επιτρέπει την ομαδοποίηση παρόμοιων ανωμαλιών και την περαιτέρω χειροκίνητη κατηγοριοποίηση με βάση τους τύπους συμπεριφοράς τους. Ακολουθώντας μια τέτοια διαδικασία, μπορούμε να χρησιμοποιήσουμε μη εποπτευόμενους αλγορίθμους μηχανικής μάθησης για να εντοπίσουμε ανωμαλίες, να τις ομαδοποιήσουμε και να παρέχουμε ετικέτες σε κάθε ομάδα χειροκίνητα. Παράδειγμα: Το φιλτράρισμα ανεπιθύμητων μηνυμάτων, όπου ο αλγόριθμος μηχανικής μάθησης αναλύει όλα τα εισερχόμενα μηνύματα, τα συγκεντρώνει και εντοπίζει τα ανεπιθύμητα ως ακραίες τιμές. Χρησιμοποιείται επίσης εκτενώς για τον εντοπισμό απάτης σε χρηματοοικονομικούς, ασφαλιστικούς, πληροφοριακούς και άλλους τομείς αλλά και στον κατασκευαστικό τομέα για την πρόβλεψη αστοχιών εξοπλισμού πριν την εμφάνισή τους. Είδη: κρυφά μοντέλα Markove (Hidden Markove Models - HMMs), δίκτυα Bayes, KNN, autoencoder, Αποκλίσεις από τους κανόνες συσχέτισης και τα συχνά σύνολα δεδομένων, κ.ά. Σύγκριση: Από την [25], η απόδοση διαφορετικών μεθόδων εξαρτάται σε μεγάλο βαθμό από το σύνολο δεδομένων και τις παραμέτρους τους αλλά και οι ίδιες οι μέθοδοι έχουν μικρά συστηματικά πλεονεκτήματα σε σχέση με μια άλλη σε σύγκριση με πολλά σύνολα δεδομένων και πολλές παραμέτρους.

Αλγόριθμος εξόρυξης συσχετίσεων: χρησιμοποιείται για τον εντοπισμό κρυφών συσχετίσεων σε μεγάλα σύνολα δεδομένων που εμφανίζονται συχνά μαζί. Μπορεί να επεξεργαστεί μη αριθμητικά, κατηγορηματικά δεδομένα, πράγμα που σημαίνει ότι απαιτεί περισσότερες ενέργειες παρά απλή μέτρηση. Επίσης, χρησιμοποιείται συνήθως για τον προσδιορισμό μοτίβων και συσχετίσεων σε συναλλαγές, σχεσιακές ή σε οποιαδήποτε παρόμοια βάση δεδομένων. Παράδειγμα: Ανάλυση καλαθιού αγορών μέσω της επεξεργασίας δεδομένων από σαρωτές γραμμωτού κώδικα, και ορισμός αγαθών που αγοράζονται μαζί. Χρήση και στον τομέα της υγείας για πρόβλεψη και συσχέτιση συμπτωμάτων, αλλά και στο δημόσιο τομέα, όπως μεταφορές, εργοστάσια κ.λ.π. Αλγόριθμοι: Apriori, Eclat, FP –growth, ASSOC κ.ά.

Σύγκριση: Από τη μελέτη [27], μια τροποποιημένη μορφή του FP-growth, με παράλληλη εκτέλεση, αποδεικνύεται ότι βελτιώνει πολύ την απόδοση του αλγορίθμου. Επιπλέον ο παράλληλος FP – growth έχει καλύτερη προσαρμοστικότητα στην αυξανόμενη ποσότητα δεδομένων και έχει καλή απόδοση εξόρυξης για μεγάλα αρχεία. Αποδεικνύεται επίσης ότι η αύξηση των υπολογιστικών κόμβων μπορεί να βελτιώσει την αποτελεσματικότητα της εξόρυξης, η οποία αποτελεί και τη μεγαλύτερη δύναμη του κατανεμημένου συστήματος συσταδοποίησης. Επικρατών αλγόριθμος και χαρακτηριστικά: Από την [27], αναδείχτηκε ο παράλληλος αλγόριθμος FP-growth. Ο αλγόριθμος αυτός σαρώνει τη βάση δεδομένων δύο φορές· η πρώτη σάρωση της βάσης δεδομένων είναι η ίδια με τον Apriori, η οποία εξάγει ένα σύνολο συχνών αντικειμένων και υπολογίζει το ποσοστό που συνεισφέρουν. Τη δεύτερη φορά που γίνεται σάρωση της βάσης δεδομένων, οι συναλλαγές στη βάση δεδομένων εισάγονται στο δέντρο συχνών μοτίβων με φθίνουσα σειρά ως προς τη συχνότητα των στοιχείων και δημιουργείται ένας κλαδί για κάθε συναλλαγή.

Αποφάσεις Ταυτοποίησης Ελέγχου: χρησιμοποιείται από ένα σύστημα ανίχνευσης εισβολής [36]. Η σύγκριση αποτυπώνεται στον παρακάτω πίνακα:

Πίνακας 3: Ανάλυση τεχνικών ανίχνευσης εισβολής με βάση την υπογραφή

Τύπος αλγορίθμου	Αλγόριθμος Δέντρων Απόφασης	Νευρωνικό Δίκτυο	Τεχνητό Νευρωνικό Δίκτυο, SVM	k-means clustering + Εποπτευόμενος αλγόριθμος	Hidden Markov Model (HMM) + Naïve Bayes	Random Forest, Βαθύ Νευρωνικό Δίκτυο (DNN), SVM	TANN, SVM, k-means+KNN
Πλεονεκτήματα	Δε χρειάζονται δεδομένα για προετοιμασία	Ικανότητα εντοπισμού ενός νέου τύπου επίθεσης με λιγότερους υπολογισμούς	Η κωδικοποίηση βάσει συχνότητας είναι πιο αποτελεσματική για την ανίχνευση εισβολής	Χαμηλή επίδραση στα κινητά όσον αφορά την κατανάλωση μπαταρίας, τη χρήση CPU / μνήμης	Πολύ ακριβής	Η ακρίβεια είναι υψηλή για την ανίχνευση επίθεσης, το ποσοστό σφάλματος είναι χαμηλό για όλα τα δοκιμασμένα σύνολα δεδομένων, ικανά να ταξινομήσουν αραιές επιθέσεις	Το ποσοστό ακρίβειας και το ποσοστό ανίχνευσης είναι υψηλό για μία κατηγορία
Περιορισμοί και προκλήσεις	Ψευδή θετικά βρέθηκαν σε γνωστές επιθέσεις στην άμυνα Εκ των προτέρων έρευνα του έργου	Η κρυφή επίθεση που κρύβει τις λέξεις-κλειδιά απαιτεί επιπλέον δράση	Η λανθασμένη επιλογή παραμέτρων οδηγεί σε overfitting στο ANN	Εφαρμόστηκε επιλεκτική πολιτική για την παρακολούθηση ύποπτων συμβάντων	Το HMM απαιτεί πιο πολλές από 5 καταστάσεις για τη διατήρηση υψηλής ακρίβειας	Απαιτείται βελτιστοποίηση παραμέτρου βάρους και κατωφλίου για κάθε επίπεδο DNN	Χαμηλή ακρίβεια για περισσότερες από μία κατηγορίες
Μετρικές απόδοσης	Αληθή θετικά, ψευδή θετικά, ψευδή αρνητικά, αληθή αρνητικά	Ψεύτικα alarm, Ρυθμός ανίχνευσης 80%	Ρυθμός ανίχνευσης επίθεσης 95,9%	Χρήση CPU, χρήση RAM, χρήση μπαταρίας, απεσταλμένα / ληφθέντα πακέτα	Ακρίβεια 100%	Ακρίβεια 91,97% Ανάκληση 92,23% Ρυθμός σφάλματων 7,9%	Ακρίβεια – 96,91%, Ρυθμός ανίχνευσης ψευδώς θετικών και εισβολής – 99,6% και 98,95%
Ταξινομητής	Μονός	Μονός	Μονός	Υβριδικός	Υβριδικός	Σύνολο	Υβριδικός

Βασικές ιδέες και αποτελέσματα

Στην αναφορά [18], έχει αναπτυχθεί αυτόματη υποστήριξη αποφάσεων για τον προσδιορισμό των ελέγχων ασφαλείας που σχετίζονται με ένα συγκεκριμένο σύστημα δεδομένου ενός πλαισίου. Αυτή η προσέγγιση, η οποία βασίζεται στη μηχανική μάθηση, αξιοποιεί παλαιότερα δεδομένα από αξιολογήσεις ασφαλείας που πραγματοποιήθηκαν σε προηγούμενα συστήματα προκειμένου να προτείνει ελέγχους ασφαλείας για ένα νέο σύστημα. Οι συγγραφείς λειτουργούν και αξιολογούν εμπειρικά την προσέγγισή τους χρησιμοποιώντας πραγματικά παλαιότερα δεδομένα από τον τραπεζικό τομέα. Τα αποτελέσματά τους δείχνουν ότι, όταν κάποιος αποκλείει ελέγχους ασφαλείας που είναι πιο σπάνιοι στα παλαιότερα δεδομένα, η προσέγγισή του έχει μέση ανάκληση ≈ 95% (μετρική ανάλογα με το πόσα αντικείμενα / στοιχεία ανιχνεύει) και μέση ακρίβεια ≈ 67%.

Οι μετρικές στις οποίες γίνεται αναφορά είναι οι εξής, [29]:

Ακρίβεια (accuracy) =(TP+TN)/(TP+FP+TN+FN)

Η ακρίβεια ορίζεται ως ο αριθμός του ακριβούς ταξινομημένου στιγμιότυπου δια του συνολικού αριθμού στιγμιότυπων στο σύνολο δεδομένων.

Θετική προβλεπόμενη τιμή (presision) =TP/(TP+FP)

Η θετική προβλεπόμενη τιμή είναι η μέση πιθανότητα σχετικής ανάκτησης.

Ανάκληση =TP/(TP+FN)

Η ανάκληση ορίζεται ως η μέση πιθανότητα πλήρους ανάκτησης.

Μετρκή F =(2·(Θετική προβλεπόμενη τιμή·Ανάκληση))/(Θετική προβλεπόμενη τιμή+Ανάκληση)

Όπου:

TP=True Positive

TN=True Negative

FP=False Positive

FN=False Negative

Συνεισφορά

Η υψηλή ανάκληση - που δείχνει ότι λείπουν μόνο λίγα σχετικά στοιχεία ασφαλείας - σε συνδυασμό με το λογικό επίπεδο ακρίβειας - που δείχνει ότι η προσπάθεια που απαιτείται για την επικύρωση των συστάσεων δεν είναι παραπάνω απ’ όσο χρειάζεται - υποδηλώνει ότι η προσέγγισή τους είναι μια χρήσιμη βοήθεια για τους αναλυτές για τον πιο αποτελεσματικό προσδιορισμό των σχετικών ελέγχων ασφαλείας, καθώς και για τη μείωση της πιθανότητας να παραλειφθούν σημαντικοί έλεγχοι.

4. Αντίμετρα απέναντι σε Επιθέσεις Υλικού

4.1 Αντίμετρα με Αλγορίθμους Μηχανικής Μάθησης

4.1.1 Αλγόριθμοι Μηχανικής Μάθησης σε Συστήματα Ανίχνευσης Εισβολών (Intrusion Detection Systems – IDS)

Στις μέρες μας, πολιτικές και εμπορικές οντότητες εμπλέκονται ολοένα και περισσότερο σε ένα εξελιγμένο κυβερνο-πόλεμο για να προκαλέσουν ζημιά, αλλοίωση ή ανίχνευση πληροφοριών σε δίκτυα υπολογιστών [17]. Κατά το σχεδιασμό πρωτοκόλλων δικτύου, υπάρχει η ανάγκη να διασφαλιστεί η αξιοπιστία έναντι των εισβολών, ιδιαίτερα ισχυρών επιθέσεων που ως αποτέλεσμα μπορούν ακόμη και να ελέγξουν ένα μέρος των μερών του δικτύου. Τα ελεγχόμενα αυτά μέρη μπορούν να προκαλέσουν τόσο παθητικές (π.χ., υποκλοπές, μη συμμετοχή) όσο και ενεργές επιθέσεις (π.χ. μπλοκάρισμα, απαλοιφή μηνυμάτων, διαφθορά και πλαστογράφηση).

Δίνοντας τον ορισμό της ανίχνευσης εισβολής, πρόκειται για τη διαδικασία της δυναμικής παρακολούθησης συμβάντων που συμβαίνουν σε ένα υπολογιστικό σύστημα ή δίκτυο, την ανάλυσή τους για ενδείξεις πιθανών γεγονότων και συχνά παρεμπόδιση της μη εξουσιοδοτημένης πρόσβασης. Αυτό επιτυγχάνεται συνήθως με την αυτόματη συλλογή πληροφοριών από πολλά συστήματα και πηγές δικτύου και, στη συνέχεια, με την ανάλυση των πληροφοριών αυτών για πιθανά προβλήματα ασφάλειας.

Ένα Σύστημα Ανίχνευσης Εισβολών (Intrusion Detection System – IDS) γενικά πρέπει να αντιμετωπίσει προβλήματα όπως δίκτυα με κίνηση μεγάλου όγκου δεδομένων, πολύ ανισοκατανομή φόρτου δεδομένων, δυσκολία στην οριοθέτηση ορίων απόφασης για περιπτώσεις φυσιολογικής και μη συμπεριφοράς και απαίτηση για συνεχή προσαρμογή σε ένα συνεχώς μεταβαλλόμενο περιβάλλον. Γενικά, η πρόκληση είναι η αποτελεσματική κατανόηση και ταξινόμηση διαφόρων συμπεριφορών σε ένα υπολογιστικό δίκτυο. Οι στρατηγικές για την ταξινόμηση των συμπεριφορών δικτύου χωρίζονται σε δύο κατηγορίες: ανίχνευση κακής χρήσης και ανίχνευση ανωμαλιών.

Οι τεχνικές ανίχνευσης κακής χρήσης εξετάζουν τόσο τη δραστηριότητα του δικτύου όσο και του συστήματος για γνωστές περιπτώσεις κατάχρησης χρησιμοποιώντας αλγόριθμους ταιριάσματος υπογραφών. Αυτή η τεχνική είναι αποτελεσματική στην ανίχνευση επιθέσεων που είναι ήδη γνωστές. Ωστόσο, συχνά δεν ανιχνεύονται νέες επιθέσεις που δημιουργούν ψεύτικα σήματα (false negatives). Οι ειδοποιήσεις μπορεί να δημιουργηθούν από το IDS, αλλά η αντίδραση σε κάθε ειδοποίηση χρησιμοποιεί χρόνο και πόρους που οδηγούν σε αστάθεια

του συστήματος. Για να ξεπεραστεί αυτό το πρόβλημα, το IDS δεν πρέπει να ξεκινήσει τη διαδικασία απομάκρυνσης της διεργασίας μόλις εντοπιστεί η ανιχνευθείσα ειδοποίηση, αλλά πρέπει να είναι περιμένει για να συλλέξει ειδοποιήσεις και να αποφασίσει με βάση τη συσχέτιση αυτών.

Τα συστήματα ανίχνευσης ανωμαλιών βασίζονται στην κατασκευή ενός μοντέλου συμπεριφοράς χρήστη που θεωρείται φυσιολογικό. Αυτό επιτυγχάνεται με τη χρήση συνδυαστικών μεθόδων στατιστικής ή μηχανικής μάθησης για την εξέταση προβλημάτων δικτύου ή ειδοποιήσεων και διαδικασιών συστήματος. Η ανίχνευση νέων επιθέσεων είναι πιο επιτυχημένη χρησιμοποιώντας την προσέγγιση ανίχνευσης ανωμαλιών, καθώς οποιαδήποτε αποκλίνουσα συμπεριφορά ταξινομείται ως εισβολή. Ωστόσο, η κανονική συμπεριφορά σε ένα μεγάλο και δυναμικό σύστημα δεν μπορεί να προσδιοριστεί καλά και αλλάζει με την πάροδο του χρόνου. Αυτό συχνά οδηγεί σε σημαντικό αριθμό ψευδών συναγερμών γνωστών ως ψευδώς θετικά (false positives). Ένα δίκτυο βασισμένο σε ένα IDS ελέγχει την εισερχόμενη κίνηση δικτύου για μοτίβα που μπορούν να υποδηλώσουν αν ένα άτομο ψάχνει το δίκτυο για ευάλωτους υπολογιστές. Δεδομένου ότι η αντίδραση σε κάθε ειδοποίηση καταναλώνει σχετικά μεγάλο χρονικό διάστημα και πόρους, το IDS δεν πρέπει να αποκρίνεται σε κάθε ειδοποίηση που δημιουργείται. Η παράβλεψη αυτού του γεγονότος μπορεί να οδηγήσει σε αυτοπροκαλούμενη άρνηση υπηρεσίας (denial-of-service). Για να ξεπεραστεί αυτό το πρόβλημα, οι ειδοποιήσεις πρέπει να συγκεντρωθούν και να συσχετιστούν προκειμένου να παράγουν λιγότερες αλλά πιο σαφείς και αξιοσημείωτες ειδοποιήσεις.

Οι προσεγγίσεις που βασίζονται στη μηχανική μάθηση για την ανίχνευση εισβολών διαιρούνται σε δύο κατηγορίες: προσεγγίσεις βασισμένες σε τεχνικές τεχνητής νοημοσύνης (AI) και προσεγγίσεις βασισμένες σε μεθόδους υπολογιστικής νοημοσύνης (Computational Intelligence - CI). Οι τεχνικές AI αναφέρονται στις μεθόδους από τον τομέα της κλασικής τεχνητής νοημοσύνης, όπως η στατιστική μοντελοποίηση και οι τεχνικές CI αναφέρονται σε μεθόδους εμπνευσμένες από τη φύση που χρησιμοποιούνται για την αντιμετώπιση πολύπλοκων προβλημάτων που οι κλασικές μέθοδοι δεν μπορούν να λύσουν. Σημαντικές μεθοδολογίες CI είναι ο εξελικτικός υπολογισμός, η ασαφής λογική, τα τεχνητά νευρωνικά δίκτυα και τα απρόσβλητα συστήματα. Η CI διαφέρει από το γνωστό πεδίο της AI· η AI χειρίζεται τη συμβολική αναπαράσταση της γνώσης, ενώ η CI χειρίζεται την αριθμητική αναπαράσταση των πληροφοριών. Παρόλο που το όριο μεταξύ αυτών των δύο κατηγοριών δεν είναι πάντα σαφές και πολλές υβριδικές μέθοδοι έχουν προταθεί στη βιβλιογραφία, οι περισσότερες προηγούμενες εργασίες έχουν σχεδιαστεί κυρίως με βάση μία από τις δύο

κατηγορίες. Επιπλέον, θα ήταν πολύ χρήσιμο να κατανοήσουμε πόσο καλά αποδίδουν οι φυσικές τεχνικές σε αντίθεση με τις κλασικές μεθόδους [36].

Πίνακας 4: προσεγγίσεις IDS που βασίζονται σε ανωμαλίες με βάση τους τύπους ταξινομητών

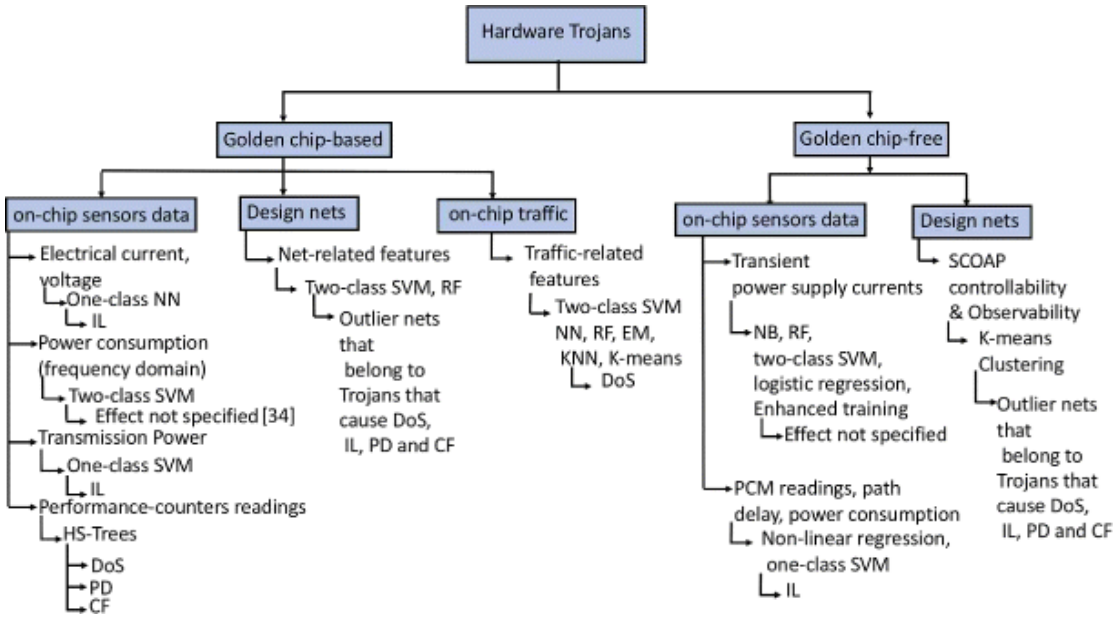
Τύπος αλγορίθμου	Αλγόριθμος Δέντρων Απόφασης	Νευρωνικό Δίκτυο	Τεχνητό Νευρωνικό Δίκτυο, SVM	k-means clustering + Εποπτευόμενος αλγόριθμος	Hidden Markov Model (HMM) + Naïve Bayes	Random Forest, Βαθύ Νευρωνικό Δίκτυο (DNN), SVM	TANN, SVM, k-means+KNN
Πλεονεκτήματα	Δε χρειάζονται δεδομένα για προετοιμασία	Ικανότητα εντοπισμού ενός νέου τύπου επίθεσης με λιγότερους υπολογισμούς	Η κωδικοποίηση βάσει συχνότητας είναι πιο αποτελεσματική για την ανίχνευση εισβολής	Χαμηλή επίδραση στα κινητά όσον αφορά την κατανάλωση μπαταρίας, τη χρήση CPU / μνήμης	Πολύ ακριβής	Η ακρίβεια είναι υψηλή για την ανίχνευση επίθεσης, το ποσοστό σφάλματος είναι χαμηλό για όλα τα δοκιμασμένα σύνολα δεδομένων, ικανά να ταξινομήσουν αραιές επιθέσεις	Το ποσοστό ακρίβειας και το ποσοστό ανίχνευσης είναι υψηλό για μία κατηγορία
Περιορισμοί και προκλήσεις	Ψευδή θετικά βρέθηκαν σε γνωστές επιθέσεις στην άμυνα Εκ των προτέρων έρευνα του έργου	Η κρυφή επίθεση που κρύβει τις λέξεις-κλειδιά απαιτεί επιπλέον δράση	Η λανθασμένη επιλογή παραμέτρων οδηγεί σε overfitting στο ANN	Εφαρμόστηκε επιλεκτική πολιτική για την παρακολούθηση ύποπτων συμβάντων	Το HMM απαιτεί πιο πολλές από 5 καταστάσεις για τη διατήρηση υψηλής ακρίβειας	Απαιτείται βελτιστοποίηση παραμέτρου βάρους και κατωφλίου για κάθε επίπεδο DNN	Χαμηλή ακρίβεια για περισσότερες από μία κατηγορίες
Μετρικές απόδοσης	Αληθή θετικά, ψευδή θετικά, ψευδή αρνητικά, αληθή αρνητικά	Ψεύτικα alarm, Ρυθμός ανίχνευσης 80%	Ρυθμός ανίχνευσης επίθεσης 95,9%	Χρήση CPU, χρήση RAM, χρήση μπαταρίας, απεσταλμένα / ληφθέντα πακέτα	Ακρίβεια 100%	Ακρίβεια 91,97% Ανάκληση 92,23% Ρυθμός σφάλματων 7,9%	Ακρίβεια – 96,91%, Ρυθμός ανίχνευσης ψευδώς θετικών και εισβολής – 99,6% και 98,95%
Ταξινομητής	Μονός	Μονός	Μονός	Υβριδικός	Υβριδικός	Σύνολο	Υβριδικός

4.1.2 Αλγόριθμοι Μηχανικής Μάθησης για την προστασία του υλικού από Δούρειους Ίππους Υλικού (Hardware Trojans Horses)

Θεωρούμε δύο τύπους μέτρων αντιμετώπισης: (1) αντίμετρα που προϋποθέτουν τη διαθεσιμότητα των "χρυσών κυκλωμάτων" και (2) αντίμετρα που προϋποθέτουν ότι η πρόσβαση σε κυκλώματα χωρίς trojan δεν είναι πρακτική.

Σημείωση: Από τη φύση του, ο χρυσός είναι ιδανικός για εφαρμογές ηλεκτρονικών ειδών διότι διαμορφώνεται και είναι εύκολα διαχειρίσιμος για συνδέσεις, καλώδια και επαφές ρελέ. Ο χρυσός παρέχει ηλεκτρική ενέργεια πολύ αποτελεσματικά (μια προφανής απαίτηση για εφαρμογές PCB). Μπορεί να μεταφέρει μικρές ποσότητες ρεύματος, απαραίτητες για τις σημερινές ηλεκτρονικές συσκευές.

Το παρακάτω σχήμα συνοψίζει τα προαναφερθέντα αντίμετρα ως προς: (1) τα εξαγόμενα χαρακτηριστικά, (2) τον τύπο των εκπαιδευμένων μοντέλων και (3) τη συμπεριφορά των ανιχνευόμενων δούρειων ίππων στο υλικό. Το ανιχνευμένο υλικό με trojans μπορεί να προκαλέσει διαρροή πληροφοριών (Information Leakage - IL), άρνηση υπηρεσίας (Denial-Of-Service - DoS), αλλαγή λειτουργικότητας (Change of Functionality - CF) ή υποβάθμιση απόδοσης (Performance Degradation - PD).



Εικόνα 4: Αντίμετρα για Trojan Horses

Τα αντίμετρα που βασίζονται σε χρυσά κυκλώματα (Golden Chip-Based Countermeasures) χρησιμοποιούν δεδομένα που εξάγονται από χρυσά κυκλώματα (που δεν έχουν trojans) για τη δημιουργία μοντέλων μηχανικής μάθησης, τα οποία κατηγοριοποιούν σε δεδομένα αισθητήρα κυκλωμάτων, netlists ή κίνηση μεταξύ πολλαπλών πυρήνων για την ανίχνευση trojan υλικού.

4.1.2.1 Δεδομένα από αισθητήρες κυκλώματος

Η ενεργοποίηση trojans υλικού οδηγεί σε εμφάνιση ανωμαλιών σε διάφορες πηγές δεδομένων στο κύκλωμα, για παράδειγμα, στις ροές δεδομένων μετρητών απόδοσης και τρέχουσες μετρήσεις. Οι μετρήσεις ηλεκτρικού ρεύματος λαμβάνονται δειγματοληπτικά από τους αισθητήρες του κυκλώματος και στη συνέχεια μετατρέπονται σε συνεχή τάση για να ταξινομηθούν από νευρωνικά δίκτυα μιας κατηγορίας για να ελέγξουν αν το κύκλωμα έχει ενεργοποιημένο trojan. Τα χαρακτηριστικά στο πεδίο της συχνότητας των ιχνών κατανάλωσης ισχύος μπορούν επίσης να ταξινομηθούν χρησιμοποιώντας ένα SVM δύο κατηγοριών για τον εντοπισμό trojan υλικού. Τα νευρωνικά δίκτυα μπορούν να χρησιμοποιηθούν τόσο για

εξαγωγή χαρακτηριστικών όσο και για εξαγωγή συμπερασμάτων. Τα σχετικά χαρακτηριστικά εξάγονται από τα ίχνη κατανάλωσης ισχύος, περνώντας τα μέσα από τα κρυμμένα επίπεδα του νευρωνικού δικτύου. Τα ίχνη κατανάλωσης ισχύος ταξινομούνται από νευρωνικά δίκτυα για την ανίχνευση μοτίβων που έχουν μολυνθεί από trojan. Η Bayesian μεθοδολογία χρησιμοποιείται για τη βαθμονόμηση της διακύμανσης της διαδικασίας για τη διευκόλυνση της ανίχνευσης trojan μέσω της ανάλυσης των ρευμάτων διαρροής.

Τα Half Space Trees (HST) ανιχνεύουν ανώμαλη συμπεριφορά σε ροές δεδομένων μετρητών απόδοσης. Οι μετρητές απόδοσης είναι ενσωματωμένοι καταχωρητές στους περισσότερους πυρήνες μικροεπεξεργαστών. Καταγράφουν τη δραστηριότητα μικρο-αρχιτεκτονικών γεγονότων, για παράδειγμα, ποσοστά απώλειας κρυφής μνήμης και της κρυφής μνήμης αναζήτησης (Translation lookaside buffer – TLB). Αυτή η ανώμαλη συμπεριφορά προκαλείται από την ενεργοποίηση του trojan υλικού που προκαλεί άρνηση υπηρεσίας, αλλαγή στη λειτουργικότητα και υποβάθμιση της απόδοσης. Επιπλέον, τα δείγματα ισχύος μετάδοσης συλλέγονται και ταξινομούνται χρησιμοποιώντας έναν ταξινομητή SVM μιας κατηγορίας με πυρήνα RBF.

4.1.2.2 Δεδομένα από netlists

Σημείωση: Στην ηλεκτρονική σχεδίαση, ένα netlist είναι μια περιγραφή της συνδεσιμότητας ενός ηλεκτρονικού κυκλώματος. Στην απλούστερη μορφή του, ένας δικτυακός κατάλογος (netlist) αποτελείται από μια λίστα των ηλεκτρονικών εξαρτημάτων σε ένα κύκλωμα και μια λίστα των κόμβων με τους οποίους συνδέονται. Ο όρος δίκτυο (net) εν προκειμένω σημαίνει μια συλλογή δύο ή περισσότερων διασυνδεδεμένων στοιχείων. Η δομή, η πολυπλοκότητα και η αναπαράσταση των netlists μπορεί να ποικίλλει σημαντικά, αλλά ο θεμελιώδης σκοπός κάθε netlist είναι να μεταφέρει πληροφορίες συνδεσιμότητας . Οι λίστες δικτύου συνήθως δεν παρέχουν τίποτα περισσότερο από παρουσίες, κόμβους και ίσως κάποια χαρακτηριστικά των σχετικών στοιχείων. Εάν εκφράζουν πολύ περισσότερα από αυτό, συνήθως θεωρούνται γλώσσα περιγραφής υλικού όπως το Verilog ή το VHDL ή μία από τις πολλές γλώσσες που έχουν σχεδιαστεί ειδικά για είσοδο σε προσομοιωτές .

Κάθε δίκτυο λοιπόν, [1], κατά το σχεδιασμό του πρέπει να προσδιοριστεί αν είναι μέρος ενός κυκλώματος που έχει μολυνθεί από trojan. Για να επιτευχθεί αυτός ο στόχος, πραγματοποιείται στατική ανάλυση της netlist και εξάγονται τα ακόλουθα χαρακτηριστικά από κάθε δίκτυο: (1) η λογική πύλη fan-in, η οποία περιγράφεται ως ο αριθμός εισόδων στη λογική πύλη που είναι n -επίπεδα μακριά από την είσοδο του δικτύου στόχου, όπου n = 2, (2)ο ελάχιστος αριθμός επιπέδων πύλης από την έξοδο οποιουδήποτε flip flop έως το δίχτυο-στόχο, (3) ο ελάχιστος

αριθμός επιπέδων πύλης από το δίκτυο-στόχο μέχρι την είσοδο οποιουδήποτε flip flop και (4) ο ελάχιστος αριθμός επιπέδων πύλης μεταξύ της κύριας εισόδου μέχρι το δίκτυο-στόχο και από το δίκτυο-στόχο έως την κύρια έξοδο της συγκεκριμένης σχεδίασης. Αυτά τα χαρακτηριστικά χρησιμοποιούνται για την εκπαίδευση ενός ταξινομητή SVM δύο κατηγοριών. Παρομοίως, 51 χαρακτηριστικά εξάγονται και μετά μειώνονται στα πιο σημαντικά 11 χρησιμοποιώντας τον ταξινομητή δύο κατηγοριών Random Forest. Ο ταξινομητής Random Forest χρησιμοποιείται επίσης για να προσδιοριστεί αν το δίκτυο στόχος λειτουργεί σε λογική trojan. Τα εξαγόμενα 51 χαρακτηριστικά περιλαμβάνουν: (1) τη λογική πύλη fan-in έως 5 επίπεδα, (2) τον αριθμό των flip flop έως και 5 επιπέδων μακριά από την είσοδο και έξοδο του δικτύου-στόχου, (3) το επίπεδο του πλησιέστερου flip-flop στην είσοδο και την έξοδο του δικτύου-στόχου, (4) τον αριθμό των πολυπλεκτών έως 5 επίπεδα μακριά από την είσοδο και την έξοδο του δικτύου-στόχου, (5) το επίπεδο του πλησιέστερου πολυπλέκτη στην είσοδο και την έξοδο του δικτύου-στόχου, (6) τον αριθμό των βρόχων m επιπέδων όπου m είναι έως 5 από την είσοδο και την έξοδο του δικτύου-στόχου, (7) τον αριθμό των δικτύων που έχουν σταθερή τιμή 0 ή 1 έως 5 επίπεδα μακριά από το δίκτυο-στόχο και (8) τον ελάχιστο αριθμό επιπέδων πύλης μεταξύ της κύριας εισόδου μέχρι το δίκτυο-στόχο και από το δίκτυο-στόχο έως την κύρια έξοδο της συγκεκριμένης σχεδίασης.

Ένας άλλος αλγόριθμος που χρησιμοποιείται προκειμένου να αφαιρεθούν ορισμένα χαρακτηριστικά που είναι λιγότερο σχετικά είναι ο μηχανισμός βαθμολόγησης του eXtreme Gradient Boosting (XGBoost). Συγκεκριμένα αναλύονται περαιτέρω τα υπάρχοντα 51 χαρακτηριστικά trojan κυκλωμάτων και προτείνονται 5 νέα χαρακτηριστικά. Ο κατάλληλος αριθμός χαρακτηριστικών αποτελεί προϋπόθεση για την αποτελεσματική εκπαίδευση των ταξινομητών. Κατά τη διαδικασία ανίχνευσης trojan υλικού, το Trojan-net πρέπει να ανιχνεύεται όσο το δυνατόν περισσότερο. Επομένως, στοχεύεται η μεγιστοποίηση της ανάκλησης και χρήσης του αλγορίθμου XGBoost για την εκπαίδευση του ταξινομητή trojan υλικού. Το πείραμα σύγκρισης δείχνει ότι ο αλγόριθμος XGBoost είναι πιο κατάλληλος για ανίχνευση trojan υλικού και δείχνει επίσης ότι οι προτεινόμενες λειτουργίες μπορούν να βελτιώσουν περαιτέρω την ανίχνευση trojan υλικού. Κατά τη διαδικασία ανίχνευσης των trojan netlists, μπορούν να επιτευχθούν σωστά αποτελέσματα ανίχνευσης χωρίς λάθη. Για τις λίστες δικτύου που έχουν εισαχθεί από τον trojan, μπορούμε προκύπτει 89,84% μέση ανάκληση και μέση ακρίβεια 99,83%.

4.1.2.3 Δεδομένα από πολλούς πυρήνες

Όταν πολλοί πυρήνες (network core) επικοινωνούν μεταξύ τους σε ένα σύστημα, οι δρομολογητές που είναι μολυσμένοι με trojan μπορούν να εκκινήσουν επιθέσεις DoS στους πυρήνες του συστήματος. Έχουν εκπονηθεί μορφές επιθέσεων DoS που σχετίζονται με την κίνηση στο κύκλωμα και προκαλούν εκτροπή κίνησης, βρόχο δρομολόγησης και επιθέσεις υποκλοπής πυρήνα. Η εκτροπή της κίνησης αντιπροσωπεύει την επίθεση όπου οι δρομολογητές που έχουν μολυνθεί από trojan εκτρέπουν τα πακέτα επικοινωνίας σε προορισμούς διαφορετικούς από τους αρχικούς προορισμούς. Ο βρόχος δρομολόγησης αντιπροσωπεύει επιθέσεις όπου τα πακέτα αποστέλλονται πίσω στον πυρήνα αποστολής. Έτσι, ο πυρήνας που δέχεται επίθεση χάνει την επικοινωνία με τους άλλους πυρήνες του συστήματος. Οι βασικές επιθέσεις spoofing αντιπροσωπεύουν επιθέσεις όπου όλα τα πακέτα μεταφέρονται σε έναν συγκεκριμένο πυρήνα για να τον καταστήσουν μη διαθέσιμο σε άλλους πυρήνες του συστήματος. Επομένως, τα χαρακτηριστικά που εξάγονται από την κυκλοφορία στο κύκλωμα αναλύονται για τον εντοπισμό αυτών των επιθέσεων. Τέτοια χαρακτηριστικά σχετίζονται με την προέλευση πακέτου και τις διευθύνσεις προορισμού, τη διαδρομή μεταφοράς και την απόσταση μεταφοράς. Τα χαρακτηριστικά που εξάγονται από "χρυσά" και μολυσμένα από trojan κυκλώματα πολλών πυρήνων χρησιμοποιούνται στην εκπαίδευση πολλών εποπτευόμενων μοντέλων μηχανικής μάθησης, για παράδειγμα, στο μοντέλο k-πλησιέστερου γείτονα (KNN), της γραμμικής παλινδρόμησης (Linear Regression - LR), του SVM και δυναμικού χρόνου (Dynamic Time – DT). Μπορούν επίσης να χρησιμοποιηθούν μη εποπτευόμενα μοντέλα μάθησης, όπως η συσταδοποίηση K-means, η μεγιστοποίηση εκτίμησης και η ιεραρχική συσταδοποίηση. Σε πραγματικές καταστάσεις, οι επιτιθέμενοι μπορούν να εκτοξεύσουν επιθέσεις που δεν αναμένονται κατά τη διάρκεια της προπόνησης. Επομένως, τα εκπαιδευμένα μοντέλα μπορούν να αναβαθμιστούν μέσω διαδικτύου χρησιμοποιώντας τον αλγόριθμο Modified Balanced Winnow (MBW) για να ληφθούν υπόψη οι νέες επιθέσεις. Ο αλγόριθμος Modified Balanced Winnow (MBW) χρησιμοποιεί ένα πολλαπλασιαστικό σχήμα που του επιτρέπει να αποδίδει πολύ καλύτερα όταν πολλές διαστάσεις είναι άσχετες (εξ ου και η ονομασία του winnow που σημαίνει ξεσκαρτάρω, κοσκινίζω, ξεδιαλέγω, διακρίνω), κλιμακώνεται καλά σε πολυδιάστατα δεδομένα και χρησιμοποιείται στην online διαδικασία μάθησης, όπου η φάση μάθησης και η φάση ταξινόμησης δεν διαχωρίζονται σαφώς.

Σημείωση: Spoofing (πλαστογράφηση): είναι μια τεχνική για να αποκτήσουμε παράνομη πρόσβαση σε υπολογιστές με την δημιουργία πακέτων TCP/IP, χρησιμοποιώντας τη διεύθυνση και τα στοιχεία κάποιου άλλου αξιόπιστου. Οι δρομολογητές (routers)

χρησιμοποιούν την διεύθυνση της IP προορισμού (destination IP) ώστε να διαδώσουν τα πακέτα μέσω διαδικτύου αγνοώντας - αλλάζοντας εικονικά την διεύθυνση της IP πηγής (source IP). Αυτή η διεύθυνση χρησιμοποιείται μόνο από το μηχάνημα προορισμού όταν απαντά πίσω στη πηγή.

Ακόμα και κατά την απουσία χρυσών μοντέλων (Golden Chip-free Countermeasures), τα δίκτυα σχεδίασης μπορούν να ομαδοποιηθούν για να εντοπιστούν δίκτυα που είναι μολυσμένα από trojans πχ μέσω της συσταδοποίηση K-means. Κάθε netlist είναι ομαδοποιημένη σύμφωνα με το SCOAP (=Sandia Controllability Observability Analysis Program). Το SCOAP είναι ένα πρόγραμμα που αναπτύχθηκε στα Sandia National Laboratories για την ανάλυση της δυνατότητας δοκιμής ενός ψηφιακού κυκλώματος (=testability measure of a circuit). Η δυνατότητα δοκιμής σχετίζεται με τη δυσκολία ελέγχου και παρατήρησης των λογικών τιμών των εσωτερικών κόμβων από εισόδους και εξόδους κυκλώματος, αντίστοιχα). Επιπλέον, ένας αλγόριθμος ομαδοποίησης με βάση την πυκνότητα (OPTICS) μπορεί να χρησιμοποιηθεί για τον εντοπισμό των εξωτερικών netlists που ανήκουν σε κυκλώματα με trojan. Ένα γράφημα κατασκευάζεται χρησιμοποιώντας το σχεδιασμό επιπέδου πύλης όπου οι είσοδοι και οι έξοδοι μιας netlist αντιπροσωπεύονται ως κορυφές και οι είσοδοι και οι έξοδοι της ίδιας netlist συνδέονται με μια ακμή που έχει ένα βάρος βάσει της συσχέτισης μεταξύ εισόδων και εξόδων. Ο αλγόριθμος ομαδοποίησης σχηματίζει ένα γράφημα δενδρογράμματος και οι netlists που ανήκουν σε ένα κύκλωμα με trojan εντοπίζονται με ακραίες τιμές.

4.1.3 Αλγόριθμοι Μηχανικής Μάθησης για την αντιμετώπιση του IC counterfeiting

Η αλλοίωση ολοκληρωμένων κυκλωμάτων αποτελεί μια αναδυόμενη απειλή για τη βιομηχανία παραγωγής IC. Η ηλεκτρονική βιομηχανία βιώνει μία αναπτυσσόμενη αγορά πλαστών προϊόντων, με αποτέλεσμα οι αλυσίδες ηλεκτρονικών προμηθειών να είναι επιρρεπείς σε πλαστά εξαρτήματα. Οι παραδοσιακές μέθοδοι για την αξιολόγηση και την αξιοπιστία ενός IC και την διάκρισή του ως πλαστού ή αυθεντικού ήταν οι εξής:

1) η χειροκίνητη οπτική επιθεώρηση των IC από εκπαιδευμένους εμπειρογνώμονες που όμως ήταν ευάλωτη σε σφάλματα και χρονοβόρα διαδικασία και
2) η χρήση υπολογιστικής τομογραφίας ακτίνων Χ που όμως πολλές φορές μπορούσε να καταστρέψει την επιφάνεια του IC λόγω της θερμότητας που προκαλούνταν από τη θερμότητα των ακτίνων Χ.

Τα τελευταία όμως χρόνια με την βοήθεια της Μηχανικής Μάθησης χρησιμοποιούνται τεχνικές επεξεργασίας εικόνας και τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks –

ANN) για την αυτοματοποίηση της διαδικασίας ελέγχου. Τα νευρωνικά δίκτυα συλλέγουν τις οπτικές διαφορές και λοιπά δεδομένα από τους αισθητήρες στα κυκλώματα (πχ καθυστέρηση όλων των διαδρομών στο κύκλωμα κλπ), τις αναλύουν και στη συνέχεια της ταξινομούν με χρήση του αλγορίθμου SVM ή του αλγορίθμου συσταδοποίησης K-means. Επιπλέον, οι ταλαντωτές δακτυλίου (Ring Oscillators - ROs) εφαρμόζονται με δοκιμές σε κυκλώματα FPGA. Οι συχνότητες του ταλαντωτή δακτυλίου και ο ρυθμός υποβάθμισης αυτών των συχνοτήτων ταξινομούνται χρησιμοποιώντας SVM μιας κατηγορίας για την ανίχνευση ανακυκλωμένων IC.

4.1.4 Αλγόριθμοι Μηχανικής Μάθησης για την αντιμετώπιση του Reverse Engineering

Η αντίστροφη μηχανική δεν έχει αντιμετωπιστεί εκτενώς μέσω της μηχανικής μάθησης με την έννοια ότι δεν έχει καταφέρει μέχρι στιγμής να ανακτήσει πληροφορίες για τα συστατικά ενός προϊόντος δηλαδή από τι ολοκληρωμένα κυκλώματα (IC) και πλακέτες τυπωμένου κυκλώματος Printed circuit board = PCB) αποτελείται για να κατανοήσει στη συνέχεια τη λειτουργικότητα της συσκευής, παρά μόνο επιτρέπει αποτελεσματικά στους ιδιοκτήτες IP να εξετάσουν αν έχει παραβιαστεί η πνευματική τους ιδιοκτησία, να επαληθεύσουν και να επικυρώσουν το design και όλο αυτό το κάνει με το να αποφανθεί αν υπάρχει έγχυση Trojans υλικού ή όχι. Έτσι λοιπόν, οι τεχνικές αντίστροφης μηχανικής που βασίζονται στη μηχανική μάθηση έχουν επικεντρωθεί στον εντοπισμό των "χρυσών" κυκλωμάτων. Τα "χρυσά" κυκλώματα μπορούν στη συνέχεια να χρησιμοποιηθούν στα αντίστοιχα αντίμετρα με βάση τα "χρυσά" κυκλώματα. Οι διαδικασίες αντίστροφης μηχανικής περιλαμβάνουν: (i) αποσυναρμολόγηση, (ii) αποεπιπεδοποίηση, (iii) απεικόνιση, (iv) σχολιασμό και (v) σχηματική δημιουργία. Τα βήματα σχολιασμού και σχηματικής δημιουργίας είναι χρονοβόρα. Επομένως, τα μοντέλα μηχανικής μάθησης εκπαιδεύονται σε χαρακτηριστικά που εξάγονται από τη φάση της απεικόνισης της διάταξης για να ταξινομήσουν τα κυκλώματα ως χωρίς trojan ή μολυσμένα με trojan. Υποτίθεται ότι τα trojan τοποθετούνται κατά τη διάρκεια της διαδικασίας κατασκευής, επομένως, η χρυσή φυσική διάταξη του κυκλώματος είναι διαθέσιμη. Η εικόνα του IC και η χρυσή του διάταξη χωρίζονται σε πλέγματα. Στη συνέχεια, τα πλέγματα της εικόνας του IC συγκρίνονται με τα πλέγματα της χρυσής φυσικής διάταξής του. Ακολούθως, οι διαφορές ανάμεσα στα πλέγματα χρυσής διάταξης και στα πλέγματα εικόνων του IC από πλευράς τοποθεσιών και κεντρικών τοποθεσιών των μερών τους χρησιμοποιούνται για την εκπαίδευση μοντέλων μηχανικής μάθησης. Χρησιμοποιείται ο SVM μιας κατηγορίας με πολυωνυμικό πυρήνα. Συνεπώς, η διαδικασία επιλύει αυτό το πρόβλημα

υποθέτοντας ότι τα trojans βρίσκονται σε ένα μικρό μόνο ποσοστό της συνολικής περιοχής διάταξης του IC, επομένως ακόμη και αν ένα IC έχει μολυνθεί από trojan, τα περιθώρια σφάλματος στον SVM μιας κατηγορίας εξαλείφουν το σφάλμα που προκαλείται από τα δίκτυα που έχουν μολυνθεί από trojan. Εναλλακτικά, ο μη εποπτευόμενος αλγόριθμος συσταδοποίησης K-means μπορεί να χρησιμοποιηθεί για την ομαδοποίηση των πλεγμάτων IC για να δείξει αν είναι μολυσμένα από trojan ή όχι.

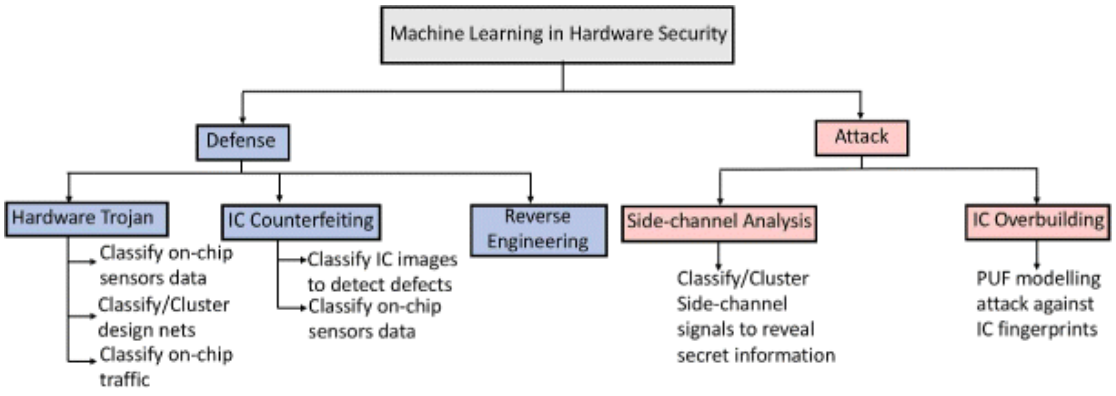
Το 2016, η startup εταιρεία ανάλυσης ασφάλειας τσιπ Texplained πραγματοποίησε τις πρώτες της επενδύσεις σε εργαστηριακό εξοπλισμό, αναλύοντας το εσωτερικό διαφόρων τσιπ και εμπορευματοποίησε τις αναφορές της (reports) ως κατάλογο με πολύ διεξοδικά IPs (Intellectual Properties) αντίστροφης μηχανικής, εντοπίζοντας αδυναμίες ασφαλείας. Το εργαλείο που έκτοτε χρησιμοποιούν εκμεταλλεύεται τη δύναμη της μηχανικής εκμάθησης για να εξάγει αυτόματα και παράλληλα τις netlists από τις εκατοντάδες εκατομμύρια τρανζίστορ που βλέπει. Με την ανάλυση των netlists εντοπίζονται τα σήματα και φιλτράρονται τα μπλοκ σχεδίασης ανά πεδίο ενδιαφέροντος, για παράδειγμα αναζητώντας μόνο clock trees, voltage rails κ.λπ. Στην συνέχεια μόλις γίνει η εξαγωγή αυτών των χαρακτηριστικών, χρησιμοποιούνται ως είσοδοι σε άλλους αλγορίθμους μηχανικής μάθησης που ανιχνεύουν πιο προχωρημένα χαρακτηριστικά και λειτουργίες. Μετά από αυτή την ανακοίνωση, η Texplained εξακολουθεί να προσφέρει τις συμβουλευτικές υπηρεσίες της και νέα reports περί τσιπ στον ιστότοπό της και απέκτησε νέο εξοπλισμό μηχανικής μάθησης, επειδή πλέον η ζήτηση από τις προ αξιολόγηση εταιρείες ήταν περισσότερη από αυτή που μπορούσε να αντιμετωπίσει.

4.1.5 Αλγόριθμοι Μηχανικής Μάθησης για την αντιμετώπιση Επιθέσεων Πλευρικού Καναλιού (Side Channels Attacks)

Η αντιμετώπιση με μηχανική μάθηση των επιθέσεων που προήλθαν από μηχανική μάθηση ήταν ένα πολύ δύσκολο εγχείρημα ακριβώς γιατί χρησιμοποιούν την ίδια «δύναμη» και της ίδιες δυνατότητες των αλγορίθμων.

Ένα από τα παραδείγματα είναι η χρήση διαφόρων μοντέλων Machine Learning (ML) για τον εντοπισμό Side Channel Attacks (CSCAs) στην μνήμη Cache στην αρχιτεκτονική x86 της Intel. Επιπλέον έγινε ποσοτική και ποιοτική αξιολόγηση της απόδοσης των 12 μοντέλων ML με βάση την ακρίβεια ανίχνευσης χρόνου εκτέλεσης, την ταχύτητα, την υπολογιστική επιβάρυνση και την κατανομή σφαλμάτων όσον αφορά ψευδώς θετικά και ψευδώς αρνητικά. Τα πειράματα εκτελέστηκαν χρησιμοποιώντας τις state-of-the-art CSCAs επιθέσεις και

συγκεκριμένα τις Flush + Reload και Flush + Flush, υπό ρεαλιστικές συνθήκες και σε κρυπτοσυστήματα RSA και AES.



Εικόνα 5: Χρήση της μηχανικής μάθησης για α) επίθεση και β)άμυνα σε θέματα ασφάλειας υλικού

Λόγω της αδυναμίας των επεξεργαστών Intel X86, η κοινή χρήση σελίδων εκθέτει διαδικασίες σε διαρροές πληροφοριών. Το Flush + Reload είναι μια τεχνική επίθεσης πλευρικού καναλιού cache που εκμεταλλεύεται αυτήν την αδυναμία για την παρακολούθηση της πρόσβασης σε γραμμές μνήμης σε κοινόχρηστες σελίδες. Σε αντίθεση με τις προηγούμενες επιθέσεις στο κανάλι της προσωρινής μνήμης, το Flush + Reload στοχεύει την προσωρινή μνήμη τελευταίου επιπέδου (δηλ. L3 σε επεξεργαστές με τρία επίπεδα προσωρινής μνήμης). Κατά συνέπεια, το πρόγραμμα επίθεσης και το θύμα δεν χρειάζεται να μοιραστούν τον πυρήνα εκτέλεσης.

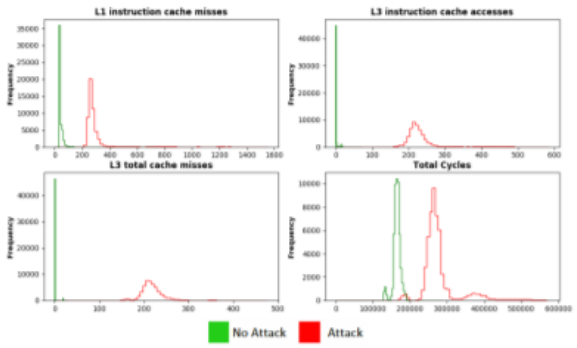
Η επίθεση Flush + Flush βασίζεται μόνο στον χρόνο εκτέλεσης της εντολής flush, η οποία εξαρτάται από το εάν τα δεδομένα είναι προσωρινά αποθηκευμένα ή όχι. Το Flush + Flush δεν κάνει πρόσβαση στη μνήμη, σε αντίθεση με οποιαδήποτε άλλη επίθεση στην κρυφή μνήμη. Έτσι, δεν προκαλεί καθόλου κρυφή μνήμη και ο αριθμός των επισκέψεων της κρυφής μνήμης μειώνεται στο ελάχιστο λόγω των συνεχών εκκενώσεων της προσωρινής μνήμης. Επομένως, οι επιθέσεις Flush + Flush είναι κρυφές, δηλαδή η διαδικασία κατασκοπείας δεν μπορεί να ανιχνευθεί με βάση τις επισκέψεις cache και τις απώλειες, ή τους προηγμένους μηχανισμούς ανίχνευσης. Η επίθεση Flush + Flush εκτελείται σε υψηλότερη συχνότητα και επομένως είναι ταχύτερη από οποιαδήποτε υπάρχουσα επίθεση στην κρυφή μνήμη. Με 496 KB / s σε ένα κρυφό κανάλι cross-core, είναι 6,7 φορές ταχύτερο από οποιοδήποτε κρυφό κανάλι cache.

Τα μοντέλα Μηχανικής Μάθησης βρίσκουν μοτίβα από τα δεδομένα εισόδου, εξόδου και την μεταξύ τους σχέση όπως instruction and data cache misses & hits for L1/L2/LLC, total CPU

cycles, branch miss-predictions, total cache accesses κλπ., τα αναλύουν, αποφαίνονται τι μπορεί να είναι attack και τι όχι και αποτρέπουν την πιθανή επίθεση. Όλα τα παραπάνω τα πραγματοποιούν για δύο ξεχωριστές μελέτες περίπτωσης: μία για επίθεση Flush + Reload και μία για Flush + Flush.

TABLE I: Selected events related to cache-based SCAs

Scope of Event	Hardware Event as Feature	Feature ID
L1 Caches	Data Cache Misses	L1-DCM
	Instruction Cache Misses	L1-ICM
	Total Cache Misses	L1-TCM
L2 Caches	Instruction Cache Accesses	L2-ICA
	Instruction Cache Misses	L2-ICM
	Total Cache Accesses	L2-TCA
	Total Cache Misses	L2-TCM
L3-Caches	Instruction Cache Accesses	L3-ICA
	Total Cache Accesses	L3-TCA
	Total Cache Misses	L3-TCM
System-wide	Total CPU Cycles	TOT_CYC
	Branch Miss-Predictions	BR_MSP



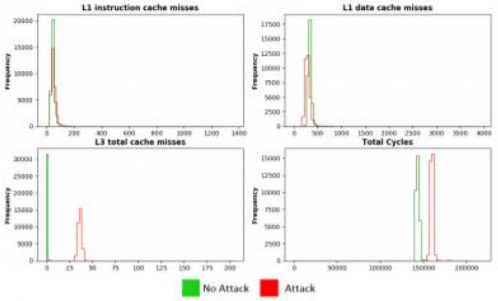


Fig. 2: Selected hardware events under Zero Load conditions for AES encryption algorithm: With & Without F+F Attack

TABLE II: List of Machine Learning Models for CSCA Detection (Non-exhaustive)

No.	Machine Learning Model	Category
1	Linear Regression (LR)	Linear
2	Linear Discriminant Analysis (LDA)	Linear
3	Support Vector Machine (SVM)	Linear
4	Quadratic Discriminant Analysis (QDA)	Non-linear
5	Random Forest (RF)	Non-linear
6	K-Nearest Neighbors (KNN)	Non-linear
7	Nearest Centroid	Linear
8	Naive Bayes	Linear
9	Perceptron	Linear
10	Decision Tree	Non-linear
11	Dummy	Non-linear
12	Neural Networks	Non-linear

Εικόνα 6: Γραφικά αποτελέσματα αλγορίθμου μηχανικής μάθησης κατά την αντιμετώπιση επίθεσης πλευρικού καναλιού (side channel attack)

4.1.6 Αλγόριθμοι Μηχανικής Μάθησης για την αντιμετώπιση επιθέσεων σε περιβάλλον Internet of Thinks

Οι υπολογιστές και οι φορητές συσκευές που λειτουργούν με διάφορα λειτουργικά συστήματα διαθέτουν πληθώρα λύσεων ασφαλείας και πρωτοκόλλων κρυπτογράφησης που μπορούν να τους προστατεύσουν από τις απειλές που αντιμετωπίζουν μόλις συνδεθούν στο Διαδίκτυο. Αυτό ωστόσο δε συμβαίνει με το IoT, [19].

Από τα δισεκατομμύρια συσκευές IoT που χρησιμοποιούνται επί του παρόντος, ένα σημαντικό ποσοστό είναι συσκευές επεξεργασίας χαμηλής ισχύος και μικρής χωρητικότητας αποθήκευσης και δεν έχουν τη δυνατότητα να επεκταθούν σε λύσεις ασφαλείας. Ωστόσο, είναι συνδεδεμένες στο Διαδίκτυο, το οποίο είναι ένα εξαιρετικά "εχθρικό" περιβάλλον. Είναι σαν να πηγαίνει κάποιος στο πεδίο της μάχης χωρίς πανοπλία.

Αυτός είναι ο λόγος για τον οποίο εμφανίζονται νέες ευπάθειες στο IoT συνεχώς και αμέτρητες συσκευές IoT πέφτουν θύματα από hacking, botnets και άλλων κακόβουλων ενεργειών καθημερινά. Χρειάζονται μόνο λίγα λεπτά για έναν κακόβουλο hacker να βρει χιλιάδες ευάλωτες συσκευές στη μηχανή αναζήτησης Shodan και οι παραβιασμένες συσκευές IoT γίνονται συχνά αιτία για πιο σοβαρές παραβιάσεις στα δίκτυα. Το σημαντικό είναι ότι πάρα πολλές από τις έξυπνες συσκευές μας είναι από τη φύση τους πολύ χαζές για να προστατεύσουν τον εαυτό τους (και εμάς) από κυβερνοεπιθέσεις.

Αλλά αυτό είναι ένα κενό που μπορεί να γεφυρωθεί με τη μηχανική μάθηση και τον τομέα ανάλυσης στοιχείων, ειδικά αφού πλέον καθίστανται όλο και πιο εύκολα διαθέσιμα στους προγραμματιστές και τους κατασκευαστές.

Οι συσκευές IoT παράγουν μεγάλους όγκους δεδομένων και η μηχανική μάθηση χρησιμοποιείται για την ανάλυση και τη διάγνωση αυτών των δεδομένων για τη βελτίωση της αποτελεσματικότητας και της εξυπηρέτησης των πελατών και τη μείωση του κόστους και της κατανάλωσης ενέργειας. Οι ίδιοι μηχανισμοί μπορούν να χρησιμοποιηθούν σε περιπτώσεις χρήσης που σχετίζονται με την ασφάλεια, όπως ο προσδιορισμός της συμπεριφοράς της ασφαλούς συσκευής και των γενικών προτύπων χρήσης, τα οποία μπορούν στη συνέχεια να βοηθήσουν στον εντοπισμό και στον αποκλεισμό μη φυσιολογικής δραστηριότητας αλλά και δυνητικά επιβλαβούς συμπεριφοράς.

Ήδη, αρκετές τεχνολογικές εταιρείες βασίζονται σε αυτό για να προσφέρουν λύσεις που βελτιώνουν την ασφάλεια του IoT, ειδικά σε έξυπνα σπίτια, όπου δεν υπάρχουν καθορισμένα πρότυπα και πρακτικές ασφάλειας.

4.1.6.1 Bitdefender

Η Bitdefender είναι μια ρουμανική εταιρεία τεχνολογίας στον κυβερνοχώρο με έδρα στο Βουκουρέστι της Ρουμανίας, με γραφεία στις Ηνωμένες Πολιτείες, την Ευρώπη, την Αυστραλία και τη Μέση Ανατολή. Η εταιρεία ιδρύθηκε το 2001 από τον τρέχοντα διευθύνοντα σύμβουλο και τον κύριο μέτοχο, Florin Talpes και έχει βραβευμένο λογισμικό ασφαλείας για το adroid τηλέφωνο και tablet.

Η προσέγγισή της είναι να συγκεντρωθούν σε δεδομένα server cloud από όλα τα τελικά σημεία που βασίζονται στα προϊόντα του· η είσοδος αναλύεται για τον προσδιορισμό των προτύπων και τον εντοπισμό κακόβουλης συμπεριφοράς. "Συγκεντρώνετε όλη την κίνηση", λέει ο Balan (επικεφαλής ερευνητής ασφάλειας και Διευθύνων Σύμβουλους της Bitdefender), "καθαρίστε

και κανονικοποιήστε το, μάθετε από αυτό, δείτε με ποιους διακομιστές μιλούν οι συσκευές, σε ποιες άλλες συσκευές μιλούν, πώς αλληλεπιδρούν κανονικά με το Διαδίκτυο και μεταξύ τους και εσείς βρείτε την ασυνήθιστη κίνηση".

Το Bitdefender χρησιμοποιεί την ευφυΐα και την αναγνώριση προτύπων που βασίζονται στην τεχνολογία cloud, μαζί με την ανάλυση του τοπικού δικτύου μέσω της πλατφόρμας λογισμικού και υλικού ασφαλείας τελικού σημείου, για τον έλεγχο της κίνησης στο Διαδίκτυο στα οικιακά δίκτυα και τον αποκλεισμό συνδέσεων με κακόβουλες διευθύνσεις URL, λήψεων κακόβουλου λογισμικού και ύποπτων πακέτων. Η αξιοποίηση των υπηρεσιών σύννεφου επέτρεψε στην εταιρεία να φέρει ευφυΐα και προστασία σε επίπεδο επιχειρήσεων στον χώρο των καταναλωτών.

4.1.6.2 PatternEx

Η PatternEx είναι μια εταιρεία που συνδυάζει τη δύναμη των ανθρώπων και των μηχανών σε ένα σύστημα AI που ανιχνεύει απειλές στον κυβερνοχώρο σε κλίμακα και σε πραγματικό χρόνο.

"Η μηχανική μάθηση είναι ένα κρίσιμο στοιχείο για την ανάπτυξη της Τεχνητής Νοημοσύνης για την ασφάλεια του IoT", λέει ο Uday Veeramachaneni, συνιδρυτής και διευθύνων σύμβουλος του PatternEx. "Το πρόβλημα είναι ότι οι συσκευές IoT θα διανεμηθούν μαζικά και αν υπάρχει επίθεση θα πρέπει να αντιδράσετε σε πραγματικό χρόνο".

Τα περισσότερα συστήματα που βασίζονται σε μηχανική μάθηση και ανάλυση συμπεριφοράς θα συλλέξουν πληροφορίες σχετικά με το δίκτυο και τις συνδεδεμένες συσκευές και στη συνέχεια θα αναζητήσουν ό,τι είναι μη φυσιολογικό. Το πρόβλημα με αυτήν την πρωτόγονη μέθοδο είναι ότι παράγει πάρα πολλά ψευδή σήματα ειδοποιήσεων για ενδεχόμενα και ψευδή θετικά.

Η προσέγγιση που προτείνει η PatternEx είναι να αναπτύξει μια λύση που ενσωματώνει τη μηχανική μάθηση και την αυξάνει με την ανθρώπινη ανάλυση στοιχείων για καλύτερη ανίχνευση επιθέσεων. "Ο τρόπος αντιμετώπισης αυτού σε πραγματικό χρόνο είναι να δημιουργήσουμε ένα μαθησιακό σύστημα που να παίρνει αυτές τις ακραίες τιμές και να ζητά ανθρώπινη ανατροφοδότηση», εξηγεί ο Veeramachaneni. "Ο άνθρωπος από μόνος του μπορεί να κάνει διάκριση μεταξύ κακόβουλης και καλοήθους ενέργειας, και αυτή η ανατροφοδότηση επιστρέφει στο σύστημα για να δημιουργήσει προγνωστικά μοντέλα που μπορούν να μιμηθούν την ανθρώπινη κρίση - αλλά σε τεράστια κλίμακα και σε πραγματικό χρόνο.

Αυτό ισχύει ιδιαίτερα για τα οικοσυστήματα IoT, όπου εμπλέκονται μεγάλοι αριθμοί συσκευών και η ανάλυση σε πραγματικό χρόνο της συντριπτικής ποσότητας δεδομένων που παράγονται είναι πέρα από τις ανθρώπινες ικανότητες.

Η PatternEx χρησιμοποιεί αλγόριθμους μηχανικής μάθησης για τον εντοπισμό ακραίων τιμών και εκπαιδεύει το μοντέλο να είναι πιο ακριβές σε πραγματικό χρόνο. Η εκπαίδευση γίνεται από έναν άνθρωπο, τον αναλυτή, που μπορεί να εντοπίσει μια νέα επίθεση που συμβαίνει. Το σύστημα δημιουργεί συμβάντα που υποδεικνύουν πιθανές επιθέσεις. Ο άνθρωπος διερευνά τα γεγονότα και καθορίζει αν το σύστημα ήταν σωστό στην εκτίμησή του ή όχι. Το σύστημα μαθαίνει από την εμπειρία αυτή και λαμβάνει πιο ακριβείς αποφάσεις την επόμενη φορά.

"Αυτό το μοντέλο συμβάλλει στη βελτίωση της ακρίβειας ανίχνευσης απειλών και στη μείωση του αριθμού των ψευδών θετικών με την πάροδο του χρόνου", λέει ο Veeramachaneni.

4.1.6.3 Dojo-Labs

Η Dojo-Labs είναι μια εταιρεία που παρέχει λύσεις ασφάλειας και απορρήτου των καταναλωτών για όλες τις συσκευές IoT σε ένα συνδεδεμένο έξυπνο σπίτι.

Οι συσκευές IoT έχουν σχεδιαστεί για να εκτελούν ένα περιορισμένο σύνολο λειτουργιών. Επομένως, με την υποστήριξη της μηχανικής μάθησης και τη χρήση δεδομένων, καθίσταται πολύ εύκολο να εντοπιστεί η μη ομαλή συμπεριφορά. Αυτή η ιδέα αξιοποιήθηκε από την startup εταιρεία τεχνολογίας Dojo-Labs για τη δημιουργία μιας λύσης ασφάλειας έξυπνου σπιτιού IoT.

"Όσον αφορά τις συσκευές IoT, έχουν σχεδιαστεί για να κάνουν μια πολύ, πολύ συγκεκριμένη λειτουργία", λέει ο Yossi Atias, συνιδρυτής και διευθύνων σύμβουλος της εταιρείας. "Έτσι, υποθέτοντας ότι έχουμε πολλούς χρήστες που χρησιμοποιούν την ίδια κάμερα ή την ίδια έξυπνη τηλεόραση ή το ίδιο έξυπνο συναγερμό ή έξυπνη κλειδαριά, δεν υπάρχει πραγματικός λόγος για τον οποίο μια συσκευή θα συμπεριφέρεται διαφορετικά από την άλλη, επειδή όλοι χρησιμοποιούν το ίδιο λογισμικό, γεγονός το οποίο δεν είναι κάτι που ο χρήστης μπορεί να αλλάξει".

Η μέθοδος της Dojo-Labs περιλαμβάνει τη συλλογή μεταδεδομένων από διαφορετικά τελικά σημεία και τον καθορισμό του εύρους συμπεριφοράς κάθε τύπου συσκευής, ώστε να είναι σε θέση να εντοπίζει και να αποκλείει την κακόβουλη συμπεριφορά. Όπως συμβαίνει με όλες τις λύσεις που αφορούν τη μηχανική μάθηση, το μοντέλο Dojo-Labs βελτιώνεται καθώς συλλέγει όλο και περισσότερα δεδομένα από τους πελάτες.

Η λύση περιλαμβάνει: μια μικρή συσκευή που μοιάζει με μικρή πέτρα και εγκαθίσταται στο οικιακό δίκτυο, μια εφαρμογή για κινητά που επιτρέπει στο χρήστη να ελέγχει τη συσκευή και να παρακολουθεί την κατάσταση του δικτύου και μια υπηρεσία σύννεφου όπου τα δεδομένα ενοποιούνται και αναλύονται χρησιμοποιώντας στατιστικά τεχνολογικά και μαθηματικά μοντέλα σε συνδυασμό με αλγόριθμους μηχανικής μάθησης.

4.1.6.4 Exabeam

Η Exabeam είναι μια μεγάλη εταιρεία ανάλυσης ασφάλειας δεδομένων που απλοποιεί τις λειτουργίες ασφαλείας και αλλάζει τον τρόπο με τον οποίο εντοπίζονται οι κυβερνοεπιθέσεις.

Σύμφωνα με τον Gartner πάνω από 8 δισεκατομμύρια συσκευές IoT χρησιμοποιήθηκαν το 2017 και έκτοτε ο αριθμός αυτός έχει αυξηθεί εκθετικά.

Η ασφάλεια ήταν πάντα ένα ζήτημα για τις συσκευές IoT κάτι που οφείλεται στους κατασκευαστές που τις έβγαλαν στην αγορά με κωδικούς πρόσβασης, με αδυναμία ενημέρωσής τους και άλλες ευπάθειες. Αυτό οδήγησε σε επιθέσεις DDoS και σε άλλες παραβιάσεις ασφάλειας με χρήση προσαρμοσμένων συστημάτων ελέγχου θερμοκρασίας (Heating, Ventilation, and Air Conditioning - HVAC), κλειστών κυκλωμάτων τηλεόρασης (Closed Circuit Television - CCTV) ή ακόμη και drone.

Ένα από τα προϊόντα της Exabeam που παρέχουν ασφάλεια από κυβερνοεπιθέσεις σε περιβάλλον IoT είναι το Exabeam Entity Analytics. Το προϊόν είναι διαθέσιμο στην αγορά, και η τιμολόγησή του εξαρτάται από τον αριθμό των συσκευών που διαχειρίζονται. Το Exabeam Entity Analytics χρησιμοποιεί τη μηχανική μάθηση για να εντοπίσει παραβιασμένες συσκευές IoT, [10]. Συγκεκριμένα μαθαίνει την κανονική συμπεριφορά βιομηχανικών, δικτυακών, κινητών, οικιακών και ιατρικών συσκευών και χρησιμοποιεί αυτές τις γνώσεις για να ειδοποιεί τα τμήματα ασφάλειας και πληροφορικής όταν εντοπίζεται ύποπτη δραστηριότητα.

Χρησιμοποιεί αρχείων καταγραφής για να αναζητήσει ύποπτη δραστηριότητα πχ συσκευές που στέλνουν πακέτα σε παράξενες τοποθεσίες, λήψη ασυνήθιστα μεγάλων ποσοτήτων δεδομένων ή προσπάθεια πρόσβασης σε ιδιόκτητους διακομιστές και δίκτυα.

Όταν εντοπίζεται τέτοια δραστηριότητα, τα τμήματα ασφάλειας ή / και πληροφορικής ενημερώνονται και τους δίνεται μια λίστα συσκευών που κινδυνεύουν προκειμένου να τις διερευνήσουν. Μπορούν επίσης να επιτρέψουν στην εφαρμογή να χειριστεί το πρόβλημα

ρυθμίζοντας εκ νέου αυτόματα τη συσκευή ή απομονώνοντας την από όλες τις άλλες συσκευές στο δίκτυο. Τα βασικά χαρακτηριστικά περιλαμβάνουν:

- Αυτόματη δημιουργία χρονοδιαγραμμάτων δραστηριότητας για συσκευές, δίνοντας στους αναλυτές μια πλήρη εικόνα για το πότε μια συσκευή άρχισε να εμφανίζει απρόσμενη συμπεριφορά.
- Υπολογισμός βαθμολογιών κινδύνου για κάθε συσκευή, με λεπτομέρεια για αναλυτική ανάλυση για επιτάχυνση της διαδικασίας.
- Μη εποπτευόμενη μηχανική εκμάθηση που ανακαλύπτει αυτόματα κανονικές συμπεριφορές όλων των συσκευών σε ένα δίκτυο.

"Οι άνθρωποι είναι πραγματικά μόνο το μισό του προβλήματος, και ίσως ούτε καν το μισό, δεδομένου του πόσο γρήγορα αυξάνεται η ρομποτοποίηση και ο αυτοματισμός", δήλωσε ο Sylvain Gil, αντιπρόεδρος προϊόντος στην Exabeam. "Για να προσδιορίσουμε τις επικίνδυνες συσκευές, πήραμε την ίδια μηχανή ανάλυσης που μοντελοποιεί τέλεια τη συμπεριφορά των χρηστών και την εφαρμόσαμε στις προβληματικές συσκευές, με τα ίδια χρονοδιαγράμματα και τις ίδιες βαθμίδες κινδύνου που έχουν πραγματικά βοηθήσει τους πελάτες μας".

4.1.6.5 Επαυξημένη νοημοσύνη (Augmented Intelligence)

Η μηχανική μάθηση είναι πολύ ελπιδοφόρα, αλλά είναι ακόμη στην αρχή της και έχει πολύ δρόμο να διανύσει. Και σε καμία περίπτωση δεν μπορεί να θεωρηθεί από μόνη της μια ολοκληρωμένη λύση. "Η μηχανική μάθηση θα είναι σχεδόν παντού", λέει ο Veeramachaneni. "Για να αποκτήσετε ασφάλεια στην επιχείρηση ή στον τομέα IoT, πρέπει να έχετε ισχυρούς μηχανισμούς που οργανώνουν δεδομένα, να αναλύουν δεδομένα και να αναζητούν μοτίβα στα δεδομένα. Αλλά χρειάζεστε επίσης τη διαίσθηση του ανθρώπου για να εντοπίσετε νέες επιθέσεις και να εκπαιδεύσετε το σύστημα για να σταματήσετε αυτές τις νέες (και παλιές) επιθέσεις".

Ο Veeramachaneni αποκαλεί αυτόν τον συνδυασμό με τον όρο "επαυξημένη νοημοσύνη" (Augmented Intelligence), μια εναλλακτική λύση για το αρκτικόλεξο AI, όπου οι δυνάμεις τόσο του ανθρώπου όσο και της μηχανής συγκλίνουν για να νικήσουν τις απειλές στον κυβερνοχώρο. "Ούτε η μηχανική μάθηση ούτε οι άνθρωποι μπορούν να το κάνουν μόνοι τους", λέει.

4.2 Αντίμετρα με best practices

4.2.1 Μέγεθος των Εκπαιδευτικών και υπό Δοκιμή Συνόλων Δεδομένων

Τα δεδομένα είναι βασικό στοιχείο για κάθε προσέγγιση που βασίζεται στη μηχανική μάθηση. Τα μοντέλα μηχανικής εκμάθησης απαιτούν επαρκή δεδομένα εκπαίδευσης για να μάθουν τα υποκείμενα πρότυπα στα δεδομένα και να γενικεύσουν για να προβλέψουν την τελική έξοδο των προηγούμενων μη ορατών εισόδων.

Δεν υπάρχει τυπικό μέγεθος για τα εκπαιδευτικά και υπό δοκιμή σύνολα δεδομένων. Το μέγεθος των εκπαιδευτικών συνόλων δεδομένων για τις προηγούμενες μεθόδους μπορεί να είναι τόσο μικρό όσο 2 στιγμιότυπα ή περισσότερα από 500 χιλιάδες στιγμιότυπα. Ωστόσο, η αύξηση του μεγέθους του εκπαιδευτικού συνόλου δεδομένων μειώνει τον κίνδυνο overfitting του μοντέλου και βελτιώνει την ικανότητα γενίκευσης του μοντέλου.

Στην ανάλυση πλευρικού καναλιού, ο αριθμός των ιχνών επίθεσης που απαιτούνται για την επίτευξη πρόβλεψης της εντροπίας με τιμή μικρότερη από 5 - σε μέτρια επίπεδα θορύβου - μειώνεται από 187 σε 9 ίχνη καθώς ο αριθμός των ιχνών κατανάλωσης ισχύος που χρησιμοποιούνται στη φάση του προφίλ αυξάνεται από 180 σε 1800 ίχνη. Στις επιθέσεις μοντελοποίησης PUF, η πρόβλεψη της απόκρισης ενός κριτή PUF χρησιμοποιώντας παλινδρόμηση ανάλυσης δεδομένων αυξάνεται από 95 σε 99,9% καθώς το μέγεθος του εκπαιδευτικού συνόλου δεδομένων αυξάνεται από 640 σε 18050 CRPs και από 1350 σε 3920, στην περίπτωση των 64-bit και 128- bit PUFs, αντίστοιχα. Στην περίπτωση των 4-XOR κριτών PUFs, το μέσο ποσοστό πρόβλεψης αυξάνεται από 98,76 σε 99,88% καθώς ο αριθμός των στιγμιότυπων των εκπαιδευτικών δεδομένων αυξάνεται από 24 χιλιάδες σε 200 χιλιάδες παρουσίες. Τα αποτελέσματα δείχνουν επίσης ότι ο απαιτούμενος αριθμός εκπαιδευτικών CRP για την επίτευξη 99% ακρίβειας αλλάζει ανάλογα με τον τύπο του PUF που δέχεται επίθεση. Για παράδειγμα, τα PUF κριτές απαιτούν τουλάχιστον 39,2 χιλιάδες CRPs, ενώ τα ελαφριά PUF απαιτούν 1000 χιλιάδες CRPs.

Επιπλέον, η καλύτερη ακρίβεια πρόβλεψης των BR και TBR PUFs - όταν χρησιμοποιούνται 50 βελτιωμένες επαναλήψεις - αυξάνεται από περίπου 86 σε 98% και από 92 σε 99%, αντίστοιχα, καθώς το μέγεθος του εκπαιδευτικού συνόλου αυξάνεται από 100 σε 1000 CRPs. Ως αποτέλεσμα, προτείνεται η σταδιακή αύξηση του μεγέθους του εκπαιδευτικού συνόλου δεδομένων έως ότου δεν παρατηρείται περαιτέρω βελτίωση στην απόδοση πρόβλεψης στο επικυρωμένο σύνολο δεδομένων.

4.2.2 Επίδραση της Εξαγωγής Χαρακτηριστικών, της Επιλογής και της Μείωσης Διαστάσεων

Τα επιλεγμένα συγκεκριμένα χαρακτηριστικά ανά περίπτωση επηρεάζουν σημαντικά την απόδοση των μοντέλων.

Όταν χρησιμοποιείται μειωμένος ρυθμός των RO (Relaxation Oscillation) συχνοτήτων RO αντί των τιμών των συχνοτήτων RO, η ακρίβεια ανίχνευσης των ανακυκλωμένων FPGA βελτιώνεται από 25 σε 100%.

Στην ανάλυση πλευρικού καναλιού, τα ίχνη ισχύος αφαιρούνται από το μέσο ίχνος ισχύος και στη συνέχεια χρησιμοποιούνται για την πρόβλεψη των μυστικών κλειδιών. Αυτή η προεπεξεργασία ιχνών ισχύος αυξάνει το ρυθμό πρόβλεψης των μυστικών κλειδιών από 85,23% σε 94%.

Σε ορισμένες περιπτώσεις, ο αριθμός των σχετικών χαρακτηριστικών είναι πολύ μεγάλος και συσχετίζονται μεταξύ τους. Αυτό μειώνει την απόδοση του μοντέλου και αυξάνει τον κίνδυνο του overfitting. Κατά συνέπεια, εφαρμόζονται πολλές μέθοδοι επιλογής χαρακτηριστικών και μείωσης διαστάσεων κατά την επίλυση προβλημάτων ασφάλειας υλικού. Οι πιο δημοφιλείς μέθοδοι που χρησιμοποιούνται στην έρευνα ασφάλειας υλικού είναι η συσχέτιση Pearson και η PCA. Το αποτέλεσμα της χρήσης αυτών των μεθόδων στην απόδοση ποικίλλει ανάλογα με τον τύπο του προβλήματος, τις χρησιμοποιούμενες δυνατότητες, το μέγεθος των εκπαιδευτικών δεδομένων και τα μοντέλα μηχανικής μάθησης. Για παράδειγμα, η απόδοση ταξινόμησης του MLP υποβαθμίζεται όταν η PCA εφαρμόζεται σε ίχνη. Ωστόσο, όταν προβλέπεται η τιμή του τέταρτου λιγότερο σημαντικού bit της εξόδου SBox (bit (4)), το ποσοστό επιτυχίας μειώνεται από 74,7 σε 52,7% καθώς ο αριθμός των χαρακτηριστικών αυξάνεται από 3 σε 6. Χρησιμοποιείται η συσχέτιση του Pearson και τα αποτελέσματα δείχνουν ότι η αύξηση του μεγέθους του εκπαιδευτικού συνόλου δεδομένων και ο αριθμός των σημείων δεδομένων σε κάθε ίχνος, αυξάνουν την ακρίβεια της ταξινόμησης. Ωστόσο, αν το μέγεθος του εκπαιδευτικού συνόλου δεδομένων είναι μικρό, η αύξηση του αριθμού των σημείων δεδομένων ανά ίχνος μειώνει την ακρίβεια της ταξινόμησης λόγω overfitting.

4.2.3 Εξισορρόπηση Κλάσης Συνόλου Δεδομένων

Σε ορισμένα εποπτευόμενα προβλήματα ταξινόμησης, μια κλάση μπορεί να έχει περισσότερα δεδομένα από μια άλλη κλάση (ή κλάσεις). Η εκπαίδευση ενός μοντέλου χρησιμοποιώντας μη ισορροπημένα δεδομένα προκαλεί "προτίμηση" του μοντέλου προς την κλάση με τον μεγαλύτερο αριθμό στιγμιότυπων εκπαιδευτικών δεδομένων. Ο ρυθμός των σωστά

προβλεπόμενων δεδομένων (True Positive Rate – TPR) είναι 0% όταν το 96% των στιγμιότυπων εκπαιδευτικών δεδομένων επισημαίνονται ως μη μολυσμένα από trojan και μόνο το 4% των στιγμιότυπων δεδομένων επισημαίνονται ως μολυσμένα από trojan. Τα δεδομένα μπορούν να εξισορροπηθούν χρησιμοποιώντας στατική και δυναμική στάθμιση. Η στατική στάθμιση βασίζεται στην επανάληψη κάθε εκπαιδευτικού στιγμιότυπου που έχει την ετικέτα μειοψηφίας N φορές, έτσι ώστε το εκπαιδευτικό σύνολο δεδομένων να εξισορροπηστεί. Ενώ η δυναμική στάθμιση εξαλείφει τα χαρακτηριστικά με παρόμοιες τιμές σε όλες τις κλάσεις και στη συνέχεια αντιγράφει τα στιγμιότυπα που ανήκουν στην κατηγορία της μειοψηφικής κλάσης N φορές για να εξισορροπήσει τα δεδομένα. Τα αποτελέσματα δείχνουν ότι η στατική στάθμιση αυξάνει το μέσο TPR στο 30% ενώ η δυναμική στάθμιση αυξάνει το μέσο TPR στο 82,6%.

Όταν το 67% του εκπαιδευτικού συνόλου δεδομένων ανήκει στην κλάση που έχει μολυνθεί από trojan, οι επιθέσεις εντοπίζονται με ακρίβεια μεγαλύτερη από 90%. Ωστόσο, τα εκπαιδευμένα μοντέλα διατρέχουν κίνδυνο "προτίμησης" λόγω του μη ισορροπημένου εκπαιδευτικού συνόλου δεδομένων. Κατά συνέπεια, είναι ζωτικής σημασίας να διασφαλιστεί ότι το χρησιμοποιημένο εκπαιδευτικό σύνολο δεδομένων είναι ισορροπημένο.

4.2.4 Ρύθμιση Παραμέτρων Μοντέλου

Οι παράμετροι του μοντέλου πρέπει να ρυθμιστούν σωστά ώστε να επιτευχθεί η καλύτερη απόδοση. Η σωστή επιλογή του πυρήνα SVM και των υπερ-παραμέτρων (hyper-parameters) έχει σημαντικό αντίκτυπο στην ακρίβεια της ταξινόμησης. Η ακρίβεια ταξινόμησης του βάρους Hamming ως άρτια ή περιττή τιμή μειώνεται από 99 σε 82,7% όταν η τιμή του σ^2 αυξάνεται από 0,1 σε 10 χρησιμοποιώντας SVM με πυρήνα RBF. Όταν χρησιμοποιείται γραμμικός πυρήνας, η ακρίβεια της ταξινόμησης μειώνεται περαιτέρω στο 49,9%.

Επιπλέον, η βέλτιστη απόδοση σε χαμηλά επίπεδα θορύβου σήματος επιτυγχάνεται όταν ένα μοντέλο SVM με πυρήνα RBF εκπαιδεύεται με παράμετρο ποινής C με τιμή 10. Το C είναι μία από τις παραμέτρους του SVM που πρέπει να ρυθμιστεί σωστά. Αντιπροσωπεύει το κόστος που σχετίζεται με τη λανθασμένη ταξινόμηση των σημείων δεδομένων. Τα αποτελέσματα δείχνουν επίσης ότι όταν το C έχει τιμή μεγαλύτερη από 10, η απόδοση του SVM υποβαθμίζεται καθώς το μοντέλο αρχίζει εμφανίζει το φαινόμενο overfitting. Τέλος, οι παράμετροι του SVM επηρεάζουν επίσης την ακρίβεια της πρόβλεψης.

Όταν το SVM, το μοντέλο δυναμικού χρόνου (DT), της γραμμικής παλινδρόμησης (LR) και του KNN χρησιμοποιούνται για την ανίχνευση των trojans που προκαλούν εκτροπή κίνησης,

βρόχο διαδρομής και πλαστογράφηση διευθύνσεων πυρήνα, η ακρίβεια ανίχνευσης εξαρτάται από τον τύπο του trojan και τον τύπο του μοντέλου μηχανικής μάθησης που χρησιμοποιείται. Για παράδειγμα, το μοντέλο DT ανιχνεύει πλαστογράφηση διεύθυνσης πυρήνα, βρόχο διαδρομής και εκτροπή κίνησης με ακρίβεια 94, 95 και 99%, αντίστοιχα. Τα εποπτευόμενα μοντέλα μάθησης όπως το SVM, DT και LR εντοπίζουν μια επίθεση εκτροπής της κίνησης με ακρίβεια μεγαλύτερη από 95%. Αντίθετα, τα μη εποπτευόμενα μοντέλα μάθησης έχουν σημαντικά χαμηλή ακρίβεια πρόβλεψης για αυτόν τον τύπο επιθέσεων. Η μέγιστη ακρίβεια ανίχνευσης της εκτροπής κίνησης είναι 72% όταν χρησιμοποιείται η μέθοδος μεγιστοποίησης προσδοκίας. Αυτά τα αποτελέσματα δείχνουν τη σημασία της επιλογής μοντέλου στην ακρίβεια ανίχνευσης των επιθέσεων. Επιπλέον, οι αλγόριθμοι KNN-1, KNN-5 και KNN-10 ανιχνεύουν την επίθεση εκτροπής της κίνησης με ακρίβεια 95, 85 και 84%, αντίστοιχα. Αυτά τα αποτελέσματα τονίζουν τη σημασία της επιλογής παραμέτρων. Όταν η τιμή της μεταβλητής K-που αντιπροσωπεύει τον αριθμό των πλησιέστερων γειτόνων-αλλάζει, η ακρίβεια ανίχνευσης αλλάζει σημαντικά.

Επίσης, εστιάζουμε στο SVM μιας κατηγορίας για την ανίχνευση κυκλωμάτων που έχουν μολυνθεί από trojan. Πιο συγκεκριμένα, μελετάται μια παραλλαγή του SVM μιας κατηγορίας, που ονομάζεται ν-SVM. Η παράμετρος βελτιστοποίησης ν είναι προσαρμοσμένη με τέτοιο τρόπο ώστε να βελτιώνει την απόδοση ταξινόμησης. Όταν το ν έχει την τιμή 5×10^(-4), τα κυκλώματα που δεν είναι μολυσμένα με trojan ανιχνεύονται με ακρίβεια 100%, ενώ τα κυκλώματα που είναι μολυσμένα με trojan ανιχνεύονται με μέγιστη ακρίβεια 9%. Όταν το ν έχει την τιμή 5×10^(-2) δεν ανιχνεύονται τα κυκλώματα που δεν είναι μολυσμένα με trojan, και τέλος, όταν το ν ισούται με 7×10^(-3), ανιχνεύονται όλα τα κυκλώματα που είτε είναι μολυσμένα με Trojan είτε όχι.

Ακολούθως, χρησιμοποιείται ο αλγόριθμος AdaBoost με τρεις διαφορετικούς αδύναμους εκπαιδευόμενους. Τα αποτελέσματα δείχνουν ότι η ακρίβεια πρόβλεψης της απόκρισης των BR PUF και TBR PUF ποικίλλει ανάλογα με τον τύπο των ασθενών εκπαιδευομένων που χρησιμοποιούνται. Για παράδειγμα, η ακρίβεια πρόβλεψης της απόκρισης των BR PUFs, χρησιμοποιώντας 1000 CRPs εκπαιδευτικά, είναι 63,7, 75,7 και 84,6% όταν χρησιμοποιούνται αντίστοιχα μονώνυμα, δέντρα αποφάσεων και λίστες αποφάσεων, αντίστοιχα, τα οποία θεωρούνται οι αδύναμοι εκπαιδευόμενοι, χωρίς κάποια ενίσχυση. Τα αποτελέσματα δείχνουν επίσης ότι η ακρίβεια της πρόβλεψης αυξάνεται σε 96,8, 95,04 και 98,32% καθώς ο αριθμός των "ενισχυμένων" επαναλήψεων αυξάνεται σε 50.

Λαμβάνοντας υπόψη αυτά τα ευρήματα όπως αναφέρθηκαν, προτείνεται η υιοθέτηση ελεγχόμενης επικύρωσης των αποτελεσμάτων με αναζήτηση πλεγμάτων από δεδομένα, για την επιλογή των πιο κατάλληλων μοντέλων μηχανικής μάθησης και τον προσδιορισμό των βέλτιστων παραμέτρων τους.

4.2.5 Επίδραση του Θορύβου και Μεταβλητότητα της Διαδικασίας

Ένα από τα σημαντικότερα προβλήματα που αντιμετωπίζονται στη συλλογή πραγματικών δεδομένων από το υλικό είναι ο θόρυβος. Ο θόρυβος επηρεάζει σημαντικά την ακρίβεια των προβλέψεων που έχουν ληφθεί. Για παράδειγμα, σε επιθέσεις μοντελοποίησης PUF, μπορούν να εισαχθούν λάθη στο εκπαιδευτικό σύνολο δεδομένων για την αναπαράσταση του θορύβου και της μεταβλητότητας της διαδικασίας. Όταν μέθοδος παλινδρόμησης ανάλυσης δεδομένων εκπαιδεύεται σε 50 χιλιάδες εκπαιδευτικά σημεία δεδομένων για να μοντελοποιήσει τον κριτή PUF από 4-XOR, η ακρίβεια της πρόβλεψης μειώνεται από 99,37 σε 88,2% όταν το ποσοστό του εγχυόμενου σφάλματος αυξάνεται από 0 σε 10%. Επομένως, τα δεδομένα πρέπει να υποβάλλονται σε κατάλληλη προεπεξεργασία και πρέπει να εφαρμόζονται μοντέλα ανθεκτικά στο θόρυβο πρόκειται για δεδομένα με υψηλά επίπεδα θορύβου. Περιγράφονται μερικές τεχνικές που εφαρμόστηκαν, για παράδειγμα, τα δεδομένα που συλλέχθηκαν μπορούν να δειγματοληπτηθούν αρκετές φορές και να προσμετρηθούν κατά μέσο όρο για να μειωθεί η επίδραση του θορύβου στην ακρίβεια της πρόβλεψης. Ο αλγόριθμος συσταδοποίησης K-means πρέπει να αποφεύγεται όσο είναι δυνατόν, καθώς έχει ευαισθησία στο θόρυβο και έχει μικρή ισχύ γενίκευσης. Πιο συγκεκριμένα, όταν χρησιμοποιείται 1 κύκλωμα για εκπαίδευση και όταν το επίπεδο θορύβου είναι χαμηλό, ο αλγόριθμος συσταδοποίησης K-means ανιχνεύει ολοκληρωμένα κυκλώματα με trojan με ακρίβεια 92,8%. Ωστόσο, καθώς το επίπεδο θορύβου αυξάνεται σε μεσαία και υψηλά επίπεδα, η ακρίβεια μειώνεται σε 9,6 ή ακόμα και 0%, αντίστοιχα.

Συνιστώμενες Βέλτιστες Πρακτικές κατά την Εφαρμογή Μηχανικής Μάθησης σε Προβλήματα Ασφάλειας Υλικού

Προτείνονται οι ακόλουθες οδηγίες κατά την εφαρμογή μηχανικής μάθησης σε οποιοδήποτε πρόβλημα που σχετίζεται με την ασφάλεια υλικού:

- Κατάλληλη προεπεξεργασία δεδομένων πρέπει να εφαρμόζεται σε δεδομένα, για παράδειγμα, εξάλειψη θορύβου και κανονικοποίηση δεδομένων.
- Το εκπαιδευτικό σύνολο δεδομένων πρέπει να είναι ισορροπημένο πριν από την εκπαίδευση.

- Οι τεχνικές επιλογής χαρακτηριστικών και μείωσης διαστάσεων πρέπει να εφαρμόζονται όταν χρειάζεται.

Η απόδοση των μοντέλων μηχανικής εκμάθησης εξαρτάται από διάφορους παράγοντες που περιλαμβάνουν (1) την αξία και το μέγεθος των επιλεγμένων χαρακτηριστικών δεδομένων, (2) το επίπεδο θορύβου στα δεδομένα, (3) το μέγεθος των εκπαιδευτικών δεδομένων, (4) την ισορροπία μεταξύ διαφορετικών κατηγοριών δεδομένων - σε περίπτωση εποπτευόμενης μάθησης και (5) την επιλογή παραμέτρων μοντέλου. Επομένως, θα πρέπει να χρησιμοποιείται ελεγχόμενη επικύρωση για την επιλογή του καλύτερου μοντέλου με τις πιο κατάλληλες παραμέτρους.

4.2.6 Θα πρέπει να χρησιμοποιούνται διαφορετικές μετρήσεις αξιολόγησης απόδοσης

Τα εκπαιδευτικά μοντέλα μηχανικής εκμάθησης πρέπει να γίνονται επαναληπτικά για τη βελτίωση της απόδοσης τους. Για παράδειγμα, τα μοντέλα θα πρέπει να αξιολογούνται σε ένα αρχικό σύνολο χαρακτηριστικών. Στη συνέχεια, θα πρέπει να επιλεγούν τα καλύτερα μοντέλα (από πλευράς απόδοσης) και να ρυθμίζονται οι παράμετροί τους μέχρι να μην παρατηρείται επιπλέον βελτίωση στην απόδοση. Στη συνέχεια, τα εξαγόμενα χαρακτηριστικά θα μπορούσαν να επανεξεταστούν και να προστεθούν περισσότερες δυνατότητες για να διερευνηθεί η επίδρασή τους στην απόδοση του μοντέλου. Αυτές οι επαναλήψεις θα πρέπει να συνεχιστούν έως ότου δεν παρατηρείται περεταίρω βελτίωση ή να επιτευχθεί η ζητούμενη απόδοση.

4.3 Αντιμέτρα με υλικό

Στην αναζήτηση λύσεων, από [35], μπαίνουμε σε μια πολύ αβέβαιη περιοχή, με πολλά άγνωστα και πολύ λίγες συγκεκριμένες απαντήσεις. Κατά συνέπεια, ενώ οι εμπειρογνώμονες ασφαλείας πρέπει να κάνουν το καλύτερο δυνατό για να αναπτύξουν μεθόδους προστασίας σε πολλές γνωστές απειλές, θα πρέπει επίσης να σχεδιάσουν συσκευές έτσι ώστε να μπορούν να παραμετροποιηθούν και να αναβαθμιστούν για να αντιμετωπίσουν τα απρόβλεπτα τρωτά σημεία τους και περιορισμούς.

Το "Patching" είναι μια οικεία έννοια στην υπολογιστικότητα, τουλάχιστον στον τομέα του λογισμικού. Είναι γνωστή η επιτακτική ανάγκη για ενημερώσεις λογισμικού για να διασφαλιστεί η συνεχής ασφαλής λειτουργία των τηλεφώνων και των υπολογιστών των χρηστών. Αυτές οι ενημερώσεις ασφαλείας είναι απαραίτητες επειδή μια τυπική υπολογιστική συσκευή εκτίθεται σε δεκάδες νέες ευπάθειες κάθε μήνα.

Μέχρι πρόσφατα, το patching γινόταν μόνο σε λογισμικό ή σε firmware, με το οποίο οι άνθρωποι συχνά αναφέρονται στον κώδικα συστήματος που εκτελείται σε μικρές συσκευές. Το υποκείμενο υλικό παραμένει αμετάβλητο. Υποστηρίζεται ότι οι μηχανικοί δεν πρέπει να επιτρέπουν μόνο στο λογισμικό αλλά και στο υλικό να μπορεί να επιδιορθωθεί σε συσκευές που προορίζονται να γίνουν μέρος του IoT. Αυτό ισχύει, επειδή μπορεί να μην είναι δυνατό να διορθωθούν όλες οι ευπάθειες ασφαλείας μόνο τροποποιώντας το λογισμικό. Για παράδειγμα, το υλικό μπορεί να εφαρμόσει έναν αλγόριθμο κρυπτογράφησης που είναι ασφαλής τώρα, αλλά θα μπορούσε ξεπεραστεί πολύ πριν από το τέλος της διάρκειας ζωής του συστήματος. Ο μόνος τρόπος αντιμετώπισης αυτής της πιθανότητας είναι να υπάρχει υλικό που μπορεί να παραμετροποιηθεί εκ νέου μετά την κατασκευή της συσκευής.

Ένας άλλος λόγος για να είναι εφικτό το υλικό να γίνει patchable είναι ότι οι μικρές συνδεδεμένες συσκευές πρέπει συχνά να λειτουργούν με πολύ μικρή κατανάλωσης ενέργειας και οι εφαρμογές λογισμικού μιας δεδομένης λειτουργικότητας καταναλώνουν συνήθως περισσότερη ισχύ από ό, τι οι υλοποιήσεις υλικού για την ίδια ενέργεια. Συνεπώς, οι μηχανικοί συχνά δεν μπορούν να σχεδιάσουν μια μικρή συσκευή χαμηλής ισχύος που κάνει αυτό που πρέπει να κάνει μόνο χρησιμοποιώντας λογισμικό που εκτελείται σε κάποιο generic υλικό - οι συσκευές πρέπει να χρησιμοποιούν υλικό ειδικής χρήσης για την εργασία τους. Ως αποτέλεσμα, η επιδιόρθωση λογισμικού πιθανότατα θα είναι ανεπαρκής για την πραγματοποίηση των απαραίτητων αναβαθμίσεων ασφαλείας.

Η πηγή [35], προβάλλει μια λύση για ένα σχέδιο υλικού με τη χρήση FPGA, ένα κύκλωμα γενικής χρήσης στο οποίο μπορεί να παραμετροποιηθεί η λογική (logic) μετά την κατασκευή του. Στην προτεινόμενη αρχιτεκτονική, ένα κεντρικό μπλοκ υλικού, που ονομάζεται μηχανή πολιτικής ασφαλείας, θα διαχειρίζεται ένα ολοκληρωμένο σύνολο κρίσιμων για την ασφάλεια γεγονότων, συμπεριλαμβανομένης της επικοινωνίας μεταξύ άλλων μπλοκ σχεδιασμού στο σύστημα και με τον εξωτερικό κόσμο. Για παράδειγμα, η μηχανή πολιτικής ασφαλείας ενδέχεται να απαιτεί ένα μυστικό κρυπτογραφικό κλειδί που θα χρησιμοποιείται για επικοινωνία και θα είναι προσβάσιμο μόνο σε συγκεκριμένα μπλοκ υλικού. Για την επιβολή αυτού του κανόνα, ο μηχανισμός πολιτικής ασφαλείας θα πρέπει να διαχειρίζεται την κοινή χρήση μυστικών κλειδιών μεταξύ μπλοκ, απαγορεύοντας ανταλλαγές που δεν πληρούν τις προκαθορισμένες απαιτήσεις ασφαλείας. Αν αυτή η μηχανή πολιτικής ασφαλείας έχει κατασκευαστεί χρησιμοποιώντας FPGA· επειδή ένα FPGA είναι αναβαθμίσιμο, μπορεί να γίνει patch. Συγκεκριμένα, αν πρέπει να ενημερωθεί η συσκευή για μια απειλή που

ανακαλύφθηκε πρόσφατα, μπορεί να γίνει patched το υλικό για την επιβολή νέου συνόλου απαιτήσεων ασφαλείας, ακόμα και όταν είναι στο ρεύμα.

5. Προκλήσεις & κίνδυνοι

Η μηχανική μάθηση και οι προγραμματιστές τεχνητής νοημοσύνης αρχίζουν να εξετάζουν την ακεραιότητα των εκπαιδευτικών δεδομένων, τα οποία σε ορισμένες περιπτώσεις θα χρησιμοποιηθούν για την εκπαίδευση εκατομμυρίων ή και δισεκατομμυρίων συσκευών. Αλλά αυτή είναι η αρχή μόνο μιας τεράστιας προσπάθειας, γιατί σήμερα κανείς δεν είναι απόλυτα σίγουρος για το πώς αυτά τα εκπαιδευτικά δεδομένα μπορούν να καταστραφούν ή τι να κάνουν σχετικά με αυτά αν είναι κατεστραμμένα, [2].

Η μηχανική μάθηση, η βαθιά μάθηση και η τεχνητή νοημοσύνη είναι ισχυρά εργαλεία για τη βελτίωση της αξιοπιστίας και της λειτουργικότητας των συστημάτων και την επιτάχυνση τους προς στην αγορά. Ωστόσο, οι αλγόριθμοι AI μπορούν επίσης να περιέχουν σφάλματα, υποκείμενες συμπεριφορές ή ακόμη και κακόβουλα προγράμματα που μπορεί να μην εντοπιστούν για χρόνια, σύμφωνα με περισσότερους από δώδεκα ειδικούς που ερωτήθηκαν τους τελευταίους μήνες. Σε ορισμένες περιπτώσεις, η αιτία μπορεί να είναι λάθη στον προγραμματισμό, κάτι που δεν είναι ασυνήθιστο καθώς αναπτύσσονται νέα εργαλεία ή τεχνολογίες. Οι αλγόριθμοι μηχανικής μάθησης και τεχνητής νοημοσύνης εξακολουθούν να είναι καλά παραμετροποιημένοι και αναβαθμισμένοι. Αλλά παράλληλα υπάρχει ένας αυξανόμενος φόβος ότι μπορεί να γίνει ένα σημείο εισόδου για κακόβουλο λογισμικό, το οποίο γίνεται μια "πίσω πόρτα" που μπορεί να ανοίξει αργότερα.

Ακόμα και όταν εντοπίζονται ελαττώματα ή κακόβουλα προγράμματα, είναι σχεδόν αδύνατο να εντοπιστεί η βασική αιτία του προβλήματος και να διορθωθούν όλες οι συσκευές που έχουν εκπαιδευτεί με αυτά τα δεδομένα. Μέχρι εκείνο το σημείο μπορεί να υπάρχουν εκατομμύρια από αυτές τις συσκευές στην αγορά. Αν αναπτυχθούν επιδιορθώσεις, δεν θα είναι όλες αυτές οι συσκευές συνδεδεμένες στο Διαδίκτυο ή ακόμη και προσβάσιμες. Το χειρότερο ακόμα σενάριο είναι ότι αυτός ο κώδικας δεν ανακαλύπτεται έως ότου ενεργοποιηθεί από κάποιον εξωτερικό κακόβουλο χρήστη, ανεξάρτητα από το αν το τοποθέτησαν εκεί ή απλά υπέπεσαν σε αυτόν.

Ένα άλλο πρόβλημα είναι μια "backdoor" που δημιουργείτε με τα ίδια τα δεδομένα. Για παράδειγμα, όταν χρησιμοποιούμε μια κάμερα ασφαλείας μηχανικής μάθησης και την εκπαιδεύουμε για να αναζητήσουμε ένα συγκεκριμένο καπέλο, πουκάμισο ή παπούτσια, τότε μπορεί εύκολα να εξαχθεί που μένουμε ή ποιες είναι οι προτιμήσεις μας. Αμέσως-αμέσως έχουμε δημιουργήσει ένα νέο πρόβλημα που δεν υπήρχε προτού αποκτήσουμε κάμερα με δυνατότητα machine learning.

Οι κίνδυνοι ασφαλείας είναι παντού και οι συνδεδεμένες συσκευές αυξάνουν την ικανότητα περισσότερων επιθέσεων από απομακρυσμένες τοποθεσίες. Αλλά η AI, και τα υποπεδία της (η μηχανική μάθηση και η βαθιά μάθηση) δημιουργούν νέες απειλές, επειδή οι μηχανές χρησιμοποιούνται για την εκπαίδευση άλλων μηχανών και κανείς δεν είναι αρκετά σίγουρος πότε ή πώς αυτά τα εκπαιδευμένα μηχανήματα θα χρησιμοποιήσουν τελικά αυτά τα δεδομένα. Καθώς οι μηχανές χρησιμοποιούνται για την εκπαίδευση άλλων μηχανών, τα ίδια τα μηχανήματα διαδίδουν το πρόβλημα. Μόνο που σε αυτήν την περίπτωση δεν είναι ένας κλασικός ιός. Είναι οι εσωτερικές λειτουργίες των αλγορίθμων που καθοδηγούν αυτά τα συστήματα, γεγονός που καθιστά το πρόβλημα πολύ πιο δύσκολο να εντοπιστεί. Αντί να αναζητούν ένα μόνο ελάττωμα ασφαλείας, οι ειδικοί ασφαλείας θα πρέπει να αναζητούν ασυνήθιστα μοτίβα, στις καλύτερες περιπτώσεις, και καλά αποδεκτά πρότυπα συμπεριφοράς στις χειρότερες περιπτώσεις. "Στην ομάδα μας για την ασφάλεια στον κυβερνοχώρο ξοδεύουμε πολύ χρόνο ανησυχώντας για την καταπολέμηση της τεχνητής νοημοσύνης με τεχνητή νοημοσύνη", δήλωσε ο Jeff Welser, αντιπρόεδρος και διευθυντής εργαστηρίου της IBM Research Almaden. "Έχοντας συστήματα όπου ο ίδιος ο εισβολέας γράφει προγράμματα με δυνατότητα AI για να μάθει τι συμβαίνει στα μοτίβα, και ως εκ τούτου να είναι καλύτερα σε θέση να διεισδύσει σε αυτά, είναι ένας τομέας στον οποίο κάνουμε έρευνα αυτή τη στιγμή. Πραγματικά πρόκειται να είναι, κατά μία έννοια, AI έναντι AI.

Αυτό δεν είναι απαραίτητα στιγμιαία αιτία και αποτέλεσμα. Μερικές φορές ο αντίκτυπος μπορεί να χρειαστεί χρόνια για να εμφανιστεί, όπως σε συστήματα πλοήγησης αυτοκινήτου ή αεροπλάνου όπου η "πίσω πόρτα" γίνεται εργαλείο για ransomware.

Υπάρχουν μερικά προφανή πράγματα από τα οποία πρέπει να ξεκινήσουμε για να κλείσουν τρύπες ασφαλείας στη μηχανική μάθηση και στα συστήματα AI. Το ένα περιλαμβάνει τον περιορισμό της πρόσβασης σε αλγόριθμους. "Υπάρχουν μερικά σημεία που πρέπει να τα κάνετε σωστά", δήλωσε ο Raik Brinkmann, Διευθύνων Σύμβουλος της OneSpin Solutions. "Ένα είναι ο έλεγχος ταυτότητας, για να βεβαιωθείτε ότι η συσκευή που στέλνει πίσω δεδομένα είναι η συσκευή με την οποία θέλετε να μιλήσετε. Σε επίπεδο κυκλωμάτων, πρέπει να γνωρίζετε ότι αυτό το συγκεκριμένο κύκλωμα είναι αυτό με το οποίο μιλάτε. Υπάρχουν ορισμένες εταιρείες πνευματικής ιδιοκτησίας (Intellectual Property - IP) που απαντούν αυτήν την ερώτηση. Πώς τοποθετείτε ένα αναγνωριστικό σε κάτι που αναπτύσσετε; Και μόνο όταν είναι ενεργοποιημένο στον πελάτη λαμβάνεται το αναγνωριστικό και όχι όταν βρίσκεται στο εργοστάσιο. Μπορείτε να συσχετίσετε την πηγή δεδομένων με αυτό το κύκλωμα. Μετέπειτα, υπάρχουν τεχνολογίες όπως το blockchain για να βεβαιωθείτε ότι τα δεδομένα που

διακινούνται από αυτή τη συσκευή είναι αυθεντικά με τα αναμενόμενα δεδομένα. Η προστασία από τυχόν παραβιάσεις είναι σημαντική στη ροή δεδομένων. Πρέπει να ελέγξετε τη ροή δεδομένων και να εγγυηθείτε την ακεραιότητα, διαφορετικά θα έχετε ένα μεγάλο πρόβλημα ασφάλειας".

Σε αντίθεση με τα πιο συμβατικά ηλεκτρονικά συστήματα, ωστόσο, τα συστήματα τεχνητής νοημοσύνης (AI) / μηχανικής μάθησης (ML) / βαθιάς μάθησης (DL) είναι από τη φύση τους πιο ανθεκτικά. Αντί για συμβατική επεξεργασία, όπου ένα σύστημα θα σταματήσει ή θα καταρρεύσει εάν δεν μπορεί να δώσει ακριβή απάντηση, τα συστήματα AI / ML / DL παράγουν αποτελέσματα που ανήκουν σε μια κατανομή. Αυτό παρέχει κάποιο είδος "προστασίας" αν κάτι δεν ταιριάζει ακριβώς, το οποίο είναι χρήσιμο στην προσαρμογή σε πραγματικές συνθήκες, όπως η αναγνώριση ενός αντικειμένου στο δρόμο. Ωστόσο, κάνει πιο δύσκολο τον εντοπισμό για το πού ακριβώς υπάρχει πρόβλημα (εφόσον εμφανιστεί).

"Πρέπει να καθορίσετε ποιο είναι το μοντέλο απειλής και, στη συνέχεια, ιδανικά χρειάζεστε κάποιο είδος μετρικής για την ασφάλεια", δήλωσε ο Arm's Aitken. "Αυτές οι μετρικές είναι πραγματικά δύσκολο να βρεθούν γιατί αν πάτε π.χ. στο αφεντικό σας και πείτε "Πρέπει να προσθέσω επιπλέον ασφάλεια σε αυτό το κύκλωμα και θα κοστίσει τρεις μήνες δουλειάς" και το αφεντικό σας πει, "Τι πραγματικά θα κερδίσω με αυτό;" Η απάντηση θα είναι, "Θα είναι πιο ασφαλές". Αλλά πόσο πιο ασφαλές και με ποιο τρόπο; Η ύπαρξη αυτών των μετρικών για την ασφάλεια είναι απαραίτητη. Εκεί υπεισέρχεται το κομμάτι του blockchain. Το επίπεδο ελέγχου ταυτότητας και ελέγχου που χρειάζεστε για διαφορετικά δεδομένα αλλάζει. Αυτό είναι παρόμοιο με όταν πηγαίνω στο κατάστημα και αγοράζω ένα στυλό, υποθέτω ότι το κατάστημα απέκτησε το στυλό νόμιμα και δεν το ενδιαφέρει τι κάνω με αυτό αφού το αγοράσω. Αλλά αν αγοράσω ένα αυτοκίνητο, πρέπει να γνωρίζετε όλους τους ιδιοκτήτες αυτού του αυτοκινήτου και την ιστορία αυτού του αυτοκινήτου και όταν το πουλήσω το κράτος θέλει να μάθει σε ποιον το πούλησα. Το παρακολουθούν σε πολύ ένα πιο λεπτομερές επίπεδο από ό,τι ένα στυλό. Η ίδια αλυσίδα τιμών ισχύει για τα δεδομένα".

Τα προβλήματα ασφάλειας αυξάνονται παντού. Εκτός από τις μακροχρόνιες απειλές που συνυπάρχουν στο λογισμικό εφαρμογών και κατά την πρόσβαση στο δίκτυο, η εισαγωγή των απειλών Meltdown και Spectre ανέδειξαν ελαττώματα στην αρχιτεκτονική x86 που δεν είχαν καν ανιχνευθεί όταν αναπτύχθηκε αυτή η αρχιτεκτονική. Και η ιδέα μόνο ότι τα αυτοκίνητα μπορούν να χακαριστούν και να δικτυωθούν φαινόταν σχεδόν παράλογη πριν από μια δεκαετία.

Όμως ολόκληρα συστήματα συνδέονται με άλλα συστήματα, και αυτό ανοίγει την πρόσβαση σε παγκόσμιο επίπεδο σε όλους, από κακόβουλους έως εξελιγμένους οργανωμένους οργανισμούς ηλεκτρονικού εγκλήματος και χώρες. "Η μετατόπιση που έχουμε ήδη κάνει είναι από την κλιμακωτή πολυπλοκότητα στη συστημική πολυπλοκότητα, όπου μια κλίμακα είναι ουσιαστικά ο Νόμος του Moore - ο κλασικός - περισσότερα τρανζίστορ σε ένα κύκλωμα", δήλωσε ο Aart de Geus, πρόεδρος και ένας από τους CEO της Synopsys. "Τώρα έχετε πολλά κυκλώματα, πολλά συστήματα, πολλά περιβάλλοντα λογισμικού που αλληλεπιδρούν όλα μεταξύ τους, οπότε βρισκόμαστε βαθιά στη συστημική πολυπλοκότητα. Το ίδιο το γεγονός ότι η ίδια η πολυπλοκότητα του συστήματος είναι ιδιαίτερα κατάλληλη για προσεγγίσεις τεχνητής νοημοσύνης - επειδή δεν είναι λογική τύπου σωστής / λανθασμένης απάντησης σε πολλά πράγματα, μοιάζει περισσότερο με το "δείτε τα μοτίβα" - είναι επίσης μια πρόκληση από την άποψη της ασφάλειας. Όλα αυτά είναι βήματα προόδου που φέρνουν τις δικές τους προκλήσεις". Η επίθεση κατανεμημένης άρνησης υπηρεσίας (Distributed Denial of Service - DDoS), με το λογισμικό Mirai τον Οκτώβριο του 2016, έδωσε μια εικόνα στο πόσο διαδεδομένη μπορεί πολύ να επηρεάσει η επίθεση. Χρησιμοποιώντας ένα botnet, τρεις φοιτητές κατάφεραν να μολύνουν αρκετές εκατοντάδες χιλιάδες συσκευές σε όλο τον κόσμο και να τις χρησιμοποιήσουν για να υπερφορτώσουν ένα μεγάλο κεντρικό κομμάτι του Διαδικτύου.

Ενώ η μηχανική μάθηση έχει χρησιμοποιηθεί ευρέως σε εφαρμογές άμυνας συστήματος, [1], τα μοντέλα μηχανικής μάθησης επιφέρουν κινδύνους ασφαλείας, οι οποίοι περιγράφονται ως εξής: Η μηχανική μάθηση μπορεί να χρησιμοποιηθεί από τους αντιπάλους για να κάνουν επιθέσεις.

Επίσης οι εισβολείς μπορούν να επιτεθούν στα μοντέλα μηχανικής μάθησης. Οι επιτιθέμενοι μπορούν να παραπλανήσουν τη διαδικασία ταξινόμησης με έγχυση επιθέσεων που θεωρούνται σκόπιμα ασφαλείς, ενώ στην πραγματικότητα είναι κακόβουλες.

Η ακρίβεια δεν είναι 100%, πράγμα που σημαίνει ότι εξακολουθούν να υπάρχουν εσφαλμένα ταξινομημένα σημεία δεδομένων στο σύνολο υπό δοκιμή. Επιπλέον, ακόμη και αν η αναφερόμενη ακρίβεια ήταν στο 100% με βάση ένα συγκεκριμένο σύνολο υπό δοκιμή, δεν είναι εγγυημένο ότι η ακρίβεια θα ήταν 100% σε πραγματικές συνθήκες. Όταν η μηχανική μάθηση εφαρμόζεται σε ρεαλιστικά σενάρια, η εσφαλμένη ταξινόμηση μιας μεμονωμένης επίθεσης αρκεί για να εκμεταλλευτεί ολόκληρο το σύστημα. Ενώ η μηχανική μάθηση είναι

μια πολλά υποσχόμενη προσέγγιση για την ασφάλεια υλικού, αν η ακρίβεια ανίχνευσης και ο πραγματικός θετικός ρυθμός δεν είναι 100%, το σύστημα κινδυνεύει από σοβαρούς κινδύνους. Οι περιπτώσεις που επισημαίνονται ως κανονικές θα πρέπει να αντιμετωπίζονται ως ύποπτες, δηλαδή, αυτές οι περιπτώσεις μπορεί να ταξινομηθούν εσφαλμένα ως κανονικές ενώ στην πραγματικότητα είναι επιθέσεις.

6. Συμπεράσματα, Τάσεις και Ερευνητικές Προτάσεις

Οι αλγόριθμοι μηχανικής μάθησης υπόσχονται πολλά ως προς την αντιμετώπιση των απειλών υλικού επιπέδου κυκλώματος. Επειδή όμως τα προβλήματα υλικού σε υψηλότερα επίπεδα είναι παρόμοιας λογικής με εκείνα στο επίπεδο κυκλώματος, συμπεραίνουμε ότι οι πληροφορίες που προκύπτουν από την έρευνα επιπέδου κυκλώματος με μηχανική μάθηση θα είναι χρήσιμες και στην επίλυση προβλημάτων ασφάλειας υλικού σε υψηλότερα επίπεδα.

Προτείνεται η διερεύνηση του τρόπου με τον οποίο η μηχανική εκμάθηση μπορεί να χρησιμοποιηθεί από επιτιθέμενους για αντίστροφη μηχανική, ανάλυση πλευρικού καναλιού για την αποσύνθεση των προγραμμάτων που εκτελούνται σε πολύπλοκες αρχιτεκτονικές πολλαπλών πυρήνων και την αξιοποίηση του υδατογραφήματος σε IC και δακτυλικών αποτυπωμάτων σε IC.

Η μελλοντική έρευνα μπορεί επίσης να επικεντρωθεί στην ανάπτυξη πιο ισχυρών μοντέλων που μπορούν να εφαρμοστούν αποτελεσματικά σε πραγματικά συστήματα. Για παράδειγμα, η ισχύς των αλγορίθμων βαθιάς μάθησης σε εφαρμογές ασφαλείας μπορεί να διερευνηθεί περαιτέρω. Τέτοιες μελέτες πρέπει να λαμβάνουν υπόψη ότι οι αλγόριθμοι βαθιάς μάθησης απαιτούν τεράστιο όγκο δεδομένων.

Επίσης, η ασφάλεια των μοντέλων μηχανικής μάθησης έναντι των επιθέσεων είναι ένα πολλά υποσχόμενο ερευνητικό θέμα. Οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται για την εκκίνηση ορισμένων επιθέσεων σε υλικό καθώς και για την προστασία του υλικού από άλλες επιθέσεις. Αυτές οι τάσεις στις επιθέσεις που βασίζονται στη μηχανική μάθηση θα πρέπει να λαμβάνονται υπόψη προσεκτικά κατά το σχεδιασμό αντιμέτρων άμυνας.

Η εφαρμογή μεθόδων μηχανικής μάθησης στις δοκιμές λογικής αξίζει περαιτέρω μελέτης, διότι προς το παρόν δεν υπάρχει σχετική εργασία.

Η έρευνα σχετικά με προβλήματα υλικού σε επιταχυντές και υπολογιστικές μονάδες που βασίζονται σε ANN δεν έχει και πρέπει να εξεταστεί. Αυτό οφείλεται κυρίως στο γεγονός ότι ως μηχανές επιτάχυνσης, τα ANN έχουν σημειώσει σημαντική πρόοδο σε πολλούς τομείς εφαρμογών και ο σχεδιασμός του επιταχυντή υλικού IP που βασίζεται σε ANN είναι μια σημαντική τάση για το μέλλον. Επομένως, δεν μπορεί να παραβλεφθεί η ασφάλειά του.

Πρόσθετες μελέτες σχετικά με την ασφαλή αρχιτεκτονική, ειδικά εκείνες που εκμεταλλεύονται τεχνολογίες μηχανικής μάθησης για την ανάπτυξη πιο ισχυρών

αρχιτεκτονικών με γνώμονα την ασφάλεια για τα σύγχρονα SoCs για να αμύνονται σε απειλές υλικού σε επίπεδο διαύλου και SoC παρά σε επίπεδο IP, παραμένουν προκλήσεις για μελέτη.

Προς το παρόν, οι προσεγγίσεις που βασίζονται στη μηχανική μάθηση επικεντρώνονται κυρίως στην ανίχνευση, τη διάγνωση και την πρόληψη των επιθέσεων υλικού (και κυρίως trojan), ενώ υπάρχουν λίγες μελέτες για τη διόρθωση και την αποκατάσταση σφαλμάτων. Επομένως, η έρευνα σχετικά με τον τρόπο εξερεύνησης μεθόδων μηχανικής μάθησης για τη διόρθωση σφαλμάτων και ανάκτηση είναι επίσης μια ενδιαφέρουσα μελλοντική κατεύθυνση.

Οι πρόσφατες προσεγγίσεις που βασίζονται σε μηχανική μάθηση στοχεύουν σε προβλήματα υλικού στο επίπεδο κυκλώματος. Ωστόσο, αρκετές πρόσφατες εξελίξεις έχουν δείξει ότι οι απειλές υλικού ενδέχεται να απειλήσουν την ασφάλεια και την αξιοπιστία του υλικού και σε υψηλότερα επίπεδα. Ως εκ τούτου, η εξερεύνηση αντιμέτρων για απειλές υλικού υψηλού επιπέδου μέσω μηχανικής μάθησης μπορεί επίσης να είναι μια ενδιαφέρουσα μελλοντική κατεύθυνση.

Διερευνήθηκαν αλγόριθμοι βελτιστοποίησης για να βελτιώσουν την αξιοπιστία των ροών σχεδιασμού IC. Ως αποτέλεσμα, τέτοιες συζητήσεις θα μπορούσαν να χρησιμοποιηθούν ως αναφορές για τη δημιουργία μιας ισχυρής αλυσίδας εφοδιασμού IC ή ακόμα και ενός οικοσυστήματος υλικού.

Υπάρχει ελάχιστη έρευνα περί εφαρμογής ML σε πραγματικές καταστάσεις. Οι μελλοντικές τάσεις και οι ερευνητικές κατευθύνσεις μπορεί να εξετάσουν και τα προβλήματα υλικού με πρακτικά σενάρια για τη δημιουργία σχετικών στρατηγικών άμυνας. Για παράδειγμα, ζητήματα trojan υλικού σε εμφυτεύσιμα ιατρικά βοηθήματα, ή απειλές υλικού που αντιμετωπίζουν έξυπνα δίκτυα και δίκτυα οικιακής περιοχής.

Οι περισσότεροι αλγόριθμοι μηχανικής μάθησης επικεντρώνονται κυρίως στη συλλογή δεδομένων και στον πειραματισμό ενώ πολύ λίγη έρευνα υπάρχει για τα αποτελέσματα της εφαρμογής.

Η ασφάλεια των ίδιων των μοντέλων μηχανικής μάθησης έναντι διαφόρων απειλών θα πρέπει να λαμβάνεται υπόψη κατά τη δημιουργία στρατηγικών υπεράσπισης υλικού.

Αντί να επικεντρώνεται στην άμυνα υλικού, το πώς θα μπορούσε να αξιοποιηθεί η μηχανική μάθηση από τους επιτιθέμενους αρχίζει να προσελκύει την προσοχή όλο και περισσότερων ερευνητών.

Η μεγαλύτερη διείσδυση των τεχνολογιών τεχνητής νοημοσύνης, μηχανικής μάθησης και βαθιάς μάθησης, είναι γεγονός που θα συμβάλει στην εμφάνιση νέων δυνατοτήτων για συστήματα παρακολούθησης, αλλά θα βοηθήσει επίσης στην καλύτερη χρήση των υπαρχουσών τεχνολογιών, όπως η ανάλυση βίντεο και ανάλυση δεδομένων. Η μεγάλη πρόκληση για τις εταιρείες που υλοποιούν μεγάλα έργα ασφαλείας, αλλά και γενικότερα για τους επαγγελματίες συστημάτων ασφαλείας, είναι να προσπαθήσουμε να μετατρέψουμε όλες αυτές τις δυνατότητες και αλλαγές σε σημαντικά οφέλη για τους χρήστες του συστήματος. Μόνο με αυτόν τον τρόπο θα επιταχυνθεί ο ρυθμός ενσωμάτωσης αυτών των τεχνολογιών στα συστήματα παρακολούθησης, καθώς η εμπειρία έχει δείξει ότι οι απαιτήσεις της αγοράς είναι αυτές που καθορίζουν σε μεγάλο βαθμό την εξέλιξη των συστημάτων και την ολοκλήρωση των νέων τεχνολογιών. Όλα τα παραπάνω σε συνδυασμό με την έρευνα για την ασφάλεια σε υλικό και λογισμικό που γίνεται μέχρι και σήμερα, θα συνεχίσουν να επηρεάζουν όλους τους τομείς.

7. Αναφορές

[1] Elnaggar, R., Chakrabarty, K., "Machine Learning for Hardware Security: Opportunities and Risks". J Electron Test 34, 183–201 (2018): https://link.springer.com/article/10.1007/s10836-018-5726-9

[2] Sperling E., "Security Holes In Machine Learning And AI", Semiconductor Engineering:https://semiengineering.com/security-holes-in-machine-learning-and-ai/

[3] Z. Huang, Q. Wang, Y. Chen and X. Jiang, "A Survey on Machine Learning Against Hardware Trojan Attacks: Recent Advances and Challenges," in IEEE Access, vol. 8, pp. 10796-10826, 2020.

[4] Huang, Zhao & Wang, Quan & Chen, Yin & Jiang, Xiaohong. (2020). "A Survey on Machine Learning against Hardware Trojan Attacks: Recent Advances and Challenges." IEEE Access. PP. 1-1: https://www.researchgate.net/publication/338467645_A_Survey_on_Machine_Learning_against_Hardware_Trojan_Attacks_Recent_Advances_and_Challenges

[5] Li, H., Q. Liu, Jiliang Zhang and Y. Lyu. "A Survey of Hardware Trojan Detection, Diagnosis and Prevention.", 2015 14th International Conference on Computer-Aided Design and Computer Graphics (CAD/Graphics) (2015): 173-180: https://www.semanticscholar.org/paper/A-Survey-of-Hardware-Trojan-Detection%2C-Diagnosis-Li-Liu/1755770e538b3dcc60bd0fd1697a768a380d9804

[6] Bizarro P., "Machine learning innovations for fighting financial crime in an Open Banking era", IT Magazine: https://itmagazineme.com/index.php/2019/09/21/machine-learning-innovations-for-fighting-financial-crime-in-an-open-banking-era/

[7] Dong, Chen, J. Chen, Wenzhong Guo and J. Zou. "A machine-learning-based hardware-Trojan detection approach for chips in the Internet of Things.", International Journal of Distributed Sensor Networks 15 (2019): https://www.semanticscholar.org/paper/A-machine-learning-based-hardware-Trojan-detection-Dong-Chen/67114eee5f3b1ae858e60c948819f60abee783c4

[8] Williams M., "Engineers develop methods for AI bottlenecks with machine-learning algorithms", Control Engineering, June 2020: https://www.controleng.com/articles/engineers-develop-methods-for-ai-bottlenecks-with-machine-learning-algorithms/

[9] Harris B., "Factorization Machines: A New Way of Looking at Machine Learning", Security Intelligence, November 2015: https://securityintelligence.com/factorization-machines-a-new-way-of-looking-at-machine-learning/

[10] Walsh S., "Exabeam to Use Machine Learning to Tackle IoT Device Security", RT Insights, March 2018: https://www.rtinsights.com/exabeam-to-use-machine-learning-to-tackle-iot-device-security/

[11] Theodoridis S., "Machine Learning: A Bayesian and Optimization Perspective" (1st. ed.). Academic Press, Inc., USA, 2015.

[12] Wikipedia, "Data mining": https://en.wikipedia.org/wiki/Data_mining

[13] "Knowledge Discovery in Databases", August 2017: https://www.techopedia.com/definition/25827/knowledge-discovery-in-databases-kdd

[14] Wikipedia, "Artificial Intelligence": https://en.wikipedia.org/wiki/Artificial_intelligence

[15] Guru 99, "Supervised vs Unsupervised Learning: Key Differences": https://www.guru99.com/supervised-vs-unsupervised-learning.html

[16] Brownlee J., "A Tour of Machine Learning Algorithms", Machine Learning Mastery, August 2019: https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/

[17] Zamani, M. "Machine Learning Techniques for Intrusion Detection", 2013.

[18] Bettaieb, Seifeddine, Seung Yeob Shin, M. Sabetzadeh, L. Briand, Grégory Nou and Michael Garceau, 2019: "Decision Support for Security-Control Identification Using Machine Learning".

[19] Dickson B., "How IoT security can benefit from machine learning", Techrunch, April 2016: https://techcrunch.com/2016/04/22/how-iot-security-can-benefit-from-machine-learning/

[20] Akinsola J. E. T., "Supervised Machine Learning Algorithms: Classification and Comparison". International Journal of Computer Trends and Technology (2017). 48. 128 - 138.

[21] Software Testing Help, "Types Of Machine Learning: Supervised Vs Unsupervised Learning", September 2020: https://www.softwaretestinghelp.com/types-of-machine-learning-supervised-unsupervised/

[22] Taranenko L., "Unsupervised Machine Learning to improve data quality": https://mobidev.biz/blog/unsupervised-machine-learning-improve-data-quality

[23] Wikipedia, "Dimensionality reduction": https://en.wikipedia.org/wiki/Dimensionality_reduction#Dimension_reduction

[24] Readthedocs.io, "Benchmarking Performance and Scaling of Python Clustering Algorithms": https://hdbscan.readthedocs.io/en/latest/performance_and_scalability.html#comparison-of-fast-implementations

[25] Wikipedia, "Anomaly detection": https://en.wikipedia.org/wiki/Anomaly_detection

[26] Wikipedia, "Association rule learning": https://en.wikipedia.org/wiki/Association_rule_learning

[27] Fu, C., Wang, X., Zhang, L., and Qiao, L., “Mining algorithm for association rules in big data based on Hadoop”, in "Advances in Materials, Machinery, Electronics II", 2018, vol. 1955, no. 1.

[28] Khadse V., Mahalle P. N., Biraris, S. V. "An Empirical Comparison of Supervised Machine Learning Algorithms for Internet of Things Data", 2018 Fourth International Conference on Computing Communication Control and Automation, Pune, India, 2018, pp. 1-6.

[29] Hassan C. A. Ul, Khan M. S., Shah M. A., "Comparison of Machine Learning Algorithms in Data classification", 2018 24th International Conference on Automation and Computing, Newcastle upon Tyne, United Kingdom, 2018, pp. 1-6.

[30] De Hoz Diego J. D., Saldana J., Fernández-Navajas J., Ruiz-Mas J., "IoTsafe, Decoupling Security From Applications for a Safer IoT," IEEE, vol. 7, pp. 29942-29962, 2019.

[31] "IoT Developer Survey", Eclipse IoT Working Group, IEEE IoT, AGILE IoT. (2017): https://ianskerrett.wordpress.com/2017/04/19/iot-developer-trends-2017-edition/

[32] "IoT Developer Survey", Eclipse IoT Working Group, IEEE IoT, AGILE IoT. (2016): https://iot.ieee.org/images/files/pdf/iot-developer-survey-2016-report-final.pdf

[33] Becker G. T., "The gap between promise and reality: On the insecurity of XOR arbiter PUFs," in Proc. Conf. Cryptograph. Hardw. Embedded Syst. (CHES). Saint Malo, France, Sep. 2015, pp. 535–555.

[34] Bellman C., Oorschot P., "Analysis, Implications, and Challenges of an Evolving Consumer IoT Security Landscape", School of Computer Science, Carleton University, Ottawa, Canada, 2019, pp.1-7.

[35] R. Sandip, B. Abhishek, B. Swarup, "To Secure the Internet of Things, We Must Build It Out of "Patchable" Hardware", IEEE Sptectrum, October 2017: https://spectrum.ieee.org/telecom/security/to-secure-the-internet-of-things-we-must-build-it-out-of-patchable-hardware

[36] R. Sagar, R. Jhaveri, C. Borrego, "Applications in Security and Evasions in Machine Learning: A Survey", 2020. Electronics. 9. 97.

[37] Huang Z., Wang Q., Chen Y., Jiang X., Xiaohong, "A Survey on Machine Learning Against Hardware Trojan Attacks: Recent Advances and Challenges". IEEE Access, 2020, PP. 1-31.

[38] P. Prinetto, and G. Roascio, "Hardware Security, Vulnerabilities, and Attacks: A Comprehensive Taxonomy." ITASEC (2020).

[39] Wikipedia, Spectre: https://en.wikipedia.org/wiki/Spectre_(security_vulnerability)

[40] Wikipedia, Meltdown: https://en.wikipedia.org/wiki/Meltdown_(security_vulnerability)

[41] D. Forte, "Design and Fabrication of Physically Unclonable Functions (PUFs), True Random Number Generators (TRNGS), and Other Hardware Security Primitives", Florida & FICS Research: https://dforte.ece.ufl.edu/research/hardware-security-primitives/