2023-05

þÿ ¤ µ Ç ½ · Ä ® ½ ¿ · ¼ ¿ Ã Í ½ · º ± ¹ · » µ º

þÿ œ À ¹ ½ - º · , ‘ ¸ ± ½ ± Ã ¯ ±

þÿ œ µ Ä ± À Ä Å Ç ¹ ± º Ì Ã Ä ± » · Á ¿ Æ ¿ Á ¹ ± º ¬ £ Å Ã Ä ® ¼ ± Ä ± º ± ¹ ¨ · Æ ¹ ± º ® š ± ¹ ½ ¿ Ä ¿ ¼ ¯ ± ,
þÿ º ± ¹ • À ¹ Ã Ä ® ¼ · Â ¥ À ¿ » ¿ ³ ¹ Ã Ä Î ½ , ± ½ µ À ¹ Ã Ä ® ¼ ¹ ¿ • µ ¬ À ¿ » ¹ Â ¬ Æ ¿ Å

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΦΟΥ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΑ ΠΛΗΡΟΦΟΡΙΚΑ ΣΥΣΤΗΜΑΤΑ ΚΑΙ ΨΗΦΙΑΚΗ ΚΑΙΝΟΤΟΜΙΑ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΟ ΕΓΚΛΗΜΑ

ΜΠΙΝΕΚΗ ΑΘΑΝΑΣΙΑ

ΑΜ:1218613093

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ :

ΖΑΧΑΡΙΟΥΔΑΚΗΣ ΕΛΕΥΘΕΡΙΟΣ

ΠΑΦΟΣ , ΜΑΙΟΣ 2023

Abstract

Η συγκεκριμένη μεταπτυχιακή εργασία πραγματοποιήθηκε στα πλαίσια του Προγράμματος Μεταπτυχιακών σπουδών «Πληροφοριακά Συστήματα και Ψηφιακή Καινοτομία» του Πανεπιστημίου Νεάπολης Πάφου και πραγματεύεται στην ανάλυση και επεξήγηση της χρήσης της τεχνητής νοημοσύνης στα ηλεκτρονικά εγκλήματα. Η εξέλιξη της τεχνολογίας έχει συμβάλει σε κορυφαίες αλλαγές οι οποίες βιώνει ο ανθρώπινος πληθυσμός σε καθημερινούς τομείς της ζωής. Πρόκειται για μια διαδεδομένη μέθοδο αναπόσπαστου κινδύνου και επιθέσεων σε οργανισμούς , εταιρίες ακόμα και κυβερνήσεις . Η αυξημένη χρήση του διαδικτύου έχει ενισχύσει μη νόμιμες δράσεις καθώς είναι εύχρηστη και παρέχει πληθώρα πληροφοριών χωρίς χρονικό περιθώριο. Στην παρούσα εργασία εκτός από τους κινδύνους αναφέρονται και εξηγούνται οι λύσεις και οι εφαρμογές που μπορούν να χρησιμοποιηθούν έχοντας την τεχνητή νοημοσύνη και την τεχνολογία ως συνοδοιπόρους για την εξάλειψη των κυβερνοεπιθέσεων.

Abstract

This thesis was carried out within the framework of the Master's Degree Programme "Information Systems and Digital Innovation" of the University of Naples Paphos and deals with the analysis and explanation of the use of artificial intelligence in electronic crimes. The evolution of technology has contributed to leading changes that the human population is experiencing in everyday areas of life. It is a widespread method of integral risk and attacks on organizations, companies and even governments. The increased use of the internet has enhanced non legal actions as it is easy to use and provides a wealth of information without time constraints. In this paper apart from the risks, the solutions and applications that can be used having artificial intelligence and technology as companions to eliminate cyber attacks are mentioned and explained.

Λέξεις Κλειδιά: Κυβερνοεγλήματα, Επιθέσεις στο διαδίκτυο , εφαρμογές τεχνητής νοημοσύνης , εισβολές

Keywords: Cyber attacks, cyber attacks, artificial intelligence applications, cyber intrusions.

Ευχαριστίες :

Με την ολοκλήρωση της Διπλωματικής μου Εργασίας θα ήθελα να ευχαριστήσω τον αρμόδιο καθηγητή Ζαχαριουδάκη Ελευθέριο για την πολύτιμη βοήθεια του στον αγώνα της εκπόνησης της εργασίας .

Στην συνέχεια, ιδιαίτερα θα ήθελα να αναφέρω την σημαντική βοήθεια της οικογένειας μου στην πορεία ολόκληρης της διαδικασίας με την ενθάρρυνση , την έμπρακτη συμπαράσταση και την υποστήριξη τους όλο το διάστημα του προγράμματος σπουδών .

Thanks to:

With the completion of my Thesis I would like to thank the responsible professor Zacharioudakis Eleftherios for his valuable help in the struggle of preparing the thesis.

Then, I would particularly like to mention the significant help of my family in the course of the whole process with their encouragement, practical support and support throughout the study program.

Περιεχόμενα

Κεφάλαιο 1ο Εισαγωγή

1.1 Ιστορικό σχετικά με το θέμα της τεχνητής νοημοσύνης και του εγκλήματος στον κυβερνοχώρο

Τα τελευταία χρόνια, το ηλεκτρονικό έγκλημα είναι μια κερδοφόρα επιχείρηση των $1,5Τ με ένα ολόκληρο σύστημα οργανισμών που παρουσιάζονται ως νόμιμοι οργανισμοί. Κάποιοι από αυτούς δίνουν ηγεσία σε τεχνικό επίπεδο και κατευθυντήριες γραμμές ανά στάδιο μέσα από την ενίσχυση της εξυπηρέτησης πελατών μέσω του λυτρισμικού (ransomware) ως υπηρεσία. Οι δράστες με θράσος προβάλουν διαφημίσεις ως αναδυόμενα παράθυρα για την προώθηση των προϊόντων τους. Ωστόσο, ενώ η βιομηχανία του εγκλήματος στον κυβερνοχώρο έχει παρουσιάσει αύξηση τα τελευταία δέκα έτη, η πραγματικότητα είναι ότι το ηλεκτρονικό έγκλημα δεν αποτελεί καινούρια απειλή (Brenner, 2007). Η αλήθεια είναι ότι έχει παρουσιαστεί ως απειλή εδώ και πολλά χρόνια ή κι αιώνες.

Η πρώτη φορά που εμφανίστηκε η ηλεκτρονική επίθεση αυτού του είδους ήταν στη Γαλλία το 1834, ενώ ακόμα δεν υπήρχε το διαδίκτυο. Οι δράστες απέσπασαν πληροφορίες χρηματοπιστωτικής αγοράς αφού είχαν παραβιάσει το γαλλικό τηλεγραφικό σύστημα. Από τότε, το έγκλημα στον κυβερνοχώρο έχει παρουσιάσει σημαντική αύξηση, με εξελισσόμενα χαρακτηριστικά από τακτικές, τεχνικές και διάφορες διαδικασίες για κέρδος εις βάρος άλλων (Schell & Martin, 2004).

Ωστόσο, το ηλεκτρονικό έγκλημα δεν κατοχυρώθηκε πραγματικά έως τα μέσα του 20ού αιώνα. Οι δράστες του ηλεκτρονικού εγκλήματος έχοντας παρακίνηση από την ψηφιακή επανάσταση, απέκτησαν σύντομα την τεχνολογία, χρησιμοποιώντας το πλεονέκτημά τους και την ευστροφία τους για να κατασκευάσουν καινούριους και πανούργους τρόπους για να βλάψουν ανθρώπους και οργανισμούς με σκοπό προσωπικά δεδομένα και χρηματικά ποσά. Εάν υπήρχε ένα μουσείο της ύβρεως για τα εγκλήματα στον κυβερνοχώρο, οι αίθουσές του θα ήταν γεμάτες με τα ονόματα και τα πρόσωπα αυτών των επιτιθέμενων, των οποίων το «πρωτοποριακό» έργο προσέλκυσε τόσο την προσοχή των ομοσπονδιακών

ερευνητών όσο και τον φθόνο των συναδέλφων χάκερ (Schjolberg, 2020).

Η σύγχρονη ιστορία του ηλεκτρονικού εγκλήματος στο διαδίκτυο ξεκίνησε το 1962 όταν ο Allen Scherr επιτέθηκε διαδικτυακά εναντίον των δικτύων υπολογιστών του MIT, κλέβοντας κωδικούς πρόσβασης από τη βάση δεδομένων τους μέσω μιας διάτρητης κάρτας. Δημιουργήθηκε ο πρώτος ιός υπολογιστών για ερευνητικούς σκοπούς από τον BobThomas στο BBNtechnology το 1971. Αναφερόμενος ως ο ιός Creeper, το πρόγραμμα αυτοαναπαραγωγής εντοπίστηκε στο ARPANET το 1971 και αποτέλεσε τον προάγγελο ως προς τη δυνατότητα μελλοντικών ιών να προκαλέσουν σημαντική ζημιά στα συστήματα υπολογιστών (Babanina et al., 2021).

Ο Ian Murphy, το 1981, έγινε ο πρώτος άνθρωπος που καταδικάστηκε για διάπραξη εγκλήματος στον κυβερνοχώρο μετά από επιτυχή επίθεση στα εσωτερικά συστήματα της AT&T και την αλλαγή των ρολογιών των υπολογιστών τους, προκαλώντας ζημιά (Brenner, 2007).

Η πρώτη μεγάλη κυβερνοεπίθεση στο Διαδίκτυο το 1988 προήλθε από τον φοιτητή του Cornell, RobertMorris. Το «MorrisWorm» εμφανίστηκε το έτος πριν από το ντεμπούτο του Παγκόσμιου Ιστού, όταν το Διαδίκτυο ήταν κατά κύριο λόγο ένας τομέας των ακαδημαϊκών ερευνητών. Μόλυνε συστήματα υπολογιστών στα πανεπιστήμια Στάνφορντ, Πρίνστον, Τζονς Χόπκινς, NASA, LawrenceLivermoreLabs και UCBerkeley, μεταξύ άλλων ιδρυμάτων (Brenner, 2007).

Κατά τη δεκαετία του 1990 παρουσιάστηκε η εξέλιξη σε πολύ σημαντικές τεχνολογίες επικοινωνίας και έδωσε τη δυνατότητα παγκόσμιας σύνδεσης μεταξύ των ανθρώπων μέσω του διαδικτύου. Παρόλα αυτά, το ηλεκτρονικό έγκλημα μέσω διαδικτύου παρουσίασε αύξηση και ενισχύθηκε βασιζόμενο σε αυτές τις τεχνολογίες παγκόσμιας επικοινωνίας. Οι χάκερ και διάφοροι εγκληματίες του διαδικτύου εκμεταλλευτήκαν αυτό το γεγονός και κατά την εξέλιξη και τη δημιουργία αυτών των σύγχρονων τεχνολογιών, οι έλεγχοι εμπιστοσύνης και ασφάλειας δεν αποτελούσαν αρχικά σημαντική σκέψη (Schell & Martin, 2004).

Η κυβερνοασφάλεια ήταν ένας όρος που δεν είχε επινοηθεί ακόμη, πόσο μάλλον ένα ενεργό πεδίο, επομένως η δημιουργία πρωτοποριακών εφαρμογών για επικοινωνίες και επιχειρηματική αποτελεσματικότητα ήταν το επίκεντρο αυτών των

πρώτων ημερών (Brenner, 2007). Ωστόσο, ήδη αναπτυσσόταν μια παραοικονομία με αργό ρυθμό. Η αύξηση των ποσοστών του εγκλήματος στον κυβερνοχώρο έδειχνε ότι οι δράστες είχαν αρκετές ευκαιρίες και ανακάλυψαν νέους τρόπους για να αποκτήσουν μη εξουσιοδοτημένη πρόσβαση σε συστήματα και χειρισμό δεδομένων σε όλο το διαδίκτυο. Παρακάτω παρουσιάζονται μερικά από τα πιο αξιοσημείωτα ηλεκτρονικά εγκλήματα αυτής της δεκαετίας:

Data stream Cowboy και Kuji - ένας 16χρονος Βρετανός μαθητής το 1994 και ο συνεργός του χρησιμοποίησαν ένα πρόγραμμα ανίχνευσης κωδικού πρόσβασης για να εξαπολύσουν μια σειρά επιθέσεων που έπληξαν το Εργαστήριο Ρώμης της Πολεμικής Αεροπορίας, με κλοπή ερευνητικών δεδομένων που χρησιμοποιήθηκαν ως οδηγίες επίθεσης για πολεμικά αεροσκάφη σε μάχη (Schjolberg, 2020).

Ο Vladimir Levin ήταν ο πρώτος γνωστός χάκερ το 1995 που επιχείρησε να ληστέψει μια τράπεζα και μάλιστα μια πολύ μεγάλη τράπεζα. Χάκαρε το δίκτυο της Citibank και πραγματοποίησε πληθώρα δόλιων συναλλαγών. Συνολικά, μετέφερε περισσότερα από 10 εκατομμύρια δολάρια σε διάφορους τραπεζικούς λογαριασμούς παγκοσμίως. Επίσης, την ίδια χρονολογία ο Kevin Mitnick — ένας από τους πλέον διαβόητους χάκερ στην ιστορία — έγινε ο πρώτος άνθρωπος που διείσδυσε σε μεγάλα δίκτυα χειραγωγώντας ανθρώπους και χρησιμοποιώντας ανθρώπους εκ των έσω για να αποσπάσει τους κωδικούς πρόσβασης στη Motorola και τη Nokia, μεταξύ άλλων (Buch et al., 2017).

Ο Max Butler, σύμβουλος ασφαλείας του FBI, μεταξύ άλλων, εισέβαλε σε ιστότοπους της κυβέρνησης των ΗΠΑ με ψευδή προσχήματα το 1998. Η Πολεμική Αεροπορία των ΗΠΑ ενημέρωσε τους αξιωματούχους για τα παραπτώματα του και του επιβλήθηκε ποινή 18 μηνών (Schjolberg, 2020). Αργότερα, για άλλη μια παράνομη επίθεση, καταδικάστηκε σε 13 χρόνια, ρεκόρ για έναν χάκερ. Το 1999 οι ιοί υπολογιστών ήταν σχετικά άγνωστοι στο ευρύ κοινό έως ότου εμφανίστηκε ο ιός Melissa τον Μάρτιο του 1999 ο οποίος επηρέασε τους χρήστες σε ολόκληρο το Διαδίκτυο, καταστρέφοντας τα αρχεία εγγράφων της Microsoft και προκαλώντας ζημιές περίπου 80 εκατομμυρίων δολαρίων.

Η νέα χιλιετία έφερε σημαντικές εξελίξεις στις ηλεκτρονικές επιθέσεις και ένα σύνολο παραγόντων προηγμένης επίμονης απειλής (APT), οι περισσότεροι από

τους οποίους έλαβαν χρηματοδότηση από έθνη-κράτη. Η εξέλιξη του ηλεκτρονικού εγκλήματος έφερε νέους ιούς και ιούς τύπου worm, που ήταν ζημιογόνοι για την παγκόσμια και ψηφιακή οικονομία. Στα τέλη της δεκαετίας, η κυβερνοασφάλεια υπήρξε σημαντική ανησυχία για τους χρήστες ηλεκτρονικών υπολογιστών παντού, ιδιαίτερα για τις κρατικές υπηρεσίες και τις μεγάλες εταιρείες που κινδύνευαν πιο πολύ (Li, 2017). Παρακάτω παρουσιάζονται μερικά από τα πιο αξιοσημείωτα ηλεκτρονικά εγκλήματα αυτή τη χρονική περίοδο:

Ο Michael Calse, 15 χρονών χάκερ με το το προσωνύμιο «Mafiaboy» πραγματοποίησε το 2000 ένα σύνολο επιθέσεων άρνησης υπηρεσίας (DDoS) σε μερικούς από τις μεγαλύτερες ιστοσελίδες εμπορίου παγκοσμίως, όπως Amazon, Yahoo, CNN και eBay. Κατάφερε να καταστρέψει τους ιστότοπους για κάποιες ώρες, γεγονός που έφερε την αναστάτωση, αλλά και την οικονομική καταστροφή πολλών εκατομμυρίων σε αυτούς τους παγκόσμιους οργανισμούς. Το 2005 μια παραβίαση ασφαλείας σε λιανοπωλητή στις ΗΠΑ οδήγησε στη διαρροή δεδομένων 1,4 εκατομμυρίων χρηστών της HSBC Bank Master Card. Μία ακόμα επίθεση πραγματοποιήθηκε το 2008 στα συστήματα πληρωμών Heartland. Στο συγκεκριμένο έγκλημα οι δράστες χρησιμοποίησαν μία σύνθεση από SQL, ανίχνευσης κωδικών πρόσβασης και κακόβουλου λογισμικού, διακινδυνεύοντας προσωπικά δεδομένα 134 εκατομμυρίων χρηστών (ALI et al., 2018).

Την επόμενη δεκαετία παρατηρήθηκε σημαντική αύξηση στο ηλεκτρονικό έγκλημα, τροποποιώντας το έγκλημα μικρού βαθμού σε μία μεγάλη βιομηχανία. Οι δράστες δημιούργησαν προγράμματα και τεχνικές που μπορούν να βλάψουν την ηλεκτρονική ασφάλεια και με αυτό τον τρόπο αυξήθηκε και το ηλεκτρονικό έγκλημα, αλλά και ο ημερήσιος αριθμός επιθέσεων και έτσι τριπλασιάστηκε ο αριθμός των χρηματικών ποσών που χάθηκαν κατά τη διάρκεια αυτών των επιθέσεων. Το έγκλημα στον κυβερνοχώρο δεν ήταν ο μόνος κλάδος που γνώρισε τεράστια ανάπτυξη. Αυτό οδήγησε τις επιχειρήσεις να προσλάβουν μεγαλύτερο αριθμό ανθρώπινου δυναμικού για την προστασία και την ασφάλεια τους και για να είναι ικανές να έρθουν αντιμέτωπες με τους κινδύνους, τις απειλές του διαδικτύου, μίας και δεν υπήρχε πλέον η αίσθηση ψηφιακής ασφάλειας. Εξαιτίας, λοιπόν, αυτών των ενεργειών, εμφανίστηκε ένας καινούριος τομέας απασχόλησης, αυτός

του ηθικού χακαρίσματος, που έχει σκοπό να ανακαλύψει τα αδύναμα σημεία πριν από μια κακόβουλη επίθεση. Η εξέλιξη και η αυξημένη πολυπλοκότητα των διαφορετικών τύπων απειλών στον κυβερνοχώρο και ο τρόπος με τον οποίο αξιοποιούνται σε επιθέσεις έθεσε τους οργανισμούς σε επισφαλείς θέσεις όσον αφορά την άμυνα εναντίον τους. Ακολουθούν οι πλέον σοβαρές επιθέσεις από αυτήν την πιο καταστροφική δεκαετία:

Το 2010 ο ιός τύπου worm Stuxnet – αποκαλούμενο ως το πρώτο «ψηφιακό όπλο» στον κόσμο - εισέβαλε σε πυρηνικά εργοστάσια στο Ιράν, σαμποτάροντας τις εγκαταστάσεις εμπλουτισμού ουρανίου της χώρας. Επίσης, ο ιός Zeus Trojan διανεμήθηκε σε όλον τον κόσμο μέσω email σε μια επίθεση που στόχευε οργανισμούς χρηματοοικονομικών υπηρεσιών. Το κύκλωμα εγκλήματος 100 και πλέον ατόμων, που εδρεύει σε μεγάλο βαθμό στις ΗΠΑ, κατάφερε να κλέψει περισσότερα από 70 εκατομμύρια δολάρια από αμερικανικές τράπεζες. Στη συνέχεια, σε μια περιβόητη επίθεση έθνους-κράτους, η Επιχείρηση Aurora ξεκίνησε από Κινέζους στρατιωτικούς χάκερ σε περισσότερες από 20 κορυφαίες εταιρείες τεχνολογίας. Το κοινό ενημερώθηκε για πρώτη φορά για τις επιθέσεις όταν η Google ενημέρωσε το κοινό ότι η πνευματική ιδιοκτησία του είχε καταληφθεί από την επίθεση (Brenner, 2007).

Η Sony Corporation ανακοίνωσε τον Απρίλιο του 2011 ότι, μέσα σε λίγες μέρες, χάκερ έκλεψαν πληροφορίες από 77 εκατομμύρια χρήστες του δικτύου PlayStation. Αυτό περιλάμβανε τα ονόματα χρήστη και τους κωδικούς πρόσβασης των παικτών, τις ημερομηνίες γέννησής τους, απαντήσεις σε ερωτήσεις ασφαλείας και πολλά άλλα. Χρειάστηκαν 23 ημέρες για να ανακτηθεί το σύστημα και να αποκατασταθεί η απειλή (Schjolberg,2020).

Στη μεγαλύτερη ίσως διαρροή δεδομένων υψηλού προφίλ όλων των εποχών, ο πληροφοριοδότης EdwardSnowdento 2013 αποκάλυψε ευαίσθητες πληροφορίες που είχαν κλαπεί από πολλές ξένες κυβερνήσεις με λογισμικό υποκλοπής (spyware) ως μέρος του προγράμματος παρακολούθησης PRISM της Υπηρεσίας Εθνικής Ασφάλειας. Τα αρχεία της πιστωτικής κάρτας σε πάνω από 110 εκατομμύρια πελάτες της Target κλάπηκαν την ίδια χρονιά σε μια επίθεση phishing. Το πρόγραμμα περιλάμβανε ένα μήνυμα ηλεκτρονικού ταχυδρομείου με

κακόβουλο λογισμικό στον υπεργολάβο HVAC της εταιρείας, επιτρέποντας στους εγκληματίες του κυβερνοχώρου να αποκτήσουν διαπιστευτήρια πρόσβασης στα δεδομένα. Επίσης, ένας ερευνητής ανακάλυψε ότι η φινλανδική εταιρεία τηλεπικοινωνιών Nokia διενεργούσε ουσιαστικά επιθέσεις ενδιάμεσης οντότητας (man-in-the-middle) στους χρήστες έξυπνων τηλεφώνων της στέλνοντας κίνηση HTTP μέσω των διακομιστών της και αποκρυπτογραφώντας δεδομένα. Η εταιρεία είπε ότι το έκανε για να βοηθήσει στη συμπίεση δεδομένων και να διατηρήσει μειωμένες τιμές και χρεώσεις.

Το 2015 εμφανίστηκαν τα πρώτα στελέχη του λυτρισμικού SamSam, τα οποία μέχρι το 2018 είχαν αποφέρει στον δημιουργό του σχεδόν 6 εκατομμύρια δολάρια. Μεταξύ των πιο δημοφιλών στόχων «ομηρίας» ήταν η πόλη της Ατλάντα και το Υπουργείο Μεταφορών του Κολοράντο. Επιπλέον, μια επιτυχημένη επίθεση spear-phishing εναντίον στόχων του Υπουργείου Άμυνας υψηλής αξίας με προσαρμοσμένα μηνύματα ηλεκτρονικού ταχυδρομείου οδήγησε σε παραβίαση δεδομένων για 4.000 στρατιωτικούς και πολιτικό προσωπικό που εργάζονταν για το Γενικό Επιτελείο. Η επίθεση ανάγκασε το Πεντάγωνο να κλείσει το σύστημα ηλεκτρονικού ταχυδρομείου του (Babaninaetal., 2021).

Την επόμενη χρονιά εμφανίστηκε το λυτρισμικό TeleCrypt το οποίο είχε ως στόχο παίκτες, οι οποίοι το κατέβαζαν ενώ έπαιζαν παιχνίδια στο διαδίκτυο. Ευτυχώς, δημιουργήθηκε γρήγορα ένα δωρεάν εργαλείο αποκρυπτογράφησης από ερευνητές στο Malwarebytes. Επίσης, από την αυστριακή εταιρεία αεροδιαστημικής, FACCAG, εκλάπησαν 50 εκατομμύρια ευρώ με ένα πρόγραμμα στοχευόμενου ψαρέματος (spear-phishing) ξεγελώντας έναν οικονομικό υπάλληλο ώστε να μεταφέρει τα χρήματα σε τραπεζικούς λογαριασμούς που ελέγχονται από τους εγκληματίες του κυβερνοχώρου. Ως αποτέλεσμα, απολύθηκε ο Διευθύνων Σύμβουλος της εταιρείας (Babanina et al., 2021).

Ίσως το πιο ύπουλο λογισμικό από όλα τα στελέχη λυτρισμικού, ο ιός WannaCry, κατάφερε να επηρεάσει το 2017 περισσότερους από 200.000 υπολογιστές με Windows σε 150 χώρες. Ήταν ιδιαίτερα επικίνδυνο - και θανατηφόρο - καθώς επλήγησαν ακόμη και τα νοσοκομεία της Εθνικής Υπηρεσίας Υγείας του Ηνωμένου Βασιλείου. Εικάζεται ευρέως ότι πίσω από την επίθεση

βρίσκονται χάκερ από τη Βόρεια Κορέα. Στη συνέχεια, μόλις ένα μήνα μετά την επιτυχία του WannaCry, εμφανίστηκε το NotPetya, μια ενημερωμένη έκδοση του προηγούμενου στελέχους λυτρισμικού. Έπληξε οργανισμούς από τον ναυτιλιακό γίγαντα Maersk μέχρι την πολυεθνική φαρμακευτική εταιρεία Merck (Schjolberg, 2020). Ακόμα, ένας Λιθουανός κυβερνοεγκληματίας υποδύθηκε έναν Ασιάτη κατασκευαστής για να εξαπατήσει τους υπαλλήλους της Google και του Facebook για να στείλουν περισσότερα από 100 εκατομμύρια δολάρια σε μη ανιχνεύσιμους υπεράκτιους τραπεζικούς λογαριασμούς. Η απάτη συνέβη κατά τη διάρκεια δύο ετών πριν από τη σύλληψή του. Από την πλευρά της, η Google ισχυρίστηκε ότι είχε ανακτήσει τα χρήματα που είχε χάσει.

Στη μεγαλύτερη επίθεση DDoS μέχρι σήμερα, η GitHub — μια δημοφιλής πλατφόρμα προγραμματιστών — εμφάνισε κίνηση 1,3 terabyte ανά δευτερόλεπτο, γεγονός που σταμάτησε όλες τις λειτουργίες στον διακομιστή του το 2018. Η GitHub είχε θεσπίσει μέτρα ασφαλείας, πολύ περισσότερα από ότι οι περισσότεροι οργανισμοί, αλλά συγκλονίστηκε από το τεράστιο μέγεθος της επίθεσης. Ίσως η πιο αξιοσημείωτη από όλες τις επιθέσεις κακόβουλης εξόρυξης κρυπτονομισμάτων ήταν το Coinhive τη χρονιά του 2018, μια δημοφιλής υπηρεσία εξόρυξης κρυπτονομισμάτων που, για κάποιο διάστημα, θεωρήθηκε από τις κορυφαίες εταιρείες ασφαλείας ως η κορυφαία κακόβουλη απειλή για τους χρήστες του Ιστού. Ο κώδικας υπολογιστή του θα μπορούσε να χρησιμοποιηθεί σε ιστότοπους που έχουν παραβιαστεί για την κλοπή της επεξεργαστικής ισχύς των συσκευών των επισκεπτών αυτού του ιστότοπου. Για 15 μήνες, οι εγκληματίες του κυβερνοχώρου χρησιμοποιούσαν το κακόβουλο πρόγραμμα για να μολύνουν εκατομμύρια συσκευές (Konieczny,2023).

Στη συνέχεια, η τραπεζική εταιρεία Capital One έπεσε θύμα επίθεσης το 2019. Εκεί πραγματοποιήθηκε μία από τις πιο σοβαρές παραβιάσεις δεδομένων στον τραπεζικό κλάδο, στην οποία οι δράστες απόκτησαν παρανόμως πρόσβαση σε περισσότερες από 100 εκατομμύρια αιτήσεις πιστωτικών καρτών και λήφθηκαν χιλιάδες αριθμοί κοινωνικής ασφάλισης και τραπεζικών λογαριασμών. Η τραπεζική εταιρεία αναγκάστηκε να ξοδέψει πάνω από 150 εκατομμύρια δολάρια για να μετριάσει τη ζημιά που υπέστη.

Από το 2020 έως σήμερα έχουν χαθεί δισεκατομμύρια δολάρια. Ο Neiman Marcus ενημέρωσε 4,6 εκατομμύρια πελάτες ότι ένας χάκερ είχε παραβιάσει διαδικτυακούς λογαριασμούς τον Μάιο του 2020, αποκτώντας πρόσβαση σε προσωπικά δεδομένα όπως ονόματα χρήστη και κωδικούς πρόσβασης, ονόματα πελατών, στοιχεία επικοινωνίας, αριθμούς πιστωτικών καρτών, καθώς και ημερομηνίες λήξης και αριθμούς εικονικών καρτών (Li, 2017). Οι ρωσικές κυβερνοεπιθέσεις σε κυβερνητικούς οργανισμούς των ΗΠΑ έχουν αυξηθεί και, σε μια από τις πιο καταστροφικές παραβιάσεις δεδομένων κατά τη διάρκεια του 2020, ξένοι πράκτορες πληροφοριών εκμεταλλεύτηκαν ένα παραβιασμένο πρόγραμμα SolarWinds και εισέβαλαν σε περίπου 18.000 ιδιωτικά και κυβερνητικά δίκτυα. Αυτές οι παραβιάσεις δεδομένων παρείχαν στους εισβολείς πρόσβαση σε πληθώρα αναγνωρίσιμων πληροφοριών, συμπεριλαμβανομένων οικονομικών πληροφοριών, πηγαίο κώδικα, κώδικες πρόσβασης και ονόματα χρηστών.

Στις αρχές Μαΐου του 2021, μια ύποπτη ρωσική ομάδα χάκερ έθεσε το Colonial Pipeline εκτός σύνδεσης για περισσότερες από τρεις ημέρες σε μια επίθεση που έκανε το λυτρισμικό δημοφιλές. Καθώς η Colonial παρέχει το 45% της προμήθειας της Ανατολικής Ακτής σε βενζίνη, καύσιμο ντίζελ και καύσιμα αεροσκαφών, αυτή η παραβίασε αποτέλεσε ένα σημαντικό πλήγμα. Οι τιμές του φυσικού αερίου εκτοξεύτηκαν σε ολόκληρη τη χώρα, ορισμένα πρατήρια τελείωσαν από καύσιμα, οι παραδόσεις στο δρόμο καθυστέρησαν και υπήρξαν ακόμη και αναφορές για συγκέντρωση βενζίνης από άτομα και πρατήρια. Επιπλέον, η διαβόητη συλλογικότητα REvil χτύπησε τον πάροχο λογισμικού Kaseya με έδρα τη Φλόριντα με μια επίθεση λυτρισμικού, απαιτώντας 70 εκατομμύρια δολάρια σε bit coin. Αυτή η επίθεση επηρέασε επιχειρήσεις σε πέντε ηπείρους - συμπεριλαμβανομένου του κλείσιμο των δημόσιων σχολείων στη Νέα Ζηλανδία, το κλείσιμο μιας μεγάλης αλυσίδας παντοπωλείων στη Σουηδία και τη διακοπή λειτουργίας εκατοντάδων επιχειρήσεων στις Η.Π.Α..Η χρονιά έκλεισε με την αποκάλυψη μιας απειλής ημέρας 0 (zero-day) που δημιούργησε τεράστια προβλήματα στον κλάδο της κυβερνοασφάλειας, όταν οι ερευνητές ασφάλειας δημοσίευσαν μια κρίσιμη εκμετάλλευση που επαληθεύει την ιδέα σε μια ευπάθεια απομακρυσμένης εκτέλεσης κώδικα στο Log4j, μια βιβλιοθήκη καταγραφής Java

που χρησιμοποιείται από σημαντικό αριθμό διαδικτυακών εφαρμογών. Το διάστημα που ακολούθησε, οι οργανισμοί σε παγκόσμιο επίπεδο έκανα ενέργειες σε μανιώδους ρυθμούς για να επιτύχουν τον εντοπισμό και την ελάττωση του αντίκτυπου της εκμετάλλευσης, ενώ ειδικοί σε θέματα ασφαλείας πρότειναν διάφορες ενημερώσεις κώδικα και εργαλεία σάρωσης και έδιναν οδηγίες και συμβουλές σχετικά με την ηλεκτρονική ασφάλεια και την προστασία από τέτοιου είδους εγκλήματα (Konieczny, 2023).

Σε μια από τις πιο τρομακτικές εκδηλώσεις της προθυμίας των εγκληματιών του κυβερνοχώρου να θέσουν σε κίνδυνο τις ζωές και τα μέσα διαβίωσης αγνώστων, η υπηρεσία που διαχειρίζεται την κοινωνική ασφάλιση για την Κόστα Ρίκα έκλεισε από μια επίθεση λυτρισμικού στα τέλη Μαΐου του 2022, μια επίθεση που εξαπλώθηκε σε άλλα γραφεία στη χώρα και προκάλεσε κατάσταση έκτακτης ανάγκης. Επίσης, μια διείσδυση στα μέσα Σεπτεμβρίου απέδωσε μια εντυπωσιακή ποσότητα υλικού από έναν τιτάνα της βιομηχανίας τυχερών παιχνιδιών. Η πολύ αναμενόμενη κυκλοφορία του Grand Theft Auto 6 της Rock star Games επλήγη όταν ένας χάκερ γνωστός ως «teapotuberhacker» παραβίασε το εσωτερικό κανάλι Slack της Rock star και έκλεισε 90 βίντεο με το game play σε εξέλιξη. Αλλά η επίθεση δεν τελείωσε εκεί.

Ο Slack ή αλλιώς teapotuberhacker στις 14 Σεπτεμβρίου πραγματοποίησε σχεδόν ίδια επίθεση κατά της εταιρείας Uber. Η παγκόσμια εταιρεία μεταφορών κοινοχρησίας δέχτηκε πιο σημαντική επίθεση από τη Rock star, με τον δράστη να καταφέρνει να έχει πρόσβαση σε όλα τα συστήματα της, όπως τα email, τις εσωτερικές επικοινωνίες, την αποθήκευση σε cloud και τα αποθετήρια κωδικών.

Οι εγκληματίες του διαδικτύου στις μέρες μας χρησιμοποιούν την ίδια εξελιγμένη τεχνολογία που χρησιμοποιείται για την ασφάλεια στον κυβερνοχώρο, όπως τα εργαλεία μηχανικής μάθησης και τεχνητής νοημοσύνης. Με αποτέλεσμα να παρουσιάζεται σημαντική πρόκληση στο ποιος θα έχει το πλεονέκτημα. Επίσης, σημειώνεται ότι οι παράγοντες απειλών έχουν αρχίσει τη συνεργασία μεταξύ τους σε ένα μοντέλο λυτρισμικού ως υπηρεσία (RaaS) για τη διείσδυση σε οργανισμούς και όλο αυτό υποδηλώνει ότι η βιομηχανία του ηλεκτρονικού εγκλήματος συνεχώς μεγαλώνει και εξελίσσεται (Konieczny, 2023).

Το μοντέλο RaaS επιτρέπει στους προγραμματιστές μιας παραλλαγής λυτρισμικού τη στρατολόγηση συνεργατών που χρησιμοποιούν αποκλειστικά το λυτρισμικό τους σε στοχευμένες επιθέσεις σε οργανισμούς. Τυχόν πληρωμές λύτρων από τα εκβιαζόμενα θύματα στη συνέχεια κατανέμονται μεταξύ των προγραμματιστών του λυτρισμικού και της θυγατρικής που διεξήγαγε την επίθεση. Δεν μπορούμε να πούμε τι νέες, δόλιες καινοτομίες μας επιφυλάσσουν οι παράγοντες απειλής για την επόμενη δεκαετία. Η προετοιμασία για την επόμενη γενιά του εγκλήματος στον κυβερνοχώρο απαιτεί από τους χρήστες να επαγρυπνούν σχετικά με τους τύπους επιθέσεων που υπάρχουν και να κατανοήσουν ξεκάθαρα τους τρόπους με τους οποίους μπορούν να αμυνθούν έναντι αυτών των απειλών (Meland, Bayoumy & Sindre, 2020).

Οι επιτυχείς προσεγγίσεις για την ασφάλεια στον κυβερνοχώρο θα περιλαμβάνουν πολλαπλές άμυνες. Και θα περιλαμβάνουν παρόχους υπηρεσιών και τεχνογνωσία τρίτων, ακόμη και για οργανισμούς αρκετά μεγάλους ώστε να χρησιμοποιούν ολοκληρωμένη τεχνολογία κυβερνοασφάλειας και εξειδικευμένο προσωπικό στο εσωτερικό. Αυτό που χρειάζεται κάθε οργανισμός για την καταπολέμηση του εγκλήματος στον κυβερνοχώρο σήμερα, όπως και αύριο, είναι οι επιχειρήσεις ασφαλείας.

1.2 Σκοπός και στόχοι της μελέτης

Η χρήση του Διαδικτύου έχει αυξηθεί ραγδαία, ιδιαίτερα την τελευταία δεκαετία. Με την ευρεία χρήση του Διαδικτύου, το έγκλημα στον κυβερνοχώρο αυξάνεται με ανησυχητικό ρυθμό στην καθημερινή μας ζωή. Ωστόσο, με την ανάπτυξη της τεχνητής νοημοσύνης (AI), οι επιχειρήσεις επικεντρώνονται περισσότερο στην πρόληψη του εγκλήματος στον κυβερνοχώρο. Η τεχνητή νοημοσύνη γίνεται ουσιαστική συνιστώσα κάθε επιχείρησης, επηρεάζοντας τα άτομα σε όλο τον κόσμο. Το έγκλημα στον κυβερνοχώρο είναι ένας από τους πιο εξέχοντες τομείς όπου η τεχνητή νοημοσύνη έχει αρχίσει να αποδεικνύει πολύτιμες εισροές. Ως αποτέλεσμα, η τεχνητή νοημοσύνη αναπτύσσεται ως η πρώτη γραμμή άμυνας στα συστήματα των περισσότερων εταιρειών. Επειδή η τεχνητή νοημοσύνη μπορεί να ανιχνεύσει νέες επιθέσεις ταχύτερα από τους ανθρώπους, είναι η

καλύτερη εναλλακτική λύση για την κατασκευή καλύτερης προστασίας από το έγκλημα στον κυβερνοχώρο. Οι τεχνολογίες τεχνητής νοημοσύνης προσφέρουν επίσης πιο σημαντικές δυνατότητες στην ανάπτυξη μιας τέτοιας τεχνολογίας. Αυτό το έγγραφο συζητά τις πρόσφατες κυβερνοεπιθέσεις και πώς η βιομηχανία που επιτρέπει την τεχνητή νοημοσύνη προετοιμάζεται να υπερασπιστεί τον εαυτό της μακροπρόθεσμα.

Με βάση τη συστηματική ανασκόπηση η παρούσα μελέτη θα προσδιορίσει τον όρο της τεχνητής νοημοσύνης, τα χαρακτηριστικά και τις εφαρμογές της. Επίσης, θα παρουσιαστεί η σύνδεση της τεχνητής νοημοσύνης με το ηλεκτρονικό έγκλημα και τις επιπτώσεις αυτής της σχέσης. Τέλος, με την διεξαγωγή της συστηματικής ανασκόπησης θα παρουσιαστούν τα συμπεράσματα και οι προτάσεις για το μέλλον.

1.3 Αντικείμενο της μελέτης

Το αντικείμενο της μελέτης είναι να προσδιορίσουμε την εφαρμογή της τεχνητής νοημοσύνης στο κυβερνοέγκλημα, αλλά και τη χρήση της για την προστασία από τους δράστες του ηλεκτρονικού εγκλήματος.

Χρησιμοποιώντας την τεχνητή νοημοσύνη, οι εγκληματίες του κυβερνοχώρου μπορούν να παραμείνουν αδρανείς και μη ανιχνεύσιμοι στο δίκτυο μιας εταιρείας για παρατεταμένες περιόδους, κατά τη διάρκεια των οποίων μπορούν να δημιουργήσουν πίσω πόρτες στην κρίσιμη υποδομή ενός οργανισμού. Στη συνέχεια, όταν είναι έτοιμοι να ξεκινήσουν μια επίθεση εναντίον της ευρύτερης επιχείρησης, μπορούν να ακούσουν σε συσκέψεις, να εξάγουν δεδομένα, να διαδίδουν κακόβουλο λογισμικό, να δημιουργήσουν προνομιούχους λογαριασμούς για να έχουν πρόσβαση σε άλλα συστήματα ή/και να εγκαταστήσουν ransomware. Η τεχνητή νοημοσύνη είναι ένα ιδιαίτερα αποτελεσματικό εργαλείο για τους εγκληματίες του κυβερνοχώρου λόγω της ικανότητάς της να μαθαίνει και να προβλέπει τι συμβαίνει τώρα και τι μπορεί να συμβεί στο μέλλον.

Αντίθετα, ο αντίκτυπος της τεχνητής νοημοσύνης στην κυβερνοασφάλεια γίνεται γρήγορα ένα σημαντικό θέμα καθώς οι οργανισμοί σε όλο τον κόσμο

αρχίζουν τον αγώνα για την υιοθέτηση της τεχνολογίας AI στα προϊόντα τους, τα επιχειρηματικά μοντέλα ή τα προγράμματα ασφαλείας. Η τεχνητή νοημοσύνη αναδύεται γρήγορα ως ένα πεδίο που έχει τη δυνατότητα να φέρει επανάσταση στον τομέα της κυβερνοασφάλειας. Ωστόσο, η χρήση της τεχνητής νοημοσύνης στην κυβερνοασφάλεια φέρνει νέες προκλήσεις και κινδύνους όσο και νέες και καινοτόμες λύσεις. Ως εκ τούτου, είναι σημαντικό να κατανοήσουμε τόσο τις ευκαιρίες όσο και τους περιορισμούς της τεχνητής νοημοσύνης στην κυβερνοασφάλεια, ώστε να μπορέσουμε να την χρησιμοποιήσουμε με υπεύθυνο και ηθικό τρόπο.

1.4 Ερευνητικά ερωτήματα και στόχοι

Με βάση την ανάλυση που προηγήθηκε θα γίνει προσπάθεια στην παρούσα ενότητα να τεθούν τα ερευνητικά ερωτήματα, τα οποία ουσιαστικά δηλώνουν και το κενό της παρούσας μελέτης. Τα ερευνητικά ερωτήματα που θα εξεταστούν είναι τα ακόλουθα:

1. Ποιες είναι οι απειλές και τα μοντέλα επιθέσεων που χρησιμοποιούν οι χάκερς για να υπονομεύσουν ένα σύστημα υπολογιστή σε συστήματα; Linux, Windows, Android, Ios, IoT κτλ;
2. Ποιες είναι οι συμβατικές προσεγγίσεις και οι μέθοδοι μετριασμού των κινδύνων από τις επιθέσεις στον κυβερνοχώρο σε συστήματα; Linux, Windows, Android, Ios, IoT κτλ ;
3. Ποιες είναι οι προσεγγίσεις που βασίζονται στην τεχνητή νοημοσύνη προκειμένου να αντιμετωπιστούν οι απειλές στον κυβερνοχώρο ή τουλάχιστον να μετριαστούν οι κίνδυνοι που σχετίζονται με τις επιθέσεις σε αυτόν;
4. Ποιο είναι το μελλοντικό πεδίο εφαρμογής της τεχνητής νοημοσύνης στην κυβερνο-ασφάλεια σε global scale;
5. Ποιες οι ασφαλιστικές δικλίδες ως προς την χρησιμοποίηση τεχνικών τεχνητής νοημοσύνης για την ασφάλεια ενός οργανισμού π.χ ενός νοσοκομείου ή/και μιας μονάδας υψίστης ασφαλείας στρατού.

1.5 Σημασία της μελέτης

Η σημαντικότητα της παρούσας μελέτης αναφέρεται στο ότι τα τελευταία χρόνια υπάρχει μεγάλη έξαρση του ηλεκτρονικού εγκλήματος, η ανάπτυξη του διαδικτύου σε όλους τους τομείς έχει ενισχύσει και τις μη νόμιμες δράσεις φέρνοντας σε δύσκολη θέση τους χρήστες. Οπότε η αναζήτηση τρόπων αντιμετώπισης καθίσταται επιτακτική και η παρούσα μελέτη θα επιδιώξει να εξετάσει πιθανές λύσεις στο θέμα κάνοντας ειδική αναφορά στην τεχνητή νοημοσύνη και στο ρόλο που μπορεί να παίξει για την αντιμετώπιση του. Οπότε η σχέση της τεχνητής νοημοσύνης και της κυβερνοασφάλειας καθίσταται ερευνητικά σημαντική.

Κεφάλαιο 2ο Βιβλιογραφική ανασκόπηση

2.1 Επιθέσεις και απειλές στο κυβερνοχώρο

Στην εποχή μας είναι δύσκολο να βρεθεί μια εταιρεία, ίδρυμα ή οικογένεια που να μην χρησιμοποιεί το διαδίκτυο και την τεχνολογία. Η ανθρωπότητα έχει κατακλυστεί από τον αριθμό των ψηφιακών συσκευών και εφαρμογών που χρησιμοποιούνται σε καθημερινή βάση, σε σημείο που δεν υπάρχει πλέον έλεγχος στη χρήση της τεχνολογίας με τον τρόπο που θα ήθελαν. Μερικοί άνθρωποι μπορεί να είναι εθισμένοι στο διαδίκτυο και να μην μπορούν να σταματήσουν να το χρησιμοποιούν, ενώ άλλοι μπορεί να μην το χρησιμοποιούν αρκετά ώστε να συμβαδίζουν με τις γρήγορες αλλαγές στην τεχνολογία. Κάποιοι μπορεί να περνούν ώρες στα τηλέφωνα ή τους υπολογιστές τους αντί να αλληλεπιδρούν με άλλους. Το διαδίκτυο είναι ένα ισχυρό εργαλείο και δεν μπορούμε να αντισταθούμε στην παρόρμηση να το χρησιμοποιούμε για τα πάντα (Al-Yaseen, Othman&Nazri, 2017).

Δεν υπάρχει αμφιβολία ότι η τεχνολογία έχει αυξήσει την παραγωγικότητα και την αποδοτικότητά μας με πολλούς τρόπους. Ωστόσο, πρέπει επίσης να εξετάσουμε το αποτύπωμα που είχε στην προσωπική και κοινωνική μας ζωή και στην ψυχική και σωματική μας ευεξία. Παράλληλα με τη φωτεινή πλευρά της εξάπλωσης των συνδεδεμένων στο Διαδίκτυο συσκευών, μια πιο σκοτεινή πλευρά του εγκλήματος στον κυβερνοχώρο έχει επισκιάσει τις ζωές μας. Σε αυτόν τον ανοιχτό και συνδεδεμένο κόσμο, η διαρκής απειλή της απώλειας της ιδιωτικής ζωής εγείρει πολλά ερωτηματικά. Αυτή η ραγδαία αύξηση της συνδεσιμότητας μας έχει καταστήσει πιο αποτελεσματικούς, αλλά και πιο ανοιχτούς στους κινδύνους του ηλεκτρονικού εγκλήματος στον κυβερνοχώρο (Faruketal., 2021).

Πρέπει να έχουμε επίγνωση της σύνδεσής μας με τον κυβερνοχώρο. Το διαδίκτυο μας έχει μετατρέψει σε κατακτητές και αιχμαλώτους ταυτόχρονα. Το Διαδίκτυο έχει δώσει στους ανθρώπους τη δυνατότητα να συνδέονται με πρωτοφανείς τρόπους. Ωστόσο, αυτό έχει επίσης δημιουργήσει τρωτά σημεία που μπορούν να εκμεταλλευτούν οι εγκληματίες του κυβερνοχώρου. Το έγκλημα στον κυβερνοχώρο είναι ένας τύπος εγκλήματος που χρησιμοποιεί ψηφιακά μέσα για τη διάπραξη απάτης, κλοπής δεδομένων ή πρόκλησης βλάβης. Το έγκλημα στον κυβερνοχώρο είναι ένας γενικός όρος που περιλαμβάνει όλες τις μορφές

εγκλημάτων που σχετίζονται με τον κυβερνοχώρο. Στην ουσία, παράνομες δραστηριότητες που ξεκινούν με τη χρήση υπολογιστών, όπως η δίκτυο παραβίαση (hacking), το ηλεκτρονικό «ψάρεμα»(phishing), η διανομή κακόβουλου λογισμικού, η διαδικτυακή παρακολούθηση και η κλοπή ταυτότητας, μεταξύ πολλών άλλων. Το κυβερνο-έγκλημα είναι ένα από τα πιο προσοδοφόρα εγκλήματα της σύγχρονης εποχής. Κάθε χρόνο, οι κυβερνο-εγκληματίες αποκομίζουν κέρδη δισεκατομμυρίων δολαρίων κλέβοντας δεδομένα, αλλοιώνοντας δεδομένα και θέτουν σε κίνδυνο κρίσιμες υποδομές. Όπως και κάθε άλλο έγκλημα, το κυβερνοέγκλημα έχει εξελιχθεί δραματικά με την πάροδο του χρόνου και θα συνεχίσει να το κάνει (Jajodia&Wu, 2001).

Η έκθεση του Παγκόσμιου Οικονομικού Φόρουμ αναφέρει ότι μια ολιστική προσέγγιση της ασφάλειας στο διαδίκτυο θα παρέχει καλύτερη προστασία για τις επιχειρήσεις, την κοινωνία, τις κυβερνήσεις και τους ιδιώτες. Η έκθεση αναφέρει επίσης ότι υπάρχει σημαντικό χάσμα μεταξύ της τωρινής κατάστασης ετοιμότητας και των αναγκών για την προστασία στο διαδίκτυο. Το κενό αυτό πρέπει να καλυφθεί επειγόντως πριν φτάσουμε σε σημείο χωρίς επιστροφή.

Η ασφάλεια στο διαδίκτυο είναι μια βασικό στοιχείο για οποιαδήποτε υποδομή για να μην διακυβεύεται από οποιοδήποτε είδος επίθεσης από εξωτερικούς παράγοντες, όπως ιούς, κακόβουλο λογισμικό ή απόπειρες επίθεσης από χάκερ. Εντούτοις, οι περισσότερες επιθέσεις στην ασφάλειας ή ηλεκτρονικά εγκλήματα είναι αποτέλεσμα ανθρώπινου χειριού και εξωτερικών παραγόντων. Η προστασία των σημαντικών υποδομών και η προστασία των ανθρώπων από το ηλεκτρονικό έγκλημα είναι κοινή υποχρέωση όλων (Bose, Barao, &Liu, 2020).

Αν και οι ιοί και το κακόβουλο λογισμικό υπήρχαν σαν έγνοια από την αρχή τις εφεύρεσης των ηλεκτρονικών υπολογιστών, η συνειδητοποίηση της σημαντικότητας της ασφάλειας των δεδομένων εμφανίστηκε μόνο όταν η χρήση του διαδικτύου έγινε δημοφιλής και έγινε εργαλείο που το χρησιμοποιούμε καθημερινά. Υπήρξε μια πρόσφατη αύξηση στο hacking και στην έκθεση αυτών των εγκληματιών του διαδικτύου με περισσότερες δυνατότητες στον ιστό. Έτσι κατανοούμε ότι υπάρχουν πολλές απειλές που προκύπτουν από αυτή την έκθεση. Οι χάκερ έχουν τη δυνατότητα να βλάψουν σε πολλά επίπεδα, όπως να διακόψουν τη λειτουργία

ιστοσελίδων, να παραβιάσουν προσωπικά δεδομένα και ακόμη και να πραγματοποιήσουν κλοπή χρημάτων. Αυτό είχε ως αποτέλεσμα να δημιουργηθεί ένας νέος κλάδος ποινικών ενεργειών που ονομάζεται έγκλημα στον κυβερνοχώρο. Αυτό έγινε από την καθολική διείσδυση του Διαδικτύου σε παγκόσμιο επίπεδο σε περίπου 5 δισεκατομμύρια χρήστες, που αποτελούν το 63% του πληθυσμού παγκοσμίως και το ηλεκτρονικό έγκλημα έχει παρουσιάσει αύξηση με τεράστιο ρυθμό ανάπτυξης (Hayward & Maas, 2021).

Σε διεθνές επίπεδο, οι υπολογιζόμενες ζημίες από το έγκλημα στον κυβερνοχώρο ήταν περίπου 6 τρισεκατομμύρια δολάρια ΗΠΑ μέχρι το 2021 και έχει μετατραπεί στην τρίτη μεγαλύτερη οικονομία στον κόσμο εάν μετρηθεί χωριστά (Coull&Gardner, 2019). Το έγκλημα στον κυβερνοχώρο αναμένεται να κοστίσει συνολικά 10,5 τρισεκατομμύρια δολάρια έως το 2025, από μόλις 3 τρισεκατομμύρια δολάρια το 2015 σύμφωνα με την Cyber security Ventures.

Αυτή είναι αναμφισβήτητα η πιο σημαντική μεταφορά πλούτου στην ανθρώπινη ιστορία (Dobrinoiu, 2019). Περιορίζει τα κίνητρα για καινοτομίες και θα είναι πιο κερδοφόρο από οτιδήποτε έχουμε δει ποτέ. Το κόστος του εγκλήματος στον κυβερνοχώρο μπορεί να περιλαμβάνει βλάβη ή καταστροφή δεδομένων, διαγραφή ή αποκατάσταση χακαρισμένων δεδομένων, κλοπή χρημάτων, μειωμένη παραγωγικότητα, κλοπή πνευματικής ιδιοκτησίας, κλοπή προσωπικών και οικονομικών δεδομένων και υπεξαίρεση (απόκτηση περιουσιακών στοιχείων για χρήση από το πιστωμένο άτομο που ευθύνεται για το έγκλημα) και βλάβη της φήμης. Η δημιουργία ενός πιο ασφαλούς συστήματος από την αρχή, η πρόληψη του εγκλήματος στον κυβερνοχώρο και η μείωση των επιπτώσεών του όταν συμβεί αποτελεί μια υπόθεση πολλών ειδικοτήτων (Hayward & Maas, 2021).

Το έγκλημα στον κυβερνοχώρο είναι ένα αυξανόμενο πρόβλημα και είναι απαραίτητο να προστατευτούμε από αυτό. Υπάρχουν πολλοί τρόποι για αυτό, αλλά το πιο σημαντικό είναι να γνωρίζουμε το περιβάλλον μας και τι κάνουμε στο διαδίκτυο με τα προσωπικά μας στοιχεία. Με το να γνωρίζουμε τις δραστηριότητές μας στο διαδίκτυο και τον κίνδυνο που μπορεί να φέρουν αυτές, ίσως μπορέσουμε να αποφύγουμε τις περισσότερες από τις απειλές που ελλοχεύουν στο διαδίκτυο.

Οι άνθρωποι μπορούν να προστατευτούν με ποικίλους τρόπους από το

ηλεκτρονικό έγκλημα. Ένας τρόπος είναι να χρησιμοποιούν λογισμικό προστασίας από ιούς, λογισμικό ασφάλειας στο διαδίκτυο και λογισμικό τείχους προστασίας για να διασφαλίσουν ότι οι συσκευές τους είναι ασφαλείς και προστατευμένες. Ένας άλλος τρόπος είναι να χρησιμοποιούν ισχυρούς κωδικούς πρόσβασης και να τους αλλάζουν τακτικά. Τέλος, μπορούν να διατηρούν τις ενημερώσεις του λειτουργικού συστήματος ενημερωμένες με τις τελευταίες εκδόσεις κώδικα. Μπορούν επίσης να παρακολουθούν την κυκλοφορία του δικτύου για ευπάθειες και να ρυθμίσουν ένα σύστημα αυτόματης απόκρισης για την αποτροπή επιθέσεων phishing (Schjolberg, 2020).

Επιπλέον, θα πρέπει να διαχειρίζονται τις ρυθμίσεις των μέσων κοινωνικής δικτύωσης και να αποφεύγουν τη χρήση μη ασφαλών δημόσιων δικτύων Wi-Fi. Η ελαχιστοποίηση του όγκου των προσωπικών πληροφοριών που μοιράζονται στο διαδίκτυο μπορεί να τους βοηθήσει να αποφύγουν τον κίνδυνο να γίνουν στόχος μιας σειράς απειλών, συμπεριλαμβανομένης της κλοπής προσωπικών στοιχειών και της διαδικτυακής παρακολούθησης. Παράλληλα με αυτούς τους παραδοσιακούς αποτρεπτικούς παράγοντες, η χρήση της τεχνητής νοημοσύνης γίνεται ένας τομέας που αποκτά όλο και μεγαλύτερη έμφαση στον κόσμο της κυβερνοασφάλειας.

Σήμερα, η τεχνητή νοημοσύνη είναι διαδεδομένη σχεδόν σε όλους τους τομείς της επιστήμης, από την ιατρική μέχρι τις επιχειρήσεις ή από τον στρατό μέχρι τις υπηρεσίες επιβολής του νόμου. Η χρήση της τεχνητής νοημοσύνης στην επιστήμη είναι σχεδόν πανταχού παρούσα. Η χρήση της τεχνητής νοημοσύνης στο έγκλημα στον κυβερνοχώρο αυξάνεται με τόσο γρήγορο ρυθμό που αποτελεί επίσης ένας από τους σημαντικούς τομείς ανησυχίας παγκοσμίως. Η ΤΝ είναι ένα ισχυρό εργαλείο που χρησιμοποιείται για την καταπολέμηση πολλών διαφορετικών τύπων εγκλήματος. Θα είναι ζωτικής σημασίας για τις υπηρεσίες επιβολής του νόμου παγκοσμίως να εντοπιστούν νέοι τρόποι χρήσης αυτής της τεχνολογίας για να συμβαδίσουν με το συνεχώς αυξανόμενο ποσοστό του εγκλήματος στον κυβερνοχώρο (Al-Masalha, Hnaif & Kanan, 2020). Η τεχνητή νοημοσύνη εφαρμόζεται στην καταπολέμηση του εγκλήματος με διάφορους τρόπους. Στην περίπτωση του εγκλήματος στον κυβερνοχώρο, η τεχνητή νοημοσύνη χρησιμοποιείται στον εντοπισμό πιθανών απειλών, στον εντοπισμό προτύπων που

μπορούν να οδηγήσουν σε προηγούμενη εγκληματική δραστηριότητα και στον εντοπισμό νέων μορφών υφιστάμενης εγκληματικής δραστηριότητας. Ωστόσο, η τεχνητή νοημοσύνη χρησιμοποιείται επίσης ως μέρος μιας ευρύτερης ερευνητικής πρωτοβουλίας για το έγκλημα στον κυβερνοχώρο και τους δράστες του. Τα δεδομένα για το έγκλημα στον κυβερνοχώρο συλλέγονται, αναλύονται και χρησιμοποιούνται για τη δημιουργία εξελιγμένων εικονικών σκηνών εγκλήματος που μπορούν να προβλέψουν εγκλήματα πριν συμβούν (Sukhodolov, Bychkov&Bychkova, 2020).

Η τεχνική νοημοσύνη μπορεί να χρησιμοποιηθεί για την εξόρυξη δεδομένων, τον εντοπισμό προτύπων και την πρόβλεψη μελλοντικών γεγονότων. Μπορεί επίσης να χρησιμοποιηθεί για τον εντοπισμό επιθέσεων στον κυβερνοχώρο και την αποτροπή τους από το να συμβούν.

Μελλοντικά, τα συστήματα τεχνητής νοημοσύνης θα είναι σε θέση να βρίσκουν μοτίβα που δεν είναι άμεσα προφανή για τους ανθρώπους, όπως πιθανές κυβερνοεπιθέσεις, αναλύοντας την κυκλοφορία του δικτύου και προσδιορίζοντας αν διαφορετικά σύνολα δεδομένων έχουν πρόσβαση στο ίδιο ασυνήθιστο μοτίβο. Η τεχνητή νοημοσύνη μπορεί να κάνει πολλά πράγματα και θα συνεχίσει να αναπτύσσεται και να ενισχύεται καθώς χρησιμοποιείται σε όλο και περισσότερους καθημερινούς τομείς στη ζωή των ανθρώπων (Radulov, 2019).

2.2 Επισκόπηση της τεχνητής νοημοσύνης και των εφαρμογών της

2.2.1 Επισκόπηση της τεχνητής νοημοσύνης

Η τεχνητή νοημοσύνη (ΤΝ) είναι ένα από τα σημαντικότερα παγκόσμια ζητήματα του 21ου αιώνα. Ο όρος "τεχνητή νοημοσύνη" (AI) επινοήθηκε το 1956 από τον John McCarthy κατά τη διάρκεια ενός συνεδρίου που πραγματοποιήθηκε για το θέμα αυτό. Η τεχνητή νοημοσύνη είναι ο κλάδος της επιστήμης των υπολογιστών που ασχολείται με τον σχεδιασμό ευφυών υπολογιστικών συστημάτων που μιμούνται την ανθρώπινη νοημοσύνη. Η ικανότητα των μηχανών να επεξεργάζονται τη φυσική γλώσσα, να μαθαίνουν, να σχεδιάζουν καθιστά

δυνατή την εκτέλεση νέων εργασιών από ευφυή συστήματα. Ο κύριος σκοπός της ΤΝ είναι να μιμηθεί τη γνωστική λειτουργία των ανθρώπων και να εκτελέσει δραστηριότητες που τυπικά θα εκτελούνταν από έναν άνθρωπο. Η ΤΝ είναι αυτόνομη ανεξάρτητη ηλεκτρονική οντότητα που λειτουργεί όπως ο ανθρώπινος εμπειρογνώμονας στον τομέα της υγειονομικής περίθαλψης. Σήμερα, η τεχνητή νοημοσύνη έχει ενσωματωθεί στην καθημερινή μας ζωή σε διάφορες μορφές, όπως οι προσωπικοί βοηθοί, τα αυτοματοποιημένα μέσα μαζικής μεταφοράς, οι αερομεταφορές, τα ηλεκτρονικά παιχνίδια, η αναγνώριση προσώπου στον έλεγχο διαβατηρίων, η αναγνώριση φωνής στους εικονικούς βοηθούς, τα αυτοκίνητα χωρίς οδηγό, τα ρομπότ-συνοδοιπόρους κ.λπ. Οι τεχνολογίες τεχνητής νοημοσύνης αποδίδουν όλο και καλύτερα στην ανάλυση δεδομένων (Mintz & Brodie, 2019).

Ένα σημαντικό χαρακτηριστικό της τεχνολογίας ΤΝ είναι ότι μπορεί να προστεθεί στις υπάρχουσες τεχνολογίες. Η τεχνητή νοημοσύνη έχει ωφελήσει πολλούς τομείς, όπως η χημεία και η ιατρική, όπου οι διαγνώσεις ρουτίνας ξεκινούν από υπολογιστές με τη βοήθεια της τεχνητής νοημοσύνης. Αγκαλιάζει ένα ευρύ φάσμα επιστημονικών κλάδων, όπως η επιστήμη των υπολογιστών, η μηχανική, η χημεία, η βιολογία, η φυσική, η αστρονομία, η νεύρο επιστήμη και οι κοινωνικές επιστήμες.

Η τεχνητή νοημοσύνη δεν είναι μια ενιαία τεχνολογία αλλά μια σειρά υπολογιστικών μοντέλων και αλγορίθμων. Οι κυριότεροι κλάδοι της ΤΝ περιλαμβάνουν τα συστήματα εμπειρογνωμόνων, την ασαφή λογική και τα τεχνητά νευρωνικά δίκτυα (ANN), τη μηχανική μάθηση, τη βαθιά μάθηση, την επεξεργασία φυσικής γλώσσας, την όραση υπολογιστών και τη ρομποτική. Τα διάφορα υπολογιστικά εργαλεία ή τεχνολογίες που έχουν χρησιμοποιηθεί για την επίτευξη των στόχων της ΤΝ είναι τα ακόλουθα (Mason, 2003):

Συστήματα εμπειρογνωμόνων: Ένα σύστημα εμπειρογνωμόνων (ή σύστημα βασισμένο στη γνώση) επιτρέπει στους υπολογιστές να λαμβάνουν αποφάσεις ερμηνεύοντας δεδομένα και επιλέγοντας μεταξύ εναλλακτικών λύσεων όπως ακριβώς θα έκανε ένας ανθρώπινος εμπειρογνώμονας. Χρησιμοποιεί μια τεχνική γνωστή ως συμπερασματολογία βασισμένη σε κανόνες, στην οποία χρησιμοποιούνται κανόνες για την επεξεργασία των δεδομένων.

Νευρωνικά δίκτυα: Αυτά τα προγράμματα υπολογιστών ταυτοποιούν αντικείμενα ή αναγνωρίζουν μοτίβα αφού εκπαιδευτούν. Τα τεχνητά νευρωνικά δίκτυα είναι παράλληλα κατανεμημένα συστήματα που αποτελούνται από μονάδες επεξεργασίας (νευρώνες) που υπολογίζουν κάποιες μαθηματικές συναρτήσεις. Το μοντέλο τεχνητών νευρωτικών δικτύων αναπαριστά μη γραμμικές σχέσεις οι οποίες μαθαίνονται άμεσα από τα δεδομένα που μοντελοποιούνται. Τα νευρωνικά δίκτυα διερευνώνται για εφαρμογές στην υγειονομική περίθαλψη στην απεικόνιση και διάγνωση, την ανάλυση κινδύνου, τη διαχείριση και παρακολούθηση του τρόπου ζωής, τη διαχείριση πληροφοριών υγείας και την εικονική βοήθεια υγείας.

Επεξεργαστές φυσικής γλώσσας (Natural Language Processors - NLP): Προγράμματα υπολογιστών που μεταφράζουν ή ερμηνεύουν τη γλώσσα όπως αυτή ομιλείται από φυσιολογικούς ανθρώπους. Οι τεχνικές NLP εξάγουν πληροφορίες από μη δομημένα δεδομένα, όπως κλινικές σημειώσεις, για να συμπληρώσουν και να εμπλουτίσουν δομημένα ιατρικά δεδομένα. Ένας NLP περιλαμβάνει εφαρμογές όπως η αναγνώριση ομιλίας, η ανάλυση κειμένου, η μετάφραση και άλλοι στόχοι που σχετίζονται με τη γλώσσα. Υπάρχουν δύο βασικές προσεγγίσεις στο NLP: η στατιστική και η σημασιολογική. Η υγειονομική περίθαλψη είναι ο μεγαλύτερος χρήστης των εργαλείων NLP (Sadiku, Zhou & Musa, 2018).

Ρομπότ: Προγραμματιζόμενες μηχανές που βασίζονται σε υπολογιστή και διαθέτουν φυσικούς χειριστές και αισθητήρες. Η εισαγωγή ευφυών ρομπότ στον τομέα της υγειονομικής περίθαλψης ενισχύει την ικανοποίηση των ασθενών, την ακρίβεια της διάγνωσης και τη λειτουργική αποτελεσματικότητα των νοσοκομείων. Τα ιατρικά ρομπότ μπορούν να βοηθήσουν στις χειρουργικές επεμβάσεις, στην αποκατάσταση, στην κοινωνική αλληλεπίδραση, στην υποβοηθούμενη διαβίωση κ.λπ. Η ρομποτική καθοδήγηση καθίσταται συνήθης στη χειρουργική της σπονδυλικής στήλης (Wilson, 2017).

Ασαφής λογική: Συλλογισμός που βασίζεται σε ασαφείς ή ελλιπείς πληροφορίες με βάση ένα εύρος τιμών και όχι σημειακές εκτιμήσεις. Η ασαφής λογική ασχολείται με την αβεβαιότητα στη γνώση που προσομοιώνει την ανθρώπινη συλλογιστική σε ελλιπή ή ασαφή δεδομένα. Το ασαφές μοντέλο είναι ανθεκτικό στις αλλαγές των παραμέτρων και ανεκτικό στις εντυπώσεις.

Μηχανική μάθηση (Machine Learning – ML): Αλγόριθμοι για να κάνουν προβλέψεις και να ερμηνεύουν δεδομένα και να "μαθαίνουν", χωρίς στατικές οδηγίες προγράμματος. Η ML είναι μια στατιστική τεχνική για την προσαρμογή μοντέλων σε δεδομένα και την εκπαίδευση μοντέλων με δεδομένα. Η ML εξάγει χαρακτηριστικά από τα δεδομένα εισόδου με την κατασκευή αναλυτικών αλγορίθμων δεδομένων και εξετάζει τα χαρακτηριστικά για τη δημιουργία προγνωστικών μοντέλων. Οι πιο συνηθισμένοι αλγόριθμοι ML είναι η μάθηση με επίβλεψη, η μάθηση χωρίς επίβλεψη, η μάθηση ενίσχυσης και η βαθιά μάθηση. Η πιο συνηθισμένη εφαρμογή της ML είναι η ιατρική ακριβείας. Οι αλγόριθμοι ML είναι κατάλληλοι για λύσεις κατά του κακόβουλου λογισμικού, επειδή η μηχανική μάθηση είναι κατάλληλη για την επίλυση "ασαφών" προβλημάτων.

Βαθιά μάθηση: Ένα υποσύνολο της μηχανικής μάθησης που βασίζεται σε μια βαθιά ιεραρχία επιπέδων, με κάθε επίπεδο να επιλύει διαφορετικά κομμάτια ενός πολύπλοκου προβλήματος. Στοχεύει στην αύξηση της ικανότητας των αλγορίθμων μάθησης με επίβλεψη και χωρίς επίβλεψη για την επίλυση σύνθετων προβλημάτων του πραγματικού κόσμου με την προσθήκη πολλαπλών επιπέδων επεξεργασίας.

Εξόρυξη δεδομένων: Ασχολείται με την ανακάλυψη κρυμμένων προτύπων και νέας γνώσης από μεγάλες βάσεις δεδομένων. Η εξόρυξη δεδομένων παρουσιάζει μια ποικιλία αλγοριθμικών εργαλείων, όπως η στατιστική, τα μοντέλα παλινδρόμησης, τα νευρωνικά δίκτυα, τα ασαφή σύνολα και τα εξελικτικά μοντέλα. Κάθε εργαλείο τεχνητής νοημοσύνης έχει τα δικά του πλεονεκτήματα. Συνιστάται η χρήση ενός συνδυασμού αυτών των μοντέλων και όχι ενός μόνο μοντέλου. Οι τεχνολογίες τεχνητής νοημοσύνης επηρεάζουν δραστικά τον κλάδο του λιανικού εμπορίου και την εμπειρία των πελατών. Οι εφαρμογές των τεχνολογιών ΤΝ για εργασίες κυβερνοασφάλειας προσελκύουν μεγαλύτερη προσοχή από τον ιδιωτικό και τον δημόσιο τομέα λόγω του ρυθμού με τον οποίο αναπτύσσονται οι απειλές.

2.2.2 Εφαρμογές της τεχνητής νοημοσύνης

Η κυβερνοασφάλεια είναι η πρακτική της προστασίας κρίσιμων συστημάτων και ευαίσθητων πληροφοριών από ψηφιακές επιθέσεις. Υπάρχουν πολλοί τρόποι

για την προστασία των δεδομένων και της οργανωσιακής υποδομής, συμπεριλαμβανομένου του εντοπισμού εισβολών, της προστασίας από κακόβουλο λογισμικό, της αυστηρής τήρησης ορθών πρακτικών ασφαλείας και πολλών άλλων. Μια απειλή ασφαλείας στον κυβερνοχώρο μπορεί να είναι μια κυβερνοεπίθεση που χρησιμοποιεί κακόβουλο λογισμικό ή ransomware για την απόκτηση πρόσβασης σε δεδομένα, για τη διακοπή των ψηφιακών λειτουργιών ή πρόκληση βλάβης σε πληροφορίες.

Υπάρχουν κάθε είδους απειλές στον κυβερνοχώρο, συμπεριλαμβανομένων των εταιρικών κατασκόπων, των χάκερ και των τρομοκρατών. Αν και όλες οι απειλές ενέχουν διαφορετικούς λόγους για την επίθεση, θα πρέπει να αντιμετωπίζονται με εξαιρετική προσοχή, καθώς αποτελούν κίνδυνο για τα δεδομένα ενός οργανισμού και τα προσωπικά δεδομένα. Η άνοδος του Διαδικτύου έχει φέρει μια νέα εποχή ανησυχιών για την ασφάλεια στον κυβερνοχώρο. Εκτός από την απειλή των εγκληματιών χάκερ και των ξένων κυβερνήσεων, νέες προκλήσεις συνδέονται με την προστασία των πληροφοριών από εσωτερικές απειλές, όπως παραβιάσεις δεδομένων και κλοπή εμπιστευτικών πληροφοριών.

Η ασφάλεια στον κυβερνοχώρο αποτελεί επίσης μια ουσιαστική ανησυχία για ευπαθείς υποδομές, κρίσιμα περιουσιακά στοιχεία και ευαίσθητες πληροφορίες. Αυτός είναι ο λόγος για τον οποίο υπήρξε μια αξιοσημείωτη άνοδος των επαγγελματιών στην κυβερνοασφάλεια και στον κλάδο εν συνόλω και διότι καθίσταται ολοένα και πιο σημαντικό να διασφαλιστεί ότι οι μηχανισμοί άμυνας κατά των επιθέσεων στον κυβερνοχώρο είναι ολοκληρωμένοι και ισχυροί (Caballero, 2012).

Η κυβερνοασφάλεια είναι ένας ευρύς όρος που περιλαμβάνει όλα τα μέτρα που λαμβάνονται σε μια προσπάθεια να προστατευθεί μια οντότητα από απειλές στον κυβερνοχώρο, συμπεριλαμβανομένης της ασφάλειας δεδομένων και του μετριασμού της ζημίας από ένα περιστατικό ασφάλειας στον κυβερνοχώρο. Ο τομέας της κυβερνοασφάλειας μπορεί γενικά να ταξινομηθεί σε πέντε διακριτούς τομείς ασφάλειας:

- Ασφάλεια υποδομών ζωτικής σημασίας

- Ασφάλεια εφαρμογών
- Ασφάλεια δικτύου
- Ασφάλεια στο νέφος
- Ασφάλεια στο Διαδίκτυο των Πραγμάτων (IoT).

Η κυβερνοασφάλεια είναι ένα σύνθετο και συνεχώς μεταβαλλόμενο πεδίο. Είναι σημαντικό να κατανοηθούν οι διαφορετικοί τύποι των απειλών στον κυβερνοχώρο και οι τρόποι που μπορούν να μετριαστούν. Οι κυβερνοεπιθέσεις αποτελούν συχνό φαινόμενο στη σημερινή κοινωνία. Ωστόσο, αυτές οι επιθέσεις μπορούν να αποτραπούν με τα κατάλληλα μέτρα ασφαλείας.

2.3 Επισκόπηση του εγκλήματος στον κυβερνοχώρο και των διαφόρων τύπων του

Κατανεμημένη Άρνηση Υπηρεσίας (Distribu tedDenial of Service, DDoS): Είναι μια μορφή κυβερνοεπίθεσης όπου ο δράστης χρησιμοποιεί πολλαπλά συστήματα για να κατακλύσει τον στόχο με κίνηση δεδομένων. Ο σκοπός είναι να δυσκολευτεί ο στόχος να παρέχει υπηρεσίες ή να έχει πρόσβαση στον ιστότοπό του. Ο πιο συνηθισμένος τύπος επίθεσης DDoS είναι μια ογκομετρική επίθεση, η οποία κατακλύζει τον στόχο με μια συντριπτική ποσότητα δεδομένων. Αυτό μπορεί να πραγματοποιηθεί με τη χρήση ενός botnet, όπου ένα δίκτυο υπολογιστών μολύνεται με κακόβουλο λογισμικό και ελέγχεται από έναν εισβολέα χωρίς να το γνωρίζει ο ιδιοκτήτης τους. Στην κατηγορία επιθέσεων όγκου, υπάρχουν επιθέσεις κατάκλυσης και ενίσχυσης/ανάκλασης. Σε μια επίθεση κατάκλυσης, η κίνηση αποστέλλεται με την ελπίδα την εξάντληση του εύρους ζώνης, της ικανότητας επεξεργασίας ή άλλων πόρων δικτύου. Οι επιθέσεις ενίσχυσης/ανάκλασης επιδιώκουν να αναγκάσουν τα θύματα να ξοδέψουν χρήματα «υπερφορτώνοντας» τα δίκτυά τους με κίνηση ανεπιθύμητης αλληλογραφίας ή αρνούμενοι την πρόσβαση σε ορισμένους πόρους χρησιμοποιώντας μηνύματα που μοιάζουν με ανεπιθύμητα μηνύματα (Zargas, Joshi&Tipper, 2013; Furfaro et al., 2015).

Υπάρχουν πολλά μέσα των οποίων η χρήση μπορεί να βοηθήσει στην εξαπόλυση μιας επίθεσης DDoS, με πιο συνηθισμένο τη χρήση ενός botnet. Οι

επιτιθέμενοι δεν χρειάζεται να ελέγχουν το botnet, μπορούν να το νοικιάσουν από μια διαδικτυακή υπηρεσία ή να το αγοράσουν από άλλον. Παραδείγματα τέτοιων υπηρεσιών είναι οι Blackhole, Stresser και NitrousDDoS (Tandon, 2020).

Επίθεση ενδιάμεσης οντότητας (man-in-the-middle): Μια μορφή κυβερνοεπίθεσης όπου ο εισβολέας αναμεταδίδει κρυφά και πιθανώς αλλάζει την επικοινωνία μεταξύ δύο μερών που πιστεύουν ότι επικοινωνούν απευθείας μεταξύ τους ονομάζεται επίθεση ενδιάμεσης οντότητας. Ο εισβολέας μπορεί να διαβάσει όλα τα μηνύματα που διέρχονται μεταξύ των δύο συστημάτων και μπορεί επίσης να εισάγει πλαστογραφημένα μηνύματα. Ο όρος «ενδιάμεση οντότητα» προέρχεται από την κατασκοπεία, όπου ένα μέρος νομίζει ότι συνομιλεί απευθείας με ένα άλλο μέρος, ενώ στην πραγματικότητα και τα δύο μηνύματά τους διαβάζονται από έναν ωτακουστή. Η επίθεση ενδιάμεσης οντότητας μπορεί να γίνει με διάφορους τρόπους. Ένας τρόπος είναι ο εισβολέας να τοποθετηθεί φυσικά μεταξύ των δύο μερών χωρίς κανένα από τα μέρη να γνωρίζει και στη συνέχεια να μεταδίδουν μηνύματα μεταξύ τους. Αυτό θα μπορούσε να πραγματοποιηθεί, για παράδειγμα, με την πρόσβαση στο δίκτυο μιας τηλεφωνικής εταιρείας και την αλλαγή δρομολόγησης κλήσεων ή σε ένα δημόσιο σημείο πρόσβασης Wi-Fi. Ένας άλλος τρόπος είναι ένας εισβολέας με δικαιώματα διαχειριστή συστήματος σε έναν απομακρυσμένο υπολογιστή να εκμεταλλευτή τη δυνατότητα της επίθεσης ενδιάμεσης οντότητας (MITM) για την πρόσβαση και έλεγχο της κίνησης μεταξύ πελάτη και διακομιστή.

Συχνά αναφέρεται ως «οντότητα στον περιηγητή»(man-in-the-browser) επειδή χρησιμοποιεί τα τρωτά σημεία σε ένα πρόγραμμα περιήγησης ή άλλο λογισμικό για να πραγματοποιηθεί η επίθεση. Η επίθεση μπορεί να χρησιμοποιεί κοινωνική μηχανική, όπου ο εισβολέας εξαπατά τον χρήστη ώστε να αποδεχτεί μια μη ασφαλή σύνδεση μέσω HTTPS ή άλλων ασφαλών πρωτοκόλλων ή εκμεταλλευόμενος γνωστές ευπάθειες στο λογισμικό, όπως το Cross-SiteScripting. Χρησιμοποιείται επίσης ως αναφορά σε επιθέσεις όπου ένα μη εξουσιοδοτημένο άτομο αποκτά πρόσβαση σε υπολογιστή που εκτελεί ένα πρόγραμμα περιήγησης, χρησιμοποιεί τη διεπαφή του προγράμματος περιήγησης μέσω webcam και μικροφώνου και δυνατότητες εγγραφής, προκειμένου να κατασκοπεύσει τον

χρήστη. Στη συνέχεια, ο εισβολέας μπορεί να χρησιμοποιήσει αυτές τις πληροφορίες για εκβιασμό ή άλλους κακόβουλους σκοπούς.

Επιθέσεις SQL: Αυτός ο τύπος κυβερνοεπίθεσης εκμεταλλεύεται την ευπάθεια ασφαλείας στη βάση δεδομένων. Είναι ένας τύπος τεχνικής έγχυσης κώδικα που μπορεί να χρησιμοποιηθεί για επίθεση σε εφαρμογές που βασίζονται σε δεδομένα. Η επίθεση SQL είναι ένας από τους πιο συνηθισμένους και επικίνδυνους τύπους κυβερνοεπιθέσεων. Μπορεί να χρησιμοποιηθεί για την κλοπή ευαίσθητων πληροφοριών από βάσεις δεδομένων, την τροποποίηση ή τη διαγραφή δεδομένων και τη διακοπή της υπηρεσίας. Η επίθεση SQL χρησιμοποιεί τη δυναμική φύση της SQL (δομημένη γλώσσα ερωτημάτων) για να παρακάμψει την επικύρωση εισόδου και να αποκτήσει πρόσβαση σε δεδομένα που διαφορετικά δεν θα ήταν προσβάσιμα. Η επίθεση SQL συνήθως περιλαμβάνει τη χρήση λανθασμένης ή σαφώς εσφαλμένης εισαγωγής σε μια εντολή SQL. Για παράδειγμα, εάν αποθηκεύσετε έναν κωδικό πρόσβασης αφού τον καταχωρίσετε στο πεδίο λογαριασμού, μπορείτε να στείλετε το μήνυμα "Επιλέξτε κωδικό πρόσβασης χρήστη" αντί για το μήνυμα "Παρακαλώ εισάγετε τις τιμές χρήστη (όνομα χρήστη, κωδικός πρόσβασης)". Αυτό θα αναγκάσει τη βάση δεδομένων να εκτελέσει το ερώτημα και να επιστρέψει τα ονόματα χρηστών. Αυτό έχει ως αποτέλεσμα να δώσει μια λίστα με ονόματα χρηστών από την οποία ο επιτιθέμενος μπορεί να εξάγει τον κωδικό πρόσβασης του χρήστη στη βάση δεδομένων (Dilek,Çakır&Aydın,2015). Η επίθεση SQL μπορεί επίσης να χρησιμοποιηθεί για την αλλαγή δεδομένων σε έναν πίνακα χωρίς την κατάλληλη εξουσιοδότηση. Αυτό μπορεί να έχει ως αποτέλεσμα την απώλεια του απορρήτου και της διαθεσιμότητας του συγκεκριμένου πίνακα ή άλλων πινάκων που αναφέρονται σε αυτόν.

Επίθεση σε κωδικό πρόσβασης: Μια μέθοδος κυβερνοεπίθεσης όπου ο εισβολέας επιχειρεί να μαντέψει τους κωδικούς πρόσβασης ή να τους κλέψει ευθέως, συνήθως μέσω εισβολής σε ένα σύστημα υπολογιστή ή σε ένα δίκτυο, αναφέρεται ως επίθεση σε κωδικό πρόσβασης. Ο εισβολέας μπορεί να χρησιμοποιήσει μια επίθεση ενδιάμεσης οντότητας για να υποκλέψει τον κωδικό πρόσβασης του θύματος και στη συνέχεια να τον χρησιμοποιήσει για να αποκτήσει πρόσβαση στον λογαριασμό του. Είναι γνωστό ότι είναι δύσκολο να αποκτηθεί

πρόσβαση σε έναν κωδικό πρόσβασης, αλλά με προηγμένα προγράμματα και τακτικές, οι χάκερ μπορούν τελικά να επιτύχουν σημαντική πρόοδο. Υπάρχουν τρεις τύποι επιθέσεων στον κωδικό πρόσβασης: ωμής δύναμης, λεξικού και καταγραφή πληκτρολόγησης. Μια επίθεση ωμής δύναμης (bruteforce) είναι μια σειρά προσπαθειών για εικασία των κωδικών πρόσβασης μέχρι ο εισβολέας να διεισδύσει στο σύστημα. Οι επιθέσεις λεξικού περιλαμβάνουν τη δοκιμή διαφορετικών συνδυασμών λέξεων λεξικού έως ότου επιτευχθεί η εικασία του κωδικού πρόσβασης. Η καταγραφή της πληκτρολόγησης είναι μια μέθοδος για την εξαγωγή ευαίσθητων δεδομένων όπως διαπιστευτήρια σύνδεσης για χρήση σε μια επίθεση ανάκτησης κωδικού πρόσβασης. Αυτός ο τύπος επίθεσης μπορεί να αποτραπεί με τη χρήση επαλήθευσης ταυτότητας δύο παραγόντων ή αποφεύγοντας κλικ σε συνδέσμους σε μηνύματα ηλεκτρονικού ταχυδρομείου από άγνωστες πηγές.

Επιθέσεις σε συσκευές IoT: Το Διαδίκτυο των πραγμάτων (IoT) είναι ένα δίκτυο φυσικών αντικειμένων που είναι ενσωματωμένα με ηλεκτρονικά, λογισμικό, αισθητήρες και συνδεσιμότητα ώστε να επιτραπεί σε αυτά τα αντικείμενα να συλλέγουν και να ανταλλάσσουν δεδομένα. Υπάρχουν πολλά δίκτυα IoT, με τα πιο κοινά να είναι το Διαδίκτυο και τα ενδοδίκτυα. Το IoT είναι ένα νέο σύνορο για τους εγκληματίες του κυβερνοχώρου. Στοχεύουν συσκευές IoT επειδή συχνά δεν έχουν κατάλληλους μηχανισμούς ασφάλειας. Αυτό τις καθιστά εύκολο στόχο για επιθέσεις στον κυβερνοχώρο. Καθώς ο αριθμός των συσκευών που είναι συνδεδεμένες στο Διαδίκτυο συνεχίζει να αυξάνεται, καθίσταται ολοένα και πιο σημαντικό να προστατεύονται αυτές οι συσκευές από κακόβουλους χάκερ και εγκληματίες στον κυβερνοχώρο. Γενικά, η ασφάλεια του IoT απαιτεί τρία κύρια στοιχεία: αρχιτεκτονική λογισμικού συσκευών IoT, μονάδα αξιόπιστης πλατφόρμας (TPM) και πρότυπα ασφαλείας (Guanetal., 2017).

Σύμφωνα με μια μελέτη που δημοσιεύτηκε από τη Cisco Systems, «Κυβερνοαπειλές στο Διαδίκτυο των πραγμάτων: Αναδυόμενοι κίνδυνοι και στρατηγικές τακτικής», έχει βρεθεί ότι οι εισβολείς στοχεύουν τον αυξανόμενο αριθμό των συνδεδεμένων συσκευών στο Διαδίκτυο προκειμένου να αποκτήσουν πρόσβαση σε ιδιωτικές πληροφορίες, να προκαλέσουν αναστάτωση ή κλοπή χρημάτων από χρήστες. Η μελέτη διαπίστωσε ότι το Διαδίκτυο των πραγμάτων έχει

σημαντικό αντίκτυπο στις προοπτικές των επιχειρήσεων και των καταναλωτών και είναι σημαντικό για τους οργανισμούς να λαμβάνουν υπόψη τους κινδύνους για την ασφάλεια στον κυβερνοχώρο. Οι οργανισμοί θα πρέπει να επανεξετάσουν τα υπάρχοντα προϊόντα και υπηρεσίες δικτύου τους σε σχέση με αυτούς τους αναδυόμενους κινδύνους.

Με την πρόσφατη πανδημία, η κουλτούρα της εργασίας από το σπίτι προωθήθηκε σημαντικά και φαίνεται ότι η τάση των γραφείων στο σπίτι ήρθε για να μείνει. Ως συνέπεια, οι κατοικημένες περιοχές μετατρέπονται πλέον σε πολύτιμο στόχο για διάφορους λόγους (Kempetal., 2021). Οι περισσότεροι οργανισμοί είναι καλώς προετοιμασμένοι να αμυνθούν από απειλές στον κυβερνοχώρο, αλλά αυτό ισχύει πρωτίστως μέσα από την υποδομή του οργανισμού. Οι συσκευές που συνδέονται με τους διακομιστές του οργανισμού από τα σπίτια των εργαζομένων επίσης αποτελούν σημείο ευπάθειας για εκμετάλλευση από τους εισβολείς. Ο όγκος αυτών των επιθέσεων αυξήθηκε κατά 35% το πρώτο εξάμηνο του 2020 σε σύγκριση με το δεύτερο εξάμηνο του 2019, όπως αναφέρεται σε έκθεση της Microsoft. Με τη δημοτικότητα της κατ' οίκον εργασίας, υπάρχει μεγαλύτερος κίνδυνος στοχοποίησης και εκμετάλλευσης αυτών των συσκευών από χάκερ. Εάν ένας χάκερ επικεντρωθεί στην παραβίαση ενός ή περισσότερων κατοικιών σε μια περιοχή θα μπορούσε να προκαλέσει όλεθρο σε έναν οργανισμό και εύλογα θα μπορούσε να θέσει σε κίνδυνο και άλλα σπίτια. Για παράδειγμα, μια επίθεση κατάργησης ταυτότητας σε ένα μη ασφαλές ασύρματο δίκτυο μπορεί να παρέχει στον εισβολέα έναν κατακερματισμένο κωδικό πρόσβασης και αυτό σημαίνει ότι είναι δυνατή η κακόβουλη χρήση αυτού του κωδικού πρόσβασης.

2.3.1 Τεχνητή νοημοσύνη ως εργαλείο αντιμετώπισης κυβερνοεπιθέσεων

Υπάρχει ένα ευρύ φάσμα διεπιστημονικών διασταυρώσεων μεταξύ της τεχνητής νοημοσύνης και της ασφάλειας στον κυβερνοχώρο. Τα εργαλεία ΤΝ (όπως τα συστήματα εμπειρογνωμόνων, η υπολογιστική νοημοσύνη, τα νευρωνικά δίκτυα, οι ευφυείς πράκτορες, τα τεχνητά ανοσοποιητικά συστήματα, η μηχανική μάθηση, η εξόρυξη δεδομένων, η αναγνώριση προτύπων, η ασαφής λογική, η ευρετική κ.λπ.

Μπορούν να χρησιμοποιηθούν για να μάθουν πώς να επιτρέπουν στους ειδικούς ασφαλείας να κατανοούν το περιβάλλον του κυβερνοχώρου προκειμένου να εντοπίζουν ανωμαλίες. Η χρήση της τεχνητής νοημοσύνης μπορεί να συμβάλει στη διεύρυνση των οριζόντων των υφιστάμενων λύσεων ασφάλειας στον κυβερνοχώρο. Για την ενίσχυση των υφιστάμενων συστημάτων κυβερνοασφάλειας, οι εταιρείες μπορούν να εφαρμόσουν την τεχνητή νοημοσύνη στους ακόλουθους τέσσερις τομείς: αυτοματοποιημένη άμυνα, γνωστική ασφάλεια, εκπαίδευση αντιπάλων, παράλληλη και δυναμική παρακολούθηση (Dilek, Çakır & Aydın, 2015).

Αυτοματοποιημένη άμυνα: Υπάρχουν δύο τύποι συστημάτων κυβερνοασφάλειας: τα συστήματα εμπειρογνωμόνων (που καθοδηγούνται από αναλυτές) και τα αυτοματοποιημένα (που καθοδηγούνται από μηχανές). Τα συστήματα εμπειρογνωμόνων αναπτύσσονται και λειτουργούν από ανθρώπους, ενώ τα αυτοματοποιημένα συστήματα χρησιμοποιούν ευφυή εργαλεία ΤΝ. Τα συστήματα που βασίζονται στην ΤΝ είναι αυτόνομοι, αυτοδιδακτικοί πράκτορες. Ένα καλό παράδειγμα αυτοματοποιημένου συστήματος είναι το CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart). Η ταχύτητα και ο όγκος των δεδομένων που απαιτούνται για την υπεράσπιση του κυβερνοχώρου δεν μπορούν να αντιμετωπιστούν από ανθρώπους χωρίς αυτοματοποίηση. Καθώς τα δίκτυα γίνονται μεγαλύτερα και πιο πολύπλοκα, η τεχνητή νοημοσύνη μπορεί να αποτελέσει τεράστιο πλεονέκτημα για την προστασία ενός οργανισμού στον κυβερνοχώρο. Μια ιδανική κυβερνοάμυνα στοχεύει στην πλήρη προστασία των χρηστών, διατηρώντας παράλληλα όλες τις λειτουργίες. Τα αυτοματοποιημένα συστήματα τεχνητής νοημοσύνης μπορούν να ενσωματωθούν στις υπάρχουσες λειτουργίες ασφάλειας στον κυβερνοχώρο. Ορισμένες από αυτές τις λειτουργίες περιλαμβάνουν (Zhang et al.m 2022):

- Δημιουργία ακριβέστερων, βιομετρικών τεχνικών σύνδεσης
- Ανίχνευση απειλών και κακόβουλων δραστηριοτήτων με χρήση προγνωστικών αναλύσεων
- Βελτίωση της μάθησης και της ανάλυσης μέσω της επεξεργασίας φυσικής γλώσσας
- Διασφάλιση της αυθεντικοποίησης υπό όρους και της πρόσβασης

- Βελτίωση της ανθρώπινης ανάλυσης - από την ανίχνευση κακόβουλων επιθέσεων έως την προστασία τελικών σημείων
- Χρήση στην αυτοματοποίηση καθημερινών εργασιών ασφαλείας
- Μη ύπαρξη ευπαθειών μηδενικής ημέρας

Η ενσωμάτωση της τεχνητής νοημοσύνης στην ασφάλεια στον κυβερνοχώρο έχει πολλά πλεονεκτήματα. Οι λύσεις κυβερνοασφάλειας που βασίζονται στην τεχνητή νοημοσύνη είναι σχεδιασμένες να λειτουργούν όλο το εικοσιτετράωρο για την προστασία σας.

Γνωστική ασφάλεια: Η γνωστική ασφάλεια συνδυάζει τα πλεονεκτήματα της τεχνητής νοημοσύνης και της ανθρώπινης νοημοσύνης. Η γνωστική υπολογιστική (Cognitive computing - CC), ένας προηγμένος τύπος τεχνητής νοημοσύνης, αξιοποιεί διάφορες μορφές τεχνητής νοημοσύνης. Αναφέρεται σε υλικό ή/και λογισμικό που μιμείται τον τρόπο με τον οποίο λειτουργεί ο ανθρώπινος εγκέφαλος. Η τεχνητή νοημοσύνη και η γνωστική υπολογιστική παραμένουν πολύ παρόμοιες ως προς την πρόθεση, αλλά διαφέρουν ως προς την τάση τους να αλληλεπιδρούν φυσικά με τον άνθρωπο. Η τεχνητή νοημοσύνη έχει περιγραφεί ως τεχνολογία ικανή να εκτελεί εργασίες που κανονικά απαιτούν ανθρώπινη νοημοσύνη. Η γνωστική υπολογιστική επιδιώκει να ξεπεράσει τα όρια των συμβατικών προγραμματιζόμενων (von Neumann) υπολογιστών. Το Watson for cyber security, το πρώτο γνωστικό σύστημα της IBM, απέδειξε μέσω ενός εκθεσιακού αγώνα Jeopardy ότι είναι ικανό να απαντά σε πολύπλοκες ερωτήσεις εξίσου αποτελεσματικά με τους παγκόσμιους πρωταθλητές. Το Watson μαθαίνει με κάθε αλληλεπίδραση, ώστε να συνδέει τις τελείες μεταξύ των απειλών και να παρέχει πληροφορίες που μπορούν να αξιοποιηθούν. Αυτό επιτρέπει σε έναν αναλυτή να ανταποκριθεί στις απειλές με μεγαλύτερη εμπιστοσύνη και ταχύτητα (IBM, 2023).

Εκπαίδευση αντιπάλων: Αυτός ο όρος χρησιμοποιείται συχνά για να αναφερθεί στην ανάπτυξη και τη χρήση της τεχνητής νοημοσύνης για κακόβουλους σκοπούς. Οι μηχανικοί κυβερνοασφάλειας δημιουργούν προληπτικά μοντέλα αντίπαλης επίθεσης για να διερευνήσουν τα τρωτά σημεία της τεχνητής νοημοσύνης. Μια επίθεση αντίπαλης μάθησης μπορεί να προκαλέσει την κακή συμπεριφορά των αλγορίθμων ή να αποκαλύψει πληροφορίες σχετικά με την

εσωτερική τους λειτουργία. Η αντίπαλη εκπαίδευση μεταξύ των συστημάτων ΤΝ μπορεί να συμβάλει στη βελτίωση της ευρωστίας τους καθώς και να διευκολύνει τον εντοπισμό των τρωτών σημείων του συστήματος. Όσο περισσότερο υιοθετούμε μια στρατηγική "αντιπολιτευτικής" και "διαρκώς ενεργής", τόσο πιο ασφαλείς γίνονται οι εφαρμογές τεχνητής νοημοσύνης (Taddeo, McCutcheon & Floridi, 2021).

Παράλληλη και δυναμική παρακολούθηση: Οι μαθησιακές ικανότητες των στοχευόμενων συστημάτων απαιτούν κάποια μορφή συνεχούς παρακολούθησης κατά τη διάρκεια της ανάπτυξης. Η παρακολούθηση είναι απαραίτητη για να διασφαλιστεί ότι οι αποκλίσεις μεταξύ της αναμενόμενης και της πραγματικής συμπεριφοράς ενός συστήματος καταγράφονται και αντιμετωπίζονται επαρκώς. Για τον σκοπό αυτό, οι πάροχοι συστημάτων τεχνητής νοημοσύνης θα πρέπει να διατηρούν ένα σύστημα κλώνο ως σύστημα ελέγχου, το οποίο χρησιμεύει ως σημείο αναφοράς με βάση το οποίο αξιολογείται η συμπεριφορά του αρχικού συστήματος (Taddeo, McCutcheon & Floridi, 2021).

2.3.2 Τεχνητή νοημοσύνη και ανίχνευση εισβολών

Ένα σύστημα ανίχνευσης εισβολών (IDS) που ορίζεται ως "μια αποτελεσματική τεχνολογία ασφάλειας, η οποία μπορεί να ανιχνεύει, να αποτρέπει και ενδεχομένως να αντιδρά στις επιθέσεις σε υπολογιστές" είναι ένα από τα βασικά στοιχεία των υποδομών ασφαλείας. Παρακολουθεί στοχευμένες πηγές δραστηριοτήτων, όπως δεδομένα ελέγχου και κίνησης δικτύου σε συστήματα υπολογιστών ή δικτύων και αναπτύσσει διάφορες τεχνικές προκειμένου να παρέχει υπηρεσίες ασφαλείας. Ο κύριος στόχος του IDS είναι να ανιχνεύει όλες τις εισβολές με αποτελεσματικό τρόπο.

Η εφαρμογή του IDS επιτρέπει στους διαχειριστές δικτύων να εντοπίζουν παραβιάσεις του στόχου ασφαλείας. Αυτές οι παραβιάσεις του στόχου ασφάλειας κυμαίνονται από εξωτερικούς επιτιθέμενους που προσπαθούν να αποκτήσουν μη εξουσιοδοτημένη πρόσβαση στην υποδομή ασφάλειας του δικτύου ή να καταστήσουν μη διαθέσιμους τους πόρους έως εσωτερικούς χρήστες που κάνουν

κατάχρηση της πρόσβασής τους στους πόρους του συστήματος. Με την πάροδο του χρόνου και την αύξηση των επιθέσεων στους υπολογιστές, προτάθηκε αρκετή αρχιτεκτονική IDS.

Σύμφωνα με τον Axelsson (2000), τα κοινά στοιχεία για IDS αποτελούνται από τα εξής: Το δίκτυο προς παρακολούθηση είναι η ταυτότητα που πρέπει να παρακολουθείται για εισβολές. Η μονάδα συλλογής και αποθήκευσης δεδομένων είναι υπεύθυνη για τη συλλογή των δεδομένων των διαφόρων συμβάντων και τη μετατροπή τους σε κατάλληλη μορφή και την αποθήκευση στο δίσκο- η μονάδα ανάλυσης και επεξεργασίας δεδομένων είναι ο εγκέφαλος του IDS. Περιέχει όλη τη λειτουργικότητα για την εύρεση της ύποπτης συμπεριφοράς της κίνησης επιθέσεων. Κατά την ανίχνευση της επίθεσης, παράγεται ένα σήμα. Ανάλογα με τον τύπο του IDS, η δράση μπορεί να αναληφθεί από το ίδιο το σύστημα για την ανακούφιση του προβλήματος ή το σήμα διαβιβάζεται στον διαχειριστή του δικτύου για να αναλάβει την κατάλληλη δράση: Αυτό το τμήμα του συστήματος διαχειρίζεται όλες τις εξόδους από το IDS. Η έξοδος μπορεί να είναι είτε μια αυτοματοποιημένη απάντηση σε μια εισβολή είτε μια ειδοποίηση κακόβουλης δραστηριότητας για έναν διαχειριστή ασφάλειας δικτύου.

Στη βιβλιογραφία έχουν εφαρμοστεί διάφοροι τύποι τεχνικών από διάφορους κλάδους για την ανίχνευση των εισβολών. Οι κυριότερες τεχνικές είναι οι στατιστικές τεχνικές, οι τεχνικές που βασίζονται στη γνώση και οι τεχνικές που βασίζονται στην τεχνητή νοημοσύνη (AI). Στα IDS που βασίζονται στη στατιστική, η συμπεριφορά του συστήματος αναπαρίσταται από μια τυχαία οπτική γωνία. Από την άλλη πλευρά, οι τεχνικές IDS που βασίζονται στη γνώση προσπαθούν να συλλάβουν την απαιτούμενη συμπεριφορά από τα διαθέσιμα δεδομένα του συστήματος (προδιαγραφές πρωτοκόλλων, περιπτώσεις κίνησης δικτύου κ.λπ.). Τέλος, οι τεχνικές IDS με βάση την τεχνητή νοημοσύνη περιλαμβάνουν τη δημιουργία ενός ρητού ή σιωπηρού μοντέλου που επιτρέπει την κατηγοριοποίηση των προτύπων (Garcia-Teodoro et al. 2009).

2.3.2.1 Επιθυμητά χαρακτηριστικά ενός συστήματος ανίχνευσης και αποτροπής εισβολών (IDPS)

Τα εργαλεία IDPS είναι κεντρικής σημασίας για την ασφάλεια του δικτύου. Προστατεύουν τις επιχειρήσεις από εξωτερικούς και εσωτερικούς εισβολείς αναζητώντας ανωμαλίες στη συμπεριφορά του δικτύου. Για να το επιτύχουν αυτό, αναλύουν την υπογραφή της κυκλοφορίας του δικτύου, κυνηγούν ανωμαλίες συμπεριφοράς ή διεξάγουν ανάλυση πρωτοκόλλου με κατάσταση, στέλνοντας ένα σήμα και μελετώντας την απόκριση. Το IDPS μπορεί να συμβάλει στην πρόληψη διαφόρων εισβολών, όπως εισβολές στο δίκτυο της επιχείρησης, διαρροή δεδομένων, κατανεμημένες επιθέσεις άρνησης παροχής υπηρεσιών (DDoS) που επιβραδύνουν το δίκτυο, κακόβουλη χρήση εύρους ζώνης ή δόλιοι χρήστες που μεταμφιέζονται σε νόμιμους (Nitin et al., 2012).

Τα εργαλεία IDPS είναι συνήθως τεσσάρων τύπων - είτε μελετούν την κυκλοφορία του δικτύου, τη συμπεριφορά του δικτύου, την ασύρματη δραστηριότητα, είτε πληροφορίες σχετικά με το περιβάλλον του κεντρικού υπολογιστή. Αυτοί οι τύποι μπορεί να επικαλύπτονται και τα εργαλεία IDPS μπορούν να καλύψουν πολλαπλές περιπτώσεις χρήσης με μία λύση.

Τα Επιθυμητά χαρακτηριστικά ενός συστήματος ανίχνευσης και αποτροπής εισβολών (IDPS) είναι τα εξής: Παρακολούθηση δικτύου 24/7 - Ο κύριος σκοπός της ανάπτυξης ενός συστήματος ανίχνευσης και πρόληψης εισβολών είναι η παρακολούθηση του δικτύου όλο το εικοσιτετράωρο. Το εργαλείο συνδέεται με πολλαπλές συσκευές δικτύου, λογισμικό, διακομιστές, συστήματα και συσκευές τελικών σημείων, εάν χρειάζεται. Θα αναλύει το 100% της ροής της κυκλοφορίας και θα τα συγκρίνει με προκαθορισμένους κανόνες. Οι κανόνες βοηθούν στη διάκριση της νόμιμης κυκλοφορίας από την κακόβουλη παρουσία (Nitin et al., 2012).

Επιβολή κανόνων εισβολής - Το εργαλείο IDPS θα πρέπει να επιτρέπει στους χρήστες να επιβάλλουν κανόνες εισβολής. Με βάση δυναμικά ενημερωμένες πληροφορίες σχετικά με τις απειλές, οι κανόνες αυτοί υποδεικνύουν ποιος τύπος συμπεριφοράς μετράει ως εισβολή και ποιος όχι. Ανάλογα με το εργαλείο που επιλέγετε, οι κανόνες μπορεί να είναι προ-ρυθμισμένοι και να διαχειρίζονται από

τον πάροχο, κάτι που αποτελεί μια προσέγγιση χαμηλής προσπάθειας αλλά μη ευέλικτη. Τα διαμορφώσιμα σύνολα κανόνων απαιτούν μεγαλύτερη προσπάθεια για την εφαρμογή, αλλά παρέχουν στους χρήστες μεγαλύτερο έλεγχο.

Αρχεία καταγραφής δραστηριοτήτων και πληροφορίες - Ένα άλλο σημαντικό χαρακτηριστικό των εργαλείων IDPS είναι η διατήρηση λεπτομερών αρχείων καταγραφής. Κάθε περιστατικό ασφαλείας (ανεξάρτητα από το πόσο μικρό ή χαμηλής σοβαρότητας) καταγράφεται για μελλοντική αναφορά και ελέγχους δικτύου. Οι λύσεις IDPS επιτρέπουν επίσης στους χρήστες να δημιουργούν ad-hoc αναφορές για την εκπλήρωση των απαιτήσεων συμμόρφωσης - για παράδειγμα, για να αποδείξουν ότι το δίκτυο είναι τμηματοποιημένο σύμφωνα με το πρότυπο ασφάλειας δεδομένων της βιομηχανίας καρτών πληρωμών (PCI DSS) (Kumar, Maurya & Misra, 2013).

Ανίχνευση κακόβουλης παρουσίας - Τα συστήματα ανίχνευσης και πρόληψης εισβολών εντοπίζουν αμέσως μια κακόβουλη παρουσία μόλις γίνει αισθητή στο δίκτυο. Το εργαλείο δεν θα περιμένει μέχρι να προκληθεί ζημιά ή να γίνει απόπειρα παραβίασης εμπιστευτικών δεδομένων ή συστημάτων λογισμικού. Οι μικρές ή οικείες εισβολές θα ανιχνεύονται, θα καταγράφονται και θα αποκλείονται αυτόματα, ενώ οι πιο σύνθετες μπορεί να προκαλέσουν μια ειδοποίηση. Ορισμένα εργαλεία χρησιμοποιούν τεχνητή νοημοσύνη (AI) και μηχανική μάθηση (ML) για να ανιχνεύουν και να ταξινομούν με ακρίβεια τις εισβολές.

Μπλοκάρισμα κακόβουλης παρουσίας - Το εργαλείο IDPS θα πρέπει να συμβάλλει στον αποκλεισμό των εισβολέων και στον μετριασμό των ζημιών που προκαλούν. Όπως αναφέρθηκε, τα γνωστά ζητήματα επιλύονται αυτόματα και μπορεί να δημιουργηθεί μια αναφορά για την ομάδα IT. Πιο σύνθετες εισβολές, όπως κακόβουλο λογισμικό ή ύποπτα αρχεία, μπορούν να τεθούν σε καραντίνα σε ένα εικονικό sandbox. Ορισμένα εργαλεία ενσωματώνονται με εξωτερικά συστήματα για τον εξορθολογισμό της διαδικασίας αποκλεισμού.

2.4 Παρεμβάσεις τεχνητής νοημοσύνης και πρόσφατες τάσεις

Με τη συνεχή ανάπτυξη της επιστήμης της τεχνητής νοημοσύνης και την εμβάθυνση της προσομοίωσης της ανθρώπινης συνείδησης και των διαδικασιών σκέψης πληροφοριών, η τεχνητή νοημοσύνη έχει λάβει ολοένα και μεγαλύτερη προσοχή σε διάφορους τομείς. Ως αποτέλεσμα, εφαρμόζεται σταδιακά σε κλάδους όπως η ρομποτική, η ιατρική περίθαλψη, η μεταποίηση, η προστασία του περιβάλλοντος και η κατασκευή δικτύων. Επιπλέον, με την ολοένα αυξανόμενη ανάπτυξη της επιστήμης και της τεχνολογίας, οι εφαρμογές AI θα συνεχίσουν να εμφανίζονται σε όλο και περισσότερες πτυχές της καθημερινής ζωής. Η τεχνητή νοημοσύνη, και συγκεκριμένα η μηχανική μάθηση, έχουν διεισδύσει σε όλες τις πτυχές της σημερινής κοινωνίας. Τέσσερις τομείς που βλέπουν το μεγαλύτερο όφελος από την "AI +" είναι οι εξής (Bates et al., 2020):

Υγεία: Η τηλεϊατρική, η έξυπνη απεικόνιση, τα ιατρικά ρομπότ και η διάγνωση με τη βοήθεια της παθολογίας έχουν βοηθήσει τους κλινικούς γιατρούς σε αυτή την επιδημία.

Περιβάλλον: Η τεχνητή νοημοσύνη μπορεί να αντικαταστήσει τις χειρωνακτικές εργασίες προστασίας του περιβάλλοντος, όπου οι εργασίες αυτές έχουν χαμηλή απόδοση, υψηλό κόστος και υψηλούς κινδύνους που συνδέονται με αυτές. Επιπλέον, η τεχνολογία και τα προϊόντα τεχνητής νοημοσύνης μπορούν να βοηθήσουν τους ανθρώπους στην πρόληψη της περιβαλλοντικής ρύπανσης και καταστροφής.

Ασύρματες επικοινωνίες 5G: Η χρήση της τεχνητής νοημοσύνης μπορεί να επιτρέψει στο δίκτυο να επιτύχει υψηλή αποδοτικότητα της λειτουργίας και της συντήρησης, προβλεψιμότητα της κυκλοφορίας και ακρίβεια του μάρκετινγκ, βοηθώντας τις λειτουργίες των δικτύων επικοινωνίας να αντιμετωπίσουν τις προκλήσεις που σχετίζονται με τις παραδοσιακές μεθόδους διαχείρισης της λειτουργίας και της συντήρησης.

Οι τάσεις που γίνονται εμφανείς, σύμφωνα με τη διαδικασία ανάπτυξης της τεχνολογίας ΤΝ είναι η επιταχυνόμενη ενσωμάτωση της τεχνητής νοημοσύνης και της βιομηχανίας. Σε γενικές γραμμές, η εφαρμογή της τεχνητής νοημοσύνης στη βιομηχανία βρίσκεται στην αρχή της και εξακολουθούν να υπάρχουν ορισμένες

δυσκολίες που εμποδίζουν την εφαρμογή των σεναρίων εφαρμογής. Ως εκ τούτου, η τεχνητή νοημοσύνη πρέπει να ενσωματωθεί στενά με τη βιομηχανία, όχι μόνο για να προωθήσει την εφαρμογή των σεναρίων εφαρμογής της τεχνητής νοημοσύνης, αλλά και για να προωθήσει τις επαναστατικές καινοτομίες στην τεχνολογία βασικών δεδομένων και πλατφόρμας και να οικοδομήσει μια γέφυρα που να συνδέεται αποτελεσματικά με την παραδοσιακή οικολογία της βιομηχανίας (Ng, 2016).

Επίσης, για να αποτραπεί η κατάχρηση και η κατάχρηση της τεχνητής νοημοσύνης, αφενός, είναι απαραίτητο να αντιμετωπιστούν τα προβληματικά συμπτώματα από διαφορετικά επίπεδα, όπως οι νόμοι και οι κανονισμοί, οι ηθικοί κανόνες και η συναίνεση του κλάδου, αλλά και να αντιμετωπιστεί η βασική αιτία από το επίπεδο της τεχνολογικής καινοτομίας. Ως εκ τούτου, είναι ολοένα και πιο σημαντικό να ενσωματώσουμε τη δεοντολογία και τη διακυβέρνηση σε ολόκληρο τον κύκλο ζωής του σχεδιασμού, της έρευνας, της ανάπτυξης, της ανάπτυξης και της χρήσης προϊόντων τεχνητής νοημοσύνης (Trolice, Curchoe & Quaas, 2021).

2.4.1 Μείωση κυβερνοαπειλών με τη βοήθεια της τεχνητής νοημοσύνης

Η τεχνητή νοημοσύνη είναι κατάλληλη για την επίλυση ορισμένων από τα πιο δύσκολα προβλήματα, συμπεριλαμβανομένης της ασφάλειας στον κυβερνοχώρο. Πρόκειται για μια τεχνολογία που μεταμορφώνει τη ζωή μας. Ενσωματωμένη στα σπίτια, τα αυτοκίνητα και τις συσκευές μας, θα κάνει τα πάντα πιο "έξυπνα" και πιο αποτελεσματικά. Η τεχνητή νοημοσύνη μπορεί να ανιχνεύσει ένα λογισμικό είτε πρόκειται για κακόβουλο λογισμικό είτε για κανονικό λογισμικό. Οι νέες δυνατότητες ΤΝ μπορούν να κάνουν τον κόσμο πιο ασφαλή, δίκαιο και φιλικό προς το περιβάλλον. Λόγω της ευελιξίας και της προσαρμοστικής συμπεριφοράς τους, οι τεχνικές που βασίζονται στην ΤΝ μπορούν να μας βοηθήσουν να ξεπεράσουμε τις ελλείψεις των συμβατικών εργαλείων κυβερνοασφάλειας (Sadiku, Ashaolu & Musa, 2019).

Οι επιχειρήσεις εξαρτώνται πλέον από τις τεχνολογίες ΤΝ, όπως η μηχανική μάθηση, η βαθιά μάθηση και η επεξεργασία φυσικής γλώσσας, για να βοηθήσουν τους αναλυτές ασφαλείας να ανταποκριθούν στην απειλή με ταχύτητα και ακρίβεια.

Βοηθούν επίσης στην προστασία των δικτύων και των ευαίσθητων δεδομένων. Η τεχνητή νοημοσύνη μπορεί να χρησιμοποιηθεί για τον εντοπισμό απειλών και άλλων κακόβουλων δραστηριοτήτων. Μπορεί να αναλύσει τη συμπεριφορά των χρηστών, να συμπεράνει ένα μοτίβο και να εντοπίσει κάθε είδους ανωμαλίες σε ένα δίκτυο υπολογιστών. Είναι ικανή να προσαρμόζεται και να ανταποκρίνεται σε έναν συνεχώς μεταβαλλόμενο κόσμο. Βελτιώνει τον τρόπο με τον οποίο οι εμπειρογνώμονες κυβερνοασφάλειας αναλύουν, μελετούν και κατανοούν τα εγκλήματα στον κυβερνοχώρο. Ένα βασικό πλεονέκτημα της μηχανικής μάθησης στην ασφάλεια στον κυβερνοχώρο είναι ότι εντοπίζει και αντιδρά άμεσα σε ύποπτα προβλήματα.

Κεφάλαιο 3ο Μεθοδολογία

3.1 Συστηματική Ανασκόπηση

Οι συστηματικές ανασκοπήσεις έχουν κερδίσει δυναμική ως βασική μέθοδος σύνθεσης στοιχείων στην έρευνα για την παγκόσμια ανάπτυξη τον τελευταίο καιρό. Όπως ορίζεται στο εγχειρίδιο Cochrane Handbook on Systematic reviews "Οι συστηματικές ανασκοπήσεις επιδιώκουν να συγκεντρώσουν στοιχεία που ανταποκρίνονται σε προκαθορισμένα κριτήρια επιλεξιμότητας προκειμένου να απαντήσουν σε ένα συγκεκριμένο ερευνητικό ερώτημα. Στοχεύουν στην ελαχιστοποίηση της μεροληψίας με τη χρήση ρητών, συστηματικών μεθόδων που τεκμηριώνονται εκ των προτέρων με ένα πρωτόκολλο" (Higgins et al, 2019). Μια συστηματική ανασκόπηση είναι μια ολοκληρωμένη ανασκόπηση και σύνθεση δεδομένων βάσει πρωτοκόλλου που εστιάζει σε ένα θέμα ή σε συναφή βασικά ερωτήματα. Συνήθως εκτελείται από έμπειρους μεθοδολόγους με τη συμβολή εμπειρογνωμόνων του τομέα.

Το πρώτο βήμα για τη διεξαγωγή μιας συστηματικής ανασκόπησης είναι η διατύπωση συγκεκριμένων βασικών ερωτημάτων. Για περιπτώσεις που περιλαμβάνουν την αντιμετώπιση περισσότερων από ένα απλό ερώτημα, είναι συχνά χρήσιμο να κατασκευάζεται ένα αναλυτικό πλαίσιο (μοντέλο τεκμηρίωσης) που απεικονίζει τα βασικά ερωτήματα που εξετάζονται, ώστε να βοηθηθούν στην

εκτίμηση των σχέσεών τους. Επιπλέον, όταν εξετάζονται πολλά ερωτήματα, μπορεί να είναι επωφελής η κατασκευή ενός "χάρτη τεκμηρίων", μιας διερευνητικής άσκησης που ενημερώνει για την ποσότητα των τεκμηρίων που ενδεχομένως σχετίζονται με διάφορα ερωτήματα. Οι πληροφορίες αυτές μπορούν να βοηθήσουν στον λεπτομερέστερο σχεδιασμό για την κατανομή των πόρων και την έγκαιρη ολοκλήρωση του έργου. Πρόσθετα ουσιώδη βήματα περιλαμβάνουν την ανάπτυξη ενός πρωτοκόλλου, την εξειδίκευση των ερωτημάτων ενδιαφέροντος, τη διενέργεια βιβλιογραφικής αναζήτησης στοιχείων, την επιλογή μελετών που πληρούν τα κριτήρια ένταξης, την κριτική αξιολόγηση των μελετών και τη σύνθεση και ερμηνεία των αποτελεσμάτων.

Είναι σημαντικό να τονιστεί ότι η συστηματική ανασκόπηση διαφέρει από τη βιβλιογραφική ανασκόπηση. Ενώ μια βιβλιογραφική ανασκόπηση συνοψίζει ποιοτικά τα αποδεικτικά στοιχεία χωρίς συγκεκριμένο πρωτόκολλο ή κριτήρια αναζήτησης, μια συστηματική ανασκόπηση βασίζεται σε ένα σαφώς διατυπωμένο ερώτημα, εντοπίζει τις σχετικές μελέτες, αξιολογεί την ποιότητά τους και συνοψίζει τα αποδεικτικά στοιχεία με τη χρήση μιας επιλεγμένης σαφούς μεθοδολογίας. Αυτή η ρητή και συστηματική προσέγγιση είναι που διακρίνει τις συστηματικές ανασκοπήσεις από τις παραδοσιακές ανασκοπήσεις και τα σχόλια.

Είναι επίσης σημαντικό να γίνεται διάκριση μεταξύ συστηματικής ανασκόπησης και μετα-ανάλυσης. Ενώ η συστηματική ανασκόπηση αναφέρεται σε ολόκληρη τη διαδικασία επιλογής, αξιολόγησης και σύνθεσης των αποδεικτικών στοιχείων, η μετά-ανάλυση είναι ένα εξειδικευμένο υποσύνολο της συστηματικής ανασκόπησης. Η μετά-ανάλυση αναφέρεται στη στατιστική προσέγγιση του συνδυασμού δεδομένων που προέρχονται από τη συστηματική ανασκόπηση. Χρησιμοποιεί στατιστικές τεχνικές για να συνδυάσει τα δεδομένα που εξετάστηκαν από μεμονωμένες ερευνητικές μελέτες και χρησιμοποιεί τα συγκεντρωτικά δεδομένα για να καταλήξει σε νέα στατιστικά συμπεράσματα. Ως εκ τούτου, δεν περιλαμβάνουν όλες οι συστηματικές ανασκοπήσεις μετά-ανάλυση, αλλά η μετά-ανάλυση είναι απαραίτητα σε μια συστηματική ανασκόπηση (Akhter, Pauyo & Khan, 2019).

3.2 Χαρακτηριστικά Συστηματικής Ανασκόπησης

Μια συστηματική ανασκόπηση επιχειρεί να συγκεντρώσει όλα τα εμπειρικά στοιχεία που ανταποκρίνονται σε προκαθορισμένα κριτήρια επιλογής προκειμένου να απαντηθεί ένα συγκεκριμένο ερευνητικό ερώτημα. Χρησιμοποιεί σαφείς, συστηματικές μεθόδους που επιλέγονται με σκοπό την ελαχιστοποίηση της μεροληψίας, παρέχοντας έτσι πιο αξιόπιστα ευρήματα από τα οποία μπορούν να εξαχθούν συμπεράσματα και να ληφθούν αποφάσεις (Mulrow, 1994). Τα βασικά χαρακτηριστικά μιας συστηματικής ανασκόπησης είναι ένα σαφώς διατυπωμένο σύνολο στόχων με προκαθορισμένα κριτήρια επιλογής των μελετών, μια ρητή, αναπαραγωγική μεθοδολογία, μια συστηματική αναζήτηση που επιχειρεί να εντοπίσει όλες τις μελέτες που θα πληρούσαν τα κριτήρια επιλογής, μια αξιολόγηση της εγκυρότητας των ευρημάτων των ενταγμένων μελετών, για παράδειγμα μέσω της αξιολόγησης του κινδύνου μεροληψίας και μια συστηματική παρουσίαση και σύνθεση των χαρακτηριστικών και των ευρημάτων των ενταγμένων μελετών. Πολλές συστηματικές ανασκοπήσεις περιέχουν μετά-αναλύσεις. Η μετά-ανάλυση είναι η χρήση στατιστικών μεθόδων για τη σύνοψη των αποτελεσμάτων ανεξάρτητων μελετών (Glass, 1976). Συνδυάζοντας πληροφορίες από όλες τις σχετικές μελέτες, οι μετά-αναλύσεις μπορούν να παρέχουν ακριβέστερες εκτιμήσεις των επιπτώσεων της υγειονομικής περίθαλψης από εκείνες που προκύπτουν από τις μεμονωμένες μελέτες που περιλαμβάνονται σε μια ανασκόπηση. Διευκολύνουν επίσης τη διερεύνηση της συνέπειας των στοιχείων μεταξύ των μελετών και τη διερεύνηση των διαφορών μεταξύ των μελετών.

3.3 Τύποι Συστηματικής Ανασκόπησης

Υπάρχουν περισσότερα από 30 διαφορετικοί τύποι συστηματικών ανασκοπήσεων και δεν υπάρχει πάντα συναίνεση στα όρια και στις διακρίσεις μεταξύ των προσεγγίσεων αυτών. Παρακάτω παρουσιάζονται ορισμένοι βασικοί τύποι Συστηματικής Ανασκόπησης.

Εννοιολογική χαρτογράφηση: Οι επισκοπήσεις χαρτογράφησης επικεντρώνονται σε μια οπτική σύνθεση των δεδομένων και βασίζονται σε ερωτήματα και όχι σε θέματα, όπως η επισκόπηση εμβέλειας. Οι επισκοπήσεις χαρτογράφησης είναι καλύτερα σχεδιασμένες: Όταν υπάρχει αφθονία και ποικιλία ερευνών, Ως πρώτο βήμα για μια συστηματική ανασκόπηση, Για τον εντοπισμό κενών σε μια θεματική περιοχή.

Μετανάλυση: Η μετανάλυση είναι μια ερευνητική διαδικασία που χρησιμοποιείται για τη συστηματική σύνθεση ή συγχώνευση των ευρημάτων μεμονωμένων, ανεξάρτητων μελετών, χρησιμοποιώντας στατιστικές μεθόδους για τον υπολογισμό ενός συνολικού ή "απόλυτου" αποτελέσματος. Η μετανάλυση δεν συγκεντρώνει απλώς δεδομένα από μικρότερες μελέτες για την επίτευξη μεγαλύτερου μεγέθους δείγματος. Οι αναλυτές χρησιμοποιούν καλά αναγνωρισμένες, συστηματικές μεθόδους για να λάβουν υπόψη τους τις διαφορές στο μέγεθος του δείγματος, τη μεταβλητότητα (ετερογένεια) στην προσέγγιση των μελετών και τα ευρήματα (αποτελέσματα της θεραπείας) και να ελέγξουν πόσο ευαίσθητα είναι τα αποτελέσματά τους στο δικό τους πρωτόκολλο συστηματικής ανασκόπησης (επιλογή μελετών και στατιστική ανάλυση).

Ποιοτική συστηματική ανασκόπηση: Η ποιοτική συστηματική ανασκόπηση συγκεντρώνει την έρευνα για ένα θέμα, αναζητώντας συστηματικά ερευνητικά στοιχεία από πρωτογενείς ποιοτικές μελέτες και συγκεντρώνοντας τα ευρήματα.

Ταχεία ανασκόπηση: Οι ταχείες ανασκοπήσεις είναι μια μορφή σύνθεσης γνώσης που ακολουθεί τη διαδικασία της συστηματικής ανασκόπησης, αλλά τα στοιχεία της διαδικασίας απλοποιούνται ή παραλείπονται για την έγκαιρη παραγωγή πληροφοριών.

Συστηματική ανασκόπηση: Οι συστηματικές ανασκοπήσεις, όπως υποδηλώνει το όνομά τους, περιλαμβάνουν συνήθως ένα λεπτομερές και ολοκληρωμένο σχέδιο και μια στρατηγική αναζήτησης που προκύπτει εκ των προτέρων, με στόχο τη μείωση της μεροληψίας μέσω του εντοπισμού, της αξιολόγησης και της σύνθεσης όλων των σχετικών μελετών για ένα συγκεκριμένο θέμα.

Συστηματική αναζήτηση και ανασκόπηση: Αυτή η μέθοδος συνδυάζει μεθόδους από μια "κριτική ανασκόπηση" με μια ολοκληρωμένη διαδικασία αναζήτησης. Αυτός ο τύπος ανασκόπησης χρησιμοποιείται συνήθως για την αντιμετώπιση γενικών ερωτημάτων για την παραγωγή της καταλληλότερης σύνθεσης των αποδεικτικών στοιχείων. Μπορεί να περιλαμβάνει αξιολόγηση της ποιότητας των πηγών (αξιολόγηση της ποιότητας) των δεδομένων.

Συστηματική ανασκόπηση: Η συστηματική ανασκόπηση διαφέρει από τη συστηματική ανασκόπηση, καθώς πληροί τις μεθοδολογικές απαιτήσεις. Ενώ περιλαμβάνονται ορισμένα από τα στοιχεία μιας συστηματικής ανασκόπησης, μια συστηματοποιημένη ανασκόπηση περιλαμβάνει γενικά μόνο έναν κριτή, μπορεί να μην περιλαμβάνει εκτεταμένη αναζήτηση σε πολλαπλές βάσεις δεδομένων και μπορεί να αποκλείει τις αξιολογήσεις του κινδύνου μεροληψίας.

3.4 Στάδια Συστηματικής Ανασκόπησης

1) Ελέγχουμε για υπάρχουσες αναθεωρήσεις/πρωτόκολλα: Χρειάζεται ακόμη η επανεξέταση; Μπορούμε να αλλάξουμε την ερώτηση ώστε να απαντήσουμε σε διαφορετικό τροποποιημένο ερώτημα; Έγινε καλά η ανασκόπηση; Πότε έγινε; Υπήρξε εξέλιξη στην έρευνα από τότε; Είναι αρκετά ευρεία;

2) Διατυπώνουμε το ερευνητικό ερώτημα, ένα σαφές, σαφώς καθορισμένο ερευνητικό ερώτημα κατάλληλης εμβέλειας. Καθορίζουμε την ορολογία σας. Βρίσκουμε υπάρχουσες ανασκοπήσεις για το θέμα μας, ώστε να ενημερωθούμε για την ανάπτυξη του ερευνητικού σας ερωτήματος, να εντοπίσουμε τα κενά και να επιβεβαιώσουμε ότι δεν επαναλαμβάνουμε τις προσπάθειες προηγούμενων ανασκοπήσεων. Εξετάζουμε το ενδεχόμενο χρήσης ενός πλαισίου για τον καθορισμό του πεδίου εφαρμογής του ερωτήματός.

3) Καθορίζουμε τα κριτήρια συμπερίληψης και αποκλεισμού. Αυτό είναι επίσης γνωστό ως δημιουργία πρωτοκόλλου ανασκόπησης. Δηλώνουμε με σαφήνεια τα κριτήρια που θα χρησιμοποιήσουμε για να καθορίσουμε εάν μια μελέτη θα συμπεριληφθεί ή όχι στην έρευνά μας. Εξετάζουμε τους πληθυσμούς

των μελετών, το σχεδιασμό των μελετών, τους τύπους παρέμβασης, τις ομάδες σύγκρισης, τα μετρούμενα αποτελέσματα.

4) Κάνουμε αναζήτηση μελετών. Εκτελούμε τις αναζητήσεις μας στις βάσεις δεδομένων που έχουμε προσδιορίσει ως σχετικές με το θέμα μας. Προσεγγίζουμε τη γκρίζα βιβλιογραφία μεθοδικά και στοχευμένα. Συγκεντρώνουμε όλες τις ανακτημένες εγγραφές από κάθε αναζήτηση σε έναν διαχειριστή αναφοράς, όπως το Endnote, και υποδιπλασιάζουμε τη βιβλιοθήκη πριν από τη διαλογή.

5) Επιλέγουμε μελέτες για συμπερίληψη βάσει προκαθορισμένων κριτηρίων. Ξεκινάμε με έναν έλεγχο τίτλου/περιλήψεων για να αφαιρέσουμε τις μελέτες που σαφώς δεν σχετίζονται με το θέμα μας. Χρησιμοποιούμε τα κριτήρια συμπερίληψης/αποκλεισμού μας για να ελέγξουμε το πλήρες κείμενο των μελετών. Συνιστάται ιδιαίτερα να ελέγχουν όλες τις μελέτες δύο ανεξάρτητοι κριτές, επιλύοντας τις περιοχές διαφωνίας με συναίνεση.

6) Εξάγουμε δεδομένα από τις μελέτες που περιλαμβάνονται. Χρησιμοποιούμε ένα λογιστικό φύλλο ή λογισμικό συστηματικής ανασκόπησης για να εξάγουμε όλα τα σχετικά δεδομένα από κάθε μελέτη που περιλαμβάνεται. Συνιστάται να δοκιμάσουμε πιλοτικά το εργαλείο εξαγωγής δεδομένων, για να καθορίσουμε εάν θα πρέπει να συμπεριληφθούν άλλα πεδία ή να διευκρινιστούν τα υπάρχοντα πεδία.

7) Αξιολογούμε τον κίνδυνο μεροληψίας των συμπεριλαμβανόμενων μελετών. Χρησιμοποιούμε ένα εργαλείο κινδύνου μεροληψίας (όπως το Cochrane RoB Tool) για να αξιολογήσουμε τις πιθανές μεροληψίες των μελετών όσον αφορά το σχεδιασμό της μελέτης και άλλους παράγοντες. Μπορούμε να προσαρμόσουμε τα υπάρχοντα εργαλεία ώστε να ανταποκρίνονται καλύτερα στις ανάγκες της ανασκόπησής μας, ανάλογα με τους τύπους των μελετών που περιλαμβάνονται.

8) Παρουσιάζουμε τα αποτελέσματα και αξιολογήστε την ποιότητα των αποδεικτικών στοιχείων. Παρουσιάζουμε με σαφήνεια τα ευρήματά μας, συμπεριλαμβανομένης της λεπτομερούς μεθοδολογίας (όπως οι στρατηγικές αναζήτησης που χρησιμοποιήθηκαν, τα κριτήρια επιλογής κ.λπ.), έτσι ώστε η ανασκόπησή μας να μπορεί εύκολα να επικαιροποιηθεί στο μέλλον με νέα

ερευνητικά ευρήματα. Πραγματοποιούμε μετά-ανάλυση, εάν το επιτρέπουν οι μελέτες. Παρέχουμε συστάσεις για την πρακτική και τη χάραξη πολιτικής, εάν υπάρχουν επαρκή και υψηλής ποιότητας στοιχεία, ή μελλοντικές κατευθύνσεις για την έρευνα ώστε να καλυφθούν τα υπάρχοντα κενά στη γνώση ή να ενισχυθεί το σύνολο των στοιχείων.

3.5 Παρουσίαση Αποτελεσμάτων

Πίνακας 1 – Συστηματική Ανασκόπηση άρθρων μελέτης

ΣΥΓΓΡΑΦΕΑΣ	ΗΜΕΡΟΜΗΝΙΑ	ΤΙΤΛΟΣ	ΣΚΟΠΟΣ	ΣΥΜΠΕΡΑΜΑ
Al-Masalha et al.	2020	Cyber-Crime Effect on Jordanian Society	Να προσδιορίσει την έννοια της συντακτικής επίθεσης και τους τύπους της, οι οποίοι οδηγούν σε Cyber-Bullying επίθεση – και την επίδραση της διείσδυσης συσκευών στην Ιορδανική κοινωνία.	Όπως φαίνεται στο αποτέλεσμα, τα πιο συνηθισμένα ηλεκτρονικά εγκλήματα είναι εγκλήματα που σχετίζονται με τη συκοφαντία, τις απειλές και τον εκβιασμό. Τα αποτελέσματα δείχνουν επίσης την επικράτηση του εγκλήματος στον κυβερνοχώρο σε πυκνοκατοικημένες περιοχές. Αντίθετα, το έγκλημα στον κυβερνοχώρο μειώνεται στις περιοχές όπου ζουν φυλές και φυλές. Ταυτόχρονα, οι ηλικιακές ομάδες που εκτίθενται σε ηλεκτρονικά εγκλήματα είναι οι ηλικίες 18-29 και 30-

				44 ετών.
Babanina et al.	2021	Cybercrime: History of formation, current state and ways of counteraction	Το άρθρο εξετάζει την ιστορία της εμφάνισης και της ανάπτυξης του κυβερνοεγκλήματος, τις ιδιαιτερότητες της τρέχουσας κατάστασης στην κοινωνία, η οποία συμβάλλει στην αύξηση του αριθμού των εγκλημάτων στον κυβερνοχώρο και των τρόπων καταπολέμησης του ηλεκτρονικού εγκλήματος.	Με βάση την ανάλυση, αναπτύχθηκαν οι κύριες κατευθύνσεις για την καταπολέμηση του κυβερνοεγκλήματος και την πρόληψη της αύξησης του αριθμού των εγκλημάτων στον κυβερνοχώρο στην κοινωνία.
Bose et al.	2020	Explaining AI for Malware Detection: Analysis of Mechanisms of MalConv	Το άρθρο εισάγει ένα πλαίσιο που παρεμβαίνει μεταξύ δειγμάτων διαφορετικών τάξεων σε διαφορετικά στρώματα για να δούμε πώς μια αρχιτεκτονική βαθιάς δικτύωσης γενικεύεται σε δείγματα που δεν είναι στο σύνολο εκπαίδευσης, εξηγώντας τα αποτελέσματα των βαθιών δικτύων σε δοκιμές πραγματικού κόσμου.	Χρησιμοποιώντας αυτό το πλαίσιο, γίνεται προσπάθεια να απομυθοποιηθούν οι μηχανισμοί πίσω από την αρχιτεκτονική MalConv αναλύοντας τα βάρη και τις κλίσεις των πολλαπλών στρωμάτων στην αρχιτέκτονά της και να αποκρυπτογραφήθουν αυτό που μαθαίνει η αρχιτεχνία αναλύοντάς τα ακατέργαστα bytes από το δυαδικό. Για αυτή την αρχιτεκτονική, η ανάλυσή δείχνει ότι το δίκτυο εκχωρεί πολύ υψηλότερα βάρη σε συγκεκριμένα μέρη του εκτελέσιμου Υποδεικνύοντας ότι αυτά τα μέρη συμβάλλουν σημαντικά περισσότερο στην ταξινόμηση από άλλα μέρη των εκτελεστών. Μέσω του προτεινόμενου πλαισίου, εξηγούν τους μηχανισμούς πίσω από τους

				αλγόριθμους μηχανικής μάθησης και να γίνεται καλύτερη επεξήγηση των αποφάσεων τους.
Buch et al.	2017	World of Cyber Security and Cybercrime	Αυτό το έγγραφο παρέχει λεπτομερείς πληροφορίες σχετικά με την κυβερνοασφάλεια και το έγκλημα στον κυβερνοχώρο. Περιλαμβάνει τύπους ασφάλειας στον κυβερνοχώρο, την ανάγκη για κυβερνοασφάλεια, ζητήματα στον τομέα της ψηφιακής ασφάλειας, τα πλεονεκτήματα και τα μειονεκτήματά της, το ιστορικό του κυβερνοεγκλήματος, τα είδη του εγκλήματος στον κυβερνοχώρο.	Κάθε έξυπνη συσκευή που μπορεί να μεταφέρει δεδομένα σε μία ή περισσότερες άλλες συσκευές (είτε μέσω ενός δικτύου είτε όχι) εμπίπτει στο πεδίο εφαρμογής της ασφάλειας στον κυβερνοχώρο, η οποία περιλαμβάνει σχεδόν όλο το θεμέλιο της σύγχρονης κοινωνίας. Όλοι πρέπει να είναι ενήμεροι για την ασφάλεια στον κυβερνοχώρο, καθώς και για τα κυβερνοεγκλήματα και τις αιτίες τους.
Connolly & Wall	2019	The rise of crypto-ransomware in a changing cybercrime landscape: Taxonomising countermeasures	Αυτό το άρθρο διερευνά εμπειρικά πώς οι οργανισμοί και οι ερευνητές έχουν αντιδράσει στην αλλαγή στο τοπίο του run-somware από τις επιθέσεις scareware και locker στην σχεδόν αποκλειστική χρήση του crypto-ransomware.	Τα ευρήματα της έρευνας δείχνουν ότι οι απαντήσεις στο crypto-ransomware γίνονται πιο περίπλοκες από τη λεπτή σχέση μεταξύ των τεχνικών (malware που κρυπτογραφεί) και των ανθρώπινων (κοινωνική μηχανική που εξακολουθεί να προκαλεί τις περισσότερες λοιμώξεις) πτυχές μιας επίθεσης. Ως αποτέλεσμα, δεν υπάρχει απλή τεχνολογική σφαίρα που θα εξαλείψει την απειλή crypto-ransomware. Αντίθετα, απαιτείται μια πολύ επίπεδη προσέγγιση, η οποία αποτελείται από κοινωνικό-τεχνικά

				μέτρα, ζήλο των διευθυντών της πρώτης γραμμής και την ενεργό υποστήριξη από την ανώτερη διοίκηση.
Coull & Gardner	2019	Activation Analysis of a Byte-Based Deep Neural Network for Malware Classification	Σε αυτό το άρθρο, εξετάστηκε τι μαθαίνουν τα μοντέλα μηχανικής μάθησης για το κακόβουλο λογισμικό.	Μέσα από τα αποτελέσματά, εντοπίστηκαν αρκετά ενδιαφέροντα χαρακτηριστικά που έμαθε το μοντέλο και τη σύνδεσή τους με χαρακτηριστικά με μη αυτόματο τρόπο που χρησιμοποιούνται συνήθως από τα παραδοσιακά μοντέλα μηχανικής μάθησης. Επιπλέον, εξετάστηκε ο αντίκτυπος του όγκου των δεδομένων κατάρτισης και της κοινωνικοποίησης στην ποιότητα των χαρακτηριστικών που μαθαίνουν και την αποτελεσματικότητα των ταξινομητών, αποκαλύπτοντας την κάπως παράδοξη αντίληψη ότι η καλύτερη γενίκευση δεν οδηγεί απαραίτητα σε καλύτερη απόδοση για ταξινομητές κακόβουλου λογισμικού που βασίζονται σε byte.
Faruk et al.	2021	Malware Detection and Prevention using Artificial Intelligence Techniques	Σε αυτή τη μελέτη, τονίζονται οι τεχνικές που βασίζονται στην τεχνητή νοημοσύνη (AI) για την ανίχνευση και την πρόληψη της δραστηριότητας κακόβουλου λογισμικού	Η μελέτη δείχνει ότι η υιοθέτηση φουτουριστικών προσεγγίσεων για την ανάπτυξη εφαρμογών ανίχνευσης κακόβουλου λογισμικού θα προσφέρει σημαντικά πλεονεκτήματα. Η κατανόηση αυτής της σύνθεσης θα βοηθήσει τους ερευνητές για

				περαιτέρω έρευνα σχετικά με την ανίχνευση και την πρόληψη κακόβουλου λογισμικού χρησιμοποιώντας την τεχνητή νοημοσύνη.
Hayward & Mass	2021	Artificial intelligence and crime: A primer for criminologists	Αυτό το άρθρο εισάγει την έννοια της τεχνητής νοημοσύνης (AI) σε ένα εγκληματολογικό ακροατήριο. Σε όλο το έγγραφο, αναπτύσσεται μια σειρά προγραμματικών παραδειγμάτων που, συλλογικά, ελπίζει να χρησιμεύσουν ως ένα χρήσιμο πρότυπο ΤΝ για τους εγκληματολόγους που ενδιαφέρονται για το «σύνδεσμο τεχνολογίας-εγκλήματος».	Η σύνοψη του επιχειρήματος του άρθρου είναι ότι για όλη τη χρησιμότητά του, η ΤΝ δεν είναι μαγεία. Όπως κάθε πρόγραμμα που βασίζεται σε δεδομένα, η αντικειμενικότητα και η αποτελεσματικότητά του εξακολουθούν να καθορίζονται από το παλιό υπολογιστικό αξίωμα «GIGO». Πράγματι, δεδομένης της ουσιαστικά εύθραυστης φύσης της λήψης αποφάσεων του νευρικού δικτύου, είναι σαφές ότι η λεπτομερής ανθρώπινη εμπειρία είναι ακόμη πιο σημαντική στον υπολογισμό σήμερα από ό, τι ήταν ποτέ. τόσο σε σχέση με το πλαίσιο παραμέτρων / υποθέσεων όσο και τη συνολική διακυβέρνηση του συστήματος.
Kemp et al.	2021	Empty Streets, Busy Internet: A Time-Series Analysis of Cybercrime and Fraud Trends During COVID-19	Αυτό το άρθρο εφαρμόζει ανάλυση σειράς χρόνου σε ιστορικά δεδομένα σχετικά με το έγκλημα στον κυβερνοχώρο και την απάτη που αναφέρθηκαν στο Action Fraud στο Ηνωμένο Βασίλειο για να εξετάσει εάν τυχόν πιθανές αυξήσεις είναι πέρα	Τα αποτελέσματα δείχνουν ότι ενώ τόσο το συνολικό έγκλημα στον κυβερνοχώρο όσο και η συνολική απάτη αυξήθηκαν πέρα από τα προβλεπόμενα επίπεδα, οι αλλαγές στη θυματοποίηση δεν ήταν ομοιογενείς μεταξύ των τύπων απάτης και των

			από την κανονική μεταβλητότητα του εγκλήματος.	θυμάτων.
Zaegar et al.	2013	A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks	Σε αυτό το άρθρο, εξετάζεται το πεδίο εφαρμογής του προβλήματος των επιθέσεων πλημμύρας DDoS και τις προσπάθειες για την καταπολέμησή του. Κατηγορεί τις επιθέσεις πλημμύρας DDoS και ταξινομεί τα υφιστάμενα αντίμετρα ανάλογα με το πού και πότε αποτρέπουν, ανιχνεύουν και ανταποκρίνονται σε αυτές. Επιπλέον, υπογραμμίζει την ανάγκη για μια ολοκληρωμένη κατανεμημένη και συνεργατική προσέγγιση στην άμυνα.	Ένας ιδανικός ολοκληρωμένος μηχανισμός άμυνας DDoS πρέπει να διαθέτει συγκεκριμένα χαρακτηριστικά για την καταπολέμηση των επιθέσεων πλημμύρας DdoS τόσο σε πραγματικό χρόνο όσο και όσο το δυνατόν πιο κοντά στις πηγές της επίθεσης.

Κεφάλαιο 4ο Εφαρμογές της τεχνητής νοημοσύνης στην άμυνα κατά των εγκλημάτων στον κυβερνοχώρο

Με τον ολοένα αυξανόμενο ρυθμό και όγκο των εγκλημάτων στον κυβερνοχώρο, η ανθρώπινη παρέμβαση έχει αποδειχθεί ανεπαρκής για την αποτελεσματική και έγκαιρη καταπολέμηση αυτών των επιθέσεων. Δεδομένου ότι οι επιθέσεις πραγματοποιούνται από ευφυείς πράκτορες, όπως οι ιοί και τα σκουλήκια υπολογιστών, υπάρχει ανάγκη ανάπτυξης τέτοιων ευφυών πρακτόρων που πρέπει να είναι σε θέση να αναλύουν και να ανταποκρίνονται σε αυτές τις επιθέσεις γρήγορα και αποτελεσματικά. Οι δυνάμεις που δημιουργούνται από υπολογιστή θα είναι υπεύθυνες για τον εντοπισμό, την ανάλυση και την αντιμετώπιση των επιθέσεων στον κυβερνοχώρο, καθώς και για τον σχεδιασμό των στρατηγικών για την αποτροπή τυχόν δευτερογενών επιθέσεων (Radulov, 2019).

Επιπλέον, τα εγκλήματα στον κυβερνοχώρο δεν περιορίζονται σε μια συγκεκριμένη περιοχή, αλλά καλύπτουν όλους τους υπολογιστές που είναι συνδεδεμένοι με το διαδίκτυο σε ολόκληρο τον κόσμο, τα εγκλήματα στον κυβερνοχώρο δεν έχουν μόνιμο τόπο. Στον σημερινό κόσμο της πληροφορικής, είναι πολύ εύκολο για ένα άτομο να έχει πρόσβαση σε γνώσεις και εργαλεία για τη διάπραξη τέτοιων εγκλημάτων χωρίς την παρέμβαση οποιουδήποτε ειδικού. Οι συμβατικοί αλγόριθμοι προγραμματισμού έχουν καταστεί αναποτελεσματικοί στην καταπολέμηση αυτών των δυναμικά εξελισσόμενων εγκλημάτων στον κυβερνοχώρο.

Ως εκ τούτου, υπάρχει ανάγκη για καινοτόμες τεχνικές τεχνητής νοημοσύνης που μπορούν να αναλύουν και να ανταποκρίνονται γρήγορα και έγκαιρα σε αυτά τα εγκλήματα στον κυβερνοχώρο. Η τεχνητή νοημοσύνη προσφέρει αυτές και διάφορες άλλες δυνατότητες. Πολυάριθμες μέθοδοι υπολογισμού της τεχνητής νοημοσύνης, δηλαδή τεχνητά νευρωνικά δίκτυα, βαθιά μάθηση, υπολογιστική νοημοσύνη, ευφυείς πράκτορες και τεχνητά ανοσοποιητικά συστήματα, μηχανική μάθηση, αναγνώριση προτύπων, ασαφής λογική κ.λπ. έχουν κερδίσει έδαφος στην ανίχνευση και πρόληψη του εγκλήματος στον κυβερνοχώρο.

Η τεχνητή νοημοσύνη μας επιτρέπει να σχεδιάζουμε λογισμικό και μηχανές που μπορούν να σκέφτονται όπως ο άνθρωπος, να ενεργούν όπως ο άνθρωπος, να σκέφτονται ορθολογικά και να ενεργούν ορθολογικά για να λάβουν την κατάλληλη απόφαση. Η τεχνητή νοημοσύνη είναι ένας τομέας ευκαιριών για την εξεύρεση πολύ πιθανών τεχνικών καταπολέμησης των δυνητικά επικίνδυνων επιθέσεων στον κυβερνοχώρο (Radulov, 2019).

4.1 Εφαρμογές Τεχνητού Νευρωνικού Δικτύου

Το Τεχνητό Νευρωτικό Δίκτυο είναι ένας υπολογιστικός μηχανισμός που προσομοιώνει δομικές και λειτουργικές πτυχές των νευρωνικών δικτύων που υπάρχουν στα βιολογικά νευρικά συστήματα. Είναι ιδανικά για καταστάσεις που απαιτούν πρόβλεψη, ταξινόμηση ή έλεγχο σε δυναμικά και πολύπλοκα υπολογιστικά περιβάλλοντα (Hou et al., 2016).

4.2 Εφαρμογές ευφυούς πράκτορα

Οι ευφυείς πράκτορες είναι αυτόνομες δυνάμεις που δημιουργούνται από υπολογιστή και επικοινωνούν μεταξύ τους για να μοιράζονται δεδομένα και να συνεργάζονται μεταξύ τους, προκειμένου να σχεδιάζουν και να εφαρμόζουν κατάλληλες αντιδράσεις σε περίπτωση απροσδόκητων γεγονότων. Η κινητικότητά τους και η προσαρμοστικότητά τους στα περιβάλλοντα στα οποία αναπτύσσονται, καθώς και η συνεργατική τους φύση, καθιστούν την τεχνολογία των ευφυών πρακτόρων κατάλληλη για την καταπολέμηση των επιθέσεων στον κυβερνοχώρο.

4.3 Εφαρμογές τεχνητού ανοσοποιητικού συστήματος

Τα τεχνητά ανοσοποιητικά συστήματα, όπως και τα βιολογικά ανοσοποιητικά συστήματα στα οποία βασίζονται, χρησιμοποιούνται για να διατηρούν τη σταθερότητα σε ένα μεταβαλλόμενο περιβάλλον. Η ανίχνευση εισβολής με βάση το ανοσοποιητικό σύστημα περιλαμβάνει την εξέλιξη των ανοσοκυττάρων (αυτοανοχή, κλώνος, παραλλαγή κ.λπ.) και την ταυτόχρονη ανίχνευση αντιγόνων. Ένα ανοσοποιητικό σύστημα παράγει αντισώματα για να αντισταθεί στα παθογόνα και η ένταση της εισβολής μπορεί να εκτιμηθεί από τη μεταβολή της συγκέντρωσης των αντισωμάτων. Ως εκ τούτου, τα AIS διαδραματίζουν σημαντικό ρόλο στην έρευνα για την ασφάλεια στον κυβερνοχώρο.

Οι Sirisanyalak και Sornil (2007) παρουσίασαν μια προσέγγιση εξαγωγής χαρακτηριστικών ηλεκτρονικού ταχυδρομείου βάσει AIS για την ανίχνευση ανεπιθύμητων μηνυμάτων. Τα αποτελέσματα της αξιολόγησης των επιδόσεων έδειξαν ότι η προτεινόμενη μέθοδος είναι πολύ πιο αποτελεσματική στην ανίχνευση ανεπιθύμητης αλληλογραφίας από άλλα υπάρχοντα συστήματα, με πολύ χαμηλά ποσοστά ψευδώς θετικών και ψευδώς αρνητικών αποτελεσμάτων (0,91% και 1,95% αντίστοιχα).

4.4 Εφαρμογές γενετικού αλγόριθμου και ασαφών συνόλων

Οι Kim et al. (2004) πρότειναν έναν αλγόριθμο μάθησης για ανιχνευτές ανωμαλιών, ο οποίος μπορεί να ανιχνεύσει επιθέσεις χρησιμοποιώντας γενετικό αλγόριθμο. Εφάρμοσαν τον αλγόριθμό τους σε ένα τεχνητό σύστημα ασφάλειας υπολογιστών και έδειξαν την αποτελεσματικότητά του στην ανίχνευση εισβολών.

Ο Sekeh (2009) πρότειναν ένα ασαφές σύστημα ανίχνευσης εισβολών με βάση τον κεντρικό υπολογιστή που χρησιμοποιεί την τεχνική εξόρυξης δεδομένων και τις υπηρεσίες των κλήσεων του υποκείμενου λειτουργικού συστήματος. Τα αποτελέσματα προσομοίωσης έδειξαν ότι το προτεινόμενο σύστημα βελτιώνει την

απόδοση και μειώνει το μέγεθος της βάσης δεδομένων, τη χρονική πολυπλοκότητα και το ποσοστό ψευδών συναγερμών.

Οι Mabu et al. (2011) περιέγραψαν μια νέα ασαφή μέθοδο ανίχνευσης εισβολών σε δίκτυα που βασίζεται στην εξόρυξη κανόνων συσχέτισης κλάσεων σε γενετικό προγραμματισμό δικτύων. Η προτεινόμενη μέθοδος είναι ευέλικτη και αποτελεσματική τόσο για την ανίχνευση κακής χρήσης όσο και για την ανίχνευση ανωμαλιών σε δίκτυα και είναι ικανή να αντιμετωπίσει τις μικτές βάσεις δεδομένων που περιέχουν τόσο διακριτά όσο και συνεχή χαρακτηριστικά για την εξόρυξη σημαντικών κανόνων συσχέτισης κλάσεων που απαιτούνται για τη βελτιωμένη ανίχνευση εισβολών. Τα πειράματα και η αξιολόγηση της προτεινόμενης μεθόδου έδειξαν ότι η προσέγγιση αυτή παρέχει ανταγωνιστικά υψηλά ποσοστά ανίχνευσης σε σύγκριση με άλλες τεχνικές μηχανικής μάθησης.

4.5 Άλλες εφαρμογές ΤΝ

Οι Machado et al. (2005) παρουσίασαν ένα νέο μοντέλο ανίχνευσης εισβολών στο δίκτυο που βασίζεται στην τεχνολογία των κινητών ευφυών πρακτόρων και των AIS. Υλοποίησαν επίσης το σχέδιό τους και έδειξαν ότι είναι ικανό να διακρίνει μεταξύ διαφόρων επιθέσεων, παραβιάσεων ασφαλείας και διαφόρων άλλων παραβιάσεων ασφαλείας. Τα πειραματικά αποτελέσματα έδειξαν ότι το μοντέλο τους προσφέρει σημαντική αναβάθμιση σε σύγκριση με προηγούμενες εργασίες στον τομέα.

Οι Pei και Song (2008) επικεντρώθηκαν στη βελτίωση της απόδοσης των ανιχνευτών εισβολής των IDS, οπότε πρότειναν μια υβριδική προσέγγιση η οποία χρησιμοποιεί την απόδοση αναζήτησης του ανοσοποιητικού αλγορίθμου για τη δημιουργία ασαφών ανιχνευτών. Τα πειράματα έδειξαν τη μεγάλη ικανότητα αναζήτησης του ανοσοποιητικού αλγορίθμου. Τα αποτελέσματα έδειξαν επίσης ότι οι ασαφείς κανόνες ανίχνευσης μειώνουν την ευθραυστότητα των ανιχνευτών και βελτιώνουν την ακρίβεια ανίχνευσης.

4.6 Πλεονεκτήματα των εφαρμογών τεχνητής νοημοσύνης σε σύστημα ανίχνευσης και αποτροπής εισβολών (IDPS)

Οι εφαρμογές τεχνητής νοημοσύνης εισάγουν πολυάριθμα πλεονεκτήματα στην ανίχνευση και την πρόληψη εισβολών (βλ. Πίνακα 1).

Πίνακας 1. Πλεονεκτήματα που επιφέρουν ορισμένες εφαρμογές τεχνητής νοημοσύνης στην ανίχνευση και πρόληψη εισβολών.

Τεχνητά νευρωνικά δίκτυα	Παραλληλισμός στην επεξεργασία πληροφοριών, Μάθηση μέσω παραδείγματος, Μη γραμμικότητα - χειρισμός σύνθετων μη γραμμικών συναρτήσεων, Υπεροχή έναντι πολύπλοκων και αινιγματικών διαφορικών εξισώσεων, Ανθεκτικότητα σε θόρυβο και ελλιπή δεδομένα, Ευελιξία και ευελιξία με μοντέλα μάθησης, Διαισθητικότητα - καθώς αποτελούν αφαίρεση των βιολογικών νευρωνικών δικτύων (Bitter, Elizondo & Watson, 2010)
Ευφυείς πράκτορες	Κινητικότητα, Βοηθητικότητα - προσπαθούν πάντα να φέρουν εις πέρας τα καθήκοντά τους έχοντας αντιφατικούς στόχους, Ορθολογισμός - για την επίτευξη των στόχων τους, Προσαρμοστικότητα - στο περιβάλλον και στις προτιμήσεις των χρηστών, Συνεργασία - επίγνωση ότι ο ανθρώπινος χρήστης μπορεί να κάνει λάθη και να παρέχει αβέβαιες ή να παραλείπει σημαντικές πληροφορίες- συνεπώς, δεν πρέπει να δέχονται οδηγίες χωρίς να εξετάζουν και να ελέγχουν τις ασυνέπειες με τον χρήστη (Nogueira, 2006).
Τεχνητά ανοσοποιητικά συστήματα	Δυναμική δομή, Παραλληλισμός και κατανεμημένη μάθηση - χρήση επικοινωνιών δικτύου δεδομένων και παραλληλισμού σε εργασίες ανίχνευσης και εξάλειψης, Αυτοπροσαρμοστικότητα και αυτοοργάνωση - ενημέρωση των σημάτων εισβολής χωρίς ανθρώπινη συμμετοχή- ανθεκτικότητα, Επιλεκτική απόκριση - απομάκρυνση της κακόβουλης δραστηριότητας με τα καλύτερα διαθέσιμα μέσα- Ποικιλομορφία - κάθε κόμβος ανιχνευτή παράγει ένα στατιστικά μοναδικό σύνολο μη αυτοανιχνευτών, Βελτιστοποίηση πόρων, Πολυεπίπεδη δομή - οι επιτιθέμενοι δεν μπορούν να πετύχουν τις κακόβουλες

	δραστηριότητές τους παρακάμπτοντας μόνο ένα επίπεδο, δεδομένου ότι πολλαπλά επίπεδα διαφορετικών δομών είναι επιφορτισμένα με την παρακολούθηση ενός και μόνο σημείου. Αναλωσιμότητα - μη εξάρτηση από ένα μόνο συστατικό το οποίο μπορεί εύκολα να αντικατασταθεί από άλλα συστατικά (EshghiShargh, 2009).
Γενετικοί Αλγόριθμοι	Ανθεκτικότητα, Προσαρμοστικότητα στο περιβάλλον, Βελτιστοποίηση - παροχή βέλτιστων λύσεων ακόμη και για πολύπλοκα υπολογιστικά προβλήματα, Παραλληλισμός - επιτρέπει την αξιολόγηση πολλαπλών σχημάτων ταυτόχρονα, Ευέλικτη και στιβαρή συνολική αναζήτηση (Aziz et al., 2012).
Ασαφή σύνολα	Διαλειτουργικότητα - φιλικότητα προς τον άνθρωπο (Stahl et al., 2010).

4.7 Περιορισμοί τρεχουσών συστημάτων ανίχνευσης/ αποτροπής ανωμαλιών

Παρόλο που τα συστήματα ανίχνευσης ανωμαλιών προσφέρουν την ευκαιρία να ανιχνεύσουν άγνωστες προηγουμένως επιθέσεις, έχουν ορισμένους σημαντικούς περιορισμούς που πρέπει να αντιμετωπιστούν. Το κύριο ζήτημα είναι η δυσκολία δημιουργίας ενός σταθερού μοντέλου για το ποια είναι η αποδεκτή συμπεριφορά και ποια η επίθεση**,** ως εκ τούτου, ενδέχεται να δίνουν μεγάλο αριθμό ψευδώς θετικών συναγερμών, οι οποίοι μπορεί να οφείλονται σε άτυπη συμπεριφορά που στην πραγματικότητα είναι φυσιολογική περιλαμβάνουν και εξουσιοδοτημένη, δεδομένου ότι η φυσιολογική συμπεριφορά μπορεί εύκολα και εύκολα να αλλάξει (Patel et al., 2013).

4.7.1 Ορισμός περιορισμών

Προκειμένου το σύστημα ανίχνευσης ανωμαλιών να είναι σε θέση να χαρακτηρίσει τα κανονικά μοτίβα και να δημιουργήσει ένα μοντέλο της κανονικής συμπεριφοράς, απαιτούνται ευρύτατα σύνολα εκπαίδευσης των κανονικών δραστηριοτήτων του συστήματος. Οποιαδήποτε αλλαγή στα κανονικά πρότυπα του συστήματος πρέπει να οδηγεί σε αναγκαία ενημέρωση της βάσης γνώσης. Εάν το σύστημα ανίχνευσης και πρόληψης ταξινομήσει ανακριβώς μια νόμιμη δραστηριότητα ως κακόβουλη, τα αποτελέσματα μπορεί να είναι πολύ ατυχή, καθώς θα επιχειρήσει να σταματήσει τη δραστηριότητα ή να την αλλάξει (Patel et al., 2013).

Ένα σύστημα ανίχνευσης εισβολών, όσο αποτελεσματικό και αν είναι, μπορεί να απενεργοποιηθεί από τους επιτιθέμενους, αν μπορέσουν να μάθουν πώς λειτουργεί το σύστημα. Σε ετερογενή περιβάλλοντα υπάρχει επίσης το ζήτημα της ενσωμάτωσης πληροφοριών από διαφορετικές τοποθεσίες. Ένα άλλο πρόβλημα αφορά την προμήθεια συστημάτων ανίχνευσης εισβολών που θα συμμορφώνονται με νομικούς κανονισμούς, απαιτήσεις ασφαλείας ή/και συμφωνίες επιπέδου υπηρεσιών στον πραγματικό κόσμο (Bitter, Elizondo & Watson, 2010).

4.7.2 Προσδιορισμός διαδικασίας ανίχνευσης και αποτροπής

Η ασφάλεια στον κυβερνοχώρο χρειάζεται πολύ μεγαλύτερη προσοχή. Δεδομένων των ανθρώπινων περιορισμών και του γεγονότος ότι παράγοντες όπως οι ιοί υπολογιστών και τα σκουλήκια είναι ευφυείς, τα δικτυοκεντρικά περιβάλλοντα απαιτούν ευφυείς κυβερνοαισθητήρες (ή δυνάμεις που δημιουργούνται από υπολογιστές) οι οποίοι θα ανιχνεύουν, θα αξιολογούν και θα ανταποκρίνονται εγκαίρως στις κυβερνοεπιθέσεις (Stytz, Lichtblau & Banks, 2005).

Η εφαρμογή τεχνικών τεχνητής νοημοσύνης στην άμυνα στον κυβερνοχώρο θα χρειαστεί σχεδιασμό και μελλοντική έρευνα. Μία από τις προκλήσεις είναι η διαχείριση της γνώσης στον δίκτυο κεντρικό πόλεμο, επομένως ένας πολλά

υποσχόμενος τομέας για έρευνα είναι η εισαγωγή αρθρωτής και ιεραρχικής αρχιτεκτονικής γνώσης στο λογισμικό λήψης αποφάσεων. Η ταχεία εκτίμηση της κατάστασης και η υπεροχή των αποφάσεων μπορούν να διασφαλιστούν μόνο με την αυτοματοποιημένη διαχείριση της γνώσης. Είναι επίσης προβλέψιμο ότι ο μεγάλος στόχος της έρευνας ΤΝ - ανάπτυξη τεχνητής γενικής νοημοσύνης - μπορεί να επιτευχθεί στο όχι και τόσο μακρινό μέλλον, γεγονός που θα οδηγήσει στη Singularity, η οποία περιγράφεται ως "η τεχνολογική δημιουργία νοημοσύνης εξυπνότερης από την ανθρώπινη". Παρ' όλα αυτά, είναι ζωτικής σημασίας να έχουμε τη δυνατότητα να χρησιμοποιήσουμε καλύτερη τεχνολογία ΤΝ στην άμυνα στον κυβερνοχώρο από αυτή που διαθέτουν οι δράστες (Tyugu, 2011).

Επιπλέον, πρέπει να γίνει πολύ περισσότερη έρευνα προτού καταφέρουμε να κατασκευάσουμε αξιόπιστα, αναπτυγμένα συστήματα ευφυών πρακτόρων που θα μπορούν να διαχειρίζονται κατανεμημένες υποδομές. Οι μελλοντικές εργασίες πρέπει να αναζητήσουν μια θεωρία ομαδικής συνάρτησης χρησιμότητας που θα επιτρέπει σε ομάδες πρακτόρων να λαμβάνουν αποφάσεις (Smith et al., 2006).

Κεφάλαιο 5^0 Εργαλεία τεχνητής νοημοσύνης στην αντιμετώπιση των κυβερνοεγκλημάτων

Οι Chen, Jakeman και Norton (2008) σχεδίασαν το Neuro Net - ένα σύστημα νευρωνικών δικτύων το οποίο συλλέγει και επεξεργάζεται κατανεμημένες πληροφορίες, συντονίζει τις δραστηριότητες των συσκευών του κεντρικού δικτύου, αναζητά παρατυπίες, προβαίνει σε προειδοποιήσεις και δρομολογεί αντίμετρα. Τα πειράματα έδειξαν ότι το Neuro Net είναι αποτελεσματικό ενάντια σε κατανεμημένες επιθέσεις DoS. Παρουσίασαν το σύστημα ανίχνευσης εισβολών με χρήση νευρωνικού δικτύου βασισμένου στη μοντελοποίηση (IDS-NNM), το οποίο αποδείχθηκε ικανό να ανιχνεύει όλες τις απόπειρες εισβολής στην επικοινωνία του δικτύου χωρίς να δίνει ψευδείς προειδοποιήσεις. Οι Hardy et al. (2016) σχεδίασαν

ένα IDS βασισμένο σε νευρωνικά δίκτυα που μπορεί να ανιχνεύει και να ταξινομεί άμεσα διάφορες επιθέσεις.

Ειδικοί, για την αναζήτηση απειλών στον κυβερνοχώρο μέσω ηλεκτρονικού ταχυδρομείου (e-mail) είναι το Emailage. Η τεχνολογία της Emailage έχει πλέον ενσωματωθεί στο Lexis Nexis, μια τεράστια εταιρεία κυβερνοασφάλειας και διαχείρισης κινδύνου, η οποία τη βοήθησε να αποκτήσει πρόσβαση σε δεδομένα από 30 εκατομμύρια συμβάντα απάτης ετησίως.

Οι Rowe και Lester (2020) ανέπτυξαν ένα σύστημα "counterplan" το οποίο μπορεί να αποτρέψει συγκεκριμένα σχέδια επιθέσεων στον κυβερνοχώρο χρησιμοποιώντας σχεδιασμό πολλαπλών πρακτόρων σε συνδυασμό με ορισμένες νέες μεθόδους εξαγωγής συμπερασμάτων.

Μια προσέγγιση κινητού ευφυούς συστήματος πολλαπλών πρακτόρων που βασίζεται στη σύνθεση για την καταπολέμηση των κυβερνοπαθειών και που χρησιμοποιεί διαδικασίες AI για τον εντοπισμό δραστηριότητας κακόβουλου λογισμικού είναι η CrowdStrike Falcon . Αυτό το καινοτόμο εργαλείο κυβερνοασφάλειας συνδυάζει τη χρήση πρακτόρων συλλογής δεδομένων σύντομα με μια μηχανή ανάλυσης που βασίζεται σε σύννεφο. Στην προσέγγισή τους, οι ευφυείς πράκτορες χρησιμοποιούν τεχνητό νευρωνικό δίκτυο για την ανίχνευση εισβολών σε ένα δίκτυο.

Οι Ojugo et al. (2012) παρουσίασαν το GAIDS - ένα σύστημα ανίχνευσης εισβολών βασισμένο σε κανόνες γενετικού αλγορίθμου για τη βελτίωση της ασφάλειας, της εμπιστευτικότητας, της ακεραιότητας και της διαθεσιμότητας πόρων του συστήματος σε δικτυακές ρυθμίσεις. Το προτεινόμενο σύστημα χρησιμοποιεί ένα σύνολο κανόνων ταξινόμησης που λαμβάνονται από δεδομένα ελέγχου δικτύου και το πλαίσιο υποστήριξης-εμπιστοσύνης που χρησιμοποιείται ως συνάρτηση καταλληλότητας για την αξιολόγηση της ποιότητας κάθε κανόνα.

Ο Barani (2014) πρότεινε την GAAIS - μια δυναμική μέθοδο ανίχνευσης εισβολών για κινητά δίκτυα που βασίζεται σε γενετικό αλγόριθμο και τεχνητό AIS. Η

GAAIS είναι αυτοπροσαρμοζόμενη στις αλλαγές της τοπολογίας του δικτύου. Η απόδοση του προτεινόμενου συστήματος αξιολογήθηκε για την ανίχνευση διαφόρων τύπων επιθέσεων δρομολόγησης, όπως οι επιθέσεις πλημμύρας, μαύρης τρύπας, γείτονα, βιασμού και σκουληκότρυπας. Τα πειραματικά αποτελέσματα έδειξαν ότι είναι πιο αποτελεσματικό σε σύγκριση με παρόμοιες προσεγγίσεις.

Τέλος , μια προσέγγιση κινητού ευφυούς συστήματος πολλαπλών πρακτόρων που βασίζεται στη σύνθεση για την καταπολέμηση των κυβερνοπαθειών και που χρησιμοποιεί διαδικασίες AI για τον εντοπισμό δραστηριότητας κακόβουλου λογισμικού είναι η CrowdStrike Falcon . Αυτό το καινοτόμο εργαλείο κυβερνοασφάλειας συνδυάζει τη χρήση πρακτόρων συλλογής δεδομένων σύντομα με μια μηχανή ανάλυσης που βασίζεται σε σύννεφο. Στην προσέγγισή τους, οι ευφυείς πράκτορες χρησιμοποιούν τεχνητό νευρωνικό δίκτυο για την ανίχνευση εισβολών σε ένα δίκτυο.

Κεφάλαιο 6^{ο} Συμπέρασμα

Η τεχνητή νοημοσύνη έχει γίνει ένας αυξανόμενος τομέας ενδιαφέροντος και επενδύσεων στην κοινότητα της ασφάλειας στον κυβερνοχώρο. Ορισμένοι πρώτοι φορείς που υιοθέτησαν την τεχνητή νοημοσύνη περιλαμβάνουν την Google, την IBM, την Juniper Networks, την Apple, την Amazon και την Balbix. Ένας αυξανόμενος αριθμός εταιρειών και οργανισμών μεταπηδά στο άρμα της τεχνητής νοημοσύνης. Καθώς οι επιθέσεις στον κυβερνοχώρο αυξάνονται, η τεχνητή νοημοσύνη βοηθά τους αναλυτές επιχειρήσεων ασφαλείας να παραμείνουν μπροστά από τις απειλές. Τα αυτοματοποιημένα συστήματα τεχνητής νοημοσύνης θα αποτελέσουν σύντομα αναπόσπαστο μέρος των λύσεων ασφάλειας στον κυβερνοχώρο, αλλά θα χρησιμοποιηθούν επίσης από τους εγκληματίες του κυβερνοχώρου για να προκαλέσουν ζημιά. Το μέλλον της ασφάλειας στον κυβερνοχώρο με τη βοήθεια της τεχνητής νοημοσύνης είναι πολλά υποσχόμενο.

Η τεχνητή νοημοσύνη μπορεί να ανιχνεύει και να σταματά τις απειλές στον

κυβερνοχώρο σε πραγματικό χρόνο με περιορισμένους πόρους. Η συνεχώς εξελισσόμενη φύση των κυβερνοεπιθέσεων σημαίνει ότι οι άνθρωποι θα δυσκολευτούν να συμβαδίσουν με τις πληροφορίες. Ωστόσο, με τη χρήση μηχανικής μάθησης, η τεχνητή νοημοσύνη μπορεί να μασήσει τα δεδομένα για γρήγορη ανάλυση και να παρέχει εξαιρετική κάλυψη ασφαλείας χωρίς να αφαιρεί πολύ χρόνο ή ενέργεια από τις υπάρχουσες εργασίες. Η μηχανική μάθηση επιτρέπει στους ανθρώπινους αναλυτές να επικεντρωθούν στην ερμηνεία των αποτελεσμάτων της βαθιάς ανάλυσης και στην επινόηση νέων τεχνικών για την καταπολέμηση του εγκλήματος στον κυβερνοχώρο.

Η τεχνητή νοημοσύνη δεν είναι το ελιξίριο για όλες τις μορφές ασφάλειας. Αν και οι προσεγγίσεις που βασίζονται στην ΤΝ γίνονται όλο και πιο κοινές και οικονομικά αποδοτικές στις περισσότερες πτυχές της ασφάλειας στον κυβερνοχώρο, δεν παρέχουν πλήρη μέτρα πρόληψης ή αποκατάστασης. Όταν ένας ανθρώπινος αντίπαλος με αμετακίνητη στάση επιτίθεται σε ένα ευφυές σύστημα, υπάρχουν όρια στο τι μπορεί να κάνει μια τεχνητή νοημοσύνη. Είναι σημαντικό να γνωρίζουμε ότι η τεχνητή νοημοσύνη δεν είναι ένας γενικός υπάλληλος που έχει πολλές και διαφορετικές δραστηριότητες ή ευθύνες και δεν θα είναι σε θέση να χειριστεί τα πάντα μόνη της, τουλάχιστον όχι αυτή τη στιγμή. Στην πραγματικότητα χρειάζεται την εκπαίδευση και την επίβλεψη από ειδικούς ανθρώπους για να βελτιωθεί με την πάροδο του χρόνου για τα καλύτερα αποτελέσματα. Οι έρευνες δείχνουν ότι η τεχνητή νοημοσύνη έχει φαινομενικά επηρεάσει θετικά την ασφάλεια στον κυβερνοχώρο και τους κινδύνους. Ως εκ τούτου, η συνέχιση της τεχνητής νοημοσύνης και της μηχανικής μάθησης θα οδηγήσει τον τομέα της κυβερνοασφάλειας σε ένα νέο επίπεδο ευφυΐας.

Επίλογος

Ενώ η τεχνολογία τεχνητής νοημοσύνης διευκολύνει τους εγκληματίες του κυβερνοχώρου να εξαπολύουν επιθέσεις, χρησιμοποιείται επίσης για την αποτροπή τους. Η τεχνολογία τεχνητής νοημοσύνης έχει τη δυνατότητα να φέρει επανάσταση στην ασφάλεια στον κυβερνοχώρο και εγείρει επίσης ηθικές ανησυχίες. Για παράδειγμα, οι λύσεις κυβερνοασφάλειας που λειτουργούν με τεχνητή νοημοσύνη μπορούν να χρησιμοποιηθούν για την παρακολούθηση των εργαζομένων και τη συλλογή ευαίσθητων πληροφοριών για αυτούς. Αυτό εγείρει ερωτήματα σχετικά με την προστασία της ιδιωτικής ζωής και τη χρήση προσωπικών δεδομένων. Μια άλλη ανησυχία είναι το ενδεχόμενο οι επιθέσεις στον κυβερνοχώρο που υποστηρίζονται από την τεχνολογία τεχνητής νοημοσύνης να προκαλέσουν εκτεταμένη ζημία. Καθώς η τεχνολογία τεχνητής νοημοσύνης γίνεται πιο προηγμένη, οι εγκληματίες του κυβερνοχώρου μπορεί να είναι σε θέση να εξαπολύουν επιθέσεις που είναι πιο εξελιγμένες και πιο δύσκολο να εντοπιστούν. Αυτό θα μπορούσε να οδηγήσει σε εκτεταμένη αναστάτωση και χάος. Αυτό θα απαιτήσει από τους εμπειρογνώμονες της ασφάλειας στον κυβερνοχώρο να προσαρμόζονται συνεχώς και να αναπτύσσουν νέες τεχνολογίες για να παραμένουν μπροστά από τους εγκληματίες του κυβερνοχώρου.

Η τεχνολογία τεχνητής νοημοσύνης αλλάζει το παιχνίδι όσον αφορά το έγκλημα στον κυβερνοχώρο. Ενώ διευκολύνει τους εγκληματίες του κυβερνοχώρου να εξαπολύουν επιθέσεις, χρησιμοποιείται επίσης για την αποτροπή τους. Καθώς η τεχνολογία τεχνητής νοημοσύνης γίνεται πιο προηγμένη, μπορούμε να περιμένουμε να δούμε τόσο πιο εξελιγμένες επιθέσεις στον κυβερνοχώρο όσο και πιο ισχυρές λύσεις κυβερνοασφάλειας. Για να παραμείνουν ένα βήμα μπροστά, οι οργανισμοί πρέπει να λάβουν σοβαρά υπόψη τους την ασφάλεια στον κυβερνοχώρο και να επενδύσουν στις πιο πρόσφατες τεχνολογίες που λειτουργούν με τεχνητή νοημοσύνη. Με τον τρόπο αυτό, μπορούν να προστατευτούν από τη διαρκώς εξελισσόμενη απειλή του εγκλήματος στον κυβερνοχώρο και να παραμείνουν μπροστά από τους εγκληματίες του κυβερνοχώρου.

Βιβλιογραφία

Akhter, S., Pauyo, T., & Khan, M. (2019). What is the difference between a systematic review and a meta-analysis?. *Basic methods handbook for clinical orthopaedic research: a practical guide and case based research approach*, 331-342.

Al-Masalha, H., Hnaif, A. A., &Kanan, T. (2020). Cyber-crime effect on jordanian society. *Int. J. Advance Soft Compu. Appl*, *12*(3).

ALI, N., Samsuri, S., SEMAN, M. A., BROHI, I., & Shah, A. (2018). Cybercrime an emerging challenge for internet users: An overview. *Sindh University Research Journal (Science Series)*, *50*(3D), 55-58.

Axelsson, S. (2000). Intrusion detection systems: A survey and taxonomy.

Aziz, A. S. A., Salama, M. A., ella Hassanien, A., & Hanafi, S. E. O. (2012). Artificial immune system inspired intrusion detection system using genetic algorithm. *Informatica*, *36*(4).

Babanina, V., Tkachenko, I., Matiushenko, O., &Krutevych, M. (2021). Cybercrime: History of formation, current state and ways of counteraction. *Amazonia Investiga*, *10*(38), 113-122.

Barani, F. (2014, February). A hybrid approach for dynamic intrusion detection in ad hoc networks using genetic algorithm and artificial immune system. In *2014 Iranian Conference on Intelligent Systems (ICIS)* (pp. 1-6). IEEE.

Barika, F., Hadjar, K., & El-Kadhi, N. (2009). Artificial neural network for mobile IDS solution. *Security and Management*, *23*(4), 61-66.

Barman, D. K., & Khataniar, G. (2012). Design of intrusion detection system based on artificial neural network and application of rough set. *International Journal of Computer Science and Communication Networks*, *2*(4), 548-552.

Bates, D. W., Auerbach, A., Schulam, P., Wright, A., & Saria, S. (2020). Reporting and implementing interventions involving machine learning and artificial intelligence. *Annals of Internal Medicine*, *172*(11_Supplement), S137-S144.

Bitter, C., Elizondo, D. A., & Watson, T. (2010, July). Application of artificial neural networks and related techniques to intrusion detection. In *The 2010 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.

Bose, S., Barao, T., & Liu, X. (2020, July). Explaining ai for malware detection: Analysis of mechanisms of malconv. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.

Brenner, S. W. (2007). History of computer crime. In *The history of information security* (pp. 705-721). Elsevier Science BV.

Buch, R., Ganda, D., Kalola, P., &Borad, N. (2017). World of cyber security and cybercrime.

Caballero, J. (2012). Understanding the role of malware in cybercrime. *Cybercrime*, 15.

Chen, S. H., Jakeman, A. J., & Norton, J. P. (2008). Artificial intelligence techniques: an introduction to their use for modelling environmental systems. *Mathematics and computers in simulation*, *78*(2-3), 379-400.

Connolly, L. Y., & Wall, D. S. (2019). The rise of crypto-ransomware in a changing cybercrime landscape: Taxonomising countermeasures. *Computers & Security*, *87*, 101568.

Coull, S. E., & Gardner, C. (2019, May). Activation analysis of a byte-based deep neural network for malware classification. In *2019 IEEE Security and Privacy Workshops (SPW)* (pp. 21-27). IEEE.

Dilek, S., Çakır, H., & Aydın, M. (2015). Applications of artificial intelligence techniques to combating cyber crimes: A review. *arXiv preprint arXiv:1502.03552*.

Dilek, S., Çakır, H., &Aydın, M. (2015). Applications of artificial intelligence techniques to combating cyber crimes: A review. *arXiv preprint arXiv:1502.03552*.

Dobrinoiu, M. (2019). The influence of artificial intelligence on criminal liability. *Lex ET Scientia Int'l J.*, *26*, 140.

Elsadig, M., Abdullah, A., & Samir, B. B. (2010, June). Immune multi agent system for intrusion prevention and self healing system implement a non-linear classification. In *2010 International Symposium on Information Technology* (Vol. 3, pp. 1-6). IEEE.

EshghiShargh, A. (2009, November). Using artificial immune system on implementation of intrusion detection systems. In *2009 Third UKSim European Symposium on Computer Modeling and Simulation* (pp. 164-168). IEEE.

Faruk, M. J. H., Shahriar, H., Valero, M., Barsha, F. L., Sobhan, S., Khan, M. A., ... & Wu, F. (2021, December). Malware detection and prevention using artificial intelligence techniques. In *2021 IEEE International Conference on Big Data (Big Data)* (pp. 5369-5377). IEEE.

Fu, H., Yuan, X., Zhang, K., Zhang, X., & Xie, Q. (2007, December). Investigating novel immune-inspired multi-agent systems for anomaly detection. In *The 2nd IEEE Asia-Pacific Service Computing Conference (APSCC 2007)* (pp. 466-472). IEEE.

Furfaro, A., Malena, G., Molina, L., & Parise, A. (2015, March). A simulation model for the analysis of DDOS amplification attacks. In *2015 17th UKSim-AMSS International Conference on Modelling and Simulation (UKSim)* (pp. 267-272). IEEE.

Garcia-Teodoro, P., Diaz-Verdejo, J., Maciá-Fernández, G., & Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. *computers & security*, *28*(1-2), 18-28.

Gianini, G., Anisetti, M., Azzini, A., Bellandi, V., Damiani, E., & Marrara, S. (2009, June). An artificial immune system approach to anomaly detection in multimedia ambient intelligence. In *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies* (pp. 502-506). IEEE.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational researcher*, *5*(10), 3-8.

Gou, X., Jin, W., & Zhao, D. (2006). Multi-agent system for worm detection and containment in metropolitan area networks. *Journal of Electronics (China)*, *23*,

259-265.

Guan, Z., Li, J., Wu, L., Zhang, Y., Wu, J., & Du, X. (2017). Achieving efficient and secure data acquisition for cloud-supported internet of things in smart grid. *IEEE Internet of Things Journal*, *4*(6), 1934-1944.

Hardy, W., Chen, L., Hou, S., Ye, Y., & Li, X. (2016). DL4MD: A deep learning framework for intelligent malware detection. In *Proceedings of the International Conference on Data Science (ICDATA)* (p. 61). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).

Hayward, K. J., & Maas, M. M. (2021). Artificial intelligence and crime: A primer for criminologists. *Crime, Media, Culture*, *17*(2), 209-233.

Herrero, Á., Corchado, E., Pellicer, M. A., & Abraham, A. (2007). Hybrid multi agent-neural network intrusion detection with mobile visualization. *Innovations in Hybrid Intelligent Systems*, 320-328.

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

Hou, S., Saas, A., Ye, Y., & Chen, L. (2016). Droiddelver: An android malware detection system using deep belief network based on api call blocks. In *Web-Age Information Management: WAIM 2016 International Workshops, MWDA, SDMMW, and SemiBDMA, Nanchang, China, June 3-5, 2016, Revised Selected Papers 17* (pp. 54-66). Springer International Publishing.

IBM. (2023). *Artificial intelligence for a smarter kind of cyber security*. Ανάκτηση από: https://www.ibm.com/security/artificial-intelligence [Πρόσβαση στις 23-5-2023].

Ioniţă, I., & Ioniţă, L. (2013, September). An agent-based approach for building an intrusion detection system. In *2013 RoEduNet International Conference 12th Edition: Networking in Education and Research* (pp. 1-6). IEEE.

Jajodia, D. B. J. C. S., & Wu, L. P. N. (2001). Adam: Detecting intrusions by data

mining. In *Workshop on Information Assurance and Security* (Vol. 1, p. 1100).

Kemp, S., Buil-Gil, D., Moneva, A., Miró-Llinares, F., & Díaz-Castaño, N. (2021). Empty streets, busy internet: A time-series analysis of cybercrime and fraud trends during COVID-19. *Journal of Contemporary Criminal Justice*, *37*(4), 480-501.

Kim, D. W., Yang, J. W., & Sim, K. B. (2004, November). Adaptive intrusion detection algorithm based on learning algorithm. In *30th Annual Conference of IEEE Industrial Electronics Society, 2004. IECON 2004* (Vol. 3, pp. 2229-2233). IEEE.

Konieczny, M. K. (2023). CYBERCRIME–A SHORT HISTORY, CONTEMPORARY FACES AND AN UNPREDICTABLE FUTURE. *Annuals of the Administration and Law*, *1*(XXIII), 29-50.

Kumar, A., Maurya, H. C., & Misra, R. (2013). A research paper on hybrid intrusion detection system. *International Journal of Engineering and Advanced Technology (IJEAT) Vol*, *2*.

Lebbe, M. A., Agbinya, J. I., Chaczko, Z., & Chiang, F. (2007). Self-Organized Classification of Dangers for Secure Wireless Mesh Networks, proceedings of Australasian Telecommunication Networks and Applications Conference 2007.

Li, J. X. (2017). Cyber Crime and Legal Countermeasures: A Historical Analysis. *International Journal of Criminal Justice Sciences*, *12*(2).

Mabu, S., Chen, C., Lu, N., Shimada, K., & Hirasawa, K. (2010). An intrusion-detection model based on fuzzy class-association-rule mining using genetic network programming. *IEEE transactions on systems, man, and cybernetics, part C (Applications and Reviews)*, *41*(1), 130-139.

Machado, R. B., Boukerche, A., Sobral, J. B. M., Jucá, K. R. L., & Notare, M. S. M. A. (2005, April). A hybrid artificial immune and mobile agent intrusion detection based model for computer network operations. In *19th IEEE international parallel and distributed processing symposium* (pp. 8-pp). IEEE.

Mason, R. O. (2003). Ethical issues in artificial intelligence.

Meland, P. H., Bayoumy, Y. F. F., & Sindre, G. (2020). The Ransomware-as-a-Service economy within the darknet. *Computers & Security*, *92*, 101762.

Mintz, Y., & Brodie, R. (2019). Introduction to artificial intelligence in medicine. *Minimally Invasive Therapy & Allied Technologies*, *28*(2), 73-81.

Mulrow, C. D. (1994). Systematic reviews: rationale for systematic reviews. *Bmj*, *309*(6954), 597-599.

Ng, A. (2016). What artificial intelligence can and can't do right now. *Harvard Business Review*, *9*(11).

Nitin, T., Singh, S. R., & Singh, P. G. (2012). Intrusion detection and prevention system (idps) technology-network behavior analysis system (nbas). *ISCA J. Engineering Sci*, *1*(1), 51-56.

Nogueira, J. H. M. (2006). Mobile intelligent agents to fight cyber intrusions. *International Journal of Forensic Computer Science, Brasília: Brazil*.

Nogueira, J. H. M. (2006). Mobile intelligent agents to fight cyber intrusions. *International Journal of Forensic Computer Science, Brasília: Brazil*.

Ojugo, A. A., Eboka, A. O., Okonta, O., Yoro, R. E., & Aghware, F. O. (2012). Genetic algorithm rule-based intrusion detection system (GAIDS). *Journal of Emerging Trends in Computing and Information Sciences*, *3*(8), 1182-1194.

Palmer, D. (2020). AI is changing everything about cybersecurity, for better and for worse. Here's what you need to know. *ZDNet, March*, *2*.

Patel, A., Qassim, Q., Shukor, Z., Nogueira, J., Júnior, J., Wills, C., & Federal, P. (2011). Autonomic agent-based self-managed intrusion detection and prevention system. In *Proceedings of the South African information security multi-conference (SAISMC 2010)* (pp. 223-234).

Patel, A., Taghavi, M., Bakhtiyari, K., & Júnior, J. C. (2013). An intrusion detection and prevention system in cloud computing: A systematic review. *Journal of network and computer applications*, *36*(1), 25-41.

Pei, Z., & Song, J. (2008, October). Application of Immune Algorithm to Generate Fuzzy-detector in Intrusion detection. In *2008 Fourth International Conference on Natural Computation* (Vol. 5, pp. 183-186). IEEE.

Radulov, N. (2019). Artificial intelligence and security. Security 4.0. *Security &*

Future, *3*(1), 3-5.

Rowe, J. P., & Lester, J. C. (2020). Artificial intelligence for personalized preventive adolescent healthcare. *Journal of Adolescent Health*, *67*(2), S52-S58.

Rui, L., & Wanbo, L. (2010, July). Intrusion response model based on AIS. In *2010 International forum on information technology and applications* (Vol. 1, pp. 86-90). IEEE.

Sadiku, M. N. O., Ashaolu, T. J., & Musa, S. M. (2019). Artificial intelligence in medicine: a primer. *International Journal of Trend in Research and Development*, *6*(1), 270-272.

Sadiku, M. N. O., Zhou, Y., & Musa, S. M. (2018). Natural language processing in healthcare. *International Journal of Advanced Research in Computer Science and Software Engineering*, *8*(5), 39-42.

Schell, B. H., & Martin, C. (2004). *Cybercrime: A reference handbook*. ABC-CLIO. Ανάκτηση από: https://books.google.gr/books?hl=en&lr=&id=AUUERdg6b94C&oi=fnd&pg=PR11&dq=Schell,+B.+H.,+%26+Martin,+C.+(2004).+Cybercrime:+A+reference+handbook&ots=8EUdo2ZQZ7&sig=Vg6I_MgEJVkirW7L0IQt3kWYiP8&redir_esc=y#v=onepage&q=Schell%2C%20B.%20H.%2C%20%26%20Martin%2C%20C.%20(2004).%20Cybercrime%3A%20A%20reference%20handbook&f=false. [Πρόσβαση στις 17-05-2023].

Schjolberg, S. (2020). *The History of Cybercrime* (Vol. 13). BoD–Books on Demand. Ανάκτηση από: https://books.google.gr/books?hl=en&lr=&id=4onVDwAAQBAJ&oi=fnd&pg=PA86&dq=Schjolberg,+S.+(2020).+The+History+of+Cybercrime+(Vol.+13)&ots=IX-jg31HMJ&sig=VydkTvIt88-2o1xEsTntRJ5eEjs&redir_esc=y#v=onepage&q=Schjolberg%2C%20S.%20(2020).%20The%20History%20of%20Cybercrime%20(Vol.%2013)&f=false. [Πρόσβαση στις 17-05-2023].

Sekeh, M. A. (2009, August). Fuzzy intrusion detection system via data mining technique with sequences of system calls. In *2009 Fifth International*

Conference on Information Assurance and Security (Vol. 1, pp. 154-157). IEEE.

Shosha, A. F., Gladyshev, P., Wu, S. S., & Liu, C. C. (2011, September). Detecting cyber intrusions in SCADA networks using multi-agent collaboration. In *2011 16th International conference on intelligent system applications to power systems* (pp. 1-7). IEEE.

Sirisanyalak, B., & Sornil, O. (2007, September). An artificial immunity-based spam detection system. In *2007 IEEE Congress on Evolutionary Computation* (pp. 3392-3398). IEEE.

Smith, R. B., Phillips, L. R., Link, H. E., & Weiland, L. (2006). *Agent-based control of distributed infrastructure resources* (No. SAND2005-7937). Sandia National Laboratories (SNL), Albuquerque, NM, and Livermore, CA (United States).

Smith, R. B., Phillips, L. R., Link, H. E., & Weiland, L. (2006). *Agent-based control of distributed infrastructure resources* (No. SAND2005-7937). Sandia National Laboratories (SNL), Albuquerque, NM, and Livermore, CA (United States).

Stahl, B., Elizondo, D., Carroll-Mayer, M., Zheng, Y., & Wakunuma, K. (2010, July). Ethical and legal issues of the use of computational intelligence techniques in computer security and computer forensics. In *The 2010 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.

Stytz, M. R., Lichtblau, D. E., & Banks, S. B. (2005). *Toward using intelligent agents to detect, assess, and counter cyberattacks in a network-centric environment*. INSTITUTE FOR DEFENSE ANALYSES ALEXANDRIA VA.

Sukhodolov, A. P., Bychkov, A. V., &Bychkova, A. M. (2020). Criminal policy for crimes committed using artificial intelligence technologies: state, problems, prospects.

Taddeo, M., McCutcheon, T., & Floridi, L. (2021). Trusting Artificial Intelligence in Cybersecurity Is a Double-Edged Sword. *Ethics, Governance, and Policies in Artificial Intelligence*, 289-297.

Tandon, R. (2020). A survey of distributed denial of service attacks and defenses. *arXiv preprint arXiv:2008.01345*.

Trolice, M. P., Curchoe, C., & Quaas, A. M. (2021). Artificial intelligence—the future is

now. *Journal of Assisted Reproduction and Genetics*, *38*, 1607-1612.

Tyugu, E. (2011, June). Artificial intelligence in cyber defense. In *2011 3rd International conference on cyber conflict* (pp. 1-11). IEEE.

Weber, M., Schmid, M., Schatz, M., & Geyer, D. (2002, December). A toolkit for detecting and analyzing malicious software. In *18th Annual Computer Security Applications Conference, 2002. Proceedings.* (pp. 423-431). IEEE.

Wilson, T. (2017). No longer science fiction, AI and robotics are transforming healthcare. *PWC.*

Ye, Y., Chen, L., Hou, S., Hardy, W., & Li, X. (2018). DeepAM: a heterogeneous deep learning framework for intelligent malware detection. *Knowledge and Information Systems*, *54*, 265-285.

Zargar, S. T., Joshi, J., & Tipper, D. (2013). A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks. *IEEE communications surveys & tutorials*, *15*(4), 2046-2069.

Zhang, Z., Ning, H., Shi, F., Farha, F., Xu, Y., Xu, J., ... & Choo, K. K. R. (2022). Artificial intelligence in cyber security: research advances, challenges, and opportunities. *Artificial Intelligence Review*, 1-25.