ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ Μ/Υ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΝΑΥΤΙΛΙΑΣ ΚΑΙ ΒΙΟΜΗΧΑΝΙΑΣ
ΤΜΗΜΑΤΟΣ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ & ΤΕΧΝΟΛΟΓΙΑΣ
ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΤΕΧΝΟ-ΟΙΚΟΝΟΜΙΚΑ ΣΥΣΤΗΜΑΤΑ»

ΔΙΕΠΙΣΤΗΜΟΝΙΚΟ – ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΤΕΧΝΟ-ΟΙΚΟΝΟΜΙΚΑ ΣΥΣΤΗΜΑΤΑ»

Ασφάλεια συστημάτων εποπτείας, ελέγχου και συλλογής δεδομένων (SCADA)

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Νικόλαος Β. Κοντογιάννης

Επιβλέπων : Δημήτριος Ασκούνης
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2023

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ Μ/Υ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΝΑΥΤΙΛΙΑΣ ΚΑΙ ΒΙΟΜΗΧΑΝΙΑΣ
ΤΜΗΜΑΤΟΣ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ & ΤΕΧΝΟΛΟΓΙΑΣ
ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΤΕΧΝΟ-ΟΙΚΟΝΟΜΙΚΑ ΣΥΣΤΗΜΑΤΑ»

ΔΙΕΠΙΣΤΗΜΟΝΙΚΟ – ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΤΕΧΝΟ-ΟΙΚΟΝΟΜΙΚΑ ΣΥΣΤΗΜΑΤΑ»

Ασφάλεια συστημάτων εποπτείας, ελέγχου και συλλογής δεδομένων (SCADA)

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Νικόλαος Β. Κοντογιάννης

Επιβλέπων : Δημήτριος Ασκούνης
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 23η Οκτωβρίου 2023.

.....	.....	.....
Δημήτριος Ασκούνης Καθηγητής Ε.Μ.Π.	Ιωάννης Ψαρράς Καθηγητής Ε.Μ.Π.	Χρυσόστομος Δούκας Αναπληρωτής Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2023

.....
Νικόλαος Β. Κοντογιάννης

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Περίληψη

Ο όρος SCADA (supervisory control and data acquisition) περιγράφει μια κατηγορία συστημάτων βιομηχανικού αυτόματου ελέγχου και τηλεμετρίας. Ένα τυπικό σύστημα SCADA αποτελείται από πολλούς τοπικούς ελεγκτές που ελέγχουν επί μέρους διεργασίες μιας εγκατάστασης και έναν κεντρικό σταθμό στον οποίο είναι συνδεδεμένοι όλοι οι τοπικοί ελεγκτές. Στον κεντρικό σταθμό γίνεται η συλλογή η εποπτεία και ο έλεγχος όλων των κατανεμημένων ελεγκτών και του δικτύου και συχνά συνδέεται με βάσεις δεδομένων όπου μπορούν να αποθηκεύονται ιστορικά στοιχεία των μεταβλητών παρακολούθησης. Έτσι προκύπτει μια αναπαράσταση όλης της διεργασίας στην οθόνη ενός υπολογιστή. Σήμερα μπορούμε να συναντήσουμε υπερσύγχρονα συστήματα SCADA σε κοινωνικές υποδομές όπως σταθμούς παραγωγής ενέργειας ή υδραγωγεία αλλά και σε απλές βιομηχανικές μονάδες. Με την ανάπτυξη της τεχνολογίας των υπολογιστών και των επικοινωνιών, τα συστήματα εποπτείας συνδέθηκαν στο internet δίνοντας έτσι στον χρήστη την δυνατότητα παρακολούθησης αυτών από οποιοδήποτε σημείο με την χρήση ενός υπολογιστή ή tablet. Η εξέλιξη αυτή, μαζί με τα πλεονεκτήματα που έφερε, έφερε και τα μειονεκτήματα-ευπάθειες που έχουν οι συνδεδεμένες στο internet συσκευές και κυρίως αυτή των δικτυακών επιθέσεων από κακόβουλες πηγές. Η ιδιαιτερότητα που παρουσιάζουν τα συστήματα SCADA σε σχέση με τα τυπικά συστήματα τεχνολογίας πληροφορίας (Information Technlogy-IT) είναι η ανάγκη για απρόσκοπτη και συνεχή λειτουργία, κυρίως όταν αυτά διαχειρίζονται κρίσιμες κοινωνικές υποδομές. Έτσι προκύπτει η ανάγκη για δημιουργία μηχανισμών προστασίας των συστημάτων αυτών αλλά και αναγνώρισης των επιθέσεων όταν αυτές συμβαίνουν, χωρίς να θέτουν το σύστημα εκτός λειτουργίας. Στα πλαίσια αυτής της διπλωματικής καταγράφηκαν οι πιθανοί τύποι επιθέσεων που μπορεί να δεχθεί ένα συνδεδεμένο στο internet σύστημα SCADA και παρουσιάζεται μια μεθοδολογία αυτόματης ταξινόμησης αυτών με την χρήση τεχνικών κατηγοριοποίησης κειμένου (Text Classification). Συγκεκριμένα δημιουργήθηκε μια εφαρμογή, χρησιμοποιώντας αλγόριθμους μηχανικής μάθησης με την βοήθεια της γλώσσας προγραμματισμού Python. Η εφαρμογή αυτή, δέχεται σαν είσοδο μια περιγραφή της επίθεσης και την κατατάσσει αυτόματα σε προκαθορισμένους τύπους γνωστών επιθέσεων. Έτσι δίνεται στον χρήστη η δυνατότητα αναγνώρισης των επιθέσεων που δέχεται το σύστημά του ώστε να λάβει τα απαιτούμενα μέτρα για την προστασία του.

Λέξεις Κλειδιά: SCADA, Συνδεδεμένες συσκευές, ευπάθειες, κατηγοριοποίηση κειμένου

Abstract

The term SCADA (supervisory control and data acquisition) describes an industrial automatic control and telemetry system. A typical SCADA system consists of several local controllers that control individual processes of a plant and a central station to which all local controllers are connected. The central station collects, monitors and controls all distributed controllers and the network and is often connected to databases where historical data of monitoring variables can be stored. This results in a representation of the entire process on a computer screen. Today we can meet state-of-the-art SCADA systems in social infrastructures such as power plants or aqueducts but also in simple industrial units. With the development of computer and communication technology, SCADA systems were connected to the internet, thus giving the user the possibility of monitoring them from any point using a computer or tablet. This development, along with the advantages it brought, also brought the disadvantages-vulnerabilities that internet-connected devices have, and mainly that of network attacks from malicious sources. The particularity that SCADA systems present in relation to standard information technology systems (Information Technology-IT) is the need for uninterrupted and continuous operation, especially when they manage critical social infrastructures. Thus arises the need to create mechanisms to protect these systems but also to recognize attacks when they occur, without putting the system out of order. In this thesis, are recorder the possible types of attacks that can be received by a SCADA system connected to the internet and is presented a methodology for their automatic classification using text classification techniques. In particular, an application was created using machine learning algorithms with the help of the Python programming language. This application accepts as input a description of the attack and automatically classifies it into predefined types of known attacks. Thus, the user is given the possibility to identify the attacks that his system receives in order to take the necessary measures for its protection.

Keywords: SCADA, connected devices, vulnerabilities, Text classification

Ευχαριστίες

Με την ολοκλήρωση της παρούσας διπλωματικής εργασίας νιώθω την ανάγκη να εκφράσω τις ιδιαίτερες ευχαριστίες μου στον Καθηγητή του Ε.Μ.Π. κ. Δημήτριο Ασκούνη για την ευκαιρία που μου έδωσε με την εκπόνηση της εργασίας αυτής και για την επίβλεψη που παρείχε.

Επίσης, θα ήθελα να ευχαριστήσω θερμά τον κύριο Κωνσταντίνο Τουλούμη για την πολύ καλή συνεργασία που είχαμε καθ' όλη τη διάρκεια εκπόνησης της παρούσας διπλωματικής εργασίας, καθώς και για την πολύτιμη βοήθεια και καθοδήγηση μου προσέφερε.

Ευχαριστώ ακόμα τον Καθηγητή κ. Ιωάννη Ψαρρά και τον Αναπληρωτή Καθηγητή Ε.Μ.Π κ. Χρυσόστομο Δούκα για την συμμετοχή τους στην επιτροπή εξέτασης της διπλωματικής εργασίας μου.

Πίνακας Περιεχομένων

Κεφάλαιο 1: Εισαγωγή

1.1 Ορισμός προβλήματος

Ο αρχικός σχεδιασμός των συστημάτων SCADA ήταν για λειτουργία σε περιβάλλοντα μη συνδεδεμένα με το internet [41]. Για το λόγο αυτό παρουσιάζουν σημαντικές ελλείψεις προστασίας απέναντι σε επιθέσεις μέσω του διαδικτύου αλλά και σε εργαλεία εντοπισμού, διαχείρισης και ανάλυσης (Φορενζικς) των επιθέσεων αυτών [38]. Οι αλλαγές που έχουν γίνει τα τελευταία χρόνια στα συστήματα εποπτείας σχετικά με την συνδεσιμότητά τους στο διαδίκτυο, αυξάνουν τους κινδύνους στους οποίους αυτά εκτίθενται. Ωστόσο αυτές οι αλλαγές δίνουν και πλεονεκτήματα στα συστήματα εποπτείας, όπως ο απομακρυσμένος έλεγχος μέσω διαδικτύου ή η σύνδεση με συστήματα γενικού σκοπού (Historian, Servers) [38]. Κάθε επίθεση σε ένα σύστημα SCADA, επιβάλει την καταγραφή της, την αξιολόγησή της καθώς και μία διερεύνηση των επιπτώσεων και των προβλημάτων που δημιούργησε στο σύστημα. Όμως ο σημαντικός παράγοντας που πρέπει να ληφθεί υπόψιν σε τέτοια συστήματα, είναι η ανάγκη για συνεχή και απρόσκοπτη λειτουργία. Επομένως δημιουργείται η ανάγκη για γρήγορη και αυτοματοποιημένη συλλογή και επεξεργασία πληροφοριών για τις επιθέσεις που δέχονται [4].

Με βάση τα παραπάνω, αντικείμενο της διπλωματικής εργασίας αποτελεί η παρουσίαση της αρχιτεκτονικής ενός συστήματος SCADA και η καταγραφή των απειλών που μπορεί να εμφανιστούν με την σύνδεση αυτού στο διαδίκτυο. Στην συνέχεια αναπτύσσεται ένας αλγόριθμος ο οποίος χρησιμοποιεί τεχνικές τεχνητής νοημοσύνης (AI) και τεχνικές επεξεργασίας φυσικής γλώσσας (NLP) ο οποίος μπορεί αυτόματα να κατατάξει μια επίθεση, βάση της περιγραφής της, σε κάποια από τις προκαθορισμένες κατηγορίες επιθέσεων. Έτσι μπορεί ο χρήστης να συλλέξει γρήγορα στοιχεία για τον τύπο των επιθέσεων που δέχεται το σύστημά του ώστε να το προστατέψει.

1.2 Ιστορική αναδρομή επιθέσεων σε συστήματα ελέγχου και εποπτείας.

Η πρώτη επίθεση σε συστήματα ελέγχου και εποπτείας αναφέρεται το 1982 ως αιτία για την έκρηξη γραμμής μεταφοράς αερίου στην Ρωσία. Ένας μη εξουσιοδοτημένος χρήστης εισήγαγε ένα κακόβουλο λογισμικό (Trojan horse) στον υπολογιστή του συστήματος και κατάφερε να αλλάξει την λειτουργία των αντλιών και των βαλβίδων του κυκλώματος [3].

Το 1994 ένας επιτιθέμενος κατάφερε μέσω τηλεφωνικής σύνδεσης να αποκτήσει πρόσβαση στο σύστημα ελέγχου του Salt River Project των Ηνωμένων Πολιτειών. Κατάφερε να τροποποιήσει στοιχεία πελατών καθώς και κρίσιμα αρχεία για την λειτουργία των διεργασιών [3] .

Το 1999 στόχος επίθεσης έγινε η μεγαλύτερη εταιρεία διακίνησης αερίου της Ρωσίας, η Gaζprom. Ο επιτιθέμενος, κατάφερε να αποκτήσει πλήρη έλεγχο του κέντρου επιτήρησης της ροής του αερίου στο κύκλωμα [3].

Το 2000 καταγράφεται η επόμενη επίθεση σε συστήματα ελέγχου και εποπτείας σε σταθμό υδροδότησης της Αυστραλίας [3].

Το 2003 η επίθεση σε πυρηνικό σταθμό στις Ηνωμένες πολιτείες, είχε ως αποτέλεσμα να παραμείνει ανενεργό στο σύστημα παρακολούθησης του σταθμού για περίπου πέντε ώρες [3].

Το 2010, στόχος επίθεσης ήταν ένας σταθμός παραγωγής ενέργειας στο Ιράν. Χρησιμοποιήθηκε κακόβουλο λογισμικό (Stuxnet) το οποίο μπήκε στο σύστημα και εξαπλώθηκε από έναν εξωτερικό σκληρό δίσκο [3].

Το 2015, έγινε επίθεση σε σύστημα ελέγχου και εποπτείας τμήματος δικτύου ηλεκτρικής ενέργειας της Ρωσίας, με συνέπεια να παραμείνουν πάνω από 225.000 καταναλωτές χωρίς ηλεκτρική ενέργεια [3].

Στην παρακάτω εικόνα παρουσιάζονται συνοπτικά οι κυριότερες επιθέσεις σε συστήματα ελέγχου και εποπτείας, τον τομέα που επηρέασαν και την επίδραση που είχαν [3].

1982	Χώρα: Ρωσία, Βιομηχανία: Φυσικό αέριο, Επίδραση: Οικονομικές απώλειες
1994	Χώρα: ΗΠΑ, Βιομηχανία: Ενέργεια, Επίδραση: Οικονομικές απώλειες, απώλεια στοιχείων πελατών
1999	Χώρα: Ρωσία, Βιομηχανία: Φυσικό αέριο, Επίδραση: Οικονομικές απώλειες, δυσλειτουργία συστήματος
2000	Χώρα: Αυστραλία, Βιομηχανία: Εγκαταστάσεις υδροδότησης, Επίδραση: Περιβαλλοντική καταστροφή
2003	Χώρα: ΗΠΑ, Βιομηχανία: Πυρηνικός σταθμός, Επίδραση: Δυσλειτουργία συστήματος
2010	Χώρα: Ιράν, Βιομηχανία: Ενέργεια, Επίδραση: Οικονομικές απώλειες, δυσλειτουργία συστήματος
2015	Χώρα: Ρωσία, Βιομηχανία: Ενέργεια, Επίδραση: Οικονομικές απώλειες, δυσλειτουργία συστήματος

Εικόνα 1: Καταγεγραμμένες επιθέσεις σε συστήματα ελέγχου και εποπτείας

Γίνεται επομένως αντιληπτό πόσο σημαντική είναι η διασφάλιση της λειτουργικότητας των συστημάτων ελέγχου κρίσιμων κοινωνικών υποδομών και η προστασία αυτών από κακόβουλες ενέργειες. Οι κακόβουλες ενέργειες μπορεί να προέρχονται είτε τοπικά από

έναν μη εξουσιοδοτημένο χρήστη ή πλέον με την διασύνδεση των συστημάτων αυτών στο διαδίκτυο, από οποιοδήποτε σημείο της γης.

Κεφάλαιο 2: Περιγραφή συστημάτων εποπτείας και ελέγχου (SCADA)

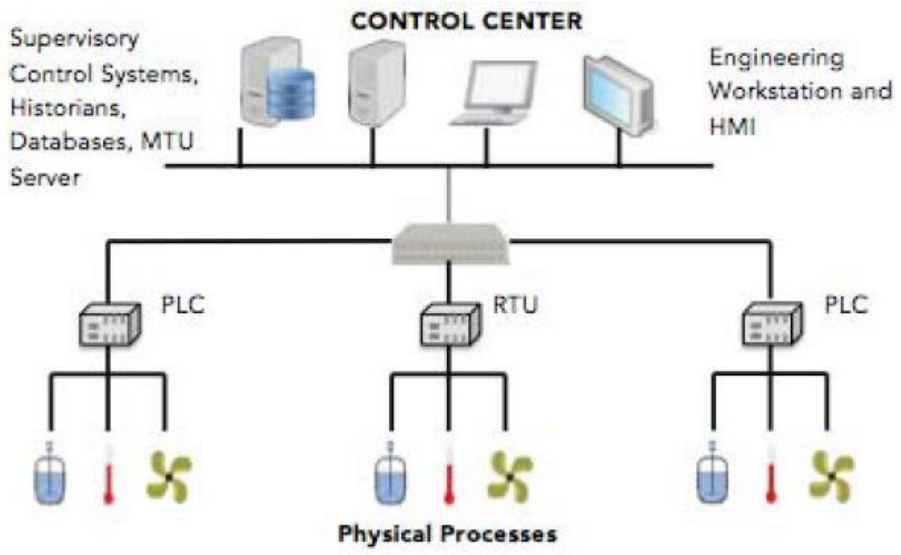
2.1 Γενική περιγραφή συστημάτων

Όπως έχει ήδη αναφερθεί, με τον όρο SCADA αναφερόμαστε σε ένα γενικότερο σύστημα εποπτείας και ελέγχου μιας διεργασίας. Η διεργασία αυτή μπορεί να είναι μια παραγωγική διαδικασία μιας βιομηχανικής μονάδας αλλά και κρίσιμες κοινωνικές υποδομές όπως είναι οι σταθμοί παραγωγής ενέργειας, διυλιστήρια ή και μεταφορές [24]. Δεδομένου ότι πολλές κρίσιμες υποδομές όπως μεταφορές, συστήματα ύδρευσης πόλεων και ενεργειακά συστήματα ελέγχονται μέσω συστημάτων SCADA, γίνεται αντιληπτό ότι η αστοχία ενός τέτοιου συστήματος έχει αρνητικό αντίκτυπο σε ένα πολύ μεγάλο ποσοστό της κοινωνίας [31]. Έτσι προκύπτει η ανάγκη συνεχή και αξιόπιστη λειτουργία των συστημάτων αυτών.
Η βασική λειτουργία ενός συστήματος SCADA είναι ο έλεγχος και η εποπτεία μιας αυτοματοποιημένης διεργασίας. Ωστόσο, έχουν ενσωματωθεί αρκετές άλλες λειτουργίες οι οποίες στοχεύουν στην βελτίωση της διεργασίας και την συλλογή δεδομένων. Τέτοιες λειτουργίες αναφέρονται επιγραμματικά παρακάτω [10].

- Ειδοποίηση του προσωπικού με οπτική ή και ηχητική σήμανση σε περίπτωση σφάλματος ώστε να επέμβει έγκαιρα.
- Αποθήκευση πληροφοριών σε βάση δεδομένων και αναπαράστασή τους μέσω γραφημάτων.
- Καταγραφή κανονικών και μη συμβάντων, ώστε να δημιουργηθεί ένα ιστορικό δεδομένων.
- Μεταφορά δεδομένων σε άλλα τμήματα του κεντρικού συστήματος και διαχείρισης.
- Έλεγχος πρόσβασης στο σύστημα και διαβάθμιση των χρηστών σε κατηγορίες ανάλογα με τις αρμοδιότητες του καθενός.

Τα δύο κύρια μέρη ενός συστήματος εποπτείας είναι οι συσκευές πεδίου (Field Sites) και το κέντρο ελέγχου (Control Centre) [12]. Οι συσκευές πεδίου, βρίσκονται κατανεμημένες σε όλη την έκταση της εποπτευόμενης διεργασίας και πάνω σε αυτές είναι συνδεδεμένοι αισθητήρες (πίεσης, θερμοκρασίας, στάθμης) αλλά και επενεργητές (κινητήρες, υδραυλικά η πνευματικά έμβολα). Οι συσκευές πεδίου μπορεί να είναι PLC (Programmable Logic Controllers-Προγραμματιζόμενοι Ελεγκτές) ή RTU (Remote Terminal Units-

Απομακρυσμένες τερματικές μονάδες), ή οποιαδήποτε άλλη IED (Intelligent Electronic Device- Έξυπνη ηλεκτρονική συσκευή). Οι συσκευές αυτές, συλλέγουν πληροφορίες από τα αισθητήρια που είναι συνδεδεμένα πάνω τους και έτσι αντιλαμβάνονται την κατάσταση της διεργασίας την οποία ελέγχουν. Στην συνέχεια βάση των τιμών που συλλέγονται από τα αισθητήρια της διεργασίας και του προγράμματος που είναι αποθηκευμένο στην μνήμη των συσκευών πεδίου, ενεργοποιούνται οι κατάλληλες έξοδοι αυτών. Επίσης οι συσκευές πεδίου συλλέγουν συνεχώς δεδομένα και τα στέλνουν στο κέντρο ελέγχου. Ένα τυπικό κέντρο ελέγχου αποτελείται από HMI (Human Machine Interface- Οθόνη αλληλεπίδρασης ανθρώπου- μηχανής), ένα Historian λογισμικό που δίνει την δυνατότητα αποθήκευσης των τιμών των μεταβλητών τις διεργασίας και ένα MTU (Master Terminal Unit). Παρακάτω παρουσιάζεται σχηματικά ένα σύστημα εποπτείας και περιγράφονται συνοπτικά τα στοιχεία του.



Εικόνα 2: Σύστημα ελέγχου και εποπτείας

- PLC (Προγραμματιζόμενοι ελεγκτές): Είναι συσκευές οι οποίες διαθέτουν υπολογιστική ισχύ και χρησιμοποιούνται για συλλογή δεδομένων μέσω αισθητήρων αλλά και για τον έλεγχο επενεργητών (Κινητήρων, εμβόλων, βαλβίδων).
- RTU (Remote Terminal Unit): Ουσιαστικά πρόκειται για συσκευές με ίδια λειτουργία με τα PLC, αλλά με μεγαλύτερη υπολογιστική ισχύ και περισσότερες δυνατότητες διασύνδεσης. Θεωρούνται επίσης περισσότερο στιβαρές και αξιόπιστες σε βιομηχανικά περιβάλλοντα.

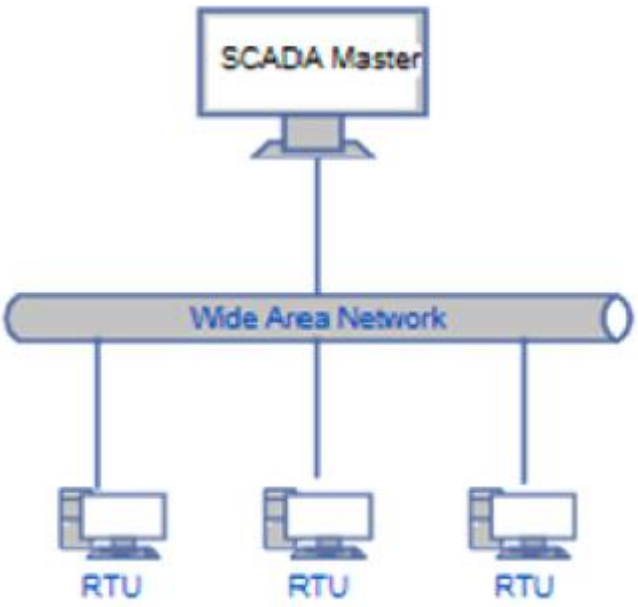
- IED (Intelligent Electronic Device): Είναι οποιαδήποτε συσκευή χρησιμοποιείται στην διεργασία και μπορεί να δώσει απευθείας δεδομένα στο κέντρο ελέγχου, χωρίς να χρειάζεται την υποστήριξη κάποιας άλλης συσκευής.
- HMI (Human Machine Interface): Χρησιμοποιούνται για την οπτικοποίηση των δεδομένων που στέλνουν οι συσκευές πεδίου στο κέντρο ελέγχου. Επιτρέπουν όχι μόνο την εποπτεία αλλά και τον έλεγχο του συστήματος, κάνοντας αλλαγές σε τιμές αναφοράς ή χειροκίνητο έλεγχο των επενεργητών. Ανάλογα με την διεργασία την οποία ελέγχει κάθε SCADA, οι συσκευές οπτικοποίησης μπορεί να έχουν διαφορετικά μεγέθη, από μεγάλες οθόνες μέχρι και το μέγεθος ενός κινητού τηλεφώνου.
- Historian: Τα Historian (Ιστορικά) είναι συστήματα διαχείρισης βάσεων δεδομένων τα οποία αποθηκεύουν όσα δεδομένα στέλνουν οι συσκευές πεδίου στο κέντρο ελέγχου. Τα δεδομένα αυτά είναι διαθέσιμα οποιαδήποτε στιγμή ο χρήστης τα χρειαστεί.
- MTU (Master Terminal Unit): Συχνά αναφέρεται και ως SCADA server και είναι υπεύθυνα για την επεξεργασία όλων των δεδομένων που φτάνουν στο κέντρο ελέγχου αλλά και για την επικοινωνία με όλες τις συσκευές πεδίου. Μπορούν να επεξεργαστούν τα δεδομένα πριν αυτά αποθηκευτούν στο Historian ενώ παρέχουν και την δυνατότητα γραφικής αναπαράστασης των δεδομένων που είναι αποθηκευμένα σε αυτό.

Η σύνθεση κάθε συστήματος SCADA μπορεί να διαφοροποιείται και να προσαρμόζεται στις ιδιαίτερες απαιτήσεις κάθε εφαρμογής.

2.2 Σύνδεση συστημάτων SCADA με internet και IT

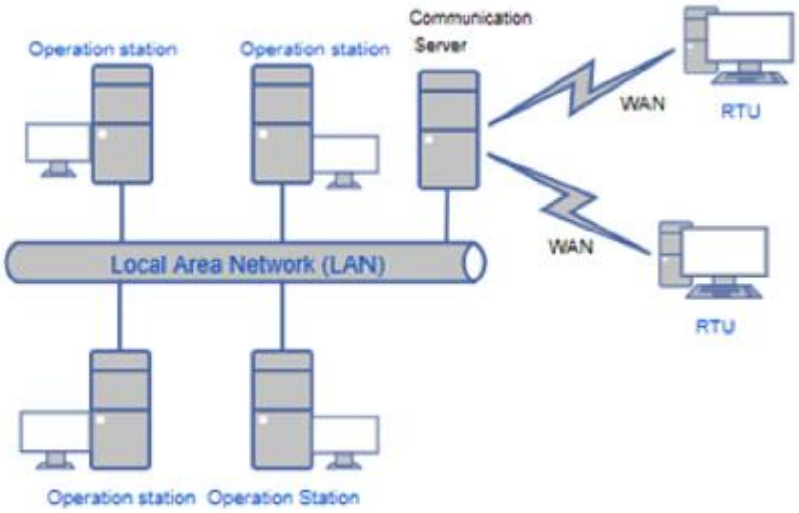
Όπως έχει ήδη αναφερθεί, τα πρώτα συστήματα SCADA είχαν αναπτυχθεί για απομονωμένη λειτουργία στον χώρο της διεργασίας την οποία ελέγχουν. Ωστόσο, με την ανάπτυξη της τεχνολογίας των υπολογιστών και επικοινωνιών τα σύγχρονα συστήματα SCADA είναι συνδεδεμένα στο internet αξιοποιώντας όλα τα πλεονεκτήματα τα οποία αυτό προσφέρει. Έτσι είναι δυνατόν ένα σύστημα SCADA να ελέγχεται πολύ πιο μακριά από την γεωγραφική του θέση [10]. Είναι επίσης δυνατόν να ελέγχεται από διαφορετικούς χρήστες αλλά και να συνδεθεί με άλλα διασυνδεδεμένα συστήματα (SAP, Server, Historian). Μαζί με τα πλεονεκτήματα που προσδίδει αυτή η διασύνδεση των SCADA στο internet, έρχεται

και ο κίνδυνος που δημιουργείται από τις διαδικτυακές επιθέσεις [10]. Ένα διασυνδεδεμένο σύστημα εποπτείες αντιμετωπίζει πλέον όλους τους κίνδυνους πιθανών επιθέσεων που αντιμετωπίζει ένα τυπικό διασυνδεδεμένο πληροφοριακό σύστημα. Έτσι μπορεί κάποιος να εκμεταλλευτεί κάποιες ευπάθειες του συστήματος και να θέσει τμήμα ή ολόκληρο το σύστημα εκτός λειτουργίας. Για να γίνει πιο κατανοητή η εξέλιξη των συστημάτων SCADA σχετικά με την συνδεσιμότητά τους στο internet, παρουσιάζονται οι παρακάτω αρχιτεκτονικές [3].



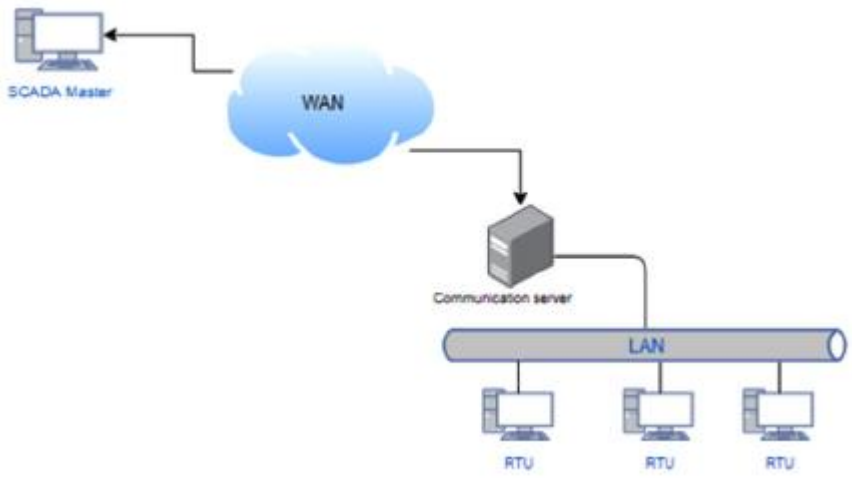
Εικόνα 3: Απομονωμένο σύστημα ελέγχου και εποπτείας

Το απομονωμένο σύστημα εποπτείας και ελέγχου, αποτελεί την πρώτη αρχιτεκτονική των συστημάτων SCADA, τότε που η έννοια της διασυνδεσιμότητας ήταν άγνωστη. Αποτελείται από ένα κεντρικό σύστημα (master) το οποίο συνδέεται με τις περιφερειακές μονάδες ελέγχου (RTU) μέσω ενός απομονωμένου ιδιόκτητου δικτύου (WAN). Συνήθως η δεύτερη και η τρίτη μονάδα ελέγχου λειτουργούν ως μονάδες ασφαλείας ενώ το δίκτυο χρησιμοποιείται αποκλειστικά για την επικοινωνία του κεντρικού συστήματος ελέγχου (master) με τις περιφερειακές μονάδες ελέγχου (RTU). Η επικοινωνία μέσω του δικτύου γινόταν μέσω πρωτοκόλλου που είχε αναπτύξει ο ίδιος ο κατασκευαστής, μην επιτρέποντας έτσι την επικοινωνία των συσκευών με συσκευές άλλου κατασκευαστή. Ήταν μια επικοινωνία κλειστή μη προσβάσιμη από τον καθένα. Ωστόσο αυτή η αρχιτεκτονική ήταν αρκετά ακριβή για να υλοποιηθεί κυρίως γιατί απαιτούσε να υπάρχουν συσκευές ασφαλείας οι οποίες αύξαναν το κόστος. Έτσι οι κατασκευαστές συστημάτων ελέγχου και εποπτείας εισήγαγαν την διανεμημένη αρχιτεκτονική των συστημάτων [3].



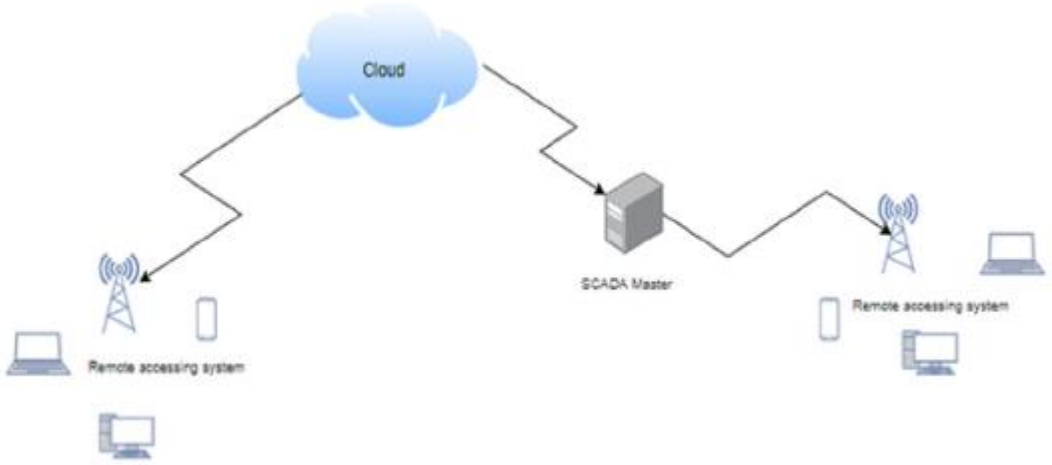
Εικόνα 4: Διανεμημένη αρχιτεκτονική συστημάτων ελέγχου και εποπτείας

Στην αρχιτεκτονική αυτή, πολλά μικρότερα αυτόνομα συστήματα επικοινωνούν μεταξύ τους μέσω ενός τοπικού δικτύου, ενώ η επικοινωνία των συστημάτων αυτών με το κεντρικό σύστημα ελέγχου (RTU), γίνεται μέσω ενός ιδιόκτητου δικτύου κλειστού πρωτοκόλλου. Η αρχιτεκτονική αυτή αυξάνει την αξιοπιστία των συστημάτων εποπτείας, εφόσον κάθε μικρό αυτόνομο σύστημα ελέγχει ένα τμήμα της συνολικής διεργασίας. Ωστόσο ακόμα και με αυτή την αρχιτεκτονική το σύστημα στηρίζεται σε ένα κλειστό πρωτόκολλο επικοινωνίας, εισάγοντας εμπόδια στην διασύνδεσή του με υλικό ή λογισμικό διαφορετικών κατασκευαστών. Η ανάπτυξη και η ζήτηση όμως της βιομηχανίας απαιτούσε την διασύνδεση των συστημάτων SCADA. Στην παρακάτω εικόνα παρουσιάζεται η αρχιτεκτονική ενός διασυνδεδεμένου συστήματος SCADA [3].



Εικόνα 5: Αρχιτεκτονική διασυνδεδεμένων συστημάτων ελέγχου και εποπτείας

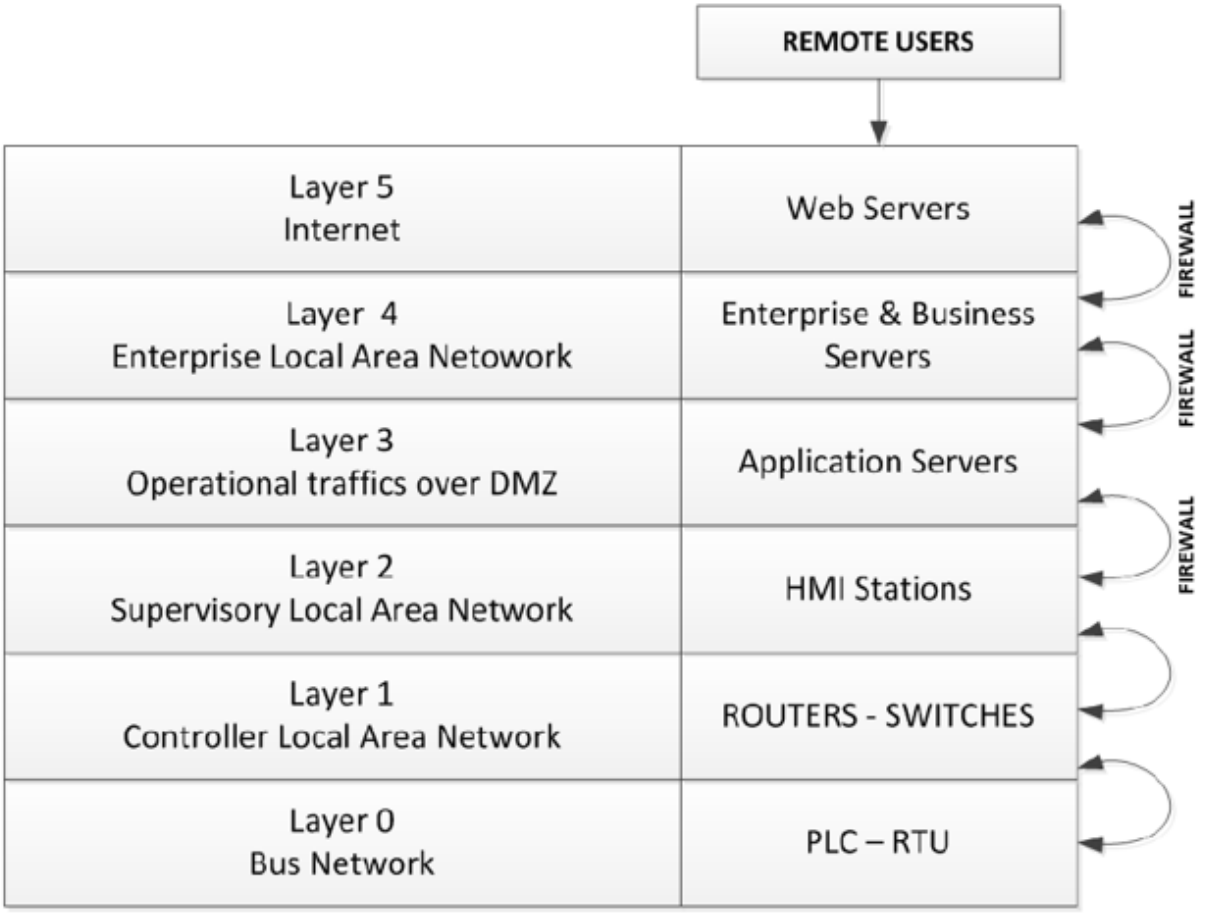
Η αρχιτεκτονική των διασυνδεδεμένων συστημάτων μοιάζει πολύ με την διανεμημένη αρχιτεκτονική αλλά διαφέρει στον τρόπο επικοινωνίας μεταξύ των τοπικών συστημάτων και του κεντρικού συστήματος ελέγχου. Για την επικοινωνία αυτή χρησιμοποιείται πλέον πρωτόκολλο διαδικτύου (internet protocol - IP) αντί του κλειστού ιδιόκτητου πρωτοκόλλου. Το πλεονέκτημα της αρχιτεκτονικής αυτής είναι ότι αυξάνεται η απόδοση των συστημάτων εφόσον μπορούν να διασυνδεθούν με υλικό και λογισμικό διαφορετικών κατασκευαστών και να εκτελούν ταυτόχρονα αρκετές λειτουργίες [36]. Το επόμενο βήμα στην διασύνδεση των συστημάτων SCADA ήταν η υιοθέτηση αρχιτεκτονικής σύννεφου (Cloud) και IoT (internet of Things), όπως παρουσιάζεται στην παρακάτω εικόνα [3].



Εικόνα 6: Αρχιτεκτονική σύννεφου συστημάτων ελέγχου και εποπτείας

Τα πλεονεκτήματα που εισάγει η αρχιτεκτονική αυτή των συστημάτων είναι ευκολία στην χρήση τους, ευελιξία, αξιοπιστία, μείωση κόστους καθώς και δυνατότητα επεξεργασίας και αποθήκευσης μεγάλου όγκου δεδομένων. Έτσι εξελίχθηκαν τα συστήματα SCADA από πλήρως απομονωμένα, σε πλήρως διασυνδεδεμένα.

Ένα σύγχρονο σύστημα SCADA μπορεί να χωριστεί σε 5 επίπεδα (Layers) όπως παρουσιάζεται στην παρακάτω εικόνα [39].



Εικόνα 7: Τα 5 επίπεδα των συστημάτων SCADA

Η διαβάθμιση των συσκευών σε διαφορετικά επίπεδα βοηθά στην κατανόηση της αρχιτεκτονικής του συστήματος, στην προστασία του από επιθέσεις αλλά και στην έρευνα των στοιχείων (φορένζικς) σε περίπτωση που δεχθεί κάποια απειλή. Στο επίπεδο 0 βρίσκονται όλες οι συσκευές πεδίου οι οποίες αλληλοεπιδρούν άμεσα με τα στοιχεία ελέγχου της διεργασίας. Οι συσκευές αυτές συνδέονται μεταξύ τους μέσω τοπικού βιομηχανικού δικτύου (bus) βασισμένο σε προκαθορισμένα πρωτόκολλα επικοινωνίας. Όλες οι πληροφορίες μεταφέρονται μέσω αυτού στο επόμενο επίπεδο (επίπεδο 1) στο οποίο ανήκουν οι συσκευές δρομολόγησης (routers) και είναι υπεύθυνα για την διασφάλιση και την εποπτεία της επικοινωνίας των συσκευών πεδίου. Στο επίπεδο 2 ανήκουν οι συσκευές οπτικοποίησης οι οποίες παρουσιάζουν τις πληροφορίες που τους παρέχονται από τα παρακάτω επίπεδα. Τα επίπεδα 3 και 4 παρουσιάζουν την δομή των πληροφοριακών συστημάτων όλου του οργανισμού όπου διαχειρίζονται εκτός από το SCADA και τις άλλες εφαρμογές που μπορεί κάθε οργανισμός να χρησιμοποιεί. Τέτοιες εφαρμογές μπορεί να είναι συστήματα διαχείρισης προσωπικού, υλικών, πωλήσεων ή και οργάνωσης παραγωγής

(Management Execution System - MES). Το επίπεδο 5, αντιπροσωπεύει το internet, δίνοντας την δυνατότητα σε όλες τις συσκευές και λογισμικά των κατώτερων επιπέδων να επικοινωνούν με οποιαδήποτε συσκευή η λογισμικό στον κόσμο.

Λαμβάνοντας υπόψιν την κρισιμότητα της λειτουργίας της διεργασίας που το SCADA ελέγχει, θα πρέπει να υπάρχει σύστημα συνεχούς εντοπισμού απειλών (live forensic) στα επίπεδα 0,1 και 2 [2]. Οι έρευνες επικεντρώνονται στην ανάπτυξη υψηλού επιπέδου λογισμικού το οποίο θα μπορεί να ανιχνεύσει ανώμαλες αλλαγές στις τιμές αισθητηρίων, στην μνήμη των συσκευών πεδίου ή ακόμα και διαφορές στον όγκο δεδομένων που διακινούνται μέσω του τοπικού δικτύου των συσκευών. Πρόκληση αποτελεί η ανάπτυξη ενός συστήματος συνεχούς ελέγχου το οποίο θα έχει την ελάχιστη επίδραση στην λειτουργία του συστήματος SCADA.

Δεδομένου ότι οι τιμές ελέγχου μιας διεργασίας η οποία ελέγχεται μέσω ενός SCADA συνεχώς αλλάζουν, θα πρέπει τα στοιχεία να συλλεχθούν όσο το δυνατόν γρηγορότερα μετά την εμφάνιση μιας απειλής πριν αντικατασταθούν από νέα δεδομένα, αν η διεργασία συνεχιστεί κανονικά μετά την επίθεση.

2.3 Ανάλυση απαιτήσεων ασφαλείας σε συστήματα SCADA και IT.

Από την οπτική των συστημάτων ασφαλείας είναι σημαντικό να καταλάβει κανείς τις διαφορετικές απαιτήσεις που υπάρχουν μεταξύ των συστημάτων SCADA και IT, εφόσον η φύση της λειτουργίας των δύο συστημάτων είναι διαφορετική [14]. Για παράδειγμα ένα σύστημα IT, αποτελούμενο από λογισμικό (software) και υλικό (hardware) μπορεί να χρησιμοποιήσει για την προστασία του ένα σύστημα firewall, χωρίς να επηρεάσει ουσιαστικά την απόδοσή του. Αντίθετα οι κατασκευαστές συστημάτων SCADA προβληματίζονται για την χρήση firewall καθώς αυτό μπορεί να εισάγει καθυστέρηση και μείωση της απόδοσης του συστήματος [35]. Επόμενος, ένας μελετητής του συστήματος ασφαλείας ενός συστήματος SCADA, θα πρέπει να εξετάσει λεπτομερώς τις απαιτήσεις του συστήματος πριν εγκαταστήσει κάποιο σύστημα ασφαλείας σε αυτό.

Οι προτεραιότητες ενός συστήματος IT είναι κατά σειρά, εμπιστευτικότητα (Confidentiality), ακεραιότητα (Integrity) και διαθεσιμότητα (Availability), (CIA), ενώ για ένα σύστημα SCADA την μεγαλύτερη προτεραιότητα την έχει η διαθεσιμότητα και ακολουθούν η ακεραιότητα και η εμπιστευτικότητα (AIC) . Ένα σύστημα SCADA το οποίο ελέγχει μια κρίσιμη διεργασία θα πρέπει να είναι πάντα λειτουργικό, και αν απαιτείται η

παύση της λειτουργίας του θα πρέπει αυτό να έχει προγραμματιστεί εκ των προτέρων. Σε αυτές τις περιπτώσεις θα πρέπει να υπάρχει ένα εφεδρικό σύστημα το οποίο θα μπορεί να αναλάβει τον έλεγχο ολόκληρης ή τμήμα της διεργασίας την οποία ελέγχει το κύριο σύστημα. Αντίθετα, ένα σύστημα IT δεν απαιτεί την ύπαρξη εφεδρικού συστήματος εφόσον η διαθεσιμότητα είναι χαμηλότερα στην ιεραρχία των προτεραιοτήτων ενός συστήματος IT σε σχέση με ένα σύστημα SCADA. Αυτό βέβαια εξαρτάται και από την φύση της εφαρμογής που υποστηρίζει το σύστημα IT [3].

Ένας άλλος παράγοντας που διαφοροποιεί τις απαιτήσεις ασφαλείας των δύο συστημάτων είναι η διαθεσιμότητα των πόρων. Συνήθως ένα σύστημα SCADA έχει λιγότερη διαθέσιμη υπολογιστική ισχύ σε σχέση με ένα σύστημα IT το οποίο μπορεί να χρησιμοποιήσει απλά εργαλεία ασφάλειας όπως η κρυπτογράφηση ή η προστασία με κωδικούς [22].

Ένας επιπλέον παράγοντας τον οποίο θα πρέπει ο σχεδιαστής ενός συστήματος ασφαλείας για SCADA είναι η δυσκολία και το κόστος για τον έλεγχο (testing) του συστήματος αυτού. Το κόστος για μια πραγματική δοκιμή ενός τέτοιου συστήματος είναι μεγάλο, εφόσον μια αστοχία θα μπορούσε να οδηγήσει σε καταστροφή στοιχείων του ελέγχου της διεργασίας [34]. Για το λόγο αυτό απαιτείται να αναπτυχθούν και προσομοιωτές για τον έλεγχο του συστήματος ασφαλείας. Για ένα σύστημα IT η μεθοδολογία ελέγχου του συστήματος ασφαλείας, δεν είναι τόσο δαπανηρή.

Τέλος, αξίζει να αναφερθεί ότι η επικοινωνία μεταξύ των στοιχείων που αποτελούν ένα σύστημα μπορεί να γίνεται με κλειστά ιδιόκτητα πρωτόκολλα του κατασκευαστή, τα οποία να μην υπακούν σε πρωτόκολλα ασφαλείας. Αντίθετα όλα τα μέρη ενός IT συστήματος, επικοινωνούν με προκαθορισμένα πρωτόκολλα τα οποία υπακούν σε διεθνή κανόνες ασφαλείας, κρυπτογράφησης και πιστοποίησης.

2.4 Εργαλεία για επιθέσεις σε SCADA

Έχουν αναπτυχθεί διάφορα εργαλεία για επιθέσεις σε συστήματα SCADA, μερικά από τα οποία αναφέρονται παρακάτω [3].

PLC-Blaster worm [37]: Ορισμένα PLC έχουν ανοιχτή την πόρτα επικοινωνίας 102 (TCP) και αυτή μπορεί να χρησιμοποιηθεί για την διεξαγωγή επίθεσης σε αυτά τα PLC. Ένα κακόβουλο λογισμικό μπορεί να μεταφερθεί στο PLC μέσω της πόρτας αυτής και να επηρεάσει τον κύκλο ζωής της συσκευής η ακόμα και το σταμάτημα της λειτουργίας της.

Μπορεί επίσης να επιτρέψει στον επιτιθέμενο να αλλάξει την κατάσταση των εξόδων του PLC.

Triton : Πρόκειται για λογισμικό (software) το οποίο μπορεί να προσβάλει το σύστημα ασφαλείας συστημάτων SCADA τα οποία χρησιμοποιούν επικοινωνία UDP μέσω της πόρτας 1502. Μπορεί να δώσει πρόσβαση στον εισβολέα στο λογισμικό που εκτελείται στο PLC, δίνοντάς του τη δυνατότητα να το αλλάξει ή να καθυστερήσει τον χρόνο εκτέλεσής του.

Havex [29]: Πρόκειται για κρυφό κακόβουλο λογισμικό το οποίο εγκαθίσταται στα συστήματα SCADA μέσω άλλου λογισμικού (Trojan horse). Χρησιμοποιεί OPC πρωτόκολλο επικοινωνίας για να επικοινωνεί με τον server του συστήματος και συλλέγει κρίσιμες πληροφορίες για την λειτουργία της διεργασίας που αυτό ελέγχει.

Industroyer : Λογισμικό το οποίο χρησιμοποιήθηκε σε επίθεση στο σύστημα ελέγχου του δικτύου ηλεκτρικής ενέργειας της Ουκρανίας το Δεκέμβριο του 2017. Είναι το πρώτο κακόβουλο λογισμικό (Malware) το οποίο σχεδιάστηκε στοχευμένα για επιθέσεις σε συστήματα ελέγχου δικτύων ηλεκτρικής ενέργειας. Νεότερη έκδοση του λογισμικού αυτού είναι το destroyer2 το οποίο χρησιμοποιήθηκε για την επίθεση στο σύστημα διανομής ηλεκτρικής ενέργειας της Ουκρανίας το 2022.

Stuxnet [8]: Είναι ίσως το πιο ισχυρό κακόβουλο λογισμικό το οποίο στοχεύει σε συστήματα τα οποία χρησιμοποιούν περιβάλλον Windows, μέσω του οποίου ανιχνεύει λογισμικό προγραμματισμού PLC. Μπορεί να τροποποιήσει το πρόγραμμα που εκτελείται στο PLC, αλλάζοντας την ροή του προγράμματος, επηρεάζοντας έτσι τις εντολές που δίνει αυτό στους επενεργητές. Μπορεί ακόμα να επηρεάσει και την επικοινωνία μεταξύ των συσκευών ενός συστήματος SCADA. Ανιχνεύτηκε πρώτη φορά το 2010 και χρησιμοποιήθηκε για την επίθεση σε πυρηνικό σταθμό στο Ιράν.

BlackEnergy [29]: Κακόβουλο λογισμικό (Malware) το οποίο αναπτύχθηκε σε γλώσσα HTTP, εμφανίστηκε για πρώτη φορά το 2007. Στοχεύει συστήματα SCADA προκαλώντας προβλήματα ή παύση της λειτουργίας τους. Μεταφέρεται μέσω κοινών αρχείων Word ή Powerpoint τα οποία επισυνάπτονται σε e-mails. Χρησιμοποιήθηκε το 2015 για την επίθεση στο σύστημα ελέγχου του δικτύου διανομής ηλεκτρικής ενέργειας της Ρωσίας.

2.5 Πρότυπα για την σχεδίαση συστημάτων ασφαλείας σε συστήματα SCADA

Δεδομένου ότι οι ευπάθειες της νέας μορφής των συστημάτων SCADA έχουν αναγνωριστεί και καταγραφεί, γίνεται προσπάθεια παγκοσμίως ώστε να δημιουργηθούν πρότυπα ασφαλείας τα οποία θα πρέπει να τηρούνται. Το διεθνές πρότυπο για την ασφάλεια των βιομηχανικών συστημάτων αυτοματισμού είναι το IEC 62443 [15]. Επίσης πολλοί εθνικοί οργανισμοί όπως ο NIST (National Institute of Standards and Technology) και ο NERC (North American Electric Reliability Corporation) στις ΗΠΑ, δημοσιεύουν προτάσεις και οδηγίες για συστήματα ασφαλείας σε SCADA [20].

Η οικογένεια προτύπων IEC 62443 προέρχεται από την επιτροπή IEC (International Electrotechnical Commission) μετά από συμφωνία όλων των εμπλεκόμενων εθνικών επιτροπών για την δημιουργία ενός κοινού πρότυπου. Το 2002 η διεθνής κοινότητα αυτοματισμού (International Society of Automation) ίδρυσε την διεθνή κοινότητα προτύπων ασφαλείας βιομηχανικού αυτοματισμού και συστημάτων ελέγχου (ISA99) [28]. Η κοινότητα αυτή δημοσίευσε τεχνικές αναφορές και οδηγίες σχετικά με την ασφάλεια των συστημάτων ελέγχου οι οποίες αποτέλεσαν τον κορμό για την δημιουργία του προτύπου IEC62443. Παράλληλα οι Γερμανικές κοινότητες μηχανικών VDI και VDE δημοσίευσε επίσης οδηγίες σχετικά με την ασφάλεια των συστημάτων βιομηχανικού ελέγχου, ο οποίες ενσωματώθηκαν στο πρότυπο IEC. Το 2021 η κοινότητα IEC δημοσίευσε την οικογένεια προτύπων IEC62443 (Industrial communications networks – Network and system security), ως παγκόσμιο πρότυπο για την ασφάλεια των συστημάτων βιομηχανικού ελέγχου [20].

Το πρότυπο αυτό χωρίζεται σε τέσσερα μέρη.

Γενικό (General): Στο μέρος αυτό καλύπτονται θέματα τα οποία είναι κοινά σε όλες τις σειρές προτύπων.

Πολιτικές και διαδικασίες (Policies and Procedures): Το μέρος αυτό επικεντρώνεται σε μεθόδους και διαδικασίες που αφορούν την ασφάλεια ενός συστήματος ελέγχου.

Σύστημα (System): Το μέρος αυτό αναφέρεται στις απαιτήσεις σε επίπεδο συστήματος.

Στοιχεία και απαιτήσεις (Components and Requirements): Στο μέρος αυτό, αναφέρονται συγκεκριμένες απαιτήσεις για προϊόντα βιομηχανικού αυτοματισμού. Στον παρακάτω πίνακα, παρουσιάζονται οι εκδόσεις για κάθε μέρος του προτύπου.

ΠΙΝΑΚΑΣ 1: Πρότυπο IEC62443

General	Policies and Procedures	System	Components and Requirements
1.1 Technical Specification Ed.1 July 2009	**2.1 Edition 1.0 November 2010**	**3.1 Technical Report, Edition 1.0, july 2009**	**4.1 Edition 1.0 January 2018**
Concepts and models	Security program requirements for IACS asset owners	Security technologies for industrial automation and control systems (IAC)	Secure product development lifecycle requirements
	2.2 Technical Report, Ed1 June 2015	**3.2 Edition 1.0, june 2020**	**4.1 Edition 1.0 February 2019**
	Patch management in the IACS enviroment	Security risk assessment and system design	Technical security requirements for IACS components
	2.3 Edition 1.1 August 2017	**3.3 Edition 1.0, August 2013**	
	Requirements for IACS service providers	System security requirements and security levels	

Το πρότυπο εισάγει επίσης την διαβάθμιση σε επίπεδα ασφαλείας (Security Level) για τα συστήματα (IEC 62443-3-3) και για τα προϊόντα (IEC 62443-4-2). Το κάθε επίπεδο ασφαλείας καθορίζει την αντοχή του συστήματος για κάθε επίπεδο απειλής που μπορεί αυτό να δεχθεί. Το πρότυπο επισημαίνει ότι το επίπεδο ασφαλείας πρέπει να καθορίζεται ανάλογα με την εφαρμογή και δεν είναι κατάλληλο για μια γενική ταξινόμηση προϊόντων. Τα επίπεδα ασφαλείας είναι τα εξής:

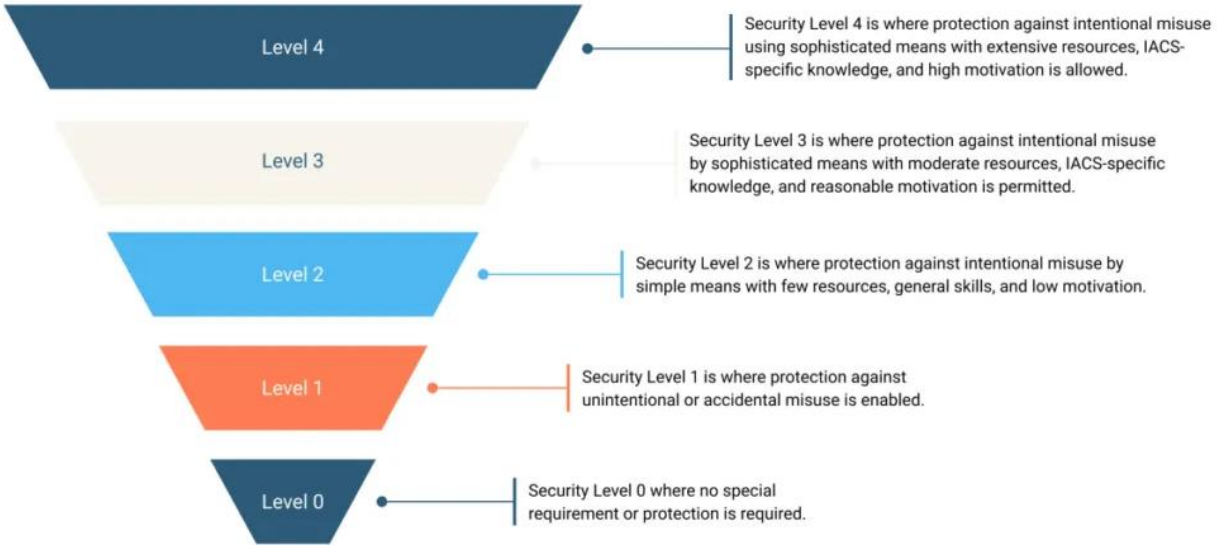
Επίπεδο 0: Δεν υπάρχουν ιδιαίτερες απαιτήσεις προστασίας.

Επίπεδο 1: Απαίτηση προστασίας κατά της απρόσεκτης ή κακής χρήσης του συστήματος.

Επίπεδο 2: Απαίτηση προστασίας από σκόπιμη κακή χρήση από ανθρώπους με περιορισμένα μέσα και δυνατότητες και χαμηλό κίνητρο.

Επίπεδο 3: Απαίτηση προστασίας από σκόπιμη κακή χρήση από ανθρώπους με αρκετά μέσα και ειδικές γνώσεις βιομηχανικού αυτοματισμού και μέτριο κίνητρο.

Επίπεδο 4: Απαίτηση προστασίας από σκόπιμη κακή χρήση (επίθεση) από ανθρώπους με εξειδικευμένα μέσα και ειδικές γνώσεις βιομηχανικού αυτοματισμού και υψηλό κίνητρο.



Εικόνα 8: Τα 5 επίπεδα ασφαλείας σύμφωνα με το πρότυπο IEC62443

Με βάση το παραπάνω πρότυπο μπορούν οι ενδιαφερόμενοι φορείς ή βιομηχανίες που χρησιμοποιούν συστήματα SCADA, να καθορίσουν το επίπεδο ασφαλείας που επιθυμούν και να προβούν στις απαραίτητες ενέργειες για να το πετύχουν.

2.6 Επιθέσεις σε SCADA, συστήματα καταγραφής και βάσεις δεδομένων.

Δεδομένου πλέον ότι τα συστήματα SCADA αποτελούν μέρος ενός γενικότερου πληροφοριακού συστήματος, στο κεφάλαιο αυτό γίνεται καταγραφή των πιθανών επιθέσεων που μπορεί ένα τέτοιο σύστημα να δεχθεί. Η διαδικασία αυτή είναι μέρος του σταδίου της προετοιμασίας μιας διερεύνησης επίθεσης όπως αυτό παρουσιάζεται αναλυτικά σε επόμενο κεφάλαιο. Μπορούμε να κατηγοριοποιήσουμε τις επιθέσεις σε τρείς τύπους [7].

1) Εξωτερικές επιθέσεις. Σε αυτό τον τύπο των επιθέσεων δεν απαιτούνται ιδιαίτερες γνώσεις ούτε κωδικοί του συστήματος το οποίο δέχεται την επίθεση.
2) Επιθέσεις ανάκτησης πληροφοριών ασφαλείας. Οι επιθέσεις αυτές σκοπό έχουν να αποσπάσουν κρίσιμους κωδικούς του συστήματος, ώστε να χρησιμοποιηθούν για την διεξαγωγή μιας εσωτερικής επίθεσης.
3) Εσωτερική επίθεση. Στις επιθέσεις αυτές ο επιτιθέμενος δρα ως εξουσιοδοτημένος χρήστης έχοντας ανακτήσει τους κωδικούς προστασίας του συστήματος.

Επίσης, οι επιθέσεις μπορούν να διαχωριστούν και σε δύο άλλες κατηγορίες.

1) Ενδιάμεση επίθεση. Κατά την επίθεση αυτή, ο επιτιθέμενος στοχεύει την ανάκτηση πληροφοριών για το επίπεδο ασφαλείας του συστήματος SCADA οι οποίες θα τον οδηγήσουν σε επιτυχημένη τελική επίθεση.
2) Τελική επίθεση. Κατά την επίθεση αυτή ο επιτιθέμενος καταφέρνει να επιτύχει τον τελικό του στόχο, δηλαδή να θέσει τμήμα η ολόκληρο το σύστημα SCADA εκτός λειτουργίας.

Παρακάτω παρουσιάζονται κάποιες από τις γνωστές εξωτερικές επιθέσεις που έχουν καταγραφεί στην βιβλιογραφία.

Spoofing attack: Κατά την επίθεση αυτή μια μη πιστοποιημένη συσκευή, αντικαθιστά μια πιστοποιημένη συσκευή του συστήματος. Αν οι συσκευές του συστήματος δεν είναι πιστοποιημένες κατάλληλα, αυτός ο τύπος εξωτερικής επίθεσης είναι πολύ εύκολο να συμβεί. Αν μια πιστοποιημένη συσκευή αντικατασταθεί από μια μη πιστοποιημένη από τον επιτιθέμενο, και το σύστημα δεν μπορέσει να το αναγνωρίσει, μπορεί να τροποποιήσει τα δεδομένα που θα έστελνε η πιστοποιημένη συσκευή, θέτοντας το σύστημα σε κίνδυνο. Κατηγορία επίθεσης: Τελική επίθεση.

Jamming attack: Πρόκειται για την παρέμβαση στις ραδιοσυχνότητες που χρησιμοποιούν οι συσκευές για την μεταξύ τους επικοινωνία. Μπορεί να επηρεάσει την διαθεσιμότητα των συσκευών στο σύστημα. Κατηγορία επίθεσης: Τελική επίθεση.

Replay attack: Κατά την επίθεση αυτή τα πακέτα πληροφοριών μεταξύ των συσκευών επαναλαμβάνονται ή καθυστερούν. Κατηγορία επίθεσης: Ενδιάμεση επίθεση.

Wormhole attack: Κατά την επίθεση αυτή ο επιτιθέμενος χρησιμοποιεί ένα δίαυλο χαμηλής καθυστέρησης για να μεταφέρει πακέτα πληροφορίας από μια συσκευή σε άλλη. Εφαρμόζεται στα ασύρματα δίκτυα επικοινωνίας και μπορεί ακόμα και να τροποποιήσει την πληροφορία που μεταφέρεται μέσω των πακέτων. Κατηγορία επίθεσης: Ενδιάμεση επίθεση.

Destroying/displacement a node: Αν οι συσκευές του συστήματος δεν βρίσκονται σε φυλασσόμενο χώρο, μπορεί ο επιτιθέμενος να καταστρέψει ολοκληρωτικά ή να αντικαταστήσει τη συσκευή. Κατηγορία επίθεσης: Τελική επίθεση.

Environmental tampering: Κατά την επίθεση αυτή, ο επιτιθέμενος αλλάζει τις συνθήκες περιβάλλοντος που η συσκευή επιτηρεί. Για παράδειγμα μπορεί να τοποθετήσει ένα μαγνήτη δίπλα σε μια συσκευή μέτρησης μαγνητικού πεδίου ή μια πηγή θερμότητας διπλά

σε μια συσκευή μέτρησης θερμοκρασίας, αλλοιώνοντας τα πραγματικά δεδομένα. Κατηγορία επίθεσης: Τελική επίθεση.

Στην συνέχεια παρουσιάζονται κάποιοι τύποι επιθέσεων ανάκτησης πληροφοριών ασφαλείας.

Cryptanalysis: Στην επίθεση αυτή, ο επιτιθέμενος προσπαθεί να αποκρυπτογραφήσει τα πακέτα πληροφοριών χωρίς να έχει περισσότερες γνώσεις για τον τρόπο κρυπτογράφησης. Στο χειρότερο σενάριο, θα μπορέσει να ανακτήσει πληροφορίες ή ακόμα και να τις τροποποιήσει. Κατηγορία επίθεσης: Ενδιάμεση επίθεση.

Exploit: Κατά την επίθεση αυτή ο επιτιθέμενος εκμεταλλεύεται κάποια ευπάθεια του λογισμικού ώστε να ανακτήσει πληροφορίες ασφαλείας για το σύστημα. Κατηγορία επίθεσης: Ενδιάμεση επίθεση.

Τέλος παρουσιάζονται κάποιοι τύποι εσωτερικών επιθέσεων για τις οποίες ο επιτιθέμενος θα πρέπει να γνωρίζει πληροφορίες ασφαλείας του συστήματος (δομή, κωδικοί).

Sybil: Κατά την επίθεση αυτή χρησιμοποιούνται μη πιστοποιημένες συσκευές οι οποίες αντιγράφουν μία ή περισσότερες ταυτότητες πιστοποιημένων συσκευών. Κατηγορία επίθεσης: Ενδιάμεση επίθεση.

Replication: Κατά την επίθεση αυτή ο επιτιθέμενος προσπαθεί να εισάγει στο δίκτυο μία η περισσότερες συσκευές οι οποίες χρησιμοποιούν την ίδια ταυτότητα (ID) με τις συσκευές του δικτύου. Κατηγορία επίθεσης: Ενδιάμεση επίθεση.

Denial of service at the link layer: Παραδείγματα τέτοιων επιθέσεων είναι η παρεμπόδιση της μετάδοσης των πακέτων πληροφοριών μεταξύ των συσκευών, εξασθένιση της μπαταρίας των συσκευών λόγω συνεχούς μετάδοσης πακέτων πληροφοριών και μη ισορροπημένη κατανομή των πόρων του δικτύου για την επικοινωνία των συσκευών. Κατηγορία επίθεσης: Τελική επίθεση.

Routing attacks: Κατά τις επιθέσεις αυτού του τύπου ο επιτιθέμενος προσπαθεί να δημιουργήσει βρόχους δρομολόγησης ή να δρομολογήσει τα πακέτα πληροφοριών σε λάθος παραλήπτη. Επίσης δημιουργεί αυξημένη κίνηση πακέτων περιορίζοντας τους πόρους του συστήματος. Κατηγορία επίθεσης: Τελική επίθεση.

Time-Synchronization attack: Οι επιθέσεις του τύπου αυτού στοχεύουν στα ρολόγια τα οποία συγχρονίζουν την επικοινωνία μεταξύ των συσκευών. Επομένως όταν οι υπάρχει μια

τέτοια επίθεση σε ένα σύστημα παρεμποδίζεται η επικοινωνία με τις συσκευές οι οποίες δεν είναι συγχρονισμένες με το δίκτυο. Κατηγορία επίθεσης: Τελική επίθεση.

Slander attack:

Wormhole attack: Αυτός ο τύπος επίθεσης καταγράφεται και ως εξωτερική επίθεση, ωστόσο ένας εσωτερικός εισβολέας μπορεί να χρησιμοποιήσει αυτή την τεχνική αποτελεσματικότερα, εφόσον μπορεί να αναγνωρίσει τα διαφορετικά είδη πακέτων πληροφορίας που διακινούνται μέσα στον δίαυλο επικοινωνίας. Κατηγορία επίθεσης: Ενδιάμεση επίθεση.

Στην συνέχεια καταγράφονται ορισμένοι ακόμα τύποι επιθέσεων οι οποίοι μπορούν να εμφανιστούν σε ένα διασυνδεδεμένο σύστημα SCADA [43].

Buffer overflow: Κατά την επίθεση αυτή επηρεάζεται η μνήμη του συστήματος, καθώς ο όγκος των δεδομένων που πρέπει να αποθηκευτούν σε αυτή είναι μεγαλύτερος από την χωρητικότητά της. Αποτέλεσμα αυτού είναι να επηρεάζεται ο αλγόριθμος που εκτελείται στο συγκεκριμένο σύστημα.

SQL Injection: Είναι ένας από τους πιο διαδεδομένους τύπους διαδικτυακών επιθέσεων. Κατά την επίθεση αυτή ο επιτιθέμενος χρησιμοποιεί SQL εντολές για την πρόσβαση σε απόρρητες πληροφορίες βάσεων δεδομένων.

Idle Scan: Κατά την μέθοδο αυτή, μια συνδεδεμένη σε δίκτυο συσκευή δέχεται πακέτα από μια άλλη συσκευή. Η συσκευή στόχος απαντάει διαφορετικά σε αυτά τα πακέτα, ανάλογα με τις λειτουργίες που είναι ενεργοποιημένες σε αυτή τη συσκευή.

Smurf attack: Κατά την μέθοδο αυτή, ο επιτιθέμενος εισάγει στο δίκτυο τροποποιημένα πακέτα εντολών τα οποία ενεργοποιούν κάποια λειτουργία των άλλων συνδεδεμένων συσκευών του δικτύου που δέχονται αυτά τα πακέτα.

SYN flood: Στην επίθεση αυτή, ο επιτιθέμενος στέλνει πακέτα πληροφοριών στις συνδεδεμένες συσκευές γρηγορότερα από τον ρυθμό με τον οποίο μπορούν αυτές να τα επεξεργαστούν. Δεδομένου ότι η επικοινωνία των SCADA συστημάτων βασίζεται στο πρωτόκολλο επικοινωνίας TCP/IP, αυτό είναι ένα είδος επίθεσης που μπορεί να εφαρμοστεί σε τέτοια συστήματα.

DNS (Domain Name System) forgery: Κατά την επίθεση αυτή αποστέλλεται μια ψεύτικη DNS απάντηση η οποία περιέχει την σωστή διεύθυνση IP και θύρα, αλλά τροποποιημένο περιεχόμενο εντολών.

Όλοι οι παραπάνω τύποι επιθέσεων, μπορούν να επηρεάσουν σημαντικούς παράγοντες ενός συστήματος SCADA όπως η ακεραιότητα (integrity), η διαθεσιμότητα (availability) και η εμπιστευτικότητα (Confidentiality).

Η παραπάνω συνοπτική καταγραφή των τύπων των επιθέσεων θα μας βοηθήσει να καταλάβουμε τις ευπάθειες και τις απειλές στις οποίες εκτίθεται ένα διασυνδεδεμένο σύστημα εποπτείας και ελέγχου. Επίσης όπως έχει ήδη αναφερθεί ισχυροποιείται η ανάγκη για συνεχή έλεγχο αλλά και καταγραφή των στοιχείων που προκύπτουν από τις επιθέσεις, κυρίως όταν πρόκειται για συστήματα που υποστηρίζουν σημαντικές κοινωνικές υποδομές όπως σταθμούς παραγωγής ενέργειας ή συστήματα μεταφορών (Critical Energy Infrastructure- CEI). Για το σκοπό αυτό έχουν αναπτυχθεί αρκετά εργαλεία διαχείρισης συμβάντων και πληροφοριών ασφαλείας (Security information and event management- SIEM). Γνωστά λογισμικά διαχείρισης συμβάντων ανοιχτού κώδικά είναι τα Wazuh και OSSIM. Ωστόσο, μια βελτίωση που θα μπορούσε να ενσωματωθεί σε τέτοιου είδους λογισμικά είναι ένα αυτοματοποιημένο σύστημα φορένζικς [40]. Με τον όρο φορένζικς εννοούμε την καταγραφή, την ταξινόμηση και την συλλογή στοιχείων σχετικά με τον τύπο της επίθεσης. Συγκεκριμένα ένα σύστημα φορένζικς θα μπορούσε να εκτελεί αυτοματοποιημένα τις παρακάτω λειτουργίες.

1) Να καταγράψει τις μεθόδους τις οποίες χρησιμοποίησε ο επιτιθέμενος ώστε να προκύψουν σημαντικές πληροφορίες σχετικά με τους πόρους, τις δεξιότητες και τα βήματα που χρησιμοποιήθηκαν για την επίθεση.
2) Να καταγράψει τις σχετιζόμενες ευπάθειες του συστήματος τις οποίες εκμεταλλεύτηκε ο επιτιθέμενος για την επίθεση.
3) Να συσχετίσει την επίθεση με παρόμοιες επιθέσεις του παρελθόντος.
4) Να κατηγοριοποιήσει, να ιεραρχήσει και παρουσιάσει τις απειλές που δέχεται ένα σύστημα.
5) Να οργανώσει τα στοιχεία τις επίθεσης σε μορφή αναφοράς, να τα ψηφιοποιήσει και να τα συνδέσει με το νομικό πλαίσιο που καλύπτει τις διαδικτυακές επιθέσεις.

Γενικά, οι βασικοί στόχοι ενός εργαλείου φορένζικς είναι η μεγιστοποίηση της διαθεσιμότητας και τις ποιότητας των στοιχείων της επίθεσης, η διατήρηση της

ακεραιότητας των στοιχείων και να συμπεριλάβει τα βασικά συστατικά της έρευνας τα οποία είναι η προετοιμασία, η έρευνα-επεξεργασία των στοιχείων και η δημιουργία της τελικής αναφοράς. Η διαδικασία της έρευνας μπορεί να χωριστεί σε επιμέρους τμήματα, όπως είναι η εξέταση του hardware και του software του συστήματος. Παρόλο που έχουν προταθεί και αναπτυχθεί διάφορες προσεγγίσεις λογισμικού φορένζικς, κανένα από αυτά δεν αξιοποιεί την γνώση παλαιότερων επιθέσεων όπως αυτή έχει καταγραφεί σε ανοιχτού λογισμικού βάσεις δεδομένων. Η γνώση αυτή μπορεί να εμπεριέχει πληροφορίες σχετικά με τον τύπο της επίθεσης, τις ευπάθειες του συστήματος που εκμεταλλεύτηκε ο επιτιθέμενος καθώς και για τις δράσεις που έγιναν ώστε να αποτραπεί μια παρόμοια επίθεση στο μέλλον. Κάποιες από τις πιο γνωστές βάσεις δεδομένων σχετικά με διαδικτυακές επιθέσεις είναι οι παρακάτω.

- CAPEC (Common Attack Pattern Enumeration and Classification) [27]: Μια ανοιχτή βάση δεδομένων η οποία περιέχει τύπους επιθέσεων και τις ευπάθειες του λογισμικού ή του υλικού που αυτές εκμεταλλεύτηκαν. Παρέχει πληροφορίες για τους πόρους που χρησιμοποιήθηκαν και προτείνει μέτρα για την αντιμετώπισή τους.
- MITRE ATT&CK (Adversial Tactics Techniques, and Common Knowledge) [11]: Επίσης ανοιχτή βάση δεδομένων η οποία περιέχει τεχνικές και τρόπους επιθέσεων που χρησιμοποιήθηκαν σε πραγματικές επιθέσεις. Επικεντρώνεται κυρίως σε επιθέσεις που αφορούν το δίκτυο.

Κατά το παρελθόν έχουν γίνει κάποιες προσπάθειες ώστε να συνδεθούν αυτές οι βάσεις δεδομένων με ένα σύστημα φορένζικς χρησιμοποιώντας διάφορες τεχνικές αναγνώρισης ομοιότητας κειμένου. Οι επικρατέστερες από τις τεχνικές αυτές είναι οι doc2vec και η tf-idf. Έχουν επίσης αναπτυχθεί τεχνικές που κάνουν χρήση νευρονικών δικτύων και μοντέλων βαθιάς μάθησης οι οποίες ανάγουν το πρόβλημα της συσχέτισης κειμένου σε πρόβλημα συσχέτισης αριθμητικών παραστάσεων. Αναλυτικότερα για τις τεχνικές αυτές και τον τρόπο που χρησιμοποιούνται θα αναφερθούμε σε επόμενο κεφάλαιο της παρούσας εργασίας.

Εκτός από τα εργαλεία διαχείρισης συμβάντων και πληροφοριών ασφαλείας (SIEM) [18] έχουν αναπτυχθεί και εργαλεία ανίχνευσης απειλών (Intrusion Detection Systems - IDS). Τα συστήματα ανίχνευσης απειλών μπορούν να επιτηρούν διάφορες λειτουργίες και καταστάσεις του συστήματος και να ενημερώνουν το χρήστη σε περίπτωση που ανιχνεύσουν κάποια απειλή. Για παράδειγμα μπορούν να ανιχνεύσουν επίμονες προσπάθειες σύνδεσης στο σύστημα (logging activities) ή κάποια βλάβη στα αρχεία του

συστήματος μετά από σάρωση. Μπορούν επίσης να ανιχνεύσουν έγκαιρα κακόβουλο λογισμικό, πολύ πιο γρήγορα από τα συνηθισμένα antivirus εργαλεία. Ακόμα, μπορούν να κάνουν επιτήρηση του δικτύου και να εντοπίσουν απειλές που σχετίζονται με την μεταφορά δεδομένων μεταξύ των σταθμών ενός SCADA συστήματος. Μειονέκτημα των συστημάτων αυτών είναι ότι καταναλώνουν υπολογιστικούς πόρους του συστήματος.

2.7 Τεχνικές ανάλυσης επιθέσεων σε συστήματα ελέγχου και εποπτείας.

Η διαδικασία των ερευνών επιθέσεων σε ένα σύστημα εποπτείας, δεν πρέπει να γίνεται μόνο μετά από μία επίθεση σε αυτό αλλά ακόμα και πριν και κατά την διάρκεια της επίθεσης. Όσο περισσότερες λεπτομέρειες γνωρίζει κανείς για το σύστημα εποπτείας που δέχεται την επίθεση, τόσα περισσότερα στοιχεία θα μπορέσει να συλλέξει για την επίθεση που αυτό δέχθηκε. Παρακάτω, παρουσιάζεται μια προτεινόμενη διαδικασία ανίχνευσης και καταγραφής στοιχείων μιας επίθεσης σε ένα σύστημα SCADA, η οποία διαχωρίζεται σε τέσσερα στάδια. Προετοιμασία, ανίχνευση, ιεράρχηση και απάντηση. Το τελευταίο στάδιο βοηθάει στην βελτίωση της διαδικασίας της έρευνας για την επόμενη φορά. Παρακάτω παρουσιάζονται αναλυτικά τα προτεινόμενα στάδια των ερευνών [12].

Στάδιο 1: Προετοιμασία.

Το πρώτο στάδιο σε μία διαδικασία ερευνών για επίθεση σε ένα σύστημα εποπτείας, υλοποιείται πριν ακόμα το σύστημα αυτό δεχθεί την επίθεση. Πρόκειται για την κατανόηση της αρχιτεκτονικής του συστήματος εποπτείας καθώς των πιθανών τύπων επιθέσεων που μπορεί αυτό να δεχθεί αλλά και για τους τρόπους που μπορεί το σύστημα να αντιμετωπίσει αυτές τις επιθέσεις.

Καθώς τα συστήματα εποπτείας διαφέρουν μεταξύ τους, είναι σημαντικό για κάθε σύστημα να έχουμε εγγράφως αποτυπωμένα την δομή του δικτύου αλλά και τα τεχνικά χαρακτηριστικά των υποσυστημάτων που το αποτελούν. Σημαντικά στοιχεία αποτελούν ο κατασκευαστής και το μοντέλο των υποσυστημάτων του συστήματος μας καθώς και το λογισμικό που είναι εγκατεστημένο σε κάθε ένα από αυτά. Η γεωγραφική αποτύπωση των υποσυστημάτων θα μπορούσε επίσης να βοηθήσει κάποιον ερευνητή στο έργο του. Σχετικά με την κρισιμότητα του κάθε υποσυστήματος, θα πρέπει να υπάρχει μία ιεράρχηση ώστε να γνωρίζει ο ερευνητής αν αυτό μπορεί να σταματήσει την λειτουργία του ή αν υπάρχει

κάποιο back up το οποίο μπορεί να επαναφέρει το υποσύστημα σε προηγούμενη λειτουργική κατάσταση. Τέλος, σχετικά με τον τύπο των επιθέσεων, θα πρέπει να προσδιοριστούν οι τύποι επιθέσεων στους οποίους είναι εκτεθειμένο το σύστημα εποπτείας που εξετάζεται. Γενικά οι τύποι επιθέσεων μπορούν να διαχωριστούν σε τρείς υποκατηγορίες και να εξεταστούν αναλυτικά. Μια επίθεση μπορεί να έχει στόχο υλικό του συστήματος (hardware) , λογισμικό του συστήματος (software) ή το δίκτυο του συστήματος (network). Αναλυτική περιγραφή των επιθέσεων θα ακολουθήσει σε επόμενο κεφάλαιο της παρούσας εργασίας.

Στάδιο 2: Αναγνώριση.

Στο στάδιο της αναγνώρισης ο ερευνητής θα πρέπει αρχικά να αναγνωρίσει τον τύπο της επίθεσης που δέχθηκε το σύστημα που εξετάζει. Αυτό μπορεί να γίνει εξετάζοντας το σύστημα και αναγνωρίζοντας τις δυσλειτουργίες τις οποίες η επίθεση έχει προκαλέσει. Στην συνέχεια θα πρέπει να αναγνωριστούν τα υποσυστήματα του συστήματος εποπτείας που έχουν δεχθεί την επίθεση και αν είναι δυνατόν να απομονωθούν από το υπόλοιπο σύστημα. Αυτό θα βοηθήσει στο επόμενο στάδιο της διαδικασίας όπου αναζητούνται οι πιθανές πηγές στοιχείων για την επίθεση.

Στάδιο 3: Ιεράρχηση.

Αξιοποιώντας στοιχεία από το πρώτο στάδιο της διαδικασίας, σχετικά με τον τύπο των στοιχείων που αποτελούν το σύστημα εποπτείας, αλλά και τις ενδείξεις σχετικά με τον τύπο της επίθεσης που το σύστημα έχει δεχθεί, μπορούν να προσδιοριστούν κάποιες πηγές στοιχείων για την έρευνα. Στην συνέχεια οι πηγές αυτές θα πρέπει να κατηγοριοποιηθούν με βάση τα στοιχεία που μπορεί να προσφέρει η κάθε μία ώστε να μεγιστοποιηθεί ο συνολικός αριθμός των στοιχείων που μπορούν να συλλεχθούν για την επίθεση στο σύστημα.

Στάδιο 4: Απάντηση.

Το στάδιο αυτό περιλαμβάνει την συλλογή των στοιχείων από τις διάφορες πηγές, την ανάλυση και την συσχέτιση αυτών ώστε να προκύψει η τελική αναφορά για την επίθεση. Οι κύριες πηγές στοιχείων είναι τα δεδομένα τα οποία είναι αποθηκευμένα σε διάφορες συσκευές του συστήματος εποπτείας αλλά και στα δεδομένα που διακινούνται μέσω του δικτύου του. Για να μπορούν τα στοιχεία που προκύπτουν από τα δεδομένα των

υποσυστημάτων να έχουν νομική υπόσταση, θα πρέπει να έχουν εξαχθεί με ενδεδειγμένες για το σκοπό αυτό μεθοδολογίες. Τέλος, κατά το στάδιο αυτό συντάσσεται η τελική αναφορά για την επίθεση η οποία περιλαμβάνει τα στοιχεία που προέκυψαν και την συσχέτιση μεταξύ τους. Πρέπει επίσης να περιέχει προτάσεις βελτίωσης της ασφάλειας του συστήματος από παρόμοιες μελλοντικές επιθέσεις.

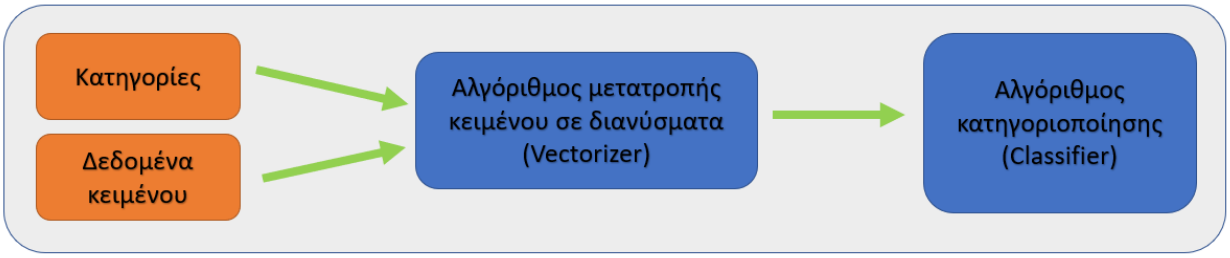
Κεφάλαιο 3: Περιγραφή ταξινόμησης κειμένου (Topic classification)

3.1 Γενική περιγραφή διαδικασίας ταξινόμησης κειμένου

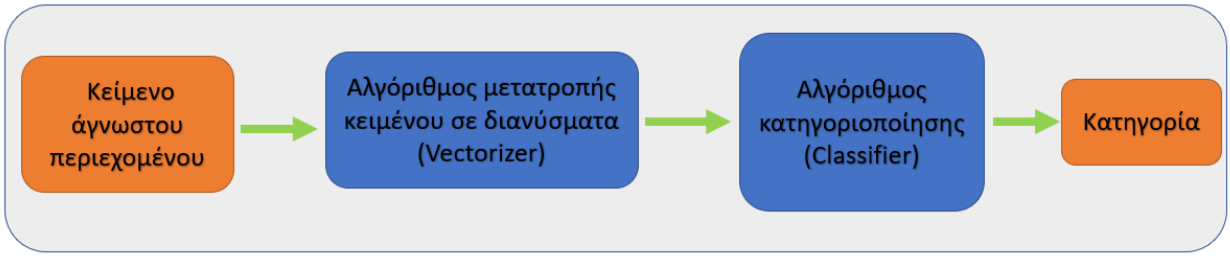
Τα δεδομένα κειμένου είναι ένας από τους πιο κοινούς τύπους δεδομένων που μπορεί κανείς να επεξεργαστεί ώστε να αντλήσει τις πληροφορίες που επιθυμεί. Όμως επειδή δεν έχουν προκαθορισμένη και σαφή δομή, μπορεί να είναι δύσκολο και χρονοβόρο να εξαχθούν οι πληροφορίες που περιέχουν. Η επεξεργασία των δεδομένων κειμένου εμπίπτει στην επεξεργασία φυσικής γλώσσας (Natural Language Processing - NLP) [6] που αποτελεί ένα μεγάλο τομέα της τεχνητής νοημοσύνης (Artificial Intelligence - AI) [5]. Ο τομέας της επεξεργασίας της φυσικής γλώσσας (NLP) ανήκει στην επιστήμη των υπολογιστών και της τεχνητής νοημοσύνης και εξετάζει τον τρόπο αλληλεπίδρασης των υπολογιστών με τις ανθρώπινες γλώσσες και τον τρόπο προγραμματισμού των υπολογιστών ώστε να επεξεργάζονται και να αναλύουν μεγάλες ποσότητες δεδομένων φυσικής γλώσσας. Κάποιες από τις εφαρμογές της τεχνικής NLP είναι οι αυτοματοποιημένες απαντήσεις σε ερωτήσεις, η δημιουργία περίληψης κειμένου ή η μετάφραση κειμένων από μια γλώσσα σε άλλη. Η ταξινόμηση κειμένου είναι επίσης μια περίπτωση χρήσης της NLP. Η κατηγοριοποίηση του κειμένου, μπορεί να γίνει με αλγορίθμους οι οποίοι δημιουργούν μόνοι τους τις κατηγορίες στις οποίες θα διαχωριστούν τα δεδομένα κειμένου. Μια τέτοια τεχνική ονομάζεται μη επιβλεπόμενη (Unsupervised – Topic Modeling) [23]. Οι μη επιβλεπόμενες τεχνικές είναι πολύ πιο γρήγορες στην εφαρμογή τους, έχουν όμως λιγότερη ακρίβεια στα αποτελέσματά τους. Τεχνικές στις οποίες οι κατηγορίες είναι εκ των προτέρων καθορισμένες και εισάγονται ως δεδομένα στο μοντέλο ονομάζονται επιβλεπόμενες (Supervised – Topic Classification). Στην επιβλεπόμενη τεχνική κάποια από τα δεδομένα συνδέονται με τις προκαθορισμένες κατηγορίες και το μοντέλο διδάσκεται ώστε να κατατάσσει δεδομένα άγνωστου περιεχομένου στις κατάλληλες κατηγορίες. Η διαδικασία αυτή απαιτεί περισσότερο χρόνο όμως έχει καλύτερα και ακριβέστερα αποτελέσματα. Όσο καλύτερα είναι τα αποτέλεσμα από μια τεχνική κατηγοριοποίησης κειμένου, τόσο πιο ακριβής θα είναι οι πληροφορίες που εξάγονται από την επεξεργασία των δεδομένων [19].

3.2 Περιγραφή επεξεργασίας κειμένου δεδομένων

Στην παρούσα διπλωματική παρουσιάζεται και εφαρμόζεται η τεχνική της επιβλεπόμενης ταξινόμησης κειμένου. Στην τεχνική αυτή οι κατηγορίες εισάγονται στον αλγόριθμο και συνδέονται με κάποια δεδομένα. Η διαδικασία αυτή αναφέρεται ως εκμάθηση του μοντέλου και τα δεδομένα που χρησιμοποιούνται ως δεδομένα εκπαίδευσης. Ωστόσο, εφόσον ο αλγόριθμος δεν μπορεί να κατανοήσει και να επεξεργαστεί κείμενο, τα δεδομένα αυτά θα πρέπει να μετατραπούν σε διανύσματα αριθμών τα οποία θα περιέχουν την πληροφορία που υπάρχει μέσα στο κείμενο. Στην συνέχεια, μπορεί ο αλγόριθμος να αναγνωρίσει ομοιότητες, να εξάγει τις πληροφορίες από το κείμενο και να το κατατάξει στην σωστή κατηγορία. Για την μετατροπή κειμένου σε διανύσματα έχουν αναπτυχθεί αρκετές τεχνικές όπως είναι όπως η bag of words, tf-idf [33] ή η WordtoVec [25]. Αφού τα δεδομένα εκπαίδευσης έχουν μετατραπεί σε διανύσματα αριθμών, συνδέονται με τις προκαθορισμένες κατηγορίες και δημιουργείται έτσι το μοντέλο ταξινόμησης. Αυτή είναι η διαδικασία εκπαίδευσης του μοντέλου ταξινόμησης ώστε να είναι ικανό να ταξινομήσει αυτόματα νέα δεδομένα στις κατηγορίες. Οι πιο γνωστοί αλγόριθμοι για μοντέλα ταξινόμησης είναι ο Naïve Bayes [26] ο K-means [30] και ο Multinimial Logistic Regression [1]. Η διαδικασία της εκπαίδευσης και η κανονική λειτουργία του μοντέλου, παρουσιάζονται γραφικά στα παρακάτω σχήματα.



Εικόνα 9: Διαδικασία εκπαίδευσης μοντέλου



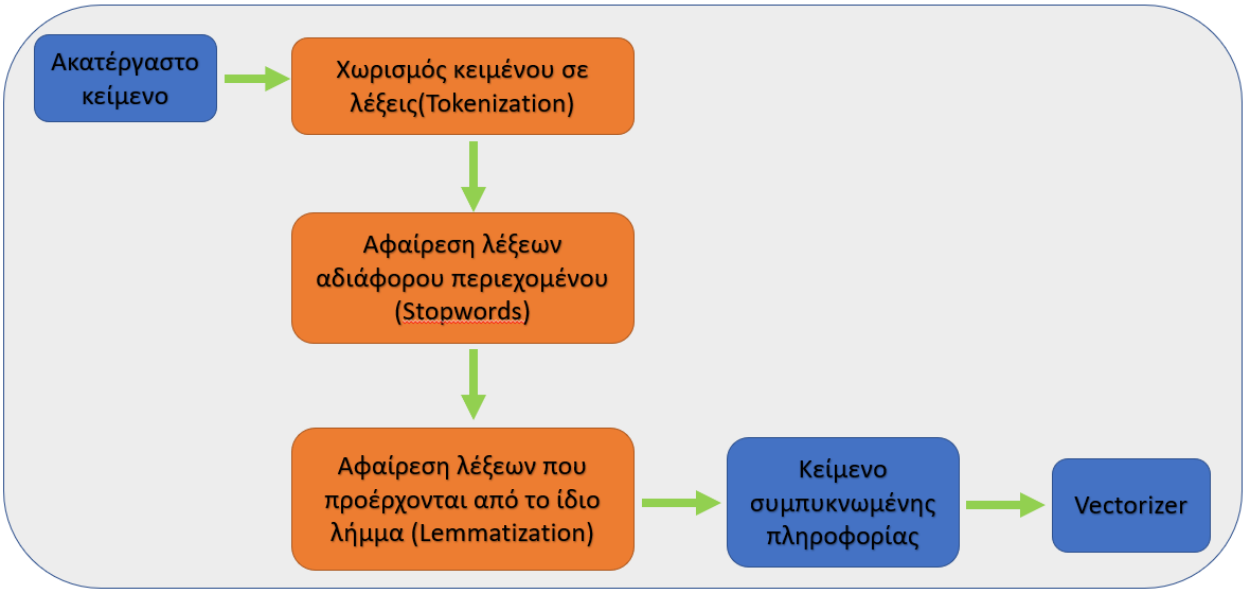
Εικόνα 10: Χρήση του μοντέλου

Για να μπορέσει να γίνει αποδοτική μετατροπή του κειμένου σε διανύσματα αριθμών, πρέπει τα δεδομένα κειμένου να περάσουν κάποια στάδια επεξεργασίας. Σκοπός της

επεξεργασίας αυτής είναι να μειωθεί ο όγκος των δεδομένων που δεν προσδίδουν κάποια επιπλέον πληροφορία στο σύστημα. Τα στάδια αυτά είναι τα παρακάτω.

- Tokenization (διαίρεση): Στο στάδιο αυτό, το κείμενο χωρίζεται σε λέξεις, και κάθε λέξη αναπαρίσταται με έναν αριθμό
- Stopwords (Λέξεις αδιάφορου περιεχομένου): Στο στάδιο αφαιρούνται από το κείμενο στοιχεία που δεν δίνουν πληροφορίες στο μοντέλο όπως είναι τα άρθρα ή τα σημεία στίξης.
- Lemmatization (Λημματοποίηση): Στο στάδιο αυτό, αφαιρούνται οι λέξεις που εμφανίζονται πολλαπλές φορές, αλλά προέρχονται από το ίδιο λήμμα.

Μετά από αυτή την επεξεργασία, το κείμενο περιέχει συμπυκνωμένη πληροφορία και είναι έτοιμο να επεξεργαστεί και να μετατραπεί σε διανύσματα. Παρακάτω παρουσιάζονται διαγραμματικά τα βήματα επεξεργασίας των κειμένων.



Εικόνα 11: Στάδια επεξεργασίας κειμένου

3.3 Περιγραφή Vectorizers και Classifiers

Όπως έχει ήδη αναφερθεί, οι πιο γνωστές τεχνικές μετατροπής κειμένου σε διανύσματα (Vectorizers) είναι οι bag of words και η tf-idf.

Bag Of Words

Η τεχνική bag of words είναι η πιο απλή τεχνική αναπαράστασης λέξεων με αριθμούς. Ουσιαστικά δημιουργείται ένα σύνολο λέξεων από τις λέξεις που χρησιμοποιούνται στα

δεδομένα, εφόσον επεξεργαστούν. Το διάνυσμα αναπαράστασης κάθε πρότασης, αποτελείται από το διάνυσμα της συχνότητας εμφάνισης κάθε λέξης στην συγκεκριμένη πρόταση. Η τεχνική αυτή δεν λαμβάνει υπόψιν την θέση της λέξης μέσα στην πρόταση, ούτε την συχνότητα εμφάνισης της κάθε λέξης σε σχέση με τις άλλες του συνόλου. Για την υλοποίηση της τεχνικής αυτής, υπάρχει έτοιμος αλγόριθμος στην γλώσσα προγραμματισμού python και εφαρμόζεται με την εντολή CountVectorizer. Για να γίνει πιο κατανοητή η μετατροπή των προτάσεων σε διανύσματα, ακολουθεί ένα σύντομο παράδειγμα.

Έστω ότι το κείμενο που πρέπει να μετατραπεί σε διανύσματα, είναι οι παρακάτω προτάσεις.

```
'coronavirus is a highly infectious disease',
'coronavirus affects older people the most',
'older people are at high risk due to this disease'
```

Χωρίς να εφαρμοστούν οι τεχνικές για να μετατραπεί το κείμενο σε κείμενο συμπυκνωμένης πληροφορίας, οι λέξεις που αποτελούν το λεξιλόγιο παραπάνω κειμένου είναι οι εξής:

```
['affects', 'are', 'at', 'coronavirus', 'disease', 'due', 'high',
'highly', 'infectious', 'is', 'most', 'older', 'people', 'risk',
'the', 'this', 'to']
```

Έτσι κάθε μία πρόταση από τις παραπάνω, θα μετατραπεί σε ένα διάνυσμα διαστάσεων [1X17], όσο δηλαδή είναι και η διάσταση του λεξιλογίου που χρησιμοποιείται. Με βάση τα παραπάνω, οι τρείς προτάσεις αναπαρίστανται διανυσματικά ως εξής:

```
[[0, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0],
[1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0],
[0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 0, 1, 1]]
```

Tf-Idf

Η τεχνική tf-idf (term frequency-inverse document frequency) στηρίζεται στην ίδια φιλοσοφία με την bag of words αλλά συσχετίζει την συχνότητα εμφάνισης μιας λέξης μέσα σε μια πρόταση με τον αριθμό των προτάσεων που περιέχουν την λέξη αυτή μέσα στον όγκο των δεδομένων. Αυτό δίνει ένα βάρος σε κάθε λέξη, ξεχωρίζοντας ποιες λέξεις είναι

πιο σημαντικές για το σύνολο των δεδομένων. Στην τεχνική αυτή, το διάνυσμα της πρότασης προκύπτει ως γινόμενο δύο παραγόντων TF και IDF. Ο όρος TF (Term Frequency), ορίζεται ως η συχνότητα εμφάνισης μιας λέξης μέσα στο κείμενο προς το συνολικό αριθμό των λέξεων του κειμένου και δίνεται από την παρακάτω σχέση.

$$TF = \frac{Frequency\ of\ word\ in\ a\ document}{Total\ number\ of\ words\ in\ that\ document}$$

Ο όρος αυτός είναι πάντα μικρότερος από 1 και εκφράζει πόσο συχνά εμφανίζεται μία λέξη στο σύνολο του κειμένου. Για την κατανόηση του όρου IDF (Inverse Document Frequency), πρέπει πρώτα να οριστεί ο όρος DF (Document Frequency) ο οποίο ορίζεται ως ο αριθμός των κειμένων που περιέχουν την συγκεκριμένη λέξη, προς το συνολικό αριθμό των κειμένων, και δίνεται από την παρακάτω σχέση.

$$DF = \frac{Documents\ containing\ word\ W}{Total\ number\ of\ documents}$$

Έτσι ο όρος DF, εκφράζει το ποσοστό των κειμένων που περιέχουν κάθε λέξη. Όμως, όσο πιο συχνά εμφανίζεται μια λέξη μέσα σε ένα κείμενο, τόσο λιγότερη είναι η σημασία της λέξης αυτής για το κείμενο. Επόμενος, χρησιμοποιείται ο αντίστροφος λόγος του DF, το IDF και συγκεκριμένα ο λογάριθμος του αντίστροφου λόγου, όπως προκύπτει από την παρακάτω σχέση.

$$IDF = \log\left(\frac{Total\ number\ of\ documents}{Documents\ Containing\ word\ W}\right)$$

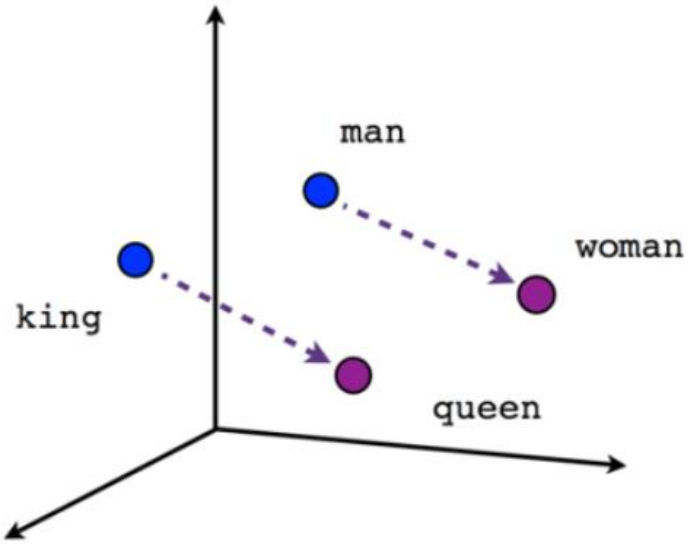
Το γινόμενο των δύο όρων TF και IDF για κάθε λέξη, αποτελεί την αριθμητική αναπαράσταση της λέξης αυτής στο κείμενο. Η τεχνική TF-IDF χρησιμοποιείται ακόμα και από μηχανές αναζήτησης ώστε να συσχετίσει την λέξη αναζήτησης με τα περιεχόμενα ιστοσελίδων. Για την υλοποίηση της τεχνικής αυτής, υπάρχει έτοιμος αλγόριθμος στην γλώσσα προγραμματισμού python και υπονοείται με την εντολή tfidfVectorizer. Εφαρμόζοντας της τεχνική TF-IDF για το κείμενο του προηγούμενο παραδείγματος, προκύπτουν τα παρακάτω διανύσματα για κάθε πρόταση.

```
[0, 0, 0, 0.37, 0.37, 0, 0, 0.49, 0.49, 0.49, 0, 0, 0, 0, 0, 0, 0],
[0.45, 0, 0, 0.34, 0, 0, 0, 0, 0, 0, 0.45, 0.34, 0.34, 0, 0.45, 0, 0],
[0, 0.33, 0.33, 0, 0.25, 0.33, 0.33, 0, 0, 0, 0, 0.25, 0.25, 0.33, 0,
0.33, 0.33]
```

Για την αντιστοίχιση των διανυσμάτων των λέξεων στην κατάλληλη κατηγορία, χρησιμοποιούνται μοντέλα ταξινόμησης (Classifiers) όπως το K-means,το Naïve Bayes και το Logistics Regression.

Word2Vec

Μια άλλη τεχνική για την αναπαράσταση των λέξεων σε διανύσματα είναι η τεχνική Word2Vec. Η λογική για την δημιουργία του διανύσματος κάθε λέξης βασίζεται στο πόσο συχνά μια λέξη εμφανίζεται μαζί με κάθε άλλη στον πυρήνα του λεξιλογίου. Με αυτή την τεχνική, τα διανύσματα εμπεριέχουν και σημασιολογική πληροφορία για τις λέξεις, ενώ λέξεις με παρόμοια σημασία έχουν παρόμοια διανύσματα. Έστω για παράδειγμα ότι τα παρακάτω γράμματα είναι λέξεις και σχηματίζουν τις παρακάτω προτάσεις (x y A z w), (x y B z k) και (x l C d m) τότε αναμένεται σημασιολογικά η λέξη A να είναι πιο κοντά στην λέξη B και όχι στην λέξη C. Αυτό μεταφράζεται μαθηματικά με διανύσματα όπου το διάνυσμα της λέξης A είναι κοντά με το διάνυσμα της λέξης B. Και βέβαια, με τα διανύσματα των λέξεων μπορούν να γίνουν και πράξεις, όπως για παράδειγμα στην πιο γνωστή φράση που συναντά κανείς κατά την μελέτη της τεχνικής Word2Vec “King-man+Woman=Queen”. Έτσι γίνεται κατανοητό ότι οι λέξεις King και Queen είναι παρόμοιες και οι διαφορά τους στο διάνυσμα αναπαράστασης οφείλεται στο διαφορετικό γένος. Για την δημιουργία του μοντέλου Word2Vec χρησιμοποιούνται πυρήνες κείμενων και επεξεργάζονται με νευρωνικά δίκτυα. Το διάνυσμα αναπαράστασης κάθε λέξης σε αυτό το μοντέλο, δεν είναι όσο το πλήθος των μοναδικών λέξεων που εμφανίζονται στον πυρήνα των δεδομένων, όπως στις δύο παραπάνω τεχνικές αλλά μπορεί να επιλεγεί και να διαμορφωθεί ανάλογα με τον πυρήνα του κειμένου ή τα επιθυμητά αποτελέσματα.



Εικόνα 12:Αναπαράσταση λέξεων σε διάνυσμα n-διαστάσεων

Naïve Bayes

Το μοντέλο Naïve Bayes είναι ένα απλό μοντέλο με πολύ καλά αποτελέσματα. Βασίζεται σε πιθανότητες, θεωρεί την κάθε λέξη της πρότασης ανεξάρτητη από την άλλη και υπολογίζει την πιθανότητα της συγκεκριμένης λέξης να ανήκει σε κάποια κατηγορία. Η πιθανότητα της κάθε πρότασης, υπολογίζεται ως το γινόμενο της πιθανότητας κάθε λέξης της πρότασης να ανήκει σε κάποια κατηγορία. Η παραδοχή αυτή είναι πολλή ισχυρή, αλλά έχει αποδειχθεί και πολλή χρήσιμη, εφόσον κάνει το μοντέλο ικανό να δουλεύει ακόμα και με λίγα δεδομένα. Σε περίπτωση που κάποια λέξη δεν εμφανίζεται σε κάποια κατηγορία, για να μην πρόκυψη μηδενική πιθανότητα, εφαρμόζεται μια τεχνική η οποία ονομάζεται κανονικοποίηση Laplace.

K-means

Το μοντέλο ταξινόμησης K-means, αντιμετωπίζει τις κατηγορίες στις οποίες πρέπει να κατηγοριοποιηθούν τα δεδομένα ως σημεία στο χώρο. Ο όρος Κ, ορίζει σε πόσες κατηγορίες θα κατηγοριοποιηθούν τα δεδομένα μας (δεδομένα εκμάθησης). Κάθε νέο δεδομένο, μετατρέπεται και αυτό σε ένα σημείο στο χώρο και κατατάσσεται στην κατηγορία από την οποία έχει την μικρότερη απόσταση.

Logistic Regression

Το μοντέλο ταξινόμησης Logistics Regression αποτελεί ένα μοντέλο ταξινόμησης μιας μεταβλητής με βάση την θεωρία πιθανοτήτων. Η σημαντική διαφορά μεταξύ της λογιστικής

παλινδρόμησης (logistic regression) και της γραμμικής παλινδρόμησης (linear regression) είναι ότι η πρώτη είναι κατηγορική και η δεύτερη ποσοτική. Η λογιστική παλινδρόμηση κατατάσσει μια μεταβλητή σε μια κατηγορία, με βάση κάποιες άλλες μεταβλητές-κριτήρια.

Για την εφαρμογή όλων των παραπάνω αλγορίθμων αυτών έχουν δημιουργηθεί έτοιμες βιβλιοθήκες σε γλώσσα Python, όπου μπορεί κανείς να εισάγει δεδομένα σε κατάλληλη μορφή και να πάρει τα αντίστοιχα αποτελέσματα.

3.4 Αξιολόγηση αλγορίθμου κατηγοριοποίησης.

Είναι γνωστό ότι οι επιβλεπόμενες (Supervised) τεχνικές έχουν καλύτερα αποτελέσματα από τις μη επιβλεπόμενες (Unsupervised). Για την αξιολόγηση όμως των αλγορίθμων ταξινόμησης έχουν αναπτυχθεί κάποιοι δείκτες βάση των οποίων μπορεί κανείς να κρίνει την αποτελεσματικότητα κάποιου μοντέλου [33]. Για την αξιολόγηση του μοντέλου χρησιμοποιείται ένα υποσύνολο των δεδομένων το οποία καλούνται δεδομένα δοκιμών (test data). Για την κατανόηση των δεικτών παρουσιάζεται ο παρακάτω πίνακας (πίνακας 2), ο οποίος παρουσιάζει τους πιθανούς συνδυασμούς των πραγματικών τιμών και των τιμών της πρόβλεψης.

ΠΙΝΑΚΑΣ 2: Συνδυασμοί πραγματικών τιμών και τιμών πρόβλεψης

ΠΙΝΑΚΑΣ ΠΙΘΑΝΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ			
		ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ	
		ΘΕΤΙΚΟ	ΑΡΝΗΤΙΚΟ
ΤΙΜΕΣ ΠΡΟΒΛΕΨΗΣ	ΘΕΤΙΚΟ	TP	FP
	ΑΡΝΗΤΙΚΟ	FN	TN

Όταν το αποτέλεσμα της πρόβλεψης είναι θετικό και η πραγματική τιμή είναι θετικό, το αποτέλεσμα χαρακτηρίζεται ως TP (True Positive).
Όταν το αποτέλεσμα της πρόβλεψης είναι θετικό και η πραγματική τιμή είναι αρνητικό, το αποτέλεσμα χαρακτηρίζεται ως FP (False Positive).
Όταν το αποτέλεσμα της πρόβλεψης είναι αρνητικό και η πραγματική τιμή είναι θετικό, το αποτέλεσμα χαρακτηρίζεται ως FN (False Negative).
Όταν το αποτέλεσμα της πρόβλεψης είναι αρνητικό και η πραγματική τιμή είναι αρνητικό, το αποτέλεσμα χαρακτηρίζεται ως TN (False Negative).
Οι δείκτες που χρησιμοποιούνται είναι οι παρακάτω.

- Accuracy: Ορίζεται ως το ποσοστό των κειμένων τα οποία ο αλγόριθμος κατέταξε στην σωστή κατηγορία.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

- Precision: Ορίζεται ως το ποσοστό των κειμένων τα οποία ο αλγόριθμος κατέταξε στην σωστή κατηγορία προς το σύνολο των κειμένων που κατέταξε στην συγκεκριμένη κατηγορία.

$$Precision = \frac{TP}{TP + FP}$$

- Recall: Ορίζεται ως το ποσοστό των συνολικών κειμένων τα οποία ο αλγόριθμος κατέταξε σε κάποια κατηγορία προς το σύνολο των κειμένων τα οποία θα έπρεπε να είχε κατατάξει στην συγκεκριμένη κατηγορία.

$$Recall = \frac{TP}{TP + FN}$$

- F1 score: Ορίζεται ως ο αρμονικός μέσος των δεικτών precision και recall.

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Μια συνήθης τακτική για την αξιοποίηση των δεδομένων είναι η χρήση του 75% αυτών για εκπαίδευση του μοντέλου και η χρήση του υπολοίπου 25% για την αξιολόγηση του μοντέλου.

3.5 Εφαρμογές αλγορίθμων κατηγοριοποίησης κειμένου σε SCADA

Η κατηγοριοποίηση κειμένου είναι μια τεχνική που εφαρμόζεται σε διάφορες δραστηριότητες των επιχειρήσεων σήμερα. Μια αυτοματοποιημένη διαδικασία κατηγοριοποίησης κειμένων μπορεί να μειώσει τον χρόνο που απαιτείται για την ταξινόμηση και σήμανση εγγράφων, την αξιολόγηση άρθρων ή την επεξεργασία των κριτικών των πελατών των επιχειρήσεων. Ωστόσο, η κατηγοριοποίησή κειμένου, μπορεί να ενταχθεί ως δευτερεύουσα λειτουργία σε συστήματα εποπτείας και να προσθέσει κάποιες επιπλέον λειτουργίες σε αυτά.

Όπως έχει ήδη αναφερθεί, μια εφαρμογή της κατηγοριοποίησης κειμένου θα μπορούσε να ενταχθεί σε μια λειτουργία online forensics σε συστήματα SCADA. Οι μηχανισμοί ανίχνευσης απειλών του συστήματος, μπορούν να συνδεθούν με βάσεις δεδομένων που περιγράφουν τύπους επιθέσεων (CAPEC/ MITRE ATT&CK). Με τον τρόπο αυτό, ο

χειριστής θα έχει μια άμεση περιγραφή για την επίθεση που δέχεται το σύστημά του ώστε να μπορεί να επέμβει άμεσα και αποτελεσματικά. Επιπλέον, με αυτό τον τρόπο οι βάσεις δεδομένων θα εμπλουτίζονται με νέα στοιχεία από επιθέσεις στα διασυνδεδεμένα SCADA, δίνοντας έτσι περισσότερες πληροφορίες στους μηχανικούς που ασχολούνται με τα συστήματα ασφαλείας αυτών.

Μια άλλη εφαρμογή θα μπορούσε να είναι η κατηγοριοποίηση των σφαλμάτων που εμφανίζονται στο σύστημα εποπτείας και ελέγχου της διεργασίας, ώστε τα σφάλματα που εμφανίζονται τα ταξινομούνται με βάση το τμήμα της διεργασίας που αυτά εμφανίζονται ή τον εξοπλισμό στον οποίο αναφέρονται. Αυτό θα έδινε αρκετά στοχευμένα στοιχεία στο τμήμα συντήρησης για την βελτίωση και αναβάθμιση του εξοπλισμού.

Τέλος, μια εξίσου σημαντική εφαρμογή μπορεί να είναι η ταξινόμηση κειμένου το οποίο εισάγεται στο SCADA από τον χειριστή και αφορά την παραγωγικότητα της διεργασίας, τους λόγους διακοπής λειτουργίας ή δυσκολίες που αντιμετωπίστηκαν. Χρησιμοποιώντας έναν αλγόριθμο ταξινόμησης κειμένου μπορούν αυτές οι πληροφορίες να εξάγουν συμπεράσματα για τον τρόπο βελτίωσης της παραγωγικότητας και της απόδοσης της παραγωγικής διαδικασίας.

Κεφάλαιο 4: Ανάπτυξη εφαρμογής κατηγοριοποίησης κειμένου

4.1 Περιγραφή μεθοδολογίας

Στα πλαίσια της παρούσας διπλωματικής αναπτύχθηκε ένας αλγόριθμος επιβλεπόμενης κατηγοριοποίησης κειμένου ο οποίος μπορεί να δεχθεί ως είσοδο μια περιγραφή μιας επίθεσης σε ένα σύστημα SCADA και να την κατηγοριοποιήσει σε μια από τις γνωστές κατηγορίες όπως παρουσιάσθηκαν στο αντίστοιχο κεφάλαιο. Τα δεδομένα προέκυψαν από αναζήτηση σε σχετική βιβλιογραφία και παρουσιάζονται στον πίνακα 3. Στην πρώτη στήλη του πίνακα αναγράφεται η κατηγορία επίθεσης και στην δεύτερη στήλη η περιγραφή της κατηγορίας όπως προκύπτει μέσα από την βιβλιογραφία. Στην τρίτη στήλη του πίνακα αναγράφεται μια διαφορετική περιγραφή της ίδιας επίθεσης όπως προκύπτει από αναζήτηση στο internet. Το κείμενο τις τρίτης στήλης χρησιμοποιείται όπως περιγράφεται στην συνέχεια για την αξιολόγηση του αλγορίθμου κατηγοριοποίησης που αναπτύχθηκε. Υλοποιήθηκαν τα βήματα της επεξεργασίας των δεδομένων, της μετατροπής τους σε διανύσματα αριθμών και χρησιμοποιήθηκαν διαφορετικά μοντέλα ταξινόμησης (Classifiers). Χρησιμοποιήθηκε η γλώσσα προγραμματισμού Python και το περιβάλλον Colab της google, για την ανάπτυξη του κώδικα. Χρησιμοποιήθηκαν επίσης έτοιμες βιβλιοθήκες της γλώσσας python, και εργαλεία για την αξιολόγηση του αποτελέσματος.

4.2 Περιγραφή εργαλείων- βιβλιοθήκες python

Για την ανάπτυξη του αλγορίθμου, χρησιμοποιήθηκε η γλώσσα προγραμματισμού python, και έτοιμα εργαλεία- βιβλιοθήκες που αυτή διαθέτει για διάφορες λειτουργίες. Στην συνέχεια περιγράφονται συνοπτικά αυτές οι βιβλιοθήκες.

Pandas: παρέχει λειτουργίες για τον χειρισμό αριθμητικών πινάκων και χρονοσειρών. Επιτρέπει την εύκολη επεξεργασία πινάκων, είτε αυτοί περιέχουν αριθμητικά δεδομένα, είτε λεκτικά.

Numpy: Παρέχει λειτουργίες μαθηματικών πράξεων μεταξύ μεγάλων πολυδιάστατων πινάκων [17].

Nltk (Natural Language Toolkit): Πρόκειται για μια σουίτα βιβλιοθηκών και προγραμμάτων για συμβολική και στατιστική επεξεργασία φυσικής γλώσσας (κείμενα γραμμένα σε αγγλική γλώσσα). Χρήσιμες εντολές της βιβλιοθήκης αυτής οι οποίες χρησιμοποιούνται και

στον κώδικα που αναπτύχθηκε είναι η stopwords και η WordNet Lemmatizer. Η εντολή stopwords, αφαιρεί λέξεις από τον πυρήνα των δεδομένων οι οποίες δεν περιέχουν κάποια σημαντική πληροφορία όπως για παράδειγμα τα άρθρα ή οι αντωνυμίες. Η εντολή WordNet Lemmatizer αφαιρεί λέξεις οι οποίες προέρχονται από το ίδιο λήμμα [21].

Re (Regular Expression): παρέχει εντολές οι οποίες μπορούν να αναγνωρίσουν σύμβολα και ειδικούς χαρακτήρες μέσα σε μια πρόταση [16].

Sklearn: Η βιβλιοθήκη αυτή αποτελεί σημαντικό κορμό για την ανάπτυξη λογισμικού εκμάθησης μηχανών (Machine Learning) με την χρήση της γλώσσας προγραμματισμού python. Εμπεριέχει αλγορίθμους μεθόδων λήψης αποφάσεων αλλά και αναγνώρισης και ταξινόμησης δεδομένων (Classifiers). Στην βιβλιοθήκη αυτή ανήκουν οι εντολές CountVectorizer και TfidfVectorizer οι οποίες υλοποιούν την μετατροπή των λέξεων σε διανύσματα όπως έχει αναλυθεί παραπάνω. Περιέχει επίσης τις εντολές MultinimialNB, LogisticRegression και KNeighborsClassifier οι οποίες υλοποιούν αλγόριθμους κατηγοριοποίησης (Classifiers). Τέλος, η βιβλιοθήκη αυτή περιέχει και τις εντολές αξιολόγησης του μοντέλου accuracy, recall, precision και f1, όπως αυτές έχουν παρουσιαστεί παραπάνω [18].

Στην συνέχεια παρουσιάζεται πως χρησιμοποιούνται οι συγκεκριμένες εντολές και βιβλιοθήκες για την ανάπτυξη της εφαρμογής της αυτόματης ταξινόμησης των περιγραφών των επιθέσεων.

4.3 Παρουσίαση κώδικα και αποτελεσμάτων

Για την ανάπτυξη του κώδικα, χρησιμοποιήθηκε η πλατφόρμα Colab της Google, η οποία είναι δωρεάν , online πλατφόρμα οπού μπορεί κανείς να γράφει πρόγραμμα σε γλώσσα python. Τα δεδομένα για την διαδικασία εκμάθησης και ελέγχου του αλγορίθμου, εισάγονται στο πρόγραμμα σε μορφή excel. Ακολουθεί η παρουσίαση των τμημάτων του προγράμματος και η επεξήγηση αυτών. Εφαρμόστηκαν οι τρείς αλγόριθμοι κατηγοριοποίησης που αναφέρθηκαν στο κεφάλαιο 3.3 (Naïve Bayes, K-means, Logistics Regression) και οι αλγόριθμοι μετατροπής λέξεων σε διανύσματα Bag of Words, TF-IDF και Word2Vec.

4.3.1 Εισαγωγή και επεξεργασία των δεδομένων

Αρχικά εισάγονται στην πλατφόρμα οι βιβλιοθήκες της γλώσσας python που παρουσιάστηκαν στο προηγούμενο κεφάλαιο, χρησιμοποιώντας την εντολή import.

```
import pandas as pd # data analysis and manipulation tool
import numpy as np # numerical data
import nltk # Natural Language toolkit , text processing libraries for classification, tokenization, stemming,
import re # regular expression
nltk.download('all')
```

Εικόνα 13: Εισαγωγή βιβλιοθηκών

Στην συνέχεια, εισάγονται τα δεδομένα από αρχείο excel, χρησιμοποιώντας την εντολή read_excel της βιβλιοθήκης pandas. Τα δεδομένα αποθηκεύονται στην μεταβλητή df.

```
df = pd.read_excel('AttacksV2.xlsx')
```

Εικόνα 14: Εισαγωγή δεδομένων

Στην συνέχεια χρησιμοποιείται η εντολή head. Η εντολή αυτή εμφανίζει τα δεδομένα που είναι αποθηκευμένα στην μεταβλητή df. Στην πρώτη στήλη είναι ο τύπος της επίθεσης, στην δεύτερη στήλη είναι η περιγραφή της κάθε επίθεσης σύμφωνα με την βιβλιογραφία και η τρίτη είναι μια δεύτερη περιγραφή της επίθεσης από κάποια πηγή του διαδικτύου. Η δεύτερη στήλη, θα χρησιμοποιηθεί για την αξιολόγηση του αλγορίθμου.

```
[4] df.head()
```

	Unnamed: 0	Unnamed: 1	Unnamed: 2
0	Attack	Description	Test Description
1	Spoofing attack	In this attack, a system entity illegitimately...	Spoofing is the act of disguising a communicat...
2	Jamming attack	Jamming is the interference with the Radio Fre...	Jamming attacks are a subset of denial of serv...
3	Replay attack	In a reply attack, a transmitted packet is mal...	A replay attack (also known as a repeat attack...
4	Wormhole attack	In this attack the adversary tunnels network m...	Wormhole attack is a grave attack in which two...

Εικόνα 15: Εμφάνιση δεδομένων

Στην συνέχεια γίνεται μετονομασία των στηλών του πίνακα δεδομένων ώστε να περιγράφουν τη χρήση κάθε στήλης και διόρθωση των δεικτών κάθε γραμμής.

```
df.columns = [ 'type', 'description','Test_description']
df = df.drop(0) # dropping the first row
new_indexes = [i for i in range(len(df))]
df.index = new_indexes
df.head()
```

	type	description	Test_description
0	Spoofing attack	In this attack, a system entity illegitimately...	Spoofing is the act of disguising a communicat...
1	Jamming attack	Jamming is the interference with the Radio Fre...	Jamming attacks are a subset of denial of serv...
2	Replay attack	In a reply attack, a transmitted packet is mal...	A replay attack (also known as a repeat attack...
3	Wormhole attack	In this attack the adversary tunnels network m...	Wormhole attack is a grave attack in which two...
4	Cryptanalysis	This attack refers to transforming encrypted d...	Cryptanalysis attack is a type of chosen plain...

Εικόνα 16: Μετονομασία στηλών-διόρθωση δεικτών

Στο παρακάτω τμήμα του κώδικα, δημιουργούνται δύο συναρτήσεις οι οποίες χρησιμοποιούν τις εντολές stopwords και WordNetLematizer της βιβλιοθήκης nltk. Οι συναρτήσεις αυτές δέχονται ως είσοδο κείμενο και το επιστρέφουν χωρίς «άχρηστες λέξεις» και λέξεις που προέρχονται από το ίδιο λήμμα.

```
# Create some functions to help us with processing our data
from nltk.corpus import stopwords
from nltk.stem import WordNetLemmatizer

def remove_stopwords(text):
  stop_words = set(stopwords.words('english')) # remove duplicates of stop words
  text_tokens = nltk.word_tokenize(text) # tokenize the text
  filtered_text = [word for word in text_tokens if word not in stop_words] # remove stop words
  text = " ".join(filtered_text) # rejoin text
  return text

def lemmatize(text):

  tokens = nltk.word_tokenize(text) # tokenize the text
  lemmatizer = WordNetLemmatizer() # create the lemmatizer object
  lemmatized_tokens = [lemmatizer.lemmatize(token) for token in tokens] # lemmatize

  # Join the lemmatized tokens back into a string
  lem_text = " ".join(lemmatized_tokens)

  return lem_text
```

Εικόνα 17: Δημιουργία συναρτήσεων

Στην συνέχεια γίνεται η επεξεργασία των κειμένων της δεύτερης στήλης, όπως έχει περιγραφεί στο αντίστοιχο κεφάλαιο ώστε να προκύψει το καθαρό κείμενο με το οποίο θα γίνει η εκπαίδευση του μοντέλου.

```
processed_descriptions = []  # Descriptions

# lowercasing the letters for every row of the description column
for desc in df['description']: # this is another way to get the column of your choice
  desc = desc.lower() # applying the .lower method for each description
  desc = re.sub(r'[^\w\s]', '', desc) # removig special characters, if there are any
  desc = re.sub(r'\d+', '', desc) # removing the numbers
  desc = remove_stopwords(desc) # renove stop words
  desc = lemmatize(desc) # lemmatize

  processed_descriptions.append(desc)

for i in range(len(df)):
  df.iloc[i, 1] = processed_descriptions[i]
```

Εικόνα 18: Επεξεργασία δεδομένων εκπαίδευσης

Η ίδια ακριβώς διαδικασία εφαρμόζεται και για τα δεδομένα της τρίτης στήλης, τα οποία θα χρησιμοποιηθούν για την αξιολόγηση του μοντέλου.

```
processed_descriptions = []    # Test Descriptions

# lowercasing the letters for every row of the description column
for desc in df['Test_description']: # this is another way to get the column of your choice
  desc = desc.lower() # applying the .lower method for each description
  desc = re.sub(r'[^\w\s]', '', desc) # removig special characters, if there are any
  desc = re.sub(r'\d+', '', desc) # removing the numbers
  desc = remove_stopwords(desc) # renove stop words
  desc = lemmatize(desc) # lemmatize

  processed_descriptions.append(desc)

for i in range(len(df)):
  df.iloc[i, 2] = processed_descriptions[i]
```

Εικόνα 19: Επεξεργασία δεδομένων αξιολόγησης

Στην συνέχεια παρουσιάζεται το καθαρό κείμενο όπως προέκυψε από την παραπάνω επεξεργασία.

```
df
```

	type	description	Test_description
0	Spoofing attack	attack system entity illegitimately assumes id...	spoofing act disguising communication identity...
1	Jamming attack	jamming interference radio frequency rf used n...	jamming attack subset denial service do attack...
2	Replay attack	reply attack transmitted packet maliciously fr...	replay attack also known repeat attack playbac...
3	Wormhole attack	attack adversary tunnel network message anothe...	wormhole attack grave attack two attacker loca...
4	Cryptanalysis	attack refers transforming encrypted data plai...	cryptanalysis attack type chosen plaintext att...

Εικόνα 20: Δεδομένα κειμένου μετά την επεξεργασία

Στο σημείο αυτό το κείμενο περιέχει συμπυκνωμένη πληροφορία, εφόσον έχουν αφαιρεθεί οι λέξεις που δεν δίνουν κάποια παραπάνω πληροφορία για το περιεχόμενο του κειμένου. Τα δεδομένα είναι έτοιμα για την εφαρμογή των αλγορίθμων μετατροπής σε διανύσματα και κατηγοριοποίησης.

4.3.2 Εφαρμογή του αλγορίθμου Naïve Bayes

Στο παρακάτω τμήμα του κώδικα, ορίζεται ότι τα δεδομένα της δεύτερης στήλης χρησιμοποιούνται για εκπαίδευση του μοντέλου (X_train) και τα δεδομένα της τρίτης για δοκιμή του μοντέλου (X_test). Ο κώδικας εκτελείται δύο φορές, μια χρησιμοποιώντας τον αλγόριθμο Bag of Words και μία χρησιμοποιώντας τον αλγόριθμο TF-IDF. Ως αλγόριθμος κατηγοριοποίησης χρησιμοποιείται και τις δύο φορές ο αλγόριθμος MultinomialNB (Naïve Bayes).

```
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import MultinomialNB
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.metrics import accuracy_score,f1_score, recall_score, precision_score

X = df['description'].to_list()

y = df['type'].to_list()

X_train, X_test, y_train, y_test = X, df['Test_description'].to_list(), y, y

vectorizer = TfidfVectorizer()
#vectorizer = CountVectorizer()

X_train_counts = vectorizer.fit_transform(X_train)

clf = MultinomialNB()
clf.fit(X_train_counts, y_train)
```

Εικόνα 21: Εφαρμογή μοντέλου Naive Bayes

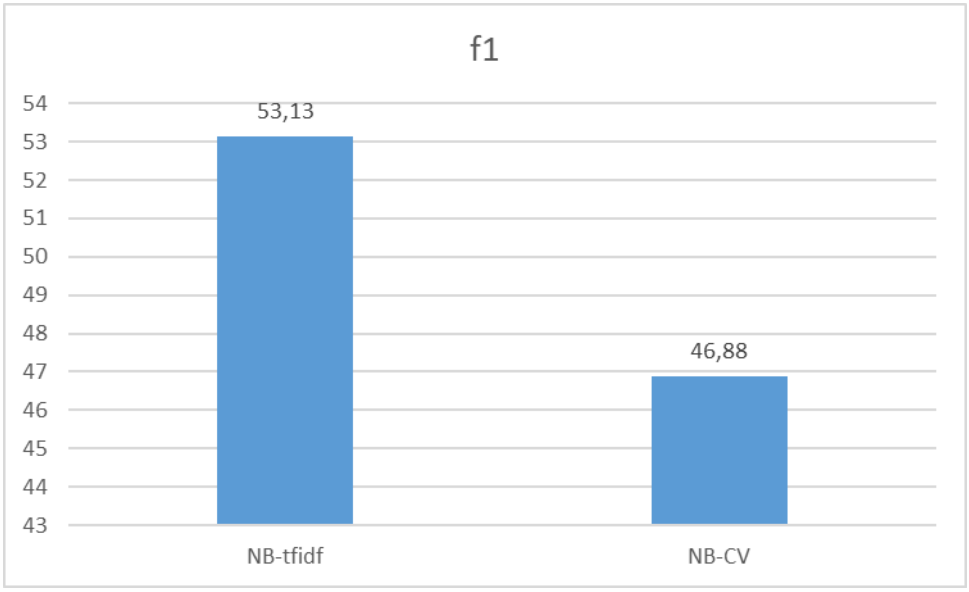
Στη συνέχεια υπολογίζονται οι δείκτες αξιολόγησης του μοντέλου, χρησιμοποιώντας τις αντίστοιχες εντολές.

```
X_test_counts = vectorizer.transform(X_test)
y_pred = clf.predict(X_test_counts)

accuracy = accuracy_score(y_test, y_pred)
print('Accuracy: ', accuracy)

precision = precision_score(y_test, y_pred, average='weighted')
print('precision: ', precision)

recall = recall_score(y_test, y_pred, average='weighted')
print('recall: ', recall)

f1 = f1_score(y_test, y_pred, average='weighted')
print('f1: ', f1)
```

```
Accuracy:  0.625
precision:  0.4895833333333333
recall:  0.625
f1:  0.53125
```

Εικόνα 22:Αποτελέσματα δεικτών αξιολόγησης μοντέλου Naive Bayes

Και αυτό το τμήμα του κώδικα εκτελείται δύο φορές, και ανάλογα με το ποιος αλγόριθμος μετατροπής λέξεων σε διανύσματα (Vectorizer) είναι ενεργός στο προηγούμενο μπλόκ, εμφανίζει και τα αντίστοιχα αποξέσματα. Στο παρακάτω διάγραμμα παρουσιάζονται τα αποτελέσματα για τους δύο διαφορετικούς Vectorizers.



Διάγραμμα 1: Δείκτης f1 - Naive Bayes

Παρατηρούμε ότι ο δείκτης αξιολόγησης f1 είναι καλύτερος όταν εφαρμόζεται η τεχνική tf-idf, γεγονός αναμενόμενο με βάση την θεωρεία και όσα έχουν αναφερθεί στο κεφάλαιο 3.4.

Το ίδιο συμβαίνει και με τους άλλους δείκτες αξιολόγησης οι οποίοι παρουσιάζονται αναλυτικά στο επόμενο κεφάλαιο.

4.3.3 Εφαρμογή του αλγορίθμου Logistic Regression

Για τα ίδια δεδομένα, μετά από την επεξεργασία των κειμένων, ακολουθείται η ίδια διαδικασία, αλλάζοντας των αλγόριθμο κατηγοριοποίησης και εφαρμόζοντας τον αλγόριθμο λογιστικής παλινδρόμησης (Logistic Regression)

```
[16] from sklearn.model_selection import train_test_split
     from sklearn.linear_model import LogisticRegression
     from sklearn.feature_extraction.text import TfidfVectorizer
     from sklearn.feature_extraction.text import CountVectorizer
     from sklearn.metrics import accuracy_score,f1_score, recall_score, precision_score

     X = df['description'].to_list()

     y = df['type'].to_list()

     X_train, X_test, y_train, y_test = X, df['Test_description'].to_list(), y, y

     vectorizer = TfidfVectorizer()
     #vectorizer = CountVectorizer()

     X_train_counts = vectorizer.fit_transform(X_train)

     clf = LogisticRegression()
     clf.fit(X_train_counts, y_train)
```

Εικόνα 23: Εφαρμογή μοντέλου Logistic Regression

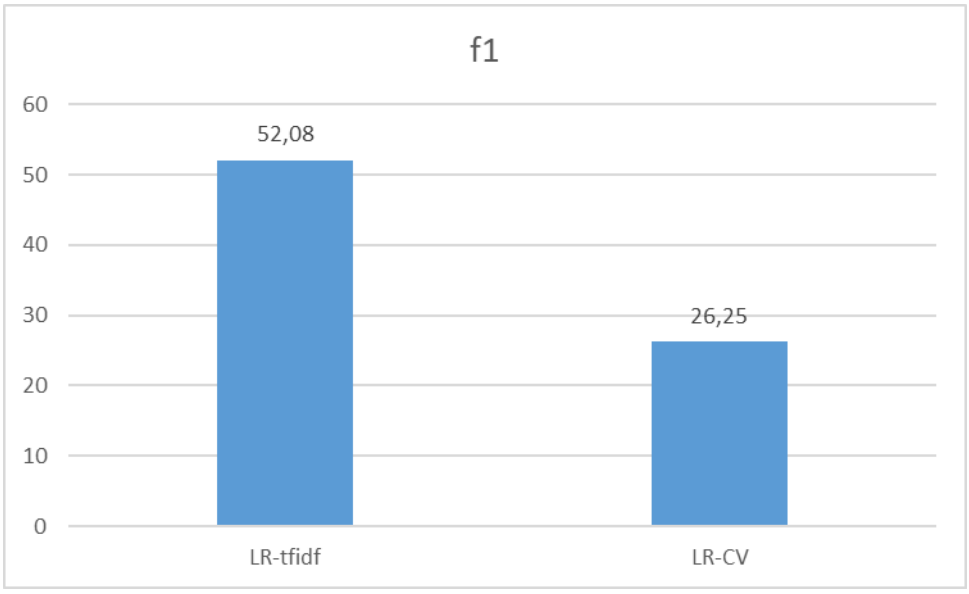
Ομοίως με την προηγούμενη περίπτωση ο αλγόριθμος αυτός εκτελείται δύο φορές, αλλάζοντας τον τρόπο μετατροπής των λέξεων σε διανύσματα και συγκρίνονται τα αποτελέσματα για τις δύο τεχνικές (tf-idf και BoW). Ο παρακάτω κώδικας δίνει τους δείκτες αξιολόγησης για τον συνδυασμό tf-idf και Logistic Regression.

```
X_test_counts = vectorizer.transform(X_test)
y_pred = clf.predict(X_test_counts)

accuracy = accuracy_score(y_test, y_pred)
print('Accuracy: ', accuracy)

precision = precision_score(y_test, y_pred, average='weighted')
print('precision: ', precision)

recall = recall_score(y_test, y_pred, average='weighted')
print('recall: ', recall)

f1 = f1_score(y_test, y_pred, average='weighted')
print('f1: ', f1)
```

```
Accuracy:  0.625
precision:  0.46875
recall:  0.625
f1:  0.5208333333333333
```

Εικόνα 24: Αποτελέσματα δεικτών αξιολόγησης μοντέλου Logistics Regression

Ο συνδυασμός των αλγορίθμων tf-idf και Logistic Regression, έχει ελαφρώς χειρότερα αποτελέσματα στους δείκτες precision και f1. Όλοι οι δείκτες για κάθε συνδυασμό μεθόδων παρουσιάζονται και συγκρίνονται στο επόμενο κεφάλαιο. Αξίζει να σημειωθεί ότι ο αλγόριθμος tf-idf έχει και πάλι καλύτερα αποτελέσματα στον δείκτη f1 σε σχέση με τον αλγόριθμο BoW. Τα αποτελέσματα για κάθε αλγόριθμο παρουσιάζονται στο παρακάτω διάγραμμα.



Διάγραμμα 2:Δείκτης f1 - Logistics Regression

4.3.4 Εφαρμογή του αλγορίθμου K-means

Και σε αυτή την περίπτωση, τα δεδομένα μένουν ίδια, ακολουθείται η ίδια διαδικασία και χρησιμοποιείται ο αλγόριθμος κατηγοριοποίησης K-means.

```
[19] from sklearn.model_selection import train_test_split
     from sklearn.neighbors import KNeighborsClassifier
     from sklearn.feature_extraction.text import TfidfVectorizer
     from sklearn.feature_extraction.text import CountVectorizer
     from sklearn.metrics import accuracy_score,f1_score, recall_score, precision_score

     X = df['description'].to_list()

     y = df['type'].to_list()


     X_train, X_test, y_train, y_test = X, df['Test_description'].to_list(), y, y

     vectorizer = TfidfVectorizer()
     #vectorizer = CountVectorizer()

     X_train_counts = vectorizer.fit_transform(X_train)

     clf = KNeighborsClassifier(n_neighbors=1)
     clf.fit(X_train_counts, y_train)
```

Εικόνα 25: Εφαρμογή μοντέλου K-means

Ο αλγόριθμος αυτός εκτελείται δύο φορές, αλλάζοντας τον τρόπο μετατροπής των λέξεων σε διανύσματα και συγκρίνονται τα αποτελέσματα για τις δύο τεχνικές. Ο παρακάτω κώδικας δίνει τους δείκτες αξιολόγησης για τον συνδυασμό tf-idf και K-means.

```
X_test_counts = vectorizer.transform(X_test)
y_pred = clf.predict(X_test_counts)

accuracy = accuracy_score(y_test, y_pred)
print('Accuracy: ', accuracy)

precision = precision_score(y_test, y_pred, average='weighted')
print('precision: ', precision)

recall = recall_score(y_test, y_pred, average='weighted')
print('recall: ', recall)

f1 = f1_score(y_test, y_pred, average='weighted')
print('f1: ', f1)

Accuracy:  0.625
precision:  0.4895833333333333
recall:  0.625
f1:  0.53125
```

Εικόνα 26: Αποτελέσματα δεικτών αξιολόγησης μοντέλου K-means

Παρατηρούμε ότι τα αποτελέσματα του αλγορίθμου αυτού είναι καλύτερα από του αλγορίθμου της λογιστικής παλινδρόμησης και ίδια με του αλγορίθμου (Naïve Bayes). Για τον αλγόριθμο κατηγοριοποίησης K-means χρησιμοποιήθηκε και η τεχνική μετατροπής λέξεων σε διανύσματα Word2Vec. Με τις παρακάτω εντολές, εισάγεται στον υπολογιστή

ένα ήδη εκπαιδευμένο μοντέλο Word2Vec το οποίο στην συνέχεια χρησιμοποιείται για να μετατρέψει τα δεδομένα που χρησιμοποιήθηκαν και στις άλλες περιπτώσεις σε διανύσματα.

```
import gensim.downloader
voc_vectors = gensim.downloader.load('word2vec-google-news-300')
```

```
[==================================================] 100.0% 1662.8/1662.8MB downloaded
```

Εικόνα 27: Εισαγωγή μοντέλου Word2Vec

Στην συνέχεια δημιουργείται μια συνάρτηση η οποία εφαρμόζει το μοντέλο Word2Vec για την μετατροπή των δεδομένων σε διανύσματα.

```
def desc_vecs(description_list):
  X = []

  for desc in description_list:
    tokens = desc.split()
    desc_vec = []
    for token in tokens:
      if token in list(voc_vectors.key_to_index.keys()):
        desc_vec.append(np.mean(voc_vectors[token]))

    X.append(desc_vec)

  lengths = []
  for desc in X:
    lengths.append(len(desc))

  desc_size = max(lengths)

  X_new = []
  for desc in X:

    desc = np.pad(desc, (0, desc_size - len(desc)),  mode='constant')
    X_new.append(desc)

  X_new = np.array(X_new)

  return X_new
```

Εικόνα 28: Συνάρτηση εφαρμογής μοντέλου Word2Vec

Και ακολούθως εφαρμόζεται αυτή η συνάρτηση για τα δεδομένα εκμάθησης και εκπαίδευσης της συγκεκριμένης εφαρμογής, όπως έχουν προκύψει μετά την επεξεργασία τους.

```
X_test = desc_vecs(df['Test_description'])
X = desc_vecs(df['description'])
```

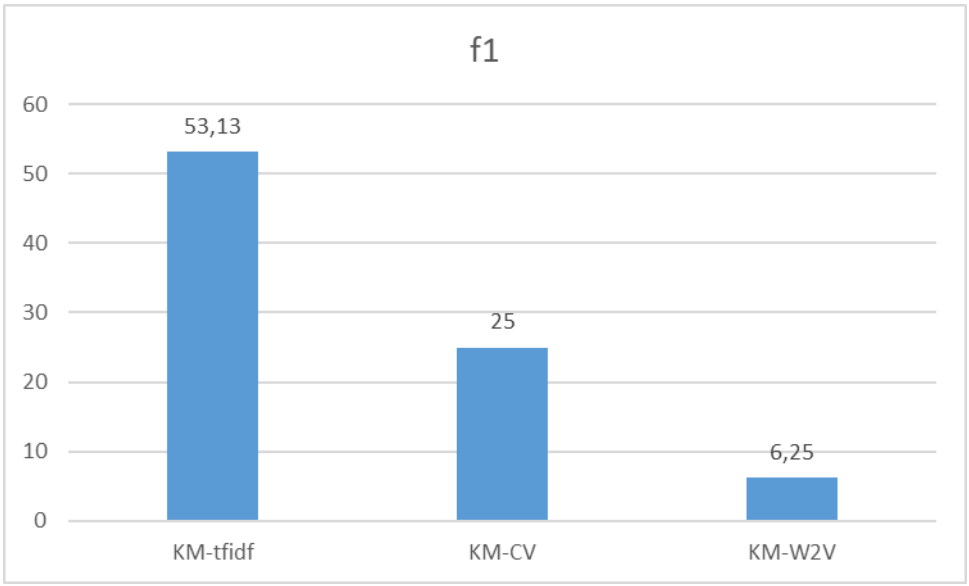
Εικόνα 29: Εφαρμογή συνάρτησης στα δεδομένα

Στο επόμενο βήμα εφαρμόζεται ο αλγόριθμος κατηγοριοποίησης K-means, με τα διανύσματα που έχουν προκύψει από την παραπάνω διαδικασία.

```
from sklearn.neighbors import KNeighborsClassifier

from sklearn.metrics import accuracy_score, f1_score, recall_score, precision_score


y = pd.factorize(df['type'])[0]

X_train, X_test, y_train, y_test = X, X_test, y, y


clf = KNeighborsClassifier(n_neighbors=1)
clf.fit(X_train, y_train)
```

Εικόνα 30: Εφαρμογή μοντέλου K-means με Word2Vec

Και το τελευταίο βήμα είναι η αξιολόγηση αυτού του συνδυασμού vectorizer και classifier μέσω των δεικτών accuracy, precision, recall και f1. Στο παρακάτω διάγραμμα παρουσιάζονται τα αποτελέσματα του δείκτη f1 για τις τρείς τεχνικές μετατροπής λέξεων σε διανύσματα που χρησιμοποιήθηκαν παραπάνω.



Διάγραμμα 3:Δείκτης f1 - K-means

Παρατηρούμε ότι η τεχνική tf-idf έχει και σε αυτή την περίπτωση καλύτερα αποτελέσματα από τις άλλες δύο τεχνικές. Επίσης παρατηρούμε ότι η τεχνική Word2Vec έχει πολύ χαμηλό αποτέλεσμα. Αυτό πιθανόν να οφείλεται στο γεγονός ότι ο πυρήνας των λέξεων που χρησιμοποιήθηκαν για την εκμάθηση του μοντέλου, ήταν ξένος με τον πυρήνα των δεδομένων ίσος δεν σχετίζεται με την ορολογία που χρησιμοποιείται στον πυρήνα των δεδομένων της παρούσας εφαρμογής.

4.4 Παρουσίαση – σχολιασμός αποτελεσμάτων.

Παρακάτω παρουσιάζονται συγκεντρωτικά με διαγράμματα τα αποτελέσματα των δεικτών αξιολόγησης για τον αλγόριθμο κατηγοριοποίησης κειμένου που αναπτύχθηκε. Τα δεδομένα εκπαίδευσης και δοκιμής του αλγορίθμου είναι ίδια, και αλλάζουν οι μέθοδοι μετατροπής των λέξεων σε διανύσματα (Vectrorizer) ή και ο αλγόριθμος κατηγοριοποίησης (Classifier). Συνοπτικά τα αποτελέσματα για διάφορους συνδυασμούς, παρουσιάζονται στον παρακάτω πίνακα.

	NB-tf-idf	NB-CV	LR-tf-idf	LR-CV	KM-tf-idf	KM-CV	KM-W2V
Accuracy	62,5	56,25	62,5	37,5	62,5	37,5	12,5
Precision	48,96	42,71	46,88	23,44	48,96	21,04	4,38
Recall	62,5	56,25	62,5	37,5	62,5	37,5	12,5
F1	53,13	46,88	52,08	26,25	53,13	25	6,25

NB-tfidf: Classifier Naïve Bayes, Vectorizer tf-idf

NB-CV: Classifier Naïve Bayes, Vectorizer CountVectrorizer

LR-tfidf: Classifier Logistic regression, Vectorizer tf-idf

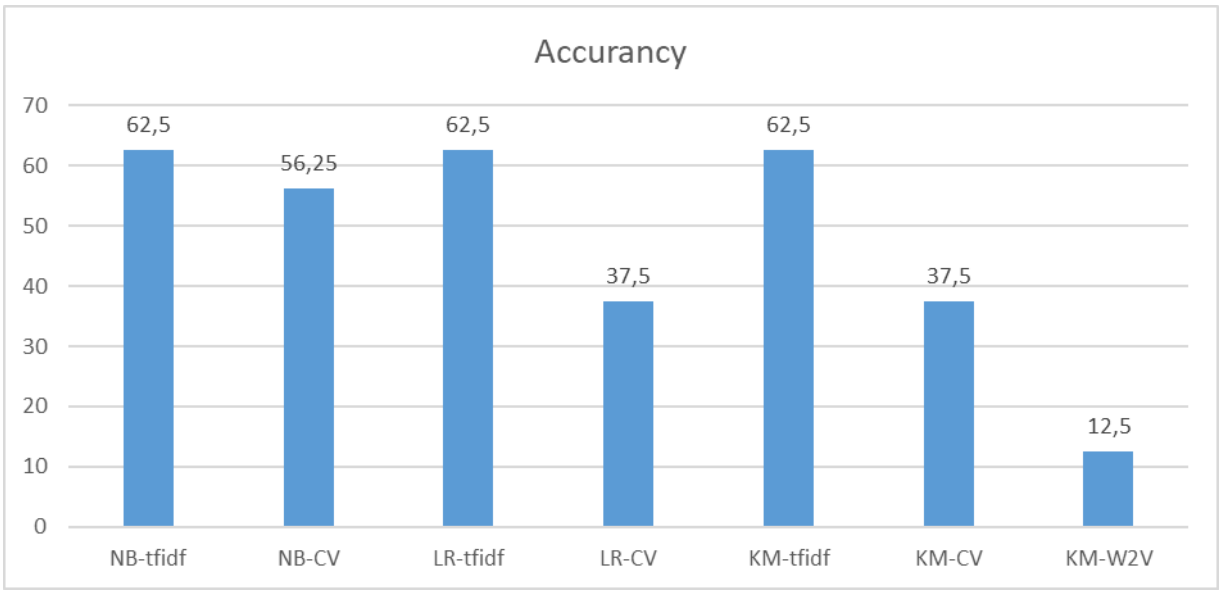
LR-CV: Classifier Logistic regression, Vectorizer CountVectrorizer

KM-tfidf: Classifier K-means, Vectorizer tf-idf

KM-CV: Classifier K-means, Vectorizer CountVectrorizer

KM-CV: Classifier K-means, Vectorizer Word2Vec

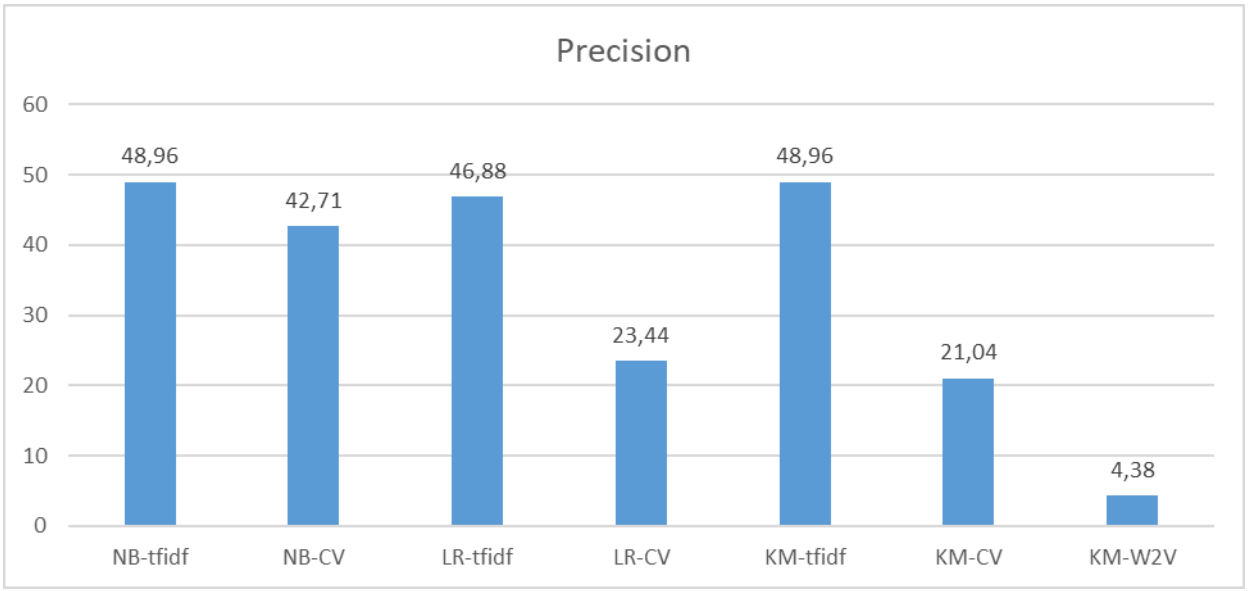
Στο παρακάτω διάγραμμα παρουσιάζονται τα αποτελέσματα για τον δείκτη accuracy για κάθε συνδυασμό classifier και vectrorizer.



Διάγραμμα 4: Δείκτης αξιολόγησης Accuracy

Παρατηρούμε ότι και στις 3 μεθόδους ο vectorizer tf-idf έχει καλύτερα αποτελέσματα από τον Vectorizer CountVectrorizer. Παρατηρούμε επίσης ότι και οι τρείς classifiers, Naïve Bayes, Logistics Regression και K-means, έχουν το ίδιο αποτέλεσμα (62,5%). Αξίζει επίσης να παρατηρήσουμε το χαμηλό αποτέλεσμα του δείκτη Accuracy που προκύπτει με την εφαρμογή του vectorizer Word2Vec.

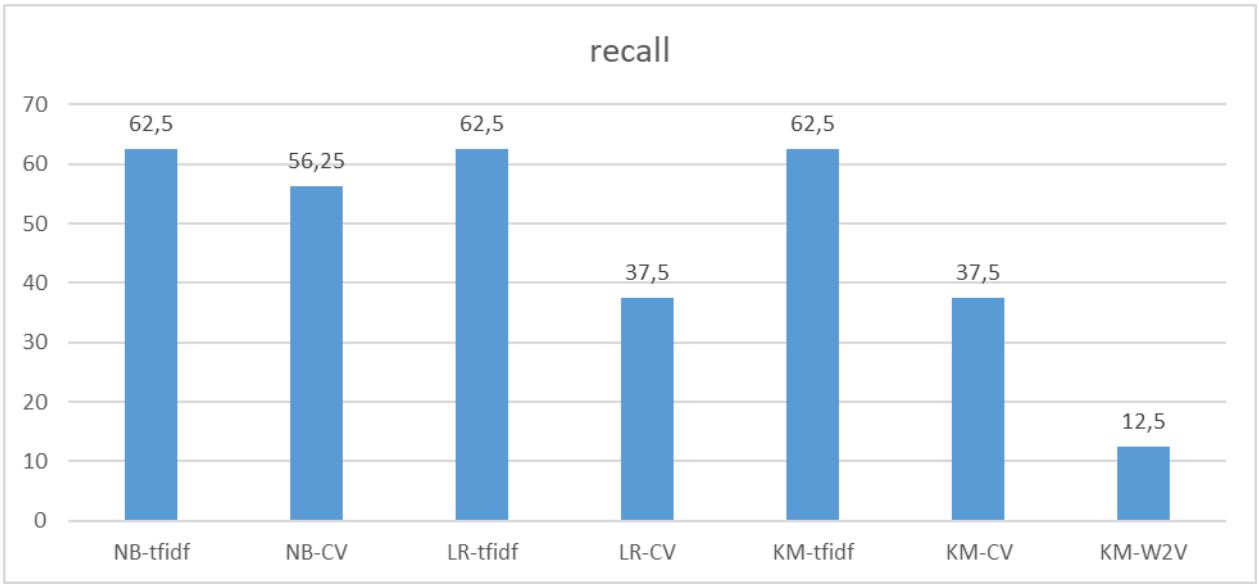
Στο παρακάτω διάγραμμα παρουσιάζονται τα αποτελέσματα του δείκτη precision για κάθε συνδυασμό classifier και vectrorizer.



Διάγραμμα 5: Δείκτης αξιολόγησης Precision

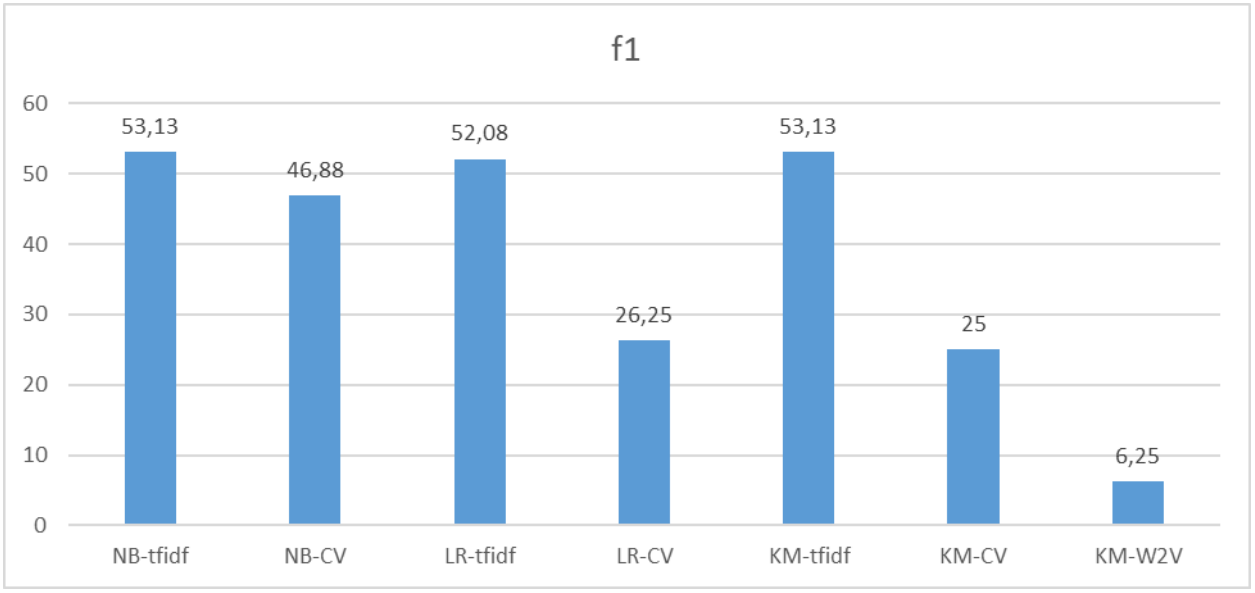
Ομοίως παρατηρούμε ότι και στις 3 μεθόδους ο vectorizer tf-idf έχει καλύτερα αποτελέσματα από τον Vectorizer CountVectrorizer. Σε αυτόν τον δείκτη ο classifier Naïve Bayes έχει καλύτερα αποτελέσματα (69,44%) συγκριτικά με τους άλλους δύο. Ακολουθεί ο classifier Logistics Regression (63,88%) ενώ το χαμηλότερο αποτέλεσμα προκύπτει με τον classifier K-means (61,11%). Και σε αυτόν το δείκτη, η εφαρμογή του vectorizer Word2Vec φέρνει χαμηλότερα αποτελέσματα (4,38%) σε σχέση με τους άλλους δύο.

Στο παρακάτω διάγραμμα παρουσιάζονται τα αποτελέσματα του δείκτη recall για κάθε συνδυασμό classifier και vectrorizer.



Διάγραμμα 6:Δείκτης αξιολόγησης Recall

Τα αποτελέσματα του δείκτη recall είναι ίδια με τα αποτελέσματα του δείκτη accuracy, γεγονός το οποίο πιθανόν να οφείλεται στον μικρό όγκο των δεδομένων.
Στην συνέχεια παρουσιάζονται τα αποτελέσματα του δείκτη f1



Διάγραμμα 7:Δείκτης αξιολόγησης f1

Ο δείκτης f1 θεωρείται πιο αξιόπιστος για την διεξαγωγή συμπερασμάτων, εφόσον εμπεριέχει τους δείκτες precision και recall. Παρατηρούμε ότι τα καλύτερα αποτελέσματα για τον αλγόριθμο κατηγοριοποίησης του κειμένου που εξετάζουμε προκύπτουν με την χρήση των αλγορίθμων Naïve Bayes ή K-means. Σε κάθε περίπτωση, είναι ξεκάθαρο ότι ο vectorizer ti-idf είναι καλύτερος από τον CountVectorizer (Bag of Words). Τα ποσοστά των αποτελεσμάτων των δεικτών αξιολόγησης είναι χαμηλά, γεγονός που πιθανόν να οφείλεται στις παρακάτω αιτίες.

1) Για κάθε κατηγορία υπάρχει μια μόνο περιγραφή που ορίζει την κατηγορία
2) Κάποιες κατηγορίες επιθέσεων μπορεί να έχουν παρόμοιες περιγραφές για τον ορισμό τους, ή ακόμα και να επικαλύπτονται οι ορισμοί τους.
3) Τα δεδομένα που χρησιμοποιήθηκαν για την αξιολόγηση του μοντέλου δεν περιέγραφαν σωστά την κατηγορία στην οποία ανήκαν.

Κεφάλαιο 5: Προτάσεις ανάπτυξης-Σενάρια χρήσης

5.1 Μέθοδοι βελτίωσης απόδοσης του συστήματος

Έχοντας παρουσιάσει τα βασικά στοιχεία που συνθέτουν έναν αλγόριθμο ταξινόμησης κειμένου, μπορούμε να δούμε σε ποια από αυτά μπορούμε να επέμβουμε ώστε να βελτιώσουμε την απόδοση του συστήματος.

Αρχικά, σημαντικό ρόλο παίζει ο όγκος των δεδομένων που θα χρησιμοποιηθούν για την εκμάθηση του αλγορίθμου. Όσο περισσότερα δεδομένα εισαχθούν κατά το στάδιο εκμάθησης τόσο καλύτερα θα είναι τα αποτελέσματα του αλγορίθμου σε νέα δεδομένα.

Στην συνέχεια, επίσης σημαντικό ρόλο παίζει ο τρόπος με τον οποίο θα μετατραπούν οι λέξεις σε διανύσματα (Vectorization). Στην παρούσα διπλωματική έχουν χρησιμοποιηθεί τρείς τέτοιες τεχνικές, η Bag of words, η tf-idf και η Word2Vec, με την τεχνική tf-idf να παρουσιάζει τα καλύτερα αποτελέσματα. Ωστόσο στον τομέα αυτό υπάρχει αρκετή έρευνα και προτάσεις για αλγορίθμους με καλύτερη απόδοση. Όπως έχει ήδη αναφερθεί μόνο η τεχνική Word2Vec μετατρέπει την λέξη σε διάνυσμα το οποίο εμπεριέχει πληροφορία για την σημασία της λέξης. Στην ίδια λογική της ενσωμάτωσης της σημασίας της λέξης στο διάνυσμα αναπαράστασης, έχουν αναπτυχθεί και άλλες τεχνικές όπως η GloVe [32] και η BERT [13]. Για την εκπαίδευση αυτών τον μοντέλων (pre-trained models) έχουν χρησιμοποιηθεί πυρήνες κειμένων όπως (Wikipedia, BookCorpus). Τέτοιες τεχνικές χρησιμοποιούνται και σε αλγορίθμους πρόβλεψης επόμενης λέξης (Next Word Prediction). Αν χρησιμοποιηθεί μια τέτοια τεχνική με πυρήνα εκπαίδευσης κείμενα με περιεχόμενο σχετικό με cyber attacks, πιθανότατα τα αποτελέσματα των δεικτών αξιολόγησης να αυξηθούν.

Επόμενος παράγοντας που επηρεάζει την αποτελεσματικότητα ενός αλγορίθμου κατηγοριοποίησης κειμένου, είναι ο αλγόριθμος ταξινόμησης (Classifier) που χρησιμοποιεί. Στην παρούσα διπλωματική έχουν χρησιμοποιηθεί οι αλγόριθμοι Naïve Bayes, K-means και Logistic Regression. Θα μπορούσε κανείς να δοκιμάσει και άλλους αλγορίθμους όπως τους Decision Tree, Stochastic Gradient Descent ή Random forest.

5.2. Πιθανές χρήσεις του συστήματος

Όλο και περισσότερα συστήματα ανίχνευσης απειλών (IDS) σε συστήματα SCADA κάνουν χρήση αλγορίθμων μηχανικής μάθησης για να μπορέσουν να ανιχνεύσουν ή ακόμα και να προβλέψουν μια επίθεση στο σύστημα. Όπως έχει ήδη αναφερθεί, μια εφαρμογή του αλγορίθμου ταξινόμησης που αναπτύχθηκε στην παρούσα διπλωματική μπορεί να είναι η σύνδεσή του με βάσεις δεδομένων επιθέσεων. Κάποιες από τις πιο γνωστές βάσεις δεδομένων για επιθέσεις σε συστήματα SCADA είναι οι CAPEC (Common Attack Pattern Enumeration and Classification) και MITRE ATT&CK (Adversarial Tactics Techniques and Common Knowledge). Στις βάσεις αυτές μπορεί να βρει κανείς πληροφορίες για επιθέσεις που έχουν γίνει κατά το παρελθόν, κατηγοριοποιημένες, ανά τύπο επίθεσης, η ανά τύπο εξοπλισμού που αυτές εμφανίστηκαν. Δίνουν μια σύντομη περιγραφή της διαδικασίας με την οποία έγινε η επίθεση, πληροφορίες για τις αδυναμίες του συστήματος που ο επιτιθέμενος εκμεταλλεύτηκε, με ποιόν άλλον τύπο επίθεσης μπορεί να σχετίζεται και για τον εξοπλισμό που χρησιμοποιήθηκε. Οι καταχωρήσεις της βάσης δεδομένων MITRE ATT&CK αφορούν κυρίως επιθέσεις στο δίκτυο των συστημάτων SCADA. Επομένως θα μπορούσε κανείς χρησιμοποιώντας την τεχνική της κατηγοριοποίησης κειμένου να αναπτύξει ένα online forensics σύστημα το οποίο θα συλλέγει πληροφορίες από την κατάσταση του συστήματος και θα ενημερώνει τον χρήστη για τυχόν επιθέσεις και τον τύπο αυτών.

5.3. Οδηγίες ανάπτυξης συστημάτων ασφαλείας σε SCADA

Οι βασικοί κανόνες προστασίας συστημάτων πληροφορικής πρέπει να προσαρμοστούν στις ιδιαιτερότητες ενός συστήματος βιομηχανικού ελέγχου και εποπτείας ώστε να είναι αποδοτικές στην προστασία αυτού. Από την εμπειρία και την εφαρμογή διάφορων συστημάτων ασφαλείας σε βιομηχανικούς αυτοματισμούς, έχουν αναπτυχθεί διάφορα πρότυπα και οδηγίες για αυτό το θέμα. Οι οδηγίες αυτές κατανοώντας τα χαρακτηριστικά και τις απαιτήσεις ενός συστήματος ελέγχου και εποπτείας, παρουσιάζουν κάποια μέτρα για την προστασία του συστήματος. Κάποιες από αυτές τις οδηγίες είναι οι παρακάτω [28].

- NIST Special Publication 800-82, *Guide to Industrial Control Systems (ICS) Security*

- ANSI/ISA-TR99.00.01-2007, *Security Technologies for Industrial Automation and Control Systems*
- North American Electric Reliability Corporation (NERC), *Critical Infrastructure Protection (CIP) Cybersecurity Standards*

- NIST Special Publication 800-53, *Recommended Security Controls for Federal Information Systems*
- Department of Homeland Security, *Catalog of Control Systems Security: Recommendations for Standards Developers*

- AMI-SEC Task Force, *AMI System Security Requirements*
- DOD Instruction 8500.2*, Information Assurance (IA) Implementation*

Παρακάτω, περιγράφονται οι δύο πρώτες οδηγίες, NIST Special Publication 800-82 και ANSI/ISA-TR99.00.01-2007.

5.3.1. Οδηγίες NIST Special Publication 800-82, Guide to Industrial Control Systems Security

Η οδηγία αυτή, ορίζει τα κύρια χαρακτηριστικά ενός συστήματος ασφαλείας για βιομηχανικά συστήματα αυτοματισμού ως ακολούθως [28].

Περιορισμός πρόσβασης στο δίκτυο του συστήματος αυτοματισμού. Αυτό επιτυγχάνεται με την χρήση ζώνης ελέγχου (demilitarized zone - DMZ) και firewalls τα οποία ελέγχουν την πρόσβαση από το δίκτυο του οργανισμού στο δίκτυο του συστήματος ελέγχου (SCADA). Επίσης μπορούν να χρησιμοποιηθούν διαφορετικοί τρόποι πιστοποίησης για τους χρήστες του δικτύου του οργανισμού και για τους χρήστες του δικτύου των συστημάτων αυτοματισμού. Επίσης, το δίκτυο του συστήματος αυτοματισμού θα πρέπει να διαχωρίζεται σε περισσότερα επίπεδα (layers) ώστε οι πιο κρίσιμες πληροφορίες να μεταφέρονται μέσω του πιο ασφαλούς και αξιόπιστου επιπέδου μεταφοράς.

Περιορισμός φυσικής πρόσβασης στο δίκτυο και τις συσκευές του συστήματος ελέγχου και εποπτείας. Δεν θα πρέπει να επιτρέπεται η πρόσβαση στις συσκευές ενός συστήματος αυτοματισμού σε μη εξουσιοδοτημένα πρόσωπα. Μπορούν να χρησιμοποιηθούν μέσα

προστασίας φυσικής πρόσβασης όπως κλειδαριές, κάρτες ελέγχου πρόσβασης ή και φρουροί.

Προστασία κάθε απομονωμένου τμήματος του συστήματος αυτοματισμού από επίθεση. Αυτό σημαίνει ότι κάθε συσκευή του αυτοματισμού, μετά τον έλεγχο και την θέση σε λειτουργία, θα πρέπει να προστατεύεται με απενεργοποίηση των θυρών επικοινωνίας για την ρύθμισή της, με χρήση κωδικών που δεν επιτρέπουν την πρόσβαση σε μη εξουσιοδοτημένα πρόσωπα ή ακόμα και με λογισμικό το οποίο μπορεί να ανιχνεύσει κακόβουλο λογισμικό (malware).

Διατήρηση λειτουργίας κάτω από δυσμενείς συνθήκες. Αυτό σημαίνει ότι το σύστημα ελέγχου και εποπτείας σχεδιάζεται έτσι ώστε για κάθε κρίσιμη συσκευή, υπάρχει εφεδρική έτοιμη να λειτουργήσει σε περίπτωση αστοχίας της πρώτης. Επιπλέον σε περίπτωση που κάποια συσκευή αστοχήσει δεν θα πρέπει να επιβαρύνει το δίκτυο επικοινωνίας ώστε να προκαλέσει προβλήματα στην επικοινωνία των υπόλοιπων συσκευών.

Επαναφορά του συστήματος μετά από επίθεση. Η ύπαρξη ενός πλάνου αντίδρασης σε περίπτωση επίθεσης είναι σημαντική , ανάλογα με την κρισιμότητα της διεργασίας που το σύστημα ελέγχει. Ένα σημαντικό χαρακτηριστικό ενός συστήματος ασφαλείας των βιομηχανικών αυτοματισμών είναι πόσο γρήγορα μπορεί αυτό να ξαναγίνει λειτουργικό μετά από κάποια επίθεση.

Η οδηγία NIST SP 800-82 κατηγοριοποιεί τους προτεινόμενους ελέγχους ασφαλείας σε τρείς κατηγορίες.

1) Έλεγχοι διοίκησης

Από μέρος της διοίκησης ενός οργανισμού που χρησιμοποιεί κάποιο σύστημα βιομηχανικού ελέγχου και εποπτείας, οι δράσεις που θα πρέπει να γίνουν είναι η μελέτη του ρίσκου, ο σχεδιασμός του συστήματος, η προμήθεια του συστήματος καθώς και η πιστοποίηση αυτού.

Η μελέτη του ρίσκου (risk assessment) αφορά τον καθαρισμό της κρισιμότητας του εξοπλισμού, με βάση την διεργασία που αυτός ελέγχει και την επίδραση που θα έχει μια αστοχία αυτού σε οικονομικά και κοινωνικό επίπεδο. Κατά τον σχεδιασμό του συστήματος, θα πρέπει να καθοριστούν τα επίπεδα ασφαλείας που αυτό θα καλύπτει, ενώ κατά την διάρκεια του χρόνου ζωής του συστήματος αυτού, θα πρέπει να γίνονται συνεχείς έλεγχοι και βελτιώσεις. Σχετικά με την προμήθεια του συστήματος, η διοίκηση έχει την ευθύνη για την προμήθεια του συστήματος, την λειτουργία, την συντήρηση- αναβάθμιση μέχρι και την

απόσυρση του συστήματος. Τέλος, η διοίκηση θα πρέπει να μεριμνήσει για την πιστοποίηση του συστήματος τόσο για την σωστή λειτουργία του όσο και για την ασφάλειά του έναντι επιθέσεων. Τα συστήματα ασφαλείας θα πρέπει να ελέγχονται και να αναβαθμίζονται τακτικά με ευθύνη της διοίκησης.

2) Έλεγχοι χρηστών

Στο πλαίσιο αυτό συμπεριλαμβάνονται όλα τα χαρακτηριστικά τα οποία πρέπει να έχει ένα σύστημα ασφαλείας ενός βιομηχανικού συστήματος ώστε να μειωθούν οι πιθανότητες πρόκλησης βλάβης από τους χρήστες. Αρχικά αναφέρεται η κατηγοριοποίηση των χρηστών σε επίπεδα πρόσβασης, ανάλογα με τις αρμοδιότητες τους. Στην συνέχεια προστασία του συστήματος από αλλαγές σε λογισμικό (software) ή υλικό (hardware) από μη εξουσιοδοτημένο προσωπικό. Επίσης η προστασία των συστημάτων και των αρχείων από λανθασμένη χρήση. Τέλος, το πλαίσιο αυτό περιλαμβάνει την ενημέρωση και την εκπαίδευση των χρηστών πάνω σε απειλές που μπορεί να δεχθεί το σύστημα καθώς και σε σχέδια δράσης σε περίπτωση που εμφανιστεί κάποια απειλή.

3) Έλεγχοι τεχνικών

Οι έλεγχοι των τεχνικών αφορούν τα στοιχεία λογισμικού (software) και υλικού (hardware) τα οποία εισάγονται στο σύστημα. Οι τεχνικοί που επεμβαίνουν στο σύστημα για την συντήρηση ή την αναβάθμιση αυτού, θα πρέπει να έχουν εξουσιοδότηση η οποία θα πιστοποιείται μέσω κωδικών η ακόμα και βιομετρικών στοιχείων ώστε να τους επιτραπεί η είσοδος πεδία μεταβολών του λογισμικού ή του υλικού. Επίσης οι αλλαγές που εφαρμόζονται από τους τεχνικούς στο σύστημα θα πρέπει να συμφωνούν με την πολιτική της διοίκησης στα θέματα ασφαλείας. Τέλος, όπως έχει ήδη αναφερθεί τα κρίσιμα μέρη του συστήματος θα πρέπει να προστατεύονται και φυσικά, με την χρήση λουκέτων η ακόμα και φρουρών.

5.3.2. Οδηγίες ANSI/ISA-TR99.00.01-2007, Security Technologies for Industrial Automation and Control Systems

Πρόκειται για άλλη μια οδηγία η οποία μέτρα τα οποία μπορεί να υιοθετήσει κάποιος χρήστης ενός βιομηχανικού συστήματος ελέγχου και εποπτείας ώστε να περιορίσει τον

κίνδυνο που αντιμετωπίζει το σύστημα έναντι απειλών. Τα μέτρα αυτά χωρίζονται στις επόμενες πέντε κατηγορίες [28].

Πιστοποίηση και εξουσιοδότηση. Η οδηγία αυτή ορίζει ως πρώτο βήμα για την προστασία ενός συστήματος τον διαχωρισμό των χρηστών ανάλογα με τις αρμοδιότητες που έχουν στο σύστημα. Στην συνέχεια περιορισμό της πρόσβασης των χρηστών στις δυνατότητες του συστήματος με την εισαγωγή κωδικών και συστημάτων πιστοποίησης. Η πιστοποίηση μπορεί να γίνει με χρήση απλών κωδικών ή ακόμα και κωδικών 2 μερών (Challenge/Response Authentication) όπου ο χρήστης και το σύστημα πρέπει να ανταλλάξουν κωδικούς και από τις δύο πλευρές. Επίσης πιστοποίηση του χρήστη μπορεί ακόμα να γίνει και από την γεωγραφική του θέση, χρησιμοποιώντας δεδομένα από GPS ή από την διεύθυνση IP.

Έλεγχος και απόρριψη πρόσβασης. Το μέρος αυτό της οδηγίας αναφέρεται στην χρήση τεχνικών ελέγχου πρόσβασης, όπως είναι τα firewalls. Ένα σύστημα firewall μπορεί να ελέγξει την πρόσβαση σε κάποιο σύστημα, δεδομένα ή λογισμικό, ελέγχοντας την πόρτα και την διεύθυνση IP του χρήστη. Όσο σημαντικό είναι να υπάρχει ένα σύστημα firewall μεταξύ του internet και το εσωτερικού δικτύου του οργανισμού, άλλο τόσο σημαντικό είναι να υπάρχει σύστημα firewall και μεταξύ του δικτύου του οργανισμού και του δικτύου το συστήματος ελέγχου και εποπτείας.

Τεχνολογίες κρυπτογράφησης και έλεγχος δεδομένων. Για την μεταφορά δεδομένων και πληροφοριών τα οποία σχετίζονται με την λειτουργία της παραγωγικής διαδικασίας προτείνεται η χρήση κρυπτογράφησης. Επίσης τεχνολογίες ελέγχου δεδομένων, μπορούν να ελέγξουν τα δεδομένα ως προς την πληρότητα και την ακρίβειά τους. Κρυπτογράφηση επίσης μπορεί να χρησιμοποιηθεί και στα εικονικά ιδιωτικά δίκτυα (Virtual Private Network- VPN) τα οποία χρησιμοποιούνται στους οργανισμούς.

Έλεγχος, μέτρηση, επιτήρηση και ανίχνευση. Για την αξιολόγηση ενός συστήματος ασφαλείας απαιτούνται συνεχής έλεγχοι του συστήματος καθώς και διαδικασίες ανίχνευσης στοιχείων (φορένζικς) σε περίπτωση που εμφανιστεί κάποια απειλή στο σύστημα. Τέτοιοι έλεγχοι μπορεί να είναι για παράδειγμα η συχνότητα εισόδου σε έναν λογαριασμό χρήστη, η πρόσβαση του κάθε χρήστη σε φακέλους και αρχεία, οι μεταβολές σε συστήματα ασφαλείας ή η μεταβολή των δικαιωμάτων κάποιου χρήστη στο σύστημα.

Λογισμικό συστημάτων ελέγχου και επιτήρησης. Ιδιαίτερη προσοχή θα πρέπει να δίνεται στο λογισμικό που χρησιμοποιείται από τις συσκευές του αυτοματισμού, εφόσον πολλές

φορές αυτά χρησιμοποιούν λειτουργικά συστήματα όπως windows, Unix ή Linux καθιστώντας τα έτσι ευπαθή σε απειλές που έχουν σχεδιαστή για τα λογισμικά αυτά. Το ίδιο φυσικά ισχύει και για τα συστήματα το λογισμικό των οποίων έχει δημιουργηθεί από τον ίδιο τον κατασκευαστή.

Φυσική προστασία. Η οδηγία επισημαίνει την ανάγκη για φυσική προστασία των συσκευών, η οποία μπορεί να γίνει με απλά μέσα, όπως λουκέτα και πόρτες ασφαλείας, είτε με πιο σύνθετα ηλεκτρονικά μέσα όπως κάμερες, αισθητήρες και βιομετρικές συσκευές.

ΠΙΝΑΚΑΣ 3: Δεδομένα εκμάθησης και αξιολόγησης αλγορίθμου

Attack	Description	Test Description
Spoofing attack	In this attack, a system entity illegitimately assumes the identity of an authorized system entity	Spoofing is the act of disguising a communication or identity so that it appears to be associated with a trusted, authorized source
Jamming attack	Jamming is the interference with the Radio Frequency (RF) used by the nodes in a network. It makes use of the broadcast nature of the communication medium	Jamming attacks are a subset of denial of service (DoS) attacks in which malicious nodes block legitimate communication by causing intentional interference in networks
Replay attack	In a reply attack, a transmitted packet is maliciously or fraudulently repeated or delayed by the adversary	A replay attack (also known as a repeat attack or playback attack) is a form of network attack in which valid data transmission is maliciously or fraudulently repeated or delayed
Wormhole attack	In this attack the adversary tunnels network messages to another part of the network through a low latency link.	Wormhole attack is a grave attack in which two attackers locate themselves strategically in the network. Then the attackers keep on listening to the network, and record the wireless information.
Cryptanalysis	This attack refers to transforming encrypted data into plaintext without having prior knowledge of the encryption parameters or processes.	Cryptanalysis attack is a type of chosen plaintext attack on block ciphers that analyzes pairs of plaintexts rather than single plaintexts, so the analyst can determine how the targeted algorithm works when it encounters different types of data
Exploit	An exploit takes advantage of a software vulnerability to compromise a system.	An exploit is a segment of code or a program that maliciously takes advantage of vulnerabilities or security flaws in software or hardware to infiltrate and initiate a denial-of-service (DoS) attack or install malware, such as spyware, ransomware, Trojan horses, worms, or viruses
Sybil	Sybil attack refers to the scenario where a malicious node pretends to have multiple identities	A Sybil attack uses a single node to operate many active fake identities (or Sybil identities) simultaneously, within a peer-to-peer network
Replication	In this attack, the adversary attempts to add one or more nodes to the network that use the same ID as another node in the network	The dangerous type of attack in which an attacker can harm the functionality of the network by injecting the clone or replica in the network

Denial of service at the link layer	Causing collision with packets in transmission, exhaustion of the node's battery due to repeated retransmission unfairness in using the wireless channel among neighboring nodes	Is a cyber-attack in which the perpetrator seeks to make a machine or network resource unavailable to its intended users by temporarily or indefinitely disrupting services of a host connected to a network
Routing attacks	Create routing loops or advertise false routes. The final objective is to degrade the availability of the system, or to receive more traffic for cryptanalysis	A cyberattack directed at an Internet service provider that aims to reduce uptime or prevent users from accessing a web-enabled system like a blockchain
Time-Synchronization attack	Time-synchronization protocols provide a mechanism for synchronizing the local clocks of the nodes in a sensor network. As a result, when there is an attack on these protocols, a fraction of the nodes in the entire network will be out-of-sync with each other	Attack that incorporates the delay of time synchronization pulses being sent from one node to. another in a network (often a sensor network). The attack relies on abusing the time. synchronization protocol (pairwise sender-receiver synchronization), between nodes A and B.
buffer overflow	Format string, integer overflow, Some general methods are stack smashing and manipulating function pointer.	Is an anomaly whereby a program writes data to a buffer beyond the buffer's allocated memory, overwriting adjacent memory locations
SQL injection	Manipulate data input into an Web application, which fails properly sanitize user-supplied input, and to insert a series of unexpected SQL statements into a query	SQL injection is a code injection technique used to attack data-driven applications, in which malicious SQL statements are inserted into an entry field for execution
Idle Scan	Is to blind port scan by bouncing off a dumb "zombie" host, often a preparation for attack	The idle scan is a TCP port scan method that consists of sending spoofed packets to a computer to find out what services are available
Smurf	Is a type of address spoofing, in general, by sending a continuous stream of modified Internet Control message Protocol(ICMP) packets to the target network with the sending address is identical to one of the target computer addresses	A Smurf attack is a distributed denial-of-service attack in which large numbers of Internet Control Message Protocol (ICMP) packets with the intended victim's spoofed source IP are broadcast to a computer network using an IP broadcast address.
DNS forgery	Send a fake DNS reply with a matching source IP, destination port, request ID, but with an attacker manipulated information inside, so that this fake reply may be processed by the client before the real reply is received from the real DNS server.	Is a cyber attack that exploits vulnerabilities in the domain name system (DNS) by diverting Internet traffic away from legitimate servers and towards fake ones.

Βιβλιογραφικές αναφορές

1. Abramovich, F. (2021). *Multiclass classification by sparse multinomial logostic regression.*
2. Ahmed, O. (2012). *SCADA Systems:Challenges for Forensic Investigators.* New Orleans, USA.
3. Alanazi, M. C. (2022). *SCADA vulnerabilities and attacks: A review of the state-of-the-art and open issues.* Melbourne, Australia.
4. Awas, R. (2018). *Tools, Techniques, and Methodologies: A Survay of Digital Forensics for SCADA Systems.*
5. Beaver, J. (2014). *An evaluation of machine learning methods to detect mallicious SCADA communications.* Tennessee, USA.
6. Cambria, E. (2014). *Jumping NLP curves:A review of Natural Language Processing research.*
7. Cardenas, R. (2009). *Rethinking security properties, threat models, and the design space in sensor networks: A case study in SCADA systems.* Berkeley, United States.
8. Chen, T. (2010). *Stuxnet, the ral start of cyber warfare?* Dortmund.
9. Cinque, M. (2018). *Challenges and directions in security information and event management (SIEM).* Italy.
10. Daneels, A. (1999). *What is SCADA.* Italy.
11. Delooze, L. (2005). *Classification of computer attacks using a self-organizing map.* USA.
12. Eden, B. (2015). *A Forensic Taxonomy of SCADA Systems and Approach to Incident Response.* Aberystwyth, UK.
13. Gao, Z. (2019). *Target-Dependent sentmen classification with BERT.*
14. Ghosh, S. (2019). *A survay of security in SCADA networks: Current issues and future challenges.*
15. Hassani, H. (2021). *Vulnerability and security risk assessment in a IIOT environment in campliance with standard IEC62443.* Leuven Belgium.

16. https://docs.python.org/3/library/re.html.

17. https://numpy.org/.

18. https://scikit-learn.org/stable/.

19. https://www.datacamp.com/tutorial/text-classification-python.

20. https://www.iec.ch/blog/understanding-iec-62443.

21. https://www.nltk.org/.

22. Igure, V. (2006). *Security issues in scada networks.* USA.

23. Ikonamakis, M. (2005). *Text classification using machine learning techniques.* Greece.

24. Irmak, E. (2018). *An overview of cyber-attack vectors on SCADA systems.*

25. Jatnika, D. (2019). *Word2Vec model analysis for semantic sililarities in english words.* Indonesia.

26. Jen Yang, F. (2018). *An implementation of Naive Bayes classifier.* USA.

27. Kotenko, I. (2015). *The CAPEC based generator of attack scenarios for network security evaluation.* Russia.

28. Krutz, R. L. (2013). *Industrial Automation and Control Systems Security Principles.* United States.

29. Lee, J. (2020). *Keeping host sanity for security of the SCADA systems.*

30. Li, Y. (2012). *A clustering method based on K-means algorithm.* China.

31. Munro, K. (2008). *SCADA - A critical situation.*

32. Pennington, J. (2014). *GloVe:Global Vectors for word representation.* Stanford.

33. Qaiser, S. (2018). *Text Mining:Use of TF-IDF to examine the relevance of word to documents.* Malaysia.

34. Queiroz, C. (2009). *Building a SCADA security testbed.*

35. Ranathunga, D. (2016). *Case studies of SCADA firewall configuration and the implications for best practices.*

36. Sajid, A. *Cloud-Assisted IOT-Based SCADA systems security:A review of the state of the art and future challenges.*

37. Spenneberg, R. *PLC-Blaster: A worm living solely in the PLC.*

38. Stirland, J. (2014). *Developing Cyber Forensics for SCADA Industrial Control Systems.* Malaysia.

39. Taveras, P. (2013). *SCADA Live Forensics:Real time data acquisition process to detect, prevent or evaluate critical situations.* Azores, Portugal.

40. Touloumis, M.-P. G. (2022). *A tool for assisting in the forensic invstigation of cyber-security incident.* Athens.

41. Ujvarosi, A. (2016). *Evolution of SCADA systems.* Brasol, Romania.

42. Wu, D. (2013). *Towards a SCADA Forensics Architecture.* Coedkernew UK.

43. Zhu, J. (2011). *A Taxonomy of Cyber Attacks on SCADA Systems.* Berkeley,United States.