Πανεπιστήμιο Αιγαίου
Τμήμα Μηχανικών Πληροφοριακών & Επικοινωνιακών Συστημάτων

Επισκόπηση Μεθόδων Προστασίας από Επιθέσεις Κοινωνικής Μηχανικής

Διπλωματική Εργασία

Κοντογεωργόπουλος Κωνσταντίνος 321/2013080

Επιβλέπων Καθηγητής: Κρητικός Κυριάκος

Μέλη 3-μελούς Επιτροπής Εξέτασης: Κρητικός Κυριάκος, Κοκολάκης Σπυρίδων, Καρύδα Μαρία

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ

ΠΙΝΑΚΑΣ ΠΙΝΑΚΩΝ

Περίληψη

Τα τελευταία χρόνια με την ολοένα και περισσότερο έντονη χρήση κοινωνικών δικτύων και υπολογιστών την αυξημένη χρήση υπολογιστών καθώς και την εργασία εξ αποστάσεως ειδικά σε περιόδους καραντίνας λόγω του Covid-19, οι επιθέσεις κοινωνικής μηχανικής είναι ένα αυξανόμενο φαινόμενο. Οι επιθέσεις αυτές είναι οι πιο συχνές, μιας και ανεξάρτητα από το πόσο προφυλαγμένο είναι ένα πληροφοριακό σύστημα από επιθέσεις σε επίπεδο ασφάλειας συστήματος, ο πιο αδύναμος κρίκος είναι ο ανθρώπινος παράγοντας.

Στην συγκεκριμένη διπλωματική εργασία θα ασχοληθούμε με τις πιο συχνές επιθέσεις της κοινωνικής μηχανικής και θα τις ταξινομήσουμε με βάση διάφορα κριτήρια. Ακόμα θα παραθέσουμε τρόπους προστασίας και μετρίασης τέτοιων επιθέσεων καθώς και εργαλεία προστασίας από αυτές που χρησιμοποιούνται ευρέως από οργανισμούς.

1. Εισαγωγή

Η κοινωνική μηχανική είναι η χειραγώγηση ατόμων με σκοπό την απόσπαση πληροφοριών, κυρίως εμπιστευτικών και ευαίσθητων δεδομένων. Αυτές οι επιθέσεις είναι τόσο διαδεδομένες, γιατί ανεξάρτητα με το πόσο δυνατή είναι η ασφάλεια ενός πληροφοριακού συστήματος και η ισχύς των προγραμμάτων προστασίας που αυτό έχει, ο τρόπος με τον οποίο επιτυγχάνεται η διείσδυση στο σύστημα οφείλεται σε εξωτερικούς παράγοντες, όπως είναι οι άνθρωποι. Οι μέθοδοι που χρησιμοποιούνται δεν απαιτούν τον ίδιο χρόνο και προσπάθεια σε σχέση με άλλα είδη επιθέσεων οι οποίες εκμεταλλεύονται μια τρωτότητα ενός συστήματος, μιας και οι άνθρωποι κυριαρχούνται και λειτουργούν τις περισσότερες φορές με βάση τα συναισθήματα. Αυτό καθιστά αυτές τις επιθέσεις από τις πιο επικίνδυνες, αφού η τρωτότητα δεν αντιμετωπίζεται με μία ολοκληρωμένη και οριστική λύση ή κάποιο μηχανισμό ασφαλείας, αλλά απαιτεί την κατάλληλη εκπαίδευση των ανθρώπων που έχουν πρόσβαση και χειρίζονται τα εκάστοτε συστήματα.

Οι κύριες συνεισφορές αυτής της διπλωματικής εργασίας είναι οι εξής:

- Μία λεπτομερή εξήγηση για το τι είναι η κοινωνική μηχανική, οι πιο σύνηθες επιθέσεις και τα στάδια που ακολουθούνται προκειμένου αυτές οι επιθέσεις να είναι επιτυχής, καθώς και παραδείγματα με τα οποία η κοινωνική μηχανική είχε επιτυχία σαν τρόπος διείσδυσης στην πραγματική ζωή.

- Μια λεπτομερή ταξινόμηση επιθέσεων κοινωνικής μηχανικής με βάση 3 διαστάσεις: τον χειριστή (operator), τις μεθόδους εξαπάτησης (methods of deception) καθώς και την φύση της επικοινωνίας (nature of communication).

- Μια εναλλακτική αλλά εξίσου λεπτομερής ταξινόμηση και ανάλυση των επιθέσεων κοινωνικής μηχανικής με βάση τις φάσεις διενέργειας των επιθέσεων αυτών.

- Μια διεξοδική ταξινόμηση και ανάλυση των τρόπων, μεθόδων και εργαλείων προστασίας που χρησιμοποιούνται για την μετρίαση, πρόληψη ή αντιμετώπιση αυτού του είδους των επιθέσεων.

- Παροχή ερευνητικών κενών και κατευθύνσεων για την πιο αποτελεσματική και ολοκληρωμένη πρόληψη και αντιμετώπιση αυτών των επιθέσεων.

Η παρούσα αναφορά έχει την εξής δομή:

1. *Θεωρητικό Υπόβαθρο*: Στο παραπάνω κεφάλαιο παρέχεται ένα πλήρες Θεωρητικό υπόβαθρο για τις επιθέσεις κοινωνικής μηχανικής καλύπτοντας:

    1. Κατηγοριοποίηση Επιθέσεων Κοινωνικής Μηχανικής.

    2. Επιπλέον Κατηγοριοποίηση επιθέσεων Κοινωνικής Μηχανικής με βάση τα βήματα που ακολουθούνται.

3. Επιθέσεις Κοινωνικής Μηχανικής.

4. Εργαλεία μετρίασης Κοινωνικής Μηχανικής.

5. Πρόληψη από επιθέσεις Κοινωνικής Μηχανικής.

2. *Μεθοδολογία Επισκόπησης*: Στο κεφάλαιο αυτό εξετάζονται τα κύρια ερωτήματα, ο τρόπος μελέτης, επισκόπησης, στρατηγικής αναζήτησης και τα κριτήρια που εφαρμόστηκαν στην παρούσα διπλωματική εργασία.
3. *Ανάλυση*: Στο κεφάλαιο αναλύονται κάποιες μεθοδολογίες και τρόποι προστασίας από επιθέσεις κοινωνικής μηχανικής σε μεγαλύτερο βάθος. Επίσης ταξινομούνται τρόποι αντιμετώπισης τέτοιων επιθέσεων με βάση κάποιες βασικές διαστάσεις

1. Μεθοδολογίες και τρόποι προστασίας από επιθέσεις Κοινωνικής Μηχανικής.

2. Διαστάσεις Τρόπων Αντιμετώπισης Επιθέσεων Κοινωνικής Μηχανικής.

    1) Κριτήρια

    2) Αξιολόγηση

3. Κατηγοριοποίηση προσεγγίσεων με βάση την δεύτερη Κατηγοριοποίηση του θεωρητικού υπόβαθρου.

4. Ανάλυση και Σχολιασμός πινάκων

4. *Προκλήσεις και Ερευνητικά κενά*: Στο κεφάλαιο αυτό εξετάζονται κάποιες προκλήσεις και ερευνητικά κενά που υπάρχουν.

5. *Μελλοντικές Κατευθύνσεις*: Στο κεφάλαιο αυτό εξετάζονται οι μελλοντικές κατευθύνσεις που μπορούν να ακολουθήσουν οι ερευνητές και οι οργανισμοί.

6. *Συμπεράσματα*: Στο κεφάλαιο αυτό παρουσιάζονται τα συμπεράσματα που βγάλαμε από την διπλωματική εργασία.

2. Θεωρητικό Υπόβαθρο

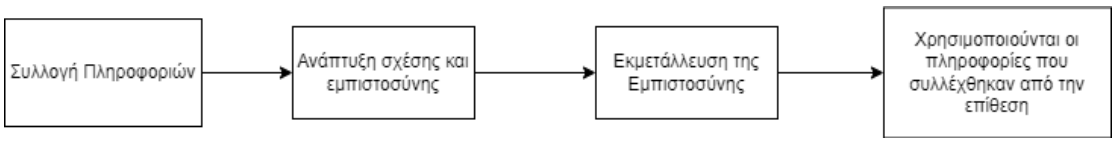
2.1 Κατηγοριοποίηση Επιθέσεων κοινωνικής μηχανικής

Κοινωνική Μηχανική είναι η πράξη της προφορικής χειραγώγησης ατόμων με σκοπό την απόσπαση πληροφοριών. Αν και είναι παρόμοια με το τέχνασμα ή την απλή απάτη, ο όρος είναι κυρίως συνδεδεμένος με την εξαπάτηση ατόμων με σκοπό την απόσπαση εμπιστευτικών πληροφοριών που είναι απαραίτητες για την πρόσβαση σε κάποιο υπολογιστικό σύστημα [168]. Η κοινωνική μηχανική χρησιμοποιώντας διάφορες τεχνικές, έχει σκοπό να χειραγωγήσει τους ανθρώπους προκειμένου να τους εξαπατήσει και να αποκτηθεί πρόσβαση σε ευαίσθητες και εμπιστευτικές πληροφορίες χωρίς τη χρήση τεχνικών μεθόδων και λογισμικό. Αυτό έχει ως αποτέλεσμα οποιοσδήποτε είναι αρκετά

πονηρός και ικανός να χειραγωγήσει τους ανθρώπους, να μπορέσει να επιτελέσει τέτοιου είδους επιθέσεις.

Οι επιθέσεις της κοινωνικής μηχανικής χωρίζονται (δείτε Εικόνα 1) σε 4 βασικά στάδια [169].

1. Η συλλογή πληροφοριών για το θύμα.
2. Η ανάπτυξη σχέσης και εμπιστοσύνης με το θύμα.
3. Η εκμετάλλευση της εμπιστοσύνης που αποκτάται με σκοπό την συλλογή πληροφοριών.
4. Η χρησιμοποίηση των πληροφοριών αυτών για να γίνει η επίθεση στο εκάστοτε σύστημα.



Εικόνα 1: 4 Βασικά Στάδια επιθέσεων κοινωνικής μηχανικής

Στην παρακάτω παράγραφο παρέχεται ένα παράδειγμα μιας απλής επίθεσης.

Ο επιτιθέμενος συλλέγει αρκετές πληροφορίες για ένα διευθυντικό στέλεχος της εταιρείας προκειμένου να αποκτήσει πρόσβαση στη λίστα εργαζομένων και τη μισθοδοσία τους. Δημιουργεί έναν ψεύτικο λογαριασμό ηλεκτρονικού ταχυδρομείου του διευθυντικού στελέχους και ζητά από κάποιον εργαζόμενο που διαθέτει την πληροφορία να του την στείλει με πρόφαση ότι χρειάζεται τα στοιχεία αυτά για μία εταιρική διαδικασία. Ο επιτιθέμενος χρησιμοποιεί αληθινές πληροφορίες του διευθυντικού στελέχους που έχει συλλέξει με αποτέλεσμα το θύμα να πειστεί ότι όντως μιλάει με το διευθυντικό στέλεχος που ο επιτιθέμενος υποδύεται και να του στείλει τις πληροφορίες.

Ας ταξινομήσουμε τις επιθέσεις κοινωνικής μηχανικής με βάση τρεις κύριες διαστάσεις [1], [169]: Χειριστής (Operator), Μέθοδοι εξαπάτησης (Methods of deception) και Φύση της επικοινωνίας (Nature of communication).

1. **Χειριστής (Operator):** Ο επιτιθέμενος έχει σαν στόχο να χειριστεί το θύμα με κάποιο τρόπο, με σκοπό να αποσπάσει τις πληροφορίες που θέλει. Η διάσταση του Χειριστή περιλαμβάνει 2 κατηγορίες επιθέσεων.

   - Επιθέσεις Κοινωνικής μηχανικής με βάση τον Άνθρωπο (Human based social engineering attacks): Ο επιτιθέμενος, εκμεταλλευόμενος τις κοινωνικές του διεξότητες ,αλληλεπιδρά με το θύμα προσωπικά χωρίς την χρήση κάποια συσκευής ή λογισμικού. [2]

   - Επιθέσεις Κοινωνικής μηχανικής με βάση τον Υπολογιστή (Computer based social engineering attacks): Ο επιτιθέμενος αλληλεπιδρά με το θύμα, μέσω μιας συσκευής, όπως ένα υπολογιστής ή ένα κινητό τηλέφωνο. [3]

2. **Μέθοδοι Εξαπάτησης (Methods of Deception):** Ο επιτιθέμενος χρησιμοποιεί μεθόδους εξαπάτησης, προκειμένου να αποσπάσει τις πληροφορίες από το θύμα. Η διάσταση των μεθόδων εξαπάτησης περιλαμβάνει τις εξής 3 κατηγορίες.

- Επιθέσεις με βάση την Ψυχολογία (Social based attacks): Ο επιτιθέμενος χρησιμοποιεί κοινωνικές και ψυχολογικές μεθόδους ώστε να εξαπατήσει το θύμα. Ένα παράδειγμα μεθόδων αυτού του είδους είναι ο Ψυχολογικός εκβιασμός: ο επιτιθέμενος παριστάνει κάποιο κοινωνικά-εταιρικά ισχυρό άτομο που ζητά άμεσα πληροφορίες από το θύμα διότι κάτι σοβαρό έχει συμβεί. Συχνότερα τέτοιες επιθέσεις γίνονται μέσω τηλεφώνου. [4]

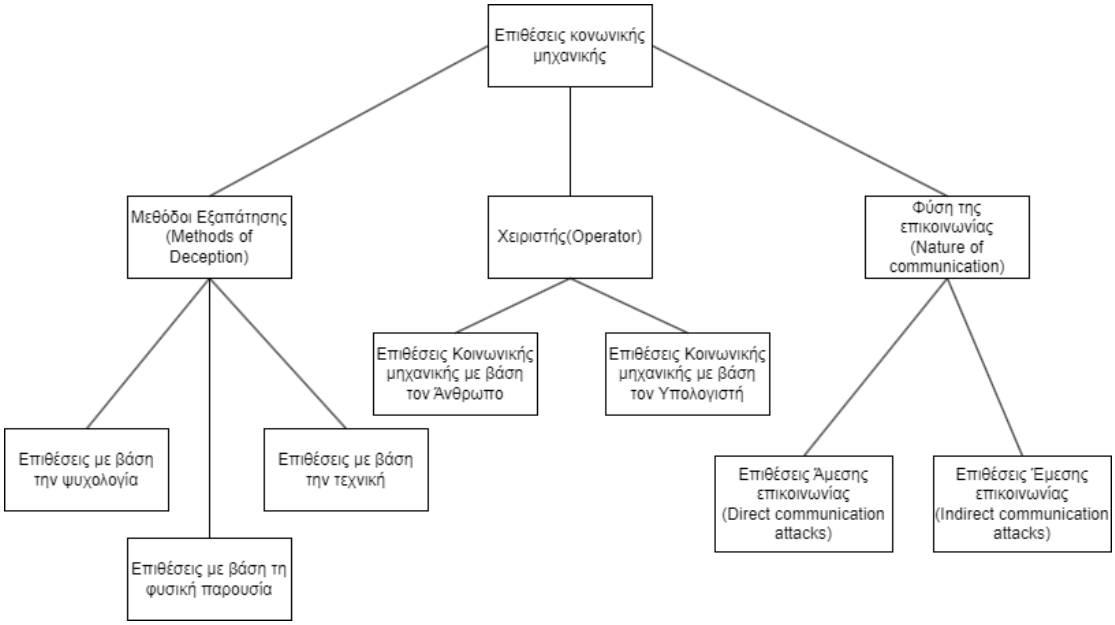
- Επιθέσεις με βάση την τεχνική (Technical based attacks): Ο επιτιθέμενος χρησιμοποιεί το διαδίκτυο. Είναι πολύ συχνό είδος επίθεσης λόγω της διάδοσης των κοινωνικών δικτύων, τα οποία χρησιμοποιούνται για βιντεοκλήσεις, chat και γενικά επικοινωνία μέσω του διαδικτύου. Ο επιτιθέμενος συνήθως αφού λάβει/υποκλέψει τις (προσωπικές) πληροφορίες του θύματος μέσω μιας ψεύτικης ιστοσελίδας που έχει δημιουργήσει ή με κάποιο άλλο τεχνικό τρόπο, εκβιάζει το θύμα για να μην τις διαρρεύσει. [4]

- Επιθέσεις με βάση την φυσική παρουσία (Physical based attacks): Ο επιτιθέμενος χρησιμοποιεί οποιαδήποτε μορφή σωματικής δράσης προκειμένου να λάβει πληροφορίες για το θύμα. Η πιο συχνή επίθεση τέτοιου τύπου είναι το dumpster diving, όπου ο επιτιθέμενος ψάχνει τα σκουπίδια μιας εταιρείας προκειμένου να βρεί τις πληροφορίες που ψάχνει, σε έγγραφα και χαρτιά που έχουν πεταχτεί. [5]

3. **Φύση της επικοινωνίας (Nature of communication):** Ο επιτιθέμενος μέσα από άμεση ή έμμεση επικοινωνία προσπαθεί να αποσπάσει τις πληροφορίες από το θύμα. Η διάσταση αυτή περιλαμβάνει τις εξής 2 υποκατηγορίες. [6]

   - Επιθέσεις Άμεσης επικοινωνίας (Direct communication attacks): Ο επιτιθέμενος επικοινωνεί απευθείας με το θύμα με τη φυσική παρουσία του είτε με την φωνή του μέσω τηλεφώνου. Ένα παράδειγμα τέτοιων επιθέσεων είναι το phone scam attack, όπου ο επιτιθέμενος επικοινωνεί κατευθείαν με το θύμα μέσω τηλεφώνου. [6]

   - Επιθέσεις Έμμεσης επικοινωνίας (Indirect communication attacks): Ο επιτιθέμενος εκτελεί την επίθεση εξ αποστάσεως χωρίς να χρειάζεται να έχει άμεση ή σωματική επαφή με το θύμα. Παραδείγματα τέτοιων επιθέσεων είναι το Ransomware ή τα pop-up windows. [7]



Εικόνα 2: Κύριες Διαστάσεις Επιθέσεων Κοινωνικής Μηχανικής

Οι κύριες διαστάσεις μιας επίθεσης κοινωνικής μηχανικής με τις κατηγορίες τους φαίνονται στην Εικόνα 2. Μάλιστα, οι επιθέσεις κοινωνικής μηχανικής που πραγματικά επιτελούνται ανήκουν αποκλειστικά σε μια από αυτές τις κατηγορίες ή σε κάποιο συνδυασμό τους. Για παράδειγμα, η επίθεση τύπου tailgaiting attack είναι συνδυασμός επίθεσης κοινωνικής μηχανικής με βάση τον άνθρωπο με επίθεση άμεσης επικοινωνίας, εφόσον ο επιτιθέμενος ως χειριστής χρειάζεται να έχει φυσική παρουσία στο χώρο όπου βρίσκεται το θύμα.

Στον παρακάτω πίνακα φαίνεται σε ποιες διαστάσεις κατατάσσονται οι τύποι επιθέσεων κοινωνικής μηχανικής με τις οποίες θα ασχοληθούμε.

Πίνακας 1: Ταξινόμηση επιθέσεων κοινωνικής μηχανικής, ανάλογα σε ποιά/ποιές διαστάσεις ανήκουν

Τύπος Επίθεσης **(Attack Type)**	**Χειριστής** **(Operator)**	**Μέθοδοι Εξαπάτησης** **(Methods of Deception)**	**Φύση της επικοινωνίας** **(Nature of communication)**
Impersonation	✓	✓	✓
Shoulder Surfing	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον άνθρωπο	Επιθέσεις με βάση την φυσική παρουσία	

Dumpster Diving	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον άνθρωπο	Επιθέσεις με βάση την φυσική παρουσία	
Being a third party	✓	✓	✓
Email Phishing	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την ψυχολογία Επιθέσεις με βάση την τεχνική	Επιθέσεις έμμεσης επικοινωνίας
Spear Phishing	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την ψυχολογία Επιθέσεις με βάση την τεχνική	Επιθέσεις έμμεσης επικοινωνίας
Whaling Phishing	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την ψυχολογία Επιθέσεις με βάση την τεχνική	✓
Vishing Phishing	✓	Επιθέσεις με βάση την φυσική παρουσία	✓
Angler Phishing	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την ψυχολογία Επιθέσεις με βάση την τεχνική	Επιθέσεις έμμεσης επικοινωνίας
Bluetooth Phishing-Snarfing	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την τεχνική Επιθέσεις με βάση την φυσική παρουσία	Επιθέσεις έμμεσης επικοινωνίας
Baiting Attacks	✓	✓	Επιθέσεις έμμεσης επικοινωνίας

Pretexting attacks	✓	✓	✓
Tailgating attacks	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον άνθρωπο	Επιθέσεις με βάση την ψυχολογία Επιθέσεις με βάση την φυσική παρουσία	Επιθέσεις Άμεσης Επικοινωνίας
Ransomware attacks	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την τεχνική	Επιθέσεις έμμεσης επικοινωνίας
Pop-up windows	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την ψυχολογία Επιθέσεις με βάση την τεχνική	Επιθέσεις έμμεσης επικοινωνίας
Scareware	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την ψυχολογία Επιθέσεις με βάση την τεχνική	✓
Phone and Email Scam attacks	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την ψυχολογία Επιθέσεις με βάση την τεχνική	✓
Quid Pro Quo	✓	✓	✓
Fake Mobile Apps	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την ψυχολογία Επιθέσεις με βάση την τεχνική	Επιθέσεις έμμεσης επικοινωνίας
HTTPS Man-In-The-Middle	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την τεχνική	

Multimedia Masquerading	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την τεχνική Επιθέσεις με βάση την ψυχολογία	Επιθέσεις έμμεσης επικοινωνίας
Peripheral Masquerading	✓	✓	Επιθέσεις Άμεσης Επικοινωνίας
Search Engine Poisoning (Spamdexing)	Επιθέσεις Κοινωνικής Μηχανικής με βάση τον υπολογιστή	Επιθέσεις με βάση την τεχνική	Επιθέσεις έμμεσης επικοινωνίας

2.2 Επιπλέον Κατηγοριοποίηση Επιθέσεων κοινωνικής μηχανικής με βάση τα βήματα που ακολουθούνται

Εισαγωγή

Στόχος της Κατηγοριοποίησης αυτής είναι να βοηθήσουμε τους ερευνητές και τους μηχανικούς να αναπτύξουν τεχνικές προσεγγίσεις άμυνας τόσο για τις τρέχουσες όσο και για τις μελλοντικές επιθέσεις κοινωνικής μηχανικής, αντιμετωπίζοντας τα βασικά χαρακτηριστικά των επιθέσεων και όχι συγκεκριμένες υλοποιήσεις [170]. Αυτό συμβαίνει επειδή ένα σύστημα άμυνας που έχει σχεδιαστεί για ένα συγκεκριμένο σύνολο χαρακτηριστικών μπορεί να είναι εφαρμόσιμο σε όλους τους τύπους επιθέσεων που μοιράζονται αυτά τα χαρακτηριστικά, καθιστώντας το έτσι μια πιο αποτελεσματική επιλογή/λύση. Σημειώστε ότι οι σημερινές επιχειρήσεις εγκλήματος στον κυβερνοχώρο μπορεί να έχουν πολλά επίπεδα και ένας αντίπαλος μπορεί να χρησιμοποιεί σύνθετες επιθέσεις που αποτελούνται από πολλαπλές φάσεις επιμέρους επιθέσεων, παράλληλα ή η μία μετά την άλλη [8].

Ένα παράδειγμα μπορεί να είναι ένα ανεπιθύμητο μήνυμα ηλεκτρονικού ταχυδρομείου που περιέχει μια διεύθυνση URL σε έναν παραποιημένο ιστότοπο, όπου ο χρήστης προωθείται σε μια drive-by λήψη που οδηγεί σε μόλυνση με μια εφαρμογή scareware. Η ταξινόμηση έχει σχεδιαστεί για να ξεχωρίζει τις επιθέσεις κοινωνικής μηχανικής ως μεμονωμένα συστατικά (π.χ. το drive-by download) και όχι τις πιθανές μεταθέσεις αυτών των στοιχείων (spam → παραποιημένος ιστότοπος → drive-by download → scareware) [170]. Αυτή είναι μια σημαντική λειτουργία της ανάπτυξης μιας γενικής ταξινόμησης των επιθέσεων που είναι πρακτικά χρησιμοποιήσιμη σε τεχνικό επίπεδο, διότι απλοποιεί τον στόχο μιας άμυνας μηχανισμού.

Υποθέτοντας ότι μια σύνθετη επίθεση απαιτεί όλες τις επιμέρους συνιστώσες της επιθέσεις για να πετύχει, οι μηχανισμοί που αντιμετωπίζουν οποιαδήποτε από τις τελευταίες θα προστατεύουν και από τη σύνθετη επίθεση. Στο παραπάνω παράδειγμα, η απειλή της εφαρμογής scareware που παρέχεται με αυτόν τον τρόπο θα μπορούσε να αποτραπεί από οποιονδήποτε τεχνικό μηχανισμό που θα εμποδίσει αποτελεσματικά το μήνυμα

ηλεκτρονικού ταχυδρομείου spam, τον παραποιημένο ιστότοπο ή την απόπειρα λήψης μέσω drive-by. Για να είναι χρησιμοποιήσιμα μακροπρόθεσμα, αυτά τα χαρακτηριστικά πρέπει να είναι καθολικά εφαρμόσιμα και ανεξάρτητα από την εμπλεκόμενη πλατφόρμα. Ειδικότερα, η ενοποίηση των πλατφορμών που οδηγείται από το Υπολογιστικό Νέφος (Cloud Computing), το Διαδίκτυο των Πραγμάτων (Internet of Things) και άλλες πρόσφατες υπολογιστικές τεχνολογίες καθιστά τη διάκριση μεταξύ κινητών, επιτραπέζιων και ενσωματωμένων συστημάτων, λιγότερο προφανή από ό,τι στο παρελθόν. Για παράδειγμα, στο εγγύς μέλλον, οι ψεύτικες ειδοποιήσεις για την πίεση των ελαστικών που εμφανίζονται στο ταμπλό ενός αυτοκινήτου [9] μπορεί να χρησιμοποιηθούν για να επιτευχθεί παραπλάνηση με τρόπο που δεν διαφέρει πολύ από τις σημερινές αναδυόμενες προειδοποιήσεις scareware για χρήστες κινητών και επιτραπέζιων υπολογιστών. Μια αμυντική προσέγγιση που επικεντρώνεται στην εξαπάτηση ως προς το χαρακτηριστικό των ψεύτικων προειδοποιήσεων και όχι στη στοχευμένη πλατφόρμα, θα μπορούσε ενδεχομένως να εφαρμοστεί.

Αυτά τα χαρακτηριστικά πρέπει επίσης να καλύπτουν όλα τα στάδια μιας επίθεσης. Για το σκοπό αυτό, υιοθετούμε το ορισμό των τριών διακριτών σταδίων ελέγχου ενορχήστρωσης, εκμετάλλευσης και εκτέλεσης που προτείνονται από την [10]. Για κάθε στάδιο, οι συγγραφείς θέτουν τα ερωτήματα που πιστεύουμε ότι θα είχαν τη μεγαλύτερη σημασία για τους προγραμματιστές τεχνικών μηχανισμών προστασίας. Επιπλέον, εξασφαλίζεται ότι για κάθε κατηγορία, οι ταξινομήσεις είναι αμοιβαία αποκλειόμενες.

Στάδιο ελέγχου 1: Ενορχήστρωση (Orchestration)

(1) *Πώς επιλέγεται ο στόχος;* Αυτό μπορεί να βοηθήσει στον προσδιορισμό των συνθηκών έκθεσης σε μία επίθεση. Για παράδειγμα, αν ένας χρήστης είναι ευάλωτος λόγω ενός συγκεκριμένου χαρακτηριστικού του ή επιλέγεται τυχαία, έχει διαφορά για τα χαρακτηριστικά του χρήστη και του συστήματος στα οποία ένα αμυντικό σύστημα πρέπει να επικεντρωθεί. Οπότε, έχουμε συγκεκριμένη στόχευση, όταν ο στόχος επιλέγεται λόγω συγκεκριμένων χαρακτηριστικών η αδυναμιών που ενδεχομένως έχει, ενώ έχουμε μη συγκεκριμένη στόχευση όταν ο στόχος επιλέγεται τυχαία.

(2) *Πώς φτάνει η επίθεση στον στόχο;* Αυτό μπορεί να βοηθήσει στον εντοπισμό των πλατφορμών που εμπλέκονται στην επίθεση, η οποία με τη σειρά της μπορεί να βοηθήσει έναν προγραμματιστή να επιλέξει ποιο απομακρυσμένο (άρα, με τη συμμετοχή ενός δικτύου) ή τοπικό σύστημα να παρακολουθεί και ενδεχομένως, πού να τοποθετηθεί ένας μηχανισμός άμυνας.

(3) *Είναι η επίθεση αυτοματοποιημένη;* Ο βαθμός αυτοματοποίησης σε μια επίθεση μπορεί να αλλάξει θεμελιωδώς τον μηχανισμό απόκρισης ή τον τύπο των δεδομένων που μπορούν να έχουν νόημα να συλλεχθούν σχετικά με αυτήν. Για παράδειγμα, μια πλήρως αυτοματοποιημένη επίθεση μπορεί να είναι δυνατό να αποτυπωθεί με βάση τα πρότυπα της συμπεριφοράς που έχει παρατηρηθεί προηγουμένως, ενώ η αποτύπωση μιας πλήρως χειροκίνητης επίθεσης μπορεί να επικεντρωθεί στη συμπεριφορά του επιτιθέμενου.

Στάδιο ελέγχου 2: Εκμετάλλευση (Exploitation)

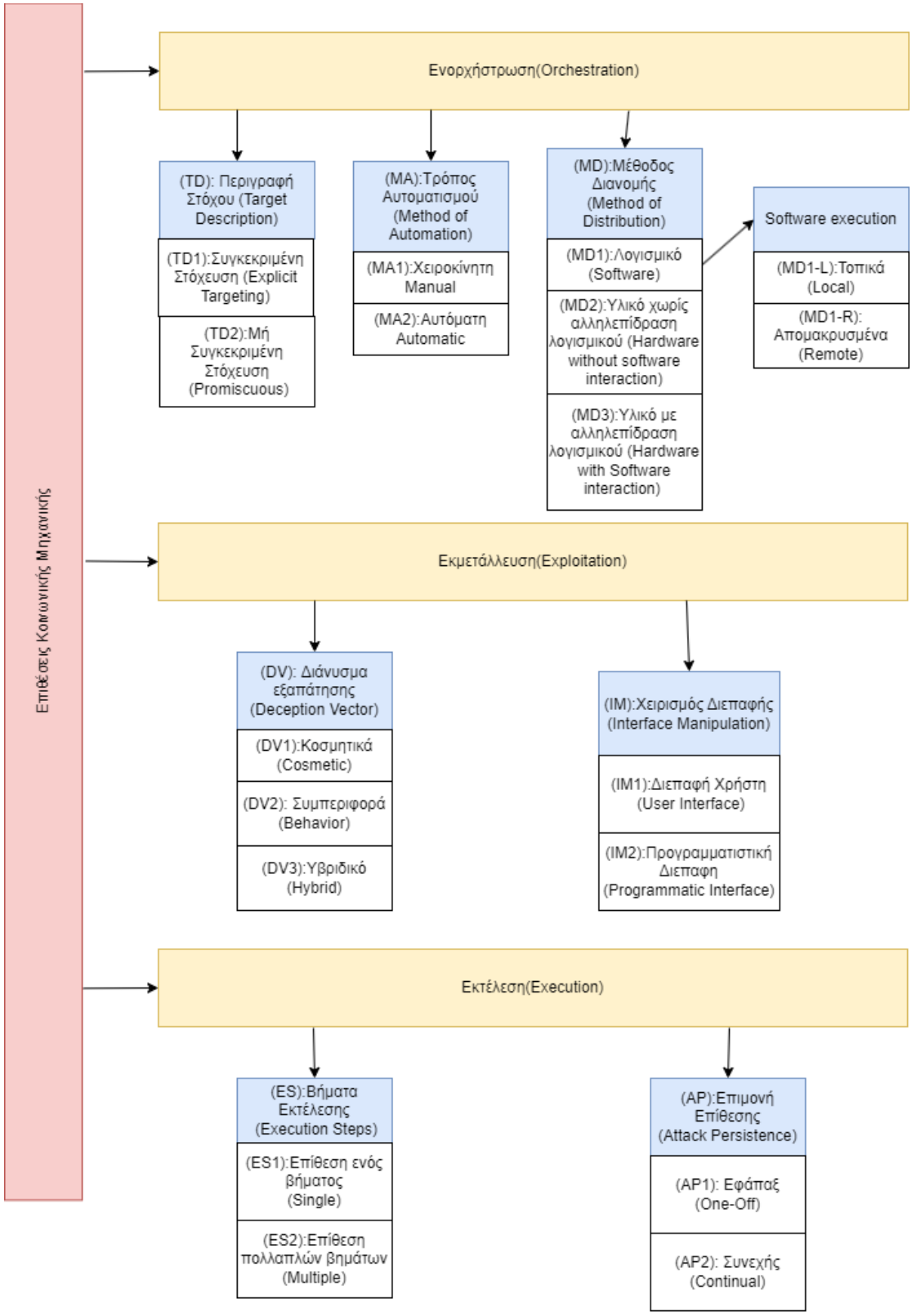
(1) *Είναι η εμφάνιση (User Interface) ή η συμπεριφορά (Behaviour) που εξαπατά τον χρήστη;* Αυτό μπορεί ενδεχομένως να βοηθήσει τον προγραμματιστή ενός αμυντικού μηχανισμού να εντοπίσει μηχανισμούς με τους οποίους ένας επιτιθέμενος μπορεί να εξαπατήσει τον χρήστη σε λανθασμένες προσδοκίες, χειραγωγώντας την οπτική ή/και τη γενική συμπεριφορά ενός συστήματος.

(2) *Η πλατφόρμα που χρησιμοποιείται για την εξαπάτηση χρησιμοποιείται μόνο ή τροποποιείται και προγραμματιστικά;* Ο προσδιορισμός του κατά πόσον η εξαπάτηση λαμβάνει χώρα σε κώδικα (ενσωματωμένο στο σύστημα ή εξωτερικά), ή με κατάχρηση της προβλεπόμενης λειτουργικότητας του χώρου χρήστη, μπορεί να βοηθήσει στη διαμόρφωση του σχεδιασμού ενός συστήματος άμυνας, περιορίζοντας το πεδίο εφαρμογής του.

Στάδιο ελέγχου 3: Εκτέλεση (Execution)

(1) *Η επίθεση ολοκληρώνει την εξαπάτηση σε ένα βήμα;* Μια επίθεση που βασίζεται σε περισσότερα από ένα βήματα, μπορεί ενδεχομένως να ανιχνευθεί ευκολότερα από ό,τι σε ένα μόνο βήμα ακόμη και πριν ολοκληρωθεί, αναζητώντας ίχνη των αρχικών βημάτων της. Μπορεί επίσης να αποτραπεί εμποδίζοντας έστω και μία από τις ενέργειες που χρειάζεται ώστε ο χρήστης για να εξαπατηθεί.

(2) *Επιμένει η εξαπάτηση;* Σε αντίθεση με τις εφάπαξ απόπειρες εξαπάτησης, οι επίμονες μπορεί να έχουν μεγάλες πιθανότητες να πετύχουν τον στόχο τους, αλλά θα μπορούσαν επίσης να βοηθήσουν ένα αμυντικό σύστημα που βασίζεται στη μάθηση να εντοπίσει σταδιακά το μοτίβο συμπεριφοράς τους και να το εμποδίσει. Κάθε κατηγορία απαντήσεων που αντιστοιχούν σε κάθε ερώτηση συμβάλλει στον καθορισμό των τμημάτων και υποτμημάτων της Κατηγοριοποίησης, όπως φαίνεται στην Εικόνα 3.



Εικόνα 3: Κατηγοριοποίηση Επιθέσεων Κοινωνικής Μηχανικής

Χρησιμοποιώντας ένα μοντέλο επίθεσης ενός στρώματος για την ταξινόμηση, η Κατηγοριοποίηση μας είναι σε θέση να προσδιορίσει τη σύνθεση μιας επίθεσης κοινωνικής μηχανικής χρησιμοποιώντας μια προσέγγιση βασισμένη σε παραμέτρους, οργανώνοντας τα κριτήρια ταξινόμησης με γραμμικό τρόπο. Σε αντίθεση με προηγούμενες προσεγγίσεις ταξινόμησης (επιθέσεων), καταγράφει πολλαπλές μεταβλητές που εμπλέκονται στην παράδοση και εκτέλεση μιας επίθεσης κοινωνικής μηχανικής, εφαρμόζοντας κριτήρια που είναι ανεξάρτητα από τους χρησιμοποιημένους φορείς επίθεσης. Ως αποτέλεσμα, μπορεί δυνητικά να αντιπροσωπεύει τόσο τις υπάρχουσες όσο και τις μελλοντικές επιθέσεις.

Στάδιο ελέγχου 1: Ενορχήστρωση(Orchestration)

TD: Περιγραφή στόχου (Target Description)

Αναφέρεται στη μεθοδολογία στόχευσης που εφαρμόζεται στη διαδικασία ενορχήστρωσης της επίθεσης: αν ο επιτιθέμενος έχει επιλέξει να στοχεύσει ένα άτομο ή μια ομάδα τυχαία ή με βάση την ταυτότητά τους ή κάποιο άλλο αποκλειστικό χαρακτηριστικό τους.

TD1: Συγκεκριμένη στόχευση (Explicit Targeting)

Με τον όρο ρητή στόχευση, αναφερόμαστε στην πρακτική της επιλογής ενός συγκεκριμένου ατόμου ή ομάδας χρηστών ως στόχων μιας επίθεσης με βάση τις συγκεκριμένη ταυτότητες ή ένός αποκλειστικού χαρακτηριστικού τους (εταιρεία, ρόλος, τοποθεσία κλπ.). Τυπικό παράδειγμα εδώ είναι το spear-phishing. Το spear-phishing είναι η στοχευμένη έκδοση του phishing, όπου ένα προσεκτικά διαμορφωμένο μήνυμα ηλεκτρονικού ταχυδρομείου τύπου phishing απευθύνεται σε ένα συγκεκριμένο άτομο ή οργανισμό. Αυτό είναι γνωστό ως whale phishing όταν ο στόχος είναι ιδιαίτερα πολύτιμος. (π.χ. ένα ανώτερο στέλεχος) [11].

Στην ίδια κατηγορία της ρητής στόχευσης με βάση ένα κοινό χαρακτηριστικό, μπορούμε να συμπεριλάβουμε την πρακτική της στόχευσης θυμάτων που βρίσκονται σε μια συγκεκριμένη χώρα, έτσι ώστε να καλλιεργούνται σκόπιμα botnets καλύτερης ποιότητας ή να γίνεται εκμετάλλευση των κοινωνικοπολιτικών υποθέσεων της εν λόγω χώρας [12]. Για λόγους έμφασης, ας διευκρινίσουμε ότι εδώ περιλαμβάνουμε μόνο τις επιθέσεις όπου ο στόχος επιλέγεται ρητά για την ταυτότητά του ή για ένα αποκλειστικό χαρακτηριστικό του. Για παράδειγμα, μια επίθεση WiFi phishing σε μια καφετέρια δεν αποτελεί στοχευμένη επίθεση (βλ., TD2 παρακάτω). Το γεγονός ότι οι άνθρωποι που πίνουν καφέ είναι στόχος είναι μόνο μια δεύτερη τάξη συνέπειας, καθώς ο επιτιθέμενος είναι εξαιρετικά απίθανο να τους έχει επιλέξει ρητά επειδή πίνουν καφέ.

Ωστόσο, θα μπορούσε κανείς να φανταστεί μια στοχευμένη περίπτωση αυτής της επίθεσης (ας πούμε, μια επίθεση "WiFi spearphishing") που εξαπολύεται ειδικά όταν ένα

συγκεκριμένο άτομο-στόχος είναι γνωστό ότι επισκέπτεται την εν λόγω καφετέρια. Αυτή η υποθετική επίθεση θα μπορούσε πράγματι να ταξινομηθεί ως TD1.

TD2: Μή Συγκεκριμένη στόχευση (Promiscuous Targeting)

Η μη συγκεκριμένη στόχευση είναι διαφανής και συνήθως επικεντρώνεται στη μέγιστη δυνατή έκθεση. Ένα τυπικό παράδειγμα θα ήταν μια μεγάλης κλίμακας εκστρατεία phishing που εξαπολύεται από ένα botnet εναντίον οποιασδήποτε διεύθυνσης ηλεκτρονικού ταχυδρομείου μπορεί να βρει ο επιτιθέμενος [13]. Τα σκουλήκια μιας επίθεσης κοινωνικής μηχανικής που χρησιμοποιούν τα άμεσα μηνύματα και την ανταλλαγή αρχείων συνήθως στοχεύουν επίσης στη μέγιστη έκθεση μέσω της άπληστης στόχευσης [14]. Οι επιθέσεις WiFi phishing είναι γενικά άδολες στη στόχευση των χρηστών. Ο επιτιθέμενος εγκαθιστά ένα κακόβουλο σημείο πρόσβασης για να εκμεταλλευτεί κάθε εισερχόμενη σύνδεση πελάτη, ανεξάρτητα από την ταυτότητα του χρήστη ή οποιοδήποτε άλλο χαρακτηριστικό [15] - [16].

MA: Τρόπος αυτοματισμού (Method of Automation)

Ο τρόπος αυτοματισμού αναφέρεται στο βαθμό εποπτείας του επιτιθέμενου στην ενεργοποίηση και διαχείρισης μιας επίθεσης κοινωνικής μηχανικής: με τη χειροκίνητη (manual) επίβλεψη κάθε στοιχείου της επίθεση καθ' όλη τη διάρκεια της ζωής της, ή με την προ-ρύθμιση κάθε πτυχής της επίθεσης και την ύπαρξη κανενός είδους ελέγχου της επίθεσης μετά την εκτόξευσή της .

MA1: Χειροκίνητη (Manual)

Απαιτείται παρέμβαση από τον επιτιθέμενο κατά την παράδοση μιας επίθεσης κοινωνικής μηχανικής. Ο επιτιθέμενος ενεργοποιεί ρητά τον μηχανισμό εξαπάτησης στο περιβάλλον του στόχου και, κατά περίπτωση, ανταποκρίνεται ρητά σε συμβάντα όταν υπάρχει αλληλεπίδραση με τον χρήστη. Η διαδικασία επίθεσης είναι ευέλικτη, επειδή ο επιτιθέμενος έχει τον πλήρη έλεγχο της διαδικασίας και μπορεί να την τροποποιεί κατά τη διάρκεια της επίθεσης. Παραδείγματα θα μπορούσαν να είναι περιλαμβάνουν την χειροκίνητη αποστολή ενός ηλεκτρονικού μηνύματος phishing από έναν παραποιημένο λογαριασμό ηλεκτρονικού ταχυδρομείου αντί ενός αυτοματοποιημένου botnet, η διανομή μολυσμένων DVD έξω από ένα κτίριο ή μία φυσική υποκλοπή νόμιμων συσκευών πολυμέσων και ενσωμάτωση σε αυτές ενός κακόβουλου ωφέλιμου φορτίου πριν φτάσουν στο στόχο τους [17].

MA2: Αυτοματοποιημένη (Automatic)

Εδώ, η επίθεση παραδίδεται με αυτοματοποιημένο τρόπο χωρίς να απαιτείται επικοινωνία ή παρέμβαση από τον επιτιθέμενο. Ο επιτιθέμενος έχει προγραμματίσει εκ των προτέρων τη διαδικασία για την παράδοση της επίθεσης. Η αυτόματη λειτουργία της παρέχει απόκρυψη του επιτιθέμενου, καθώς η επίθεση λειτουργεί χωρίς εξωτερική επιρροή αλλά περιορίζει συνάμα τον έλεγχο και την ευελιξία.

Για παράδειγμα, στην περίπτωση του σκουληκιού (worm) Anna Kournikova, το κακόβουλο μήνυμα ηλεκτρονικού ταχυδρομείου και το συνημμένο αρχείο περιείχαν το πλήρες σύνολο οδηγιών για την επίθεση ενώ ο επιτιθέμενος δεν είχε κανέναν έλεγχο της στόχευσης ή των παραμέτρων της επίθεσης [18]. Ένα πιο πρόσφατο παράδειγμα είναι το "Selfmite," ένα κακόβουλο λογισμικό SMS για συσκευές Android, το οποίο περιέχει μια σκληρά κωδικοποιημένη διαδικασία για την αποστολή μηνυμάτων κειμένου που

περιέχουν έναν σύνδεσμο σε μια εφαρμογή. Αφού εγκατασταθεί μία φορά, στη συνέχεια αναπαράγει το μήνυμα στις πρώτες 20 επαφές της συγκεκριμένης συσκευής, μεταμφιεσμένο ουσιαστικά σε κείμενο από έναν γνωστό συνεργάτη με αυτοματοποιημένο τρόπο [19].

Στην ίδια κατηγορία, θα μπορούμε να συμπεριλάβουμε τις επιθέσεις drive-by download. Μόλις ένας επιτιθέμενος ενσωματώσει κακόβουλο κώδικα σε ένα ευάλωτο δικτυακό τόπο, η drive-by λήψη εκτελείται αυτόματα όταν ένας χρήστης επισκέπτεται τον ιστότοπο [20]. Γενικά, η πλήρης αυτοματοποίηση μπορεί όχι μόνο να βοηθήσει στην απόκρυψη της προέλευσης μιας επίθεσης κοινωνικής μηχανικής και να μειώσει την προσπάθεια που απαιτείται για την αναπαραγωγή της, αλλά περιορίζει επίσης και τον έλεγχο της διαδικασίας. Οι περισσότερες αυτόματες επιθέσεις κοινωνικής μηχανικής διαθέτουν χονδροειδείς μηχανισμούς εξαπάτησης που μπορούν να γίνουν αντιληπτές από έναν έμπειρο χρήστη.

MD: Μέθοδος διανομής (Method of Automation)

Αναφέρεται στον μηχανισμό με τον οποίο μια επίθεση φτάνει στο σύστημα-στόχο: μέσω διεπαφής λογισμικού (λογισμικό), που εκτελείται στον κεντρικό υπολογιστή-στόχο (λογισμικό-τοπικό) ή σε ένα κατανεμημένο περιβάλλον κεντρικού υπολογιστή (software-remote) είτε μέσω του λειτουργικού συστήματος του κεντρικού υπολογιστή (υλικό με λογισμικό αλληλεπίδραση) είτε πρόκειται για μια φυσική συσκευή (υλικό) που διασυνδέεται με τον κεντρικό υπολογιστή-στόχο μέσω άμεσης πρόσβασης στο υλικό (υλικό χωρίς αλληλεπίδραση λογισμικού) .

Οι διεπαφές λογισμικού είναι γενικά πιο συνηθισμένες. Επιθέσεις με βάση το λογισμικό μπορούν να είναι σε μεγάλο βαθμό αυτοματοποιημένες και να επαναχρησιμοποιηθούν καθώς και να αναπαραχθούν εύκολα με ελάχιστη επίβλεψη, ενώ η ενσωμάτωση μιας συσκευής υλικού σε ένα περιβάλλον-στόχο μπορεί να είναι εντατική σε χειρωνακτικό επίπεδο, υψηλού κινδύνου, ενώ είναι δύσκολο να στοχευθεί μεγάλος αριθμός χρηστών.

MD1: Λογισμικό (Software)

Πρόκειται για επιθέσεις κοινωνικής μηχανικής που διανέμουν έναν μηχανισμό εξαπάτησης μέσω ενός λογισμικού, εκμεταλλευόμενοι συνήθως την διεπαφή χρήστη (User Interface) νόμιμων εφαρμογών και όχι τεχνικές ατέλειες. Μπορούν να ταξινομηθούν ως τοπικές (Local) ή απομακρυσμένες (Remote).

MD1-L: Τοπικά (Local)

Οι τοπικές επιθέσεις λογισμικού διανέμονται από το τοπικό λειτουργικό σύστημα, όπως δείχνουν οι επιθέσεις που χρησιμοποιούν τη λογική του δούρειου ίππου (Trojan Horse) [21]. Ενώ ένας δούρειος ίππος που μεταφέρει ένα ωφέλιμο φορτίο επίθεσης κοινωνικής μηχανικής μπορεί ο ίδιος να έχει διανεμηθεί εξ αποστάσεως (π.χ., εκφορτώνεται εν αγνοία του χρήστη από έναν ιστότοπο), η επίθεση βρίσκεται ήδη μέσα στο μηχάνημα-στόχο όταν παρουσιάζεται στο στόχο της [22]. Αυτή η συμπεριφορά παρατηρείται συνήθως στη μεταμφίεση αρχείων PDF καθώς και σε επιθέσεις scareware [23], [15], [22].

MD1-R: Απομακρυσμένα (Remote)

Εδώ η επίθεση προέρχεται από ένα περιβάλλον κεντρικού υπολογιστή διαφορετικό από το σύστημα του χρήστη-στόχου. Μπορεί να αφορά διακομιστές ιστού ή κατανεμημένες

εφαρμογές, όπως εφαρμογές υπολογιστικού νέφους (για αποθήκευση, φιλοξενία διακομιστών, ηλεκτρονικό ταχυδρομείο) και πλατφόρμες peer-to-peer (δίκτυα P2P, instant messaging, torrents). Η απομακρυσμένη διανομή ποικίλλει μεταξύ διαφορετικών στοιχείων σε μια απομακρυσμένη πλατφόρμα.

Για παράδειγμα, μια κακόβουλη διεύθυνση URL μπορεί να παρουσιαστεί μέσω ενός προγράμματος άμεσων μηνυμάτων [24-25]. Κακόβουλα αρχεία μπορούν να μεταφερθούν μέσω συγχρονισμένων υπηρεσιών αποθήκευσης υπολογιστικού νέφους, κακόβουλες διαφημίσεις ενσωματωμένες σε παραβιασμένες ιστοσελίδες, αυτοματοποιημένα bots σε ιστότοπους κοινωνικής δικτύωσης, ή ένα κλεμμένο online gaming [26-29].

MD2: Υλικό χωρίς αλληλεπίδραση λογισμικού (Hardware without Software Interaction).

Αυτό αναφέρεται σε επιθέσεις που διανέμονται στον στόχο μέσω μιας τοπικής διεπαφής υλικού χωρίς αλληλεπίδραση με το λογισμικό του κεντρικού υπολογιστή. Ένα παράδειγμα αποτελούν οι εξωτερικές συσκευές υλικού που υποκλέπτουν παθητικά δεδομένα χρήστη. Οι επιτιθέμενοι μπορούν να εγκαταστήσουν συσκευές υποκλοπής man-in-the-middle με τη μορφή σημείων εξόδου δικτύου Ethernet, να χρησιμοποιήσουν το τηλέφωνο ώστε να επικοινωνήσουν με το θύμα ή καταγραφείς κλειδιών υλικού ενσωματωμένους σε πληκτρολόγια [30].

MD3: Υλικό με αλληλεπίδραση λογισμικού (Hardware with Software Interaction)

Οι περισσότερες επιθέσεις με κατανεμημένο υλικό βασίζονται σε ένα στοιχείο λογισμικού προκειμένου να παρακολουθήσουν ενέργειες του χρήστη, όπως την ανάκτηση Ονόματος χρήστη και Κωδικού, όταν αυτά πληκτρολογούνται, με κάποιον keylogger που είναι εγκατεστημένος σε αντίστοιχο πληκτρολόγιο. Για παράδειγμα, οι [31] περιγράφουν μια σειρά επιθέσεων που περιλαμβάνουν υλικό USB, το οποίο παρουσιάζεται ως μια μεταμφιεσμένη συσκευή υλικού που περιλαμβάνει ένα λογισμικό συστατικό, όπως ένα κακόβουλο αρχείο που ενεργοποιεί τη λειτουργία αυτόματης εκκίνησης. Στην ίδια κατηγορία μπορούμε να συμπεριλάβουμε εκμεταλλεύσεις που εξαρτώνται από πρωτόκολλα όπως το WiFi και το Bluetooth, καθώς και άλλα λιγότερο προφανή υλικά και πρωτόκολλα για τη διανομή τους.

……

Στάδιο ελέγχου 2: Εκμετάλλευση (Exploitation)

Η κατασκευή και εφαρμογή φορέων επίθεσης με σκοπό την παράκαμψη της ασφάλειας πληροφοριών του συστήματος.

DV: Διάνυσμα εξαπάτησης (Deception Vector)

Το διάνυσμα εξαπάτησης αποτελεί κεντρικό σημείο αυτής της ταξινόμησης. Καθορίζει τον μηχανισμό με τον οποίο ο χρήστης εξαπατάται ώστε να διευκολύνει την παραβίαση της ασφάλειας. Το διάνυσμα εξαπάτησης μπορεί να κατηγοριοποιηθεί ως εξής: κοσμητικό, με βάση τη συμπεριφορά ή σε συνδυασμό των δύο.

DV1: Κοσμητικό (Cosmetic)

Κατά τη χρήση γραφικών διεπαφών χρήστη (GUI), υπάρχει μια έμμεση εμπιστοσύνη μεταξύ του σχεδιαστή GUI και του χρήστη. Ο σχεδιαστής του GUI εμπιστεύεται ότι το GUI θα φαίνεται πιστικό και ο χρήστης εμπιστεύεται ότι κάθε GUI στοιχείο είναι αυτό που φαίνεται να είναι. Οι αισθητικές σημασιολογικές επιθέσεις εκμεταλλεύονται αυτή την εμπιστοσύνη χειραγωγώντας την εμφάνιση των συστατικών του GUI. Για παράδειγμα, ένα αρχείο περιέχει ένα όνομα, επέκταση τύπου, και εικονίδιο ενώ μπορεί να γίνει επεξεργασία του από σχετικό πρόγραμμα. Τα ονόματα αρχείων συνήθως εμφανίζονται ως ευανάγνωστες, σαφείς ενδείξεις για το τι ένα αρχείο κάνει ή περιέχει, όπως συχνά απαιτείται από τον συγγραφέα. Ως εκ τούτου, τα επίσημης εμφάνισης και αναγνωρίσιμα ονόματα αρχείων προκαλούν ένα επίπεδο εμπιστοσύνης.

Οι επιθέσεις στο [32] και [33] αποτελούν κορυφαία παραδείγματα όπου οι τεχνικές κοσμητικής εξαπάτησης έχουν αποδειχθεί ιδιαίτερα αποτελεσματικές σε εφαρμογές P2P, είτε με την αντιστοίχιση των ονομάτων αρχείων σε ένα κοινόχρηστο φάκελο του χρήστη (π.χ “Εικόνες”, “Βίντεο”, “Αρχεία”) είτε με τη δημιουργία ονομάτων αρχείων που αναζητούνται συχνά στο δίκτυο αλλά περιέχουν κακόβουλο περιεχόμενο. Οι επεκτάσεις αρχείων ενέχουν έναν υποτιθέμενο κίνδυνο που συνδέεται με τη συμπεριφορά τους. Για παράδειγμα, ".exe" είναι μια γνωστή επέκταση εκτελέσιμου αρχείου και συνεπώς μια απειλή για το σύστημα του χρήστη εάν το αρχείο είναι κακόβουλο. Ωστόσο, η ίδια προσοχή δεν ασκείται από τον χρήστη για άλλους εκτελέσιμους τύπους, όπως ".bat", ".ini", ".lnk" ή ".scr" [34].

Επιπλέον, σε Microsoft Windows, ο τύπος αρχείου και το κατάλληλο εικονίδιο που θα εμφανιστεί καθορίζονται από την επέκταση. Καθώς η Εξερεύνηση των Windows αποκρύπτει τις επεκτάσεις των αρχείων από προεπιλογή και τα εκτελέσιμα προγράμματα μπορούν να ρυθμιστούν ώστε να εμφανίζουν οποιοδήποτε εικονίδιο [35], οι χρήστες μπορούν να εξαπατηθούν από ένα εικονίδιο εάν το όνομα αρχείου είναι συγκαλυμμένο, όπως στο "document.txt.exe" που εμφανίζεται ως "document.txt". Με τον ίδιο τρόπο, τα μηνύματα ηλεκτρονικού ταχυδρομείου περιέχουν διευθύνσεις αποστολέα, τίτλους θεμάτων, σώμα κειμένου που περιέχει εικόνες και συνημμένα [36], ενώ οι ιστότοποι χρησιμοποιούν ένα όνομα τομέα URL και κοινά χαρακτηριστικά της σελίδας, όπως κουμπιά και εικόνες, τα οποία μπορούν να χειραγωγηθούν για να εξαπατήσουν έναν χρήστη.

Οι επιτιθέμενοι μπορούν να δημιουργήσουν ιστότοπους phishing που αντικαθιστούν την ταυτότητα των νόμιμων ομολόγων τους αντιγράφοντας τον πηγαίο κώδικα και εικόνες ώστε να φαίνονται εντελώς πανομοιότυπες [37]. Η καλλωπιστική παραπλάνηση είναι διαδεδομένη σε επιθέσεις κατά διεπαφών χρήστη που βασίζονται στον ιστό, όπως οι πλατφόρμες κοινωνικής δικτύωσης, όπου ένας επιτιθέμενος δημοσιεύει κακόβουλους συνδέσμους που μιμούνται οπτικά βίντεο, δελεάζοντας τους χρήστες να τους επιλέξουν για να δουν το περιεχόμενο [38]. Σε WiFi phishing, ο επιτιθέμενος ανακατευθύνει τους χρήστες σε έναν ιστότοπο ελέγχου ταυτότητας που μοιάζει με την σελίδα σύνδεσης για δωρεάν υπηρεσία διαδικτύου [39]. Τα αναμενόμενα στοιχεία GUI για μια τέτοια σελίδα περιλαμβάνονται για να προσδώσουν ακεραιότητα και συμμόρφωση στην εκμετάλλευση. Οι Felt και Wagner [40] παρέχουν παραδείγματα τεχνικών για την ανάπτυξη ψεύτικων εφαρμογών για κινητά που φαίνονται σαν τις αντίστοιχες νόμιμες, αντιγράφοντας την οθόνη σύνδεσης, συμπεριλαμβανομένων το λογότυπο, τη διάταξη του GUI και το κείμενο.

DV2: Συμπεριφορά (Behavior)

Εδώ, η παραπλάνηση επιτυγχάνεται με τη μίμηση της συμπεριφοράς ενός νόμιμου συστήματος ή κάποιου ατόμου και όχι την εμφάνιση. Οι χρήστες εξαπατώνται από την υποτιθέμενη λειτουργικότητα έναντι της τυπικής προσέγγισης που χρησιμοποιείται σε γνωστές υλοποιήσεις. Για παράδειγμα, σε μία επίθεση σημείου πρόσβασης WiFi, ο χρήστης εξαπατάται βλέποντας ένα φαινομενικά νόμιμο δίκτυο στο κατάλογο των διαθέσιμων δικτύων WiFi [15]. Επίσης, στις επιθέσεις URL phishing σε ιστότοπους κοινωνικής δικτύωσης [41-42], η εξαπάτηση βασίζεται αποκλειστικά στην αποδεκτή συμπεριφορά που σχετίζεται με την προέλευση της επίθεσης, όπως αυτή λαμβάνεται από έναν χρήστη που βρίσκεται στη λίστα φίλων του στόχου. Στην πραγματικότητα, το μήνυμα δημιουργείται αυτόματα από μια κακόβουλη εφαρμογή που έχει εγκατασταθεί στο λογαριασμό του εν λόγω φίλου.

DV3: Υβριδικό (Hybrid)

Οι υβριδικές προσεγγίσεις συνδυάζουν πτυχές τόσο της "όψης" (DV1) όσο και της "συμπεριφοράς" (DV2) για να δημιουργήσουν μια πειστική εξαπάτηση. Οι Corrons [43] και Giles [44] περιγράφουν πως οι εφαρμογές scareware μεταμφιέζονται σε προγράμματα προστασίας από ιούς χρησιμοποιώντας παρόμοια λογότυπα, κείμενο και διεπαφές, όπως τα νόμιμα αντίστοιχα, με μηχανές σάρωσης και εργαλεία αφαίρεσης μολύνσεων. Ο Paganini [45] έχει αναφέρει μια επίθεση όπου ένα περικάλυμμα (wrapper) επισυνάφθηκε σε μια νόμιμη εφαρμογή τραπεζικών συναλλαγών μέσω κινητού τηλεφώνου. Μετά την εγκατάσταση, φόρτωνε μια in-app HTML σελίδα που μιμείται τόσο την εμφάνιση όσο και τη συμπεριφορά της νόμιμης φόρμας σύνδεσης της εφαρμογής. Κατά την εισαγωγή των διαπιστευτηρίων του χρήστη, εμφανιζόταν ένα σφάλμα που ζητούσε την χρήστη να εγκαταστήσει ξανά την εφαρμογή. Σε αυτό το σημείο, η νόμιμη έκδοση θα είχε πράγματι εγκατασταθεί και θα λειτουργούσε όπως αναμενόταν από εκεί και πέρα.

Άλλα παραδείγματα υβριδικής εξαπάτησης μπορούν να παρατηρηθούν σε επιθέσεις σε αφαιρούμενα μέσα όπως USB flash drives. Μια συσκευή αποθήκευσης USB μπορεί να προ-φορτωθεί με μια νόμιμη εμφάνιση αρχείου Microsoft Word που στην πραγματικότητα περιέχει μια κακόβουλη μακροεντολή Visual Basic. Μόλις ο χρήστης κάνει κλικ στο αρχείο, το Microsoft Word ανοίγει όπως αναμένει ο χρήστης, ενώ παράλληλα εκτελεί τον κακόβουλο κώδικα που είναι προσαρτημένος στο έγγραφο word. Μία πρόσφατη επίδειξη από τους Nohl και Lehl [46] έδειξε πώς οι ελεγκτές υλικολογισμικού USB μπορούν να επαναπρογραμματιστούν για να παραποιήσουν άλλες συσκευές ώστε να πάρουν τον έλεγχο ενός συστήματος, να διαρρεύσουν δεδομένα, ή να κατασκοπεύουν τον χρήστη. Η εκμετάλλευση δείχνει πως μια μονάδα USB flash μπορεί να αναλάβει διαφορετικές ή πολλαπλές ταυτότητες σε έναν υπολογιστή, όπως μια κάμερα, ένα πληκτρολόγιο ή μια μονάδα USB flash, με το λογισμικό να χειραγωγεί τα αναγνωριστικά διεπαφής USB που χρησιμοποιούνται από το λειτουργικό σύστημα για την αλληλεπίδραση με τη λειτουργικότητα των συσκευών USB (π.χ. επεξεργασία καταστάσεων πλήκτρων που αποστέλλονται από ένα περιφερειακό πληκτρολόγιο USB). Εδώ, ο φορέας εξαπάτησης είναι τόσο κοσμητικός όσο και συμπεριφορικός. Η συσκευή USB εμφανίζεται και λειτουργεί όπως θα περίμενε ο χρήστης, αλλά επιπλέον εξαπατά το σύστημα ξενιστή ώστε να νομίζει ότι είναι μια διαφορετική συσκευή και εν τέλει να επιτρέψει την εκτέλεση αυθαίρετου κώδικα.

IM: Χειρισμός διεπαφής (Interface manipulation)

Ένας φορέας εξαπάτησης μπορεί να αξιοποιηθεί μόνο με την κακόβουλη χρήση της υπάρχουσας λειτουργικότητας του συστήματος-στόχου (διεπαφή χρήστη) ή και με την

προγραμματιστική τροποποίηση (προγραμματιστική διεπαφή). Για παράδειγμα, ένας επιτιθέμενος μπορεί να κατευθύνει τους χρήστες σε μια επίθεση διακομιστή-στόχου με την απλή ανάρτηση ενός υπερσυνδέσμου χρησιμοποιώντας τη διεπαφή χρήσης ενός ιστότοπου κοινωνικής δικτύωσης ή προγραμματιστικά εισάγοντας κακόβουλο κώδικα JavaScript στον κώδικα HTML του ιστότοπου αυτού.

IM1: Διεπαφή χρήστη (User Interface)

Εδώ, αναφερόμαστε σε επιθέσεις που περιορίζονται στην κατάχρηση υφιστάμενων λειτουργικοτήτων που παρέχει μια διεπαφή χρήστη ως μέσο εξαπάτησης του χρήστη. Αυτό περιλαμβάνει τόσο τις διεπαφές χρήστη υλικού όσο και, συνηθέστερα, τις διεπαφές χρήστη λογισμικού. Στις εκμεταλλεύσεις κακόβουλης διαφήμισης [27], οι επιτιθέμενοι χρησιμοποιούν την διεπαφή χρήσης ενός συστήματος διαφήμισης σε έναν ιστότοπο, δημοσιεύοντας αρχικά μια νόμιμη διαφήμιση και στη συνέχεια αντικαθιστώντας την με μια κακόβουλη έκδοση μόλις αυτή αποκτήσει δημοτικότητα/εμπιστοσύνη στην πλατφόρμα υποδοχής.

Στη δηλητηρίαση μηχανών αναζήτησης, οι επιτιθέμενοι τοποθετούν επανειλημμένα άσχετες φράσεις στο σώμα των κείμενων ή τις διευθύνσεις URL σε έναν ιστότοπο του επιτιθέμενου, οι οποίες όταν ανιχνεύονται από μια μηχανή αναζήτησης, επιστρέφουν σίγουρα τον ιστότοπο αυτόν στα αποτελέσματα της αναζήτησης ενός χρήστη [47]. Στα σημεία πρόσβασης που λειτουργούν με τη μέθοδο man-in-the-middle, ο επιτιθέμενος προσαρμόζει το ρυθμιζόμενο από τον χρήστη αναγνωριστικό συνόλου υπηρεσιών (SSID) ώστε να ταιριάζει με εκείνο ενός νόμιμου σημείου πρόσβασης στο περιβάλλον στο οποίο έχει εγκατασταθεί [15].

IM2: Προγραμματική διεπαφή (Programmatic Interface)

Εδώ, αναφερόμαστε σε επιθέσεις που δεν περιορίζονται μόνο στο να χρησιμοποιούν την υπάρχουσα λειτουργικότητα, αλλά μπορούν επίσης να την τροποποιήσουν, συνήθως εκμεταλλευόμενες ευπάθειες στο σύστημα-στόχο. Αυτό είναι διαδεδομένο στις επιθέσεις drive-by download, όπου ο επιτιθέμενος εισάγει κακόβουλα σενάρια (π.χ. JavaScript) σε ευάλωτους ιστότοπους για να ανακατευθύνει τους χρήστες σε μια κακόβουλη λήψη/εγκατάσταση [48].

Μια τεχνική που περιγράφεται λεπτομερώς από τον Krishna [49] περιλαμβάνει τη συμπίεση σε ένα αρχείο zip ενός εκτελέσιμου αρχείου με ειδικούς δεκαεξαδικούς χαρακτήρες στο όνομα του αρχείου, ώστε να φαίνεται να μοιάζει με έγγραφο του Microsoft Word όταν η πρόσβαση γίνεται με ένα συγκεκριμένο δημοφιλές πρόγραμμα προβολής αρχείων zip που έχει αυτό το τεχνικό ελάττωμα. Ένα άλλο παράδειγμα παρουσιάστηκε στο [50] και [51], όπου η γλώσσα PDF χειραγωγείται για να να δημιουργηθεί ένα αρχείο που εμφανίζεται και συμπεριφέρεται όπως ένα νόμιμο αρχείο PDF, αλλά όταν εκτελείται μπορεί να ξεκινήσει μια επίθεση κοινωνικής μηχανικής ζητώντας διαπιστευτήρια εισόδου και ανακατευθύνοντας το χρήστη σε έναν κακόβουλο ιστότοπο. Αυτό επιτυγχάνεται με τη χρήση ορισμένων ειδικών λειτουργιών της γλώσσας PDF που μόλις μεταγλωτιστούν εκτελούνται κατά την εκτέλεση του αρχείου.

……

Στάδιο ελέγχου 3: Εκτέλεση (Execution)

Είναι η διαδικασία κατά την οποία εκτελείται η επίθεση.

ES: Βήματα εκτέλεσης (Execution Steps)

Μια ενιαία επίθεση κοινωνικής μηχανικής μπορεί να αποτελείται από ένα ή περισσότερα διακριτά βήματα.

ES1: Επίθεση ενός βήματος (Single)

Οι επιθέσεις ενός βήματος απαιτούν από τον χρήστη-στόχο να εκτελέσει μόνο μία ενέργεια και ο χρήστης χρειάζεται να εξαπατηθεί μόνο μία φορά για να πραγματοποιηθεί η επίθεση (π.χ. άνοιγμα ενός αρχείου, εκτέλεση ενός προγράμματος, κλικ σε ένα web σύνδεσμο ή κουμπί). Όταν λέμε ένα βήμα εννοούμε ότι ο χρήστης χρειάζεται να κάνει μία μοναδική ενέργεια ώστε να πέσει θύμα της επίθεσης. Μόλις πραγματοποιηθεί η αρχική ενέργεια από το θύμα, η επίθεση είναι σε θέση να αποκτήσει τις απαραίτητες πληροφορίες.

Παραλλαγές των επιθέσεων phishing που περιγράφονται στο [52] βασίζονται σε ένα μόνο βήμα. Όσο λιγότερη αλληλεπίδραση του χρήστη απαιτείται, τόσο μικρότερος είναι ο κίνδυνος εντοπισμού. Αυτός είναι ένας από τους λόγους για την αύξηση πλήθους των drive-by κακόβουλων λογισμικών, όπου το σύστημα του χρήστη μολύνεται με ένα μόνο κλικ [48].

ES2: Επίθεση πολλαπλών βημάτων (Multiple)

Σε μια επίθεση πολλαπλών βημάτων, ένας χρήστης πρέπει να εξαπατηθεί περισσότερες από μία φορά για να επιτύχει η επίθεση ένα ουσιαστικό αποτέλεσμα. Επιθέσεις που περιλαμβάνουν πολλαπλά βήματα στο πλαίσιο της εκμετάλλευσης χρησιμοποιούν γενικά ένα βήμα για την απόκτηση πρόσβασης και επιπλέον βήματα (για την επεξεργασία ή την αποστολή των δεδομένων) για την κλοπή πληροφοριών ή την παράδοση κακόβουλου λογισμικού. Για παράδειγμα, στο WiFi phishing, όπου ο επιτιθέμενος δημιουργεί ένα παραποιημένο SSID του τοπικού παρόχου WiFi (π.χ. σε μια καφετέρια ή ξενοδοχείο), ο χρήστης πρέπει πρώτα να συνδεθεί σε αυτό το παραπλανητικό σημείο πρόσβασης (βήμα 1) και στη συνέχεια να παράσχει τα διαπιστευτήριά του για τη σύνδεσή του στο αντίστοιχο δίκτυο WiFi (βήμα 2) αφού ανακατευθυνθεί σε μια σελίδα σύνδεσης phishing για πρόσβαση στο Διαδίκτυο [53].

Ομοίως, στη συνηθισμένη επίθεση typosquatting [54-55], ο χρήστης πρέπει πρώτα να πλοηγηθεί στην πλαστογραφημένη ιστοσελίδα με το λανθασμένο όνομα τομέα (βήμα 1) και στη συνέχεια να κάνει κλικ σε ένα κακόμορφο σύνδεσμο λήψης (βήμα 2).Οι περισσότερες επιθέσεις κοινωνικής μηχανικής που παρουσιάζουν συνεχή επιμονή της επίθεσης (AP2) είναι πολλαπλών βημάτων. Για παράδειγμα, τα scareware [56] συνήθως μολύνουν το σύστημα-στόχο μέσω παραπλάνησης (βήμα 1) και στη συνέχεια επιχειρούν περιοδικά περαιτέρω εξαπατήσεις μέσω ψεύτικων ειδοποιήσεων για κακόβουλο λογισμικό που εντοπίζεται στο σύστημα (επαναλαμβανόμενο βήμα 2).

AP: Επιμονή επίθεσης (Attack Persistence)

Διαφορετικές επιθέσεις κοινωνικής μηχανικής εφαρμόζουν ποικίλα επίπεδα παραπλάνησης που είναι εφάπαξ ή επιμένουν ανάλογα με τις προθέσεις του επιτιθέμενου.

AP1: Εφάπαξ (One-Off)

Εδώ, μόλις ο χρήστης ενεργοποιήσει το ωφέλιμο φορτίο της επίθεσης, για παράδειγμα, εκτελώντας μια ενέργεια που είτε παρέχει προνόμια πρόσβασης, είτε παρέχει ευαίσθητα δεδομένα χρήστη στον επιτιθέμενο, η επίθεση παύει κάθε περαιτέρω ενέργεια. Η μεγάλης κλίμακας ανεπιθύμητη αλληλογραφία είναι χαρακτηριστική αυτής της προσέγγισης. Οι επιτιθέμενοι αποστέλλουν μαζικά ένα μεγάλο σύνολο διευθύνσεων με στόχο να συλλάβουν στοιχεία των χρηστών μέσω ενός πλαστογραφημένου ιστότοπου, τα οποία, αφού συγκεντρωθούν, ανακατευθύνουν τους χρήστες στον νόμιμο ιστότοπο [57].

Ομοίως, μια επίθεση drive-by download ολοκληρώνεται με τη λήψη του κακόβουλου λογισμικού όταν ο χρήστης έχει πρόσβαση στον μολυσμένο ιστότοπο. Ο χρήστης παραμένει ευάλωτος μιας και η επίθεση πραγματοποιείται μια φορά, ενώ οι καρποί της επίθεσης συλλέγονται συνεχώς ακόμη και στα πλαίσια του ίδιου χρήστη θύματος.

AP2: Συνεχής (Continual)

Στις συνεχείς σημασιολογικές επιθέσεις, μια συγκεκριμένη περίπτωση επίθεσης δεν λήγει μετά την επιτυχή εκμετάλλευση οπότε ο χρήστης συνεχίζει να είναι εκτεθειμένος στις προσπάθειες εξαπάτησης. Αυτό μπορεί να περιλαμβάνει επαναλαμβανόμενη και άμεση επικοινωνία με τον στόχο μέσω της εκμετάλλευσης των μηχανισμών ανταλλαγής μηνυμάτων που παρέχονται από το ηλεκτρονικό ταχυδρομείο, τα άμεσα μηνύματα, δίκτυα P2P, φόρουμ IRC και άλλες υπηρεσίες που βασίζονται στο Διαδίκτυο.

Χαρακτηριστικό παράδειγμα αποτελούν οι εφαρμογές scareware, οι οποίες εμφανίζουν περιοδικά ψεύτικες ειδοποιήσεις κακόβουλου λογισμικού για όσο διάστημα παραμένουν σε λειτουργία στο σύστημα [56]. Η συνεχής συμπεριφορά είναι κοινή στις επιθέσεις αυτές που έχουν σχεδιαστεί για να αποσπάσουν οικονομικό κέρδος από έναν στόχο (π.χ αφού ο επιτιθέμενος αποκτήσει τους κωδικούς του χρήστη να υποκλέπτει χρηματικά ποσά για καιρό και όχι μόνο μία φορά ή να συνεχίσει να λαμβάνει νέες πληροφορίες από το μολυσμένο υπολογιστή με σκοπό νέες παραβιάσεις).

Πίνακας 2: Συσχέτιση τύπων επιθέσεων κοινωνικής μηχανικής με τις κατηγορίες της περαιτέρω κατηγοριοποίησης.

Τύπος Επίθεσης **(Attack Type)**	**Επιπλέον Κατηγοριοποίηση-Στάδια**
Impersonation	**TD1** **MA** **DV** **IM1** **ES**

	AP1
Shoulder Surfing	**TD1** **MA1** **ES1** **AP**
Dumpster Diving	**TD1** **MA1** **ES1** **AP2**
Being a third party	**TD1** **MA1** **DV** **IM** **ES** **AP1**
Email Phishing	**TD** **MA** **MD1 -> MD1-R** **DV** **IM1** **ES2** **AP2**

Spear Phishing	**TD** **MA** **MD1->MD1-R** **DV1-DV3** **IM** **ES2** **AP2**
Whaling Phishing	**TD1** **MA** **MD1->MD1-R** **DV** **IM1** **AP** **ES**
Vishing Phishing	**TD1** **MA1** **MD2** **DV** **ES1** **AP2**
Angler Phishing	**TD** **MA** **MD1->MD1-R** **DV** **IM1**

	ES **AP2**
Bluetooth Phishing-Snarfing	**TD2** **MA** **MD1->Software Execution** **DV1** **IM1** **ES1** **AP2**
Baiting Attacks	**TD2** **MA** **MD** **DV** **IM** **ES** **AP2**
Pretexting attacks	**TD** **MA** **MD1->MD1-R** **DV3** **IM1** **ES1** **AP2**
Tailgating attacks	**TD1** **MA1**

	DV2 **IM** **ES2** **AP1**
Ransomware attacks	**TD2** **MA2** **MD1->MD1-L** **DV1** **IM2** **ES2** **AP2**
Pop-up windows	**TD2** **MA2** **MD1->MD1R** **DV1** **IM** **ES2** **AP2**
Scareware	**TD2** **MA2** **MD1->MD1-R** **DV** **IM1** **ES2** **AP2**

Phone and Email Scam attacks	**TD** **MA** **MD1- MD2 -> MD1-R** **DV3** **IM1** **ES2** **AP2**
Quid Pro Quo	**TD1** **MA1** **MD1-MD2 -> MD1-R** **DV2** **IM1** **ES1** **AP1**
Fake Mobile Apps	**TD2** **MA2** **MD1->MD1-L** **DV3** **IM2** **ES2** **AP2**
HTTPS Man-In-The-Middle	**TD1** **MA1** **MD1** **DV**

	IM2 **ES2** **AP2**
Multimedia Masquerading	**TD2** **MA2** **MD1, MD1-R** **DV1** **IM2** **ES2** **AP2**
Peripheral Masquerading	**TD1** **MA1** **MD3** **DV** **IM** **ES1** **AP**
Search Engine Poisoning (Spamdexing)	**TD2** **MA2** **MD1, MD1-R** **DV1** **IM1** **ES2** **AP2**

2.3 Τύποι Επιθέσεων κοινωνικής μηχανικής

Θα αναλύσουμε κάποιους από τους πιο συχνούς τύπους επιθέσεων κοινωνικής μηχανικής και θα τους κατηγοριοποιήσουμε ανάλογα με τις διαστάσεις (Κατηγοριοποίηση της Ενότητας 1.2) τις οποίες χρησιμοποιούνται σε αυτούς. Κατά την ανάλυση, για κάθε τύπο θα αναφέρουμε και κάποια βασικά παραδείγματα επιθέσεών του.

Επιθέσεις οι οποίες ανήκουν στις παραπάνω διαστάσεις είναι:

- **Impersonation:** Ο επιτιθέμενος υποδύεται ή παίζει τον ρόλο ενός άλλου προσώπου, ώστε να μπορέσει να κερδίσει την εμπιστοσύνη του θύματος και να αποκτήσει την απαραίτητη γνώση, ώστε να εισέλθει σε ένα πληροφοριακό σύστημα ή δίκτυο. Ένα παράδειγμα είναι ότι ο επιτιθέμενος υποδύεται τον υπάλληλο μίας εταιρείας και επικοινωνεί με πρόθεση να αλλάξει τον κωδικό πρόσβασης τον οποίο ξέχασε και θέλει να αλλάξει [58].

- **Shoulder Surfing:** Ο επιτιθέμενος κοιτάει την οθόνη του θύματος την ώρα που γράφει δεδομένα στην εκάστοτε συσκευή που χρησιμοποιεί. Αφού απομνημονεύσει τα δεδομένα αυτά (username, password), μπορεί να αποκτήσει πρόσβαση στο σύστημα (του θύματος). Αυτό γίνεται ακόμα πιο εύκολα σε μια κινητή συσκευή στην οποία η πληκτρολόγηση, καθότι πιο δύσκολη από έναν υπολογιστή και του διαφορετικού περιβάλλοντος πληκτρολόγησης, είναι πιο εύκολα εμφανής στον επιτιθέμενο. [59]

- **Dumpster Diving:** Ο επιτιθέμενος αναζητά εμπιστευτικές πληροφορίες που μπορεί να ανήκουν στο θύμα, ψάχνοντας τα σκουπίδια του. Το θύμα μπορεί να είναι μία εταιρία ή ένας οργανισμός όπου πετάει μαζικά τα σκουπίδια του σε συγκεκριμένες τοποθεσίες που μπορεί να είναι εύκολα προσβάσιμες και να ανακαλυφθούν με σχετική ευκολία απο τον επιτιθέμενο [60].

- **Being a third party:** Ο επιτιθέμενος ισχυρίζεται ότι έχει αποκτήσει άδεια χρήσης του συστήματος ή κάποιας συσκευής που ανήκει στο δίκτυο, με σκοπό να αποκτήσει τις πληροφορίες που αναζητά όσο απουσιάζει το όντως εξουσιοδοτημένο πρόσωπο ή κάποιο πρόσωπο το οποίο το θύμα πιστεύει ότι έχει την άδεια [61].

- **Phishing attacks:** Ο επιτιθέμενος προσπαθεί να αποκτήσει πρόσβαση πληροφοριών με διάφορους τρόπους. Επειδή είναι από τις συχνότερες και πιο χρησιμοποιημένες επιθέσεις κοινωνικής μηχανικής θα αναλύσουμε κάποιες υποκατηγορίες της [62]:

  1. **Email Phishing:** Οι πιο συχνές επιθέσεις phishing γίνονται μέσω μηνυμάτων ηλεκτρονικού ταχυδρομείου. Ο επιτιθέμενος υποδύεται νόμιμες οργανώσεις (συνδυασμός με impersonation), όπως τράπεζες ή φιλανθρωπικά ιδρύματα. Για παράδειγμα στέλνει σαν μία τράπεζα, όπου ζητά από τον χρήστη να επιβεβαιώσει τα στοιχεία του λογαριασμού του. Με αυτό το τρόπο αποκτά πρόσβαση σε αυτές τις ευαίσθητες πληροφορίες, εφόσον ο χρήστης του στείλει τα στοιχεία αυτά [63].

  2. **Spear Phishing:** Συνήθως και αυτές οι επιθέσεις επιτελούνται μέσω μηνυμάτων ηλεκτρονικού ταχυδρομείου, τα οποία περιέχουν κακόβουλα αρχεία και ιστοσελίδες. Ο

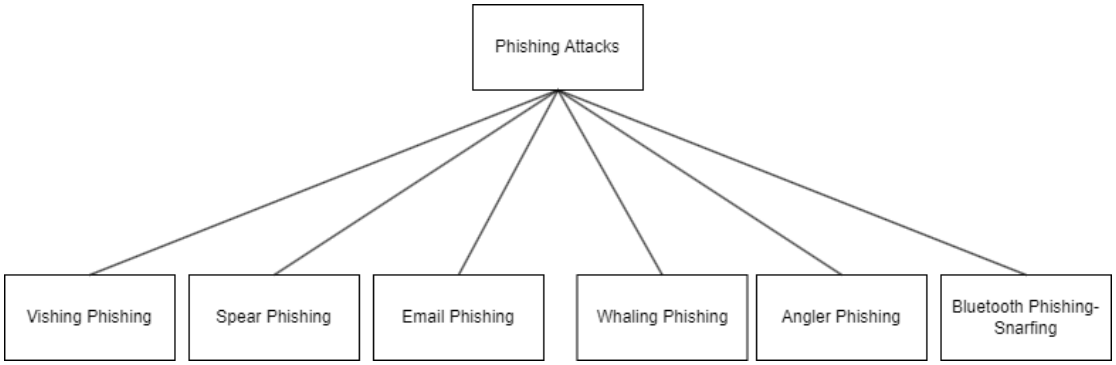
επιτιθέμενος στοχεύει στην εγκατάσταση του κακόβουλου λογισμικού σε ένα εταιρικό σύστημα ή στο σύστημα του ίδιου του θύματος μιας και με το που το θύμα επιλέξει τον σύνδεσμο ή εκτελέσει, ο επιτιθέμενος θα έχει πρόσβαση στις πληροφορίες που αναζητούσε. Το κακόβουλο αρχείο μπορεί να είναι ένα πρόγραμμα ransomware ή κάποιο keylogger. Η ιστοσελίδα θα τον παραπέμπει να κατεβάσει κάποιο πρόγραμμα για την ασφάλεια του ή για κάποια διαδικασία που θέλει να επιτελέσει είτε θα μοιάζει με μία επίσημη ιστοσελίδα όπου ο χρήστης πρέπει να ταυτοποιηθεί οδηγώντας στην αποκάλυψη των διαπιστευτηρίων πρόσβασής τους στον επιτιθέμενο [64].

3. **Whaling Phishing:** Ο επιτιθέμενος στοχεύει υπαλλήλους με περισσότερες αρμοδιότητες σε μία εταιρία, όπως ένας προϊστάμενος ή ένα στέλεχος μιας εταιρείας χρησιμοποιώντας μηνύματα ηλεκτρονικού ταχυδρομείου, τηλεφωνική επικοινωνία ή ισττοπους τους οποίους έχει ο ίδιος διαμορφώσει. Αυτές οι επιθέσεις διαφοροποιούνται με το spear phishing, στο ότι ο επιτιθέμενος στοχεύει υψηλόβαθμα μέλη μιας εταιρείας, διευθυντές και προϊσταμένους, οι οποίοι έχουν περισσότερες αρμοδιότητες και πληροφορίες για την εταιρία [65].

4. **Vishing Phishing:** Ο επιτιθέμενος μέσω τηλεφωνικής επικοινωνίας και υποδυόμενος κάποιο άλλο πρόσωπο, προσπαθεί να αποσπάσει πληροφορίες. Πολλοί εργαζόμενοι είναι απρόθυμοι να απορρίψουν ένα αίτημα από κάποιον τον οποίο θεωρούν σημαντικό [66].

5. **Angler Phishing:** Ο επιτιθέμενος χρησιμοποιεί τα social media και προσποιείται κάποιον υπεύθυνο ή υπάλληλο της πλατφόρμας αυτής, με πρόσχημα να βοηθήσει το θύμα (συνδυασμός με impersonation). Έπειτα ζητάει από το θύμα τους κωδικούς πρόσβασης οπότε το θύμα μιας και νομίζει ότι μιλάει με κάποιο στέλεχος-υπάλληλο της πλατφόρμας του δίνει τους κωδικούς [67].

6. **Bluetooth Phishing-Snarfing:** Ο επιτιθέμενος στοχεύει οποιονδήποτε χρήστη αφήνει ανοιχτές τις κινητές συσκευές του για αποδοχή Bluetooth αρχείων ως προεπιλεγμένη ρύθμιση. Ο επιτιθέμενος συνήθως στέλνει στην συσκευή του θύματος ένα αρχείο κακόβουλου λογισμικού, με όνομα αρχείου που μοιάζει με όνομα μιας δημοφιλής εφαρμογής (Facebook, Twitter, κ.α.) [68].



Εικόνα 4: Ταξινόμηση επιθέσεων Phishing και υποκατηγορίες

· **Baiting Attacks:** Ο επιτιθέμενος χειραγωγεί ψυχολογικά το θύμα και του δίνει ψεύτικες υποσχέσεις προκειμένου να του διεγείρει την περιέργεια, ώστε να του κλέψει τα προσωπικά του στοιχεία. Συνήθως αυτές οι επιθέσεις πραγματοποιούνται στον ιστό μέσω διαφημίσεων

ή ελκυστικών προσφορών, που προκαλούν το θύμα είτε να κάνει κλικ στον εκάστοτε σύνδεσμο και έπειτα να εισάγει τα προσωπικά του διαπιστευτήρια, είτε να προχωρήσει σε ηλεκτρονικές λήψεις από ιστότοπους που προσφέρουν δωρεάν μουσική ή παιχνίδια και λογισμικό, είτε να αγοράσει συσκευές μολυσμένες με κακόβουλο λογισμικό όπου το θύμα θα συνδέσει στον υπολογιστή του (π.χ ένα USB stick, με τάχα δωρεάν προγράμματα ή μουσική ώστε το θύμα να το συνδέσει στον υπολογιστή του). Η διαφορά αυτής της επίθεσης με phishing attacks είναι ότι εκμεταλλεύεται κυρίως την ανθρώπινη περιέργεια, ενώ τα phishing attacks βασίζονται σε μεγάλο βαθμό στην εμπιστοσύνη, το φόβο και την αίσθηση ότι κάτι επείγον έχει συμβεί [69-72].

· **Pretexting attacks:** Ο επιτιθέμενος δημιουργεί μία καινοτόμα ιστορία για να κερδίσει την εμπιστοσύνη του θύματος με σκοπό να το εξαπατήσει. Η επίθεση αυτή βασίζεται στην εμπιστοσύνη και στην εξουσία που έχει ο επιτιθέμενος, προσποιούμενος ένα εξουσιοδοτημένο άτομο το οποίο έχει το δικαίωμα πρόσβασης στις πληροφορίες του θύματος με σκοπό να ξεγελάσει το θύμα να κάνει μία ανεπιθύμητη ενέργεια, όπως να εισέλθει στον χώρο της επιχείρησης που εργάζεται με κάποιο συγκεκριμένο σκοπό.

Οι επιθέσεις αυτές συνήθως απαιτούν έρευνα πριν την αρχική αλληλεπίδραση με το θύμα ώστε ο επιτιθέμενος να έχει το κατάλληλο πρόσχημα. Όσο περισσότερες πληροφορίες ξέρει ο επιτιθέμενος τόσο το καλύτερο, μιας και θα είναι πιο πειστικός απέναντι στο θύμα. Συνήθως οι επιτιθέμενοι χρησιμοποιούν μηνύματα ηλεκτρονικού ταχυδρομείου, τηλεφωνικές κλήσεις ή φυσικά μέσα για να διεκπεραιώσουν την επίθεση αυτή [73-75].

- **Tailgating attacks:** Ο επιτιθέμενος παριστάνει ένα εξουσιοδοτημένο άτομο, ώστε να εισέλθει σε μία συγκεκριμένη περιοχή ή εταιρία, προκειμένου να αποκτήσει κάποιες ή συγκεκριμένες πληροφορίες [76].

- **Ransomware attacks:** Κακόβουλο λογισμικό το οποίο κρυπτογραφεί τα αρχεία του θύματος, με σκοπό τον εκβιασμό του. Συνήθως ο επιτιθέμενος ζητάει κάποιο χρηματικό ποσό η κάποιες συγκεκριμένες πληροφορίες, προκειμένου να αποκρυπτογραφήσει τα αρχεία του θύματος. Το κακόβουλο λογισμικό αυτό συνήθως εγκαθίσταται μέσω μηνυμάτων ηλεκτρονικού ταχυδρομείου ή μέσω επίσκεψης σε κακόβουλο ιστότοπο [77].

- **Pop-up windows:** Παράθυρα που εμφανίζονται στην οθόνη του χρήστη, ενημερώνοντας τον ότι έχασε την σύνδεση στο δίκτυο, παροτρύνοντας τον να εισάγει τα προσωπικά του διαπιστευτήρια είτε παρέχοντας κάποιο μήνυμα με κάποια ψεύτικη προειδοποίηση σχετικά με την ασφάλεια του υπολογιστή του. Αφού το θύμα κάνει κλικ στο παράθυρο, είτε θα δώσει τα προσωπικά του στοιχεία σε κάποια φόρμα, είτε θα του ζητηθεί να κατεβάσει κάποιο πρόγραμμα, το οποίο στην πραγματικότητα είναι κακόβουλο [78].

- **Scareware:** Τα προγράμματα αυτά εξαπατούν το θύμα, πείθοντας το ότι η συσκευή του έχει μολυνθεί από κάποιο ιό και παροτρύνοντάς το να κατεβάσει κάποιο antivirus, το οποίο στην πραγματικότητα είναι κακόβουλο λογισμικό. Το κακόβουλο αυτό λογισμικό επιτρέπει στον επιτιθέμενο να αποκτήσει πρόσβαση σε ευαίσθητα δεδομένα ή κρυπτογραφεί τα δεδομένα και εκβιάζει το θύμα με σκοπό κάποιο χρηματικό αντίτιμο ώστε το θύμα να ανακτήσει την πρόσβαση σε αυτά [79].

- **Phone and Email Scam attacks:** Ο επιτιθέμενος μέσω τηλεφώνου, μηνυμάτων ηλεκτρονικού ταχυδρομείου ή SMS επικοινωνεί με το θύμα, δημιουργώντας ένα πρόσχημα ότι έχει συμβεί κάτι επείγον (scareware) είτε παριστάνοντας ότι είναι κάποιος άλλος (impersonation), ζητώντας του πληροφορίες [80].

- **Quid Pro Quo:** Ο επιτιθέμενος ζητάει από το θύμα κάποια πολύτιμη πληροφορία είτε να κάνει κάποια ανεπιθύμητη πράξη-ενέργεια, προσφέροντας του κάποιο αντάλλαγμα είτε είναι χρηματικό είτε κάποια υπηρεσία. Όμως, ο επιτιθέμενος λαμβάνει τα δεδομένα χωρίς να δώσει κάποιο αντάλλαγμα. Για παράδειγμα, ο επιτιθέμενος επικοινωνεί με το θύμα μέσω τηλεφώνου, υποδυόμενος έναν υπάλληλο του δήμου, ζητώντας από το θύμα το ΑΦΜ του [81-83].

- **Fake Mobile Apps:** Ο επιτιθέμενος ανεβάζει μία ψεύτικη εφαρμογή κινητών τηλεφώνων, η οποία παρουσιάζεται ως νόμιμο πρόγραμμα κινητού, σε κάποιο marketplace. Κατά την εγκατάσταση της από το θύμα η εφαρμογή ζητά πλήρη πρόσβαση (root access) στη συσκευή. Συνήθως όταν εκτελείται η ψεύτικη εφαρμογή, δεν παρουσιάζεται κάποια διαφορά στο θύμα με την κανονική εφαρμογή, με αποτέλεσμα το θύμα να μην καταλαβαίνει ότι έχει πέσει θύμα επίθεσης [84].

- **HTTPS Man-In-The-Middle:** Ο επιτιθέμενος παρεμποδίζει αιτήματα HTTPS του θύματος και τοποθετεί ένα πλαστό πιστοποιητικό μέσα στο root πιστοποιητικό του συστήματος του θύματος. Αυτό έχει ως αποτέλεσμα όταν το θύμα στέλνει ένα HTTPS αίτημα, ο επιτιθέμενος να τον ανακατευθύνει σε ιστότοπους της επιλογής του, όπου φαίνονται εσφαλμένα νόμιμοι μέσω του πλαστού πιστοποιητικού. Αυτή η ευπάθεια είχε εντοπιστεί στο παρελθόν σε υπολογιστές Lenovo λόγω κακού adware που είχαν [85].

- **Multimedia Masquerading:** Ο επιτιθέμενος ενσωματώνει έναν κακόβουλο σύνδεσμο με τη μορφή προγράμματος αναπαραγωγής ήχου ή βίντεο. Μόλις επιλεγεί ο σύνδεσμος, το θύμα ανακατευθύνεται στην εγκατάσταση μιας εφαρμογής λογισμικού για την αναπαραγωγή του υποτιθέμενου αρχείου, η οποία ζητά από τον χρήστη να εισάγει ευαίσθητες πληροφορίες ώστε να αποκτήσει μία premium συνδρομή ή να συμπληρώσει ένα ερωτηματολόγιο. Συχνά αυτή η επίθεση γίνεται μέσα από πλατφόρμες κοινωνικής δικτύωσης [86].

- **Peripheral Masquerading:** Ο επιτιθέμενος χρησιμοποιεί κάποια συσκευή hardware, όπως ένα usb stick, ένα πληκτρολόγιο ή ένα ποντίκι, το οποίο περιέχει κακόβουλο λογισμικό. Συνδέει το hardware ή κοροϊδεύει το θύμα να το συνδέσει το ίδιο. Αφού το hardware συνδεθεί με τον υπολογιστή εγκαθίσταται κακόβουλο λογισμικό στο σύστημα [87].

- **Search Engine Poisoning (Spamdexing):** Σε αυτή την επίθεση η μηχανή αναζήτησης χειραγωγείται με τέτοιο τρόπο, ώστε να προτείνει στο θύμα τον κακόβουλο ιστότοπο μεταξύ των κορυφαίων αποτελεσμάτων της. Το θύμα μιας και η ιστοσελίδα εμφανίζεται σε μία αξιόπιστη μηχανή αναζήτησης, εμπιστεύεται τον ιστότοπο. Για να γίνει αυτή η επίθεση εφικτή ο ιστότοπος θα πρέπει να είναι πολύ δημοφιλής και πολύ σχετικός με τις λέξεις κλειδιά που χρησιμοποίησε το θύμα στην αναζήτηση του [88].

2.4 Εργαλεία μετρίασης επιθέσεων κοινωνικής μηχανικής

Σε αυτό το κεφάλαιο θα δούμε κάποια εργαλεία που αναπτύχθηκαν προκειμένου να μετριάσουν τους κινδύνους των επιθέσεων κοινωνικής μηχανικής και να γεμίσουν τα κενά ασφαλείας καθώς και τα τρωτά σημεία ενός συστήματος.

- **IDPS:** Ένα σύστημα Intrusion Detection Prevention System (IDPS) μπορεί να εντοπίσει μία παράβαση που έχει κατηγοριοποιηθεί ως κακόβουλη και να δράσει ώστε να σταματήσει την επίθεση και να αναφέρει την απειλή στον εκάστοτε οργανισμό. Για παράδειγμα, ένα IDPS μπορεί να ενωματώσει κάποια flags σε μηνύματα ηλεκτρονικού ταχυδρομείου, τα οποία ενδέχεται να περιέχουν κακόβουλο λογισμικό ή να προτρέπουν τον χρήστη σε κάποια περίεργη ιστοσελίδα, είτε να έχει εκπαιδεύσει τον χρήστη σε καθημερινές διαδικασίες όπου πιθανές επιθέσεις ελέγχονται (Ο χρήστης να επικοινωνεί με κάποιο υπεύθυνο ασφάλειας ή να σαρώνει τα αρχεία με κάποιο πρόγραμμα ασφαλείας αλλά και ο χρήστης εξοικειώνεται στη χρήση του κα στο να καταλαβαίνει πιο εύκολα πότε πρόκειται για επίθεση). Όμως ο θόρυβος μπορεί να μειώσει την αποτελεσματικότητα του συστήματος - ειδικότερα μπορεί να γίνεται ενημέρωση για ψευδούς συναγερμούς / ψευδώς κακόβουλες ενέργειες που δεν είναι πράγματι κακόβουλες (false alarms) [89].
- **Biometrics:** Συμπεριλαμβάνει τεχνολογίες ανίχνευσης προσώπου (face detection), δακτυλικού αποτυπώματος, γεωμετρίας χεριών και αποτυπώματος παλάμης (palm print). Για την φυσική πρόσβαση σε ένα χώρο θα μπορούσε να χρησιμοποιηθεί ένα μηχάνημα palm print ώστε να ταυτοποιείται ο χρήστης και να εισέρχεται στον προαναφερθέντα χώρο. Για μία απομακρυσμένη πρόσβαση μέσω μιας ηλεκτρονικής συσκευής σε μία εφαρμογή θα μπορούσε να χρησιμοποιηθεί ένα face detection λογισμικό, το οποίο θα ανιχνεύει και θα ταυτοποιεί τον χρήστη με βάση τα χαρακτηριστικά του προσώπου του [90].

Η χρήση βιομετρικών μηχανισμών βοηθάει στην επαλήθευση ταυτότητας σε μεγάλο βαθμό και θεωρείται ένας πολύ ασφαλής τρόπος για να ταυτοποιηθεί κάποιο άτομο, αλλά έρχεται με υψηλό κόστος και ψευδή θετικά (false positives), τα οποία επιτρέπουν σε μη εξουσιοδοτημένους χρήστες πρόσβαση στο σύστημα. Επίσης, με tailgating attacks ή κάποιο εκβιασμό από κάποιο phishing attack ο επιτιθέμενος μπορεί να ξεπεράσει αυτή την τεχνολογία ασφαλείας, μιας και θα «χειρίζεται» ή θα έχει ξεγελάσει το εξουσιοδοτημένο άτομο που έχει πρόσβαση μέσα από αυτές τις τεχνολογίες.

- **Τεχνητή Νοημοσύνη (Artificial Intelligence):** Βασίζεται σε υπολογιστικά συστήματα με την ικανότητα εκτέλεσης εργασιών που συνήθως συνδέονται με ευφυή όντα ώστε να μπορούν να μαθαίνουν και να βελτιώνονται στην εξάλειψη απειλών χωρίς να προγραμματίζονται. Η Τεχνητή Νοημοσύνη είναι αποτελεσματικ΄ξ και ακριβής, εξαλείφει σε ένα μεγάλο μέρος το ανθρώπινο λάθος και βελτιώνει τις ανθρώπινες αποφάσεις. Μπορεί να αναγνωρίζει πιο σύνθετες απειλές με αρκετά αποτελεσματικό τρόπο, αλλά χρειάζεται μεγάλος όγκος πληροφορίας, προκειμένου το σύστημα να εκπαιδευτεί [91].

Παρόλα αυτά έχει υψηλό κόστος και είναι αρκετά χρονοβόρο προκειμένου να υλοποιηθεί και να εφαρμοστεί σε ένα σύστημα. Επίσης, είναι αρκετά πολύπλοκο και με υψηλή ευαισθησία σε σφάλματα (error susceptibility).

- **Web security tools:** Είναι τύποι προγραμμάτων σχεδιασμένα να προστατεύουν από ιούς, κακόβουλο λογισμικό και adwares. Περιορίζουν τους χρήστες από το να επισκεφτούν ιστοσελίδες που ενδέχεται να περιέχουν κινδύνους ή να εκτελέσουν κάποιο αρχείο που ενδεχομένως περιέχει κακόβουλο λογισμικό. Παρόλα αυτά κάποιες φορές μπορεί να

αποκλείουν την επίσκεψη σε ασφαλής ιστοσελίδες και την εκτέλεση ασφαλών αρχείων [92].

- **Nmap:** Το λογισμικό είναι δωρεάν και χρησιμοποιείται για σάρωση δικτύων και τρωτών σημείων ασφαλείας. Δημιουργεί ένα χάρτη του δικτύου, ο οποίος περιλαμβάνει λεπτομερής πληροφορίες σχετικά με το τί κάνει κάθε θύρα και ποιά συσκευή την χρησιμοποιεί καθώς και πώς συνδέονται οι κεντρικοί υπολογιστές, τί περνάει από το τείχος προστασίας καθώς και παραθέτει τυχόν ζητήματα ασφαλείας που προκύπτουν. Αλλά αυτός ο έλεγχος είναι αρκετά χρονοβόρος, εφόσον το δίκτυο ή οι συσκευές του είναι παλαιωμένα ή δεν έχουν κατάλληλες υπολογιστικές ικανότητες [93-94].

2.5 Πρόληψη από επιθέσεις Κοινωνικής Μηχανικής

Η πρόληψη και η προστασία από επιθέσεις κοινωνικής μηχανικής είναι πολύ σημαντικές για όλους τους χρήστες ηλεκτρονικών συσκευών. Υπάρχουν αρκετές εκπαιδεύσεις, τεχνικές και σωστές συμπεριφορές, που έχουν σχεδιαστεί για την πρόληψη και τη μείωση τέτοιων επιθέσεων, τις οποίες θα αναλύσουμε παρακάτω:

- **Impersonation:** Αποφυγή ανοίγματος συνημμένων στα μηνύματα ηλεκτρονικού ταχυδρομείου που αποστέλλονται από άγνωστα άτομα, πριν επαληθευτεί η ταυτότητα του αποστολέα. Αποφυγή κλικ σε διαφημίσεις και άγνωστους ιστότοπους.

- **Shoulder Surfing:** Δεν εισάγουμε στοιχεία πιστωτικής κάρτας ή κωδικούς πρόσβασης λογαριασμών σε δημόσιους χώρους και χρησιμοποιούμε δυνατούς κωδικούς πρόσβασης, ώστε να δυσκολέψουμε τους επιτιθέμενους. Κρατάμε πάντα τον προσωπικό υπολογιστή κλειδωμένο και δεν αφήνουμε κανέναν να χρησιμοποιήσει τις προσωπικές μας συσκευές, ακόμα και αν είναι φίλος η συγγενής.

- **Dumpster Diving:** Τεμαχισμός η κάψιμο τυπωμένων αντιγράφων που περιέχουν ευαίσθητες πληροφορίες πριν τα πετάξουμε στους κάδους απορριμμάτων. Καθαρισμός συσκευών από προσωπικές πληροφορίες και ευαίσθητα δεδομένα πριν πεταχθούν ή πουληθούν.

- **Phishing attacks:**

    1. Κατάλληλη εκπαίδευση και ανάπτυξη τεχνολογικών και γενικών γνώσεων (Χρήση προγραμμάτων ασφαλείας, Συμβουλευτική με κάποιον ειδικό ασφάλειας ώστε να ενημερωθεί ο χρήστης για τους πιο συχνούς τρόπους επιθέσεων) ώστε να μπορείτε να αντιμετωπίσετε τέτοιες επιθέσεις.

    2. Μην κάνετε κλικ σε μηνύματα ηλεκτρονικού ταχυδρομείου ή προσωπικά μηνύματα τα οποία έρχονται από άγνωστους αποστολείς.

    3. Χρησιμοποιείτε πάντα HTTPS ιστοσελίδες για να εισάγετε ευαίσθητες πληροφορίες όπως username και password.

    4. Αλλάζετε κωδικούς πρόσβασης σε τακτά χρονικά διαστήματα χωρίς να επαναχρησιμοποιείτε τους παλιούς, ούτε να χρησιμοποιείτε ίδιους σε πολλούς λογαριασμούς. Οι κωδικοί πρόσβασης που επιλέγετε προσπαθείτε να είναι

σύνθετοι ώστε ο επιτιθέμενος να μην μπορεί να τους μαντέψει εύκολα. Χρησιμοποιείτε, επομένως, σύμβολα, γράμματα και αριθμούς αντί για προσωπικές πληροφορίες.

5. Συνεχής ενημέρωση λογισμικού ασφαλείας, λειτουργικών συστημάτων και εγκατάσταση ενημερωμένων προγραμμάτων τοιχών προστασίας (up to date firewall programs).

6. Δεν αποκαλύπτετε σε κανένα προσωπικές πληροφορίες, όπως ονόματα σημαντικών ανθρώπων, αριθμούς τηλεφώνου, ημερομηνίες, τόπους γέννησης, ονόματα κατοικιδίων ή οποιαδήποτε προσωπική πληροφορία μέσω διαδικτύου.

- **Baiting attacks:** Προσέξτε ποιους συνδέσμους ανοίγετε, που σας έχουν σταλεί σε κάποιο μήνυμα. Επίσης, αποφεύγετε να κάνετε κλικ σε ειδοποιήσεις ενημέρωσης λογισμικού και εφαρμογών χωρίς να σιγουρευτείτε ότι είναι εφαρμογές που έχετε εγκαταστήσει και το μήνυμα έρχεται από τις ίδιες και όχι από κάποια εξωτερική πηγή. Αποφύγετε κλικ σε δώρα και προσφορές, καθώς τα περισσότερα στοχεύουν σε απάτη.

- **Pretexting attacks:** Βεβαιωθείτε ότι τα άτομα και οι εργαζόμενοι, εκπαιδεύονται συνεχώς σχετικά με τις σχετικές επιθέσεις και πώς θα τις αντιμετωπίσουν. Επίσης, χρησιμοποιήστε το φίλτρο SPAM, ώστε να φιλτράρονται τα μηνύματα ηλεκτρονικού ταχυδρομείου και να ανιχνεύονται κακόβουλα λογισμικά. Ακόμα πρέπει να γίνει εγκατάσταση ενός web-filter με σκοπό τον αποκλεισμό επιβλαβών ιστότοπων.

- **Tailgating attacks:**

    1. Διατήρηση συνεχούς επιτήρησης και αυστηρής ασφάλειας στις εισόδους και τις εξόδους του οργανισμού-εταιρίας, ώστε να αποτραπεί είσοδος μη εργαζομένων αν δεν έχουν πάρει κάποια σχετική άδεια.

    2. Καθιέρωση των απαραίτητων πολιτικών ασφαλείας, διαδικασιών και μέτρων για την καθοδήγηση των εργαζομένων σε σωστό χειρισμό ευαίσθητων πληροφοριών και δεδομένων της εταιρείας.

    3. Επαλήθευση ταυτότητας οποιουδήποτε ύποπτου ατόμου και επαλήθευση δικαιώματος εξουσιοδότησης για τυχόν πρόσβαση σε κάποιο χώρο όπου το άτομο αυτό θέλει να εισέλθει.

    4. Αποσύνδεση από τον λογαριασμό σας, όταν δεν χρησιμοποιείτε τις συσκευές ή είστε μακριά τους.

- **Ransomware attacks:**
    1. Αντίγραφα ασφαλείας (backup) όλων των δεδομένων.
    2. Εκπαίδευση ατόμων σχετικά με αυτές τις επιθέσεις.
    3. Αποφυγή κλικ σε μη αξιόπιστους συνδέσμους και συνημμένα.
    4. Έλεγχος ταυτότητας δύο παραγόντων (Two factor authentication), για μεγαλύτερη ασφάλεια του λογαριασμού μιας και παρέχει πρόσθετα επίπεδα στην επαλήθευση ταυτότητας κατά τη σύνδεση και καθιστά πολύ πιο δύσκολο στον επιτιθέμενο να έχει πρόσβαση στο λογαριασμό

σας, ακόμα και αν έχει καταφέρει να έχει πρόσβαση στο Όνομα Χρήστη και τον Κωδικό σας.

- **Pop-up Windows:** Χρήση προγραμμάτων ασφαλείας και προστασίας καθώς και συνεχής αναβάθμισή τους (update). Αποφύγετε να κάνετε κλικ σε αναδυόμενα παράθυρα και ύποπτους ιστότοπους.

- **Scareware:** Αποφυγή κλικ σε ειδοποιήσεις ενημέρωσης λογισμικού ή εφαρμογών και ύποπτους ιστότοπους καθώς και χρήση προγραμμάτων ασφαλείας που ενημερώνονται συχνά.

- **Phone and Email scam attacks:** Έλεγχος πηγής τηλεφωνικών κλήσεων προτού απαντήσετε σε αυτές ενώ θα πρέπει να κάνετε ερωτήσεις για να επαληθεύσετε την ταυτότητα αυτού που σας καλεί. Επίσης, μην ανοίγετε μηνύματα ηλεκτρονικού ταχυδρομείου με συνημμένα, που αποστέλλονται από άγνωστα άτομα πριν τον έλεγχο τους.

- **Quid Pro Quo:** Συχνή αλλαγή κωδικών πρόσβασης στους λογαριασμούς σας. Ακόμα μην αποκαλύπτετε ποτέ πληροφορίες που σχετίζονται με εσάς ή με κάποιον από τους λογαριασμούς σας.

- **HTTPS Man-In-The-Middle:** Χρήση προγραμμάτων ασφαλείας για αποφυγή τοποθέτησης πλαστού πιστοποιητικού μέσα στο root πιστοποιητικό.

- **Multimedia Masquerading:** Έλεγχος πηγής λογισμικού πριν προβείτε σε κάποια εγκατάσταση ή εκτελέσετε κάποιο αρχείο. Καχυποψία σε σελίδες ιστού που ζητάνε προσωπικές πληροφορίες.

- **Peripheral Masquerading:** Προσοχή σε τι υλικό χρησιμοποιούμε και συνδέουμε στον υπολογιστή. Καχυποψία αν το υλικό προέρχεται από κάποιο άτομο και όχι από κάποια μεγάλη εταιρία σφραγισμένο σε κάποια συσκευασία.

- **Search Engine Poisoning (Spamdexing):** Χρήση μεγάλων και ασφαλών μηχανών αναζήτησης που είναι δύσκολο κάποιος επιτιθέμενος να μεταχειριστεί.

3. Μεθοδολογία Επισκόπησης

Οι βασικές ερωτήσεις που θέλει να απαντήσει η συγκεκριμένη διπλωματική εργασία είναι:

- Ποιά είναι τα κύρια είδη μεθόδων προστασίας από επιθέσεις κοινωνικής μηχανικής;

- Ποια είναι η αποτελεσματικότερη μέθοδος προστασίας από επιθέσεις κοινωνικής μηχανικής;

- Ποιος είναι ο μεγαλύτερος βαθμός προστασίας που μπορεί να επιτευχθεί από επιθέσεις κοινωνικής μηχανικής και τί πρέπει να γίνει ώστε να μεγιστοποιηθεί;

Στην παρούσα διπλωματική εργασία έγινε συστηματική μελέτη ανασκόπησης της βιβλιογραφίας με βάση κατάλληλη στρατηγική αναζήτησης για την αξιολόγηση των υφισταμένων μέτρων του κλάδου και των μεθόδων του (πολιτικές και εργαλεία) ως προς την αντιμετώπιση των απειλών και επιθέσεων κοινωνικής μηχανικής.

Η στρατηγική αναζήτησης που εφαρμόστηκε χρησιμοποίησε μοτίβα λέξεων-κλειδιών προκειμένου να αναζητηθεί σχετική βιβλιογραφία στο Google Scholar, Science Direct και στο Web of Science. Οι λέξεις-κλειδιά που χρησιμοποιήθηκαν είναι ”Social Engineering | Phishing | Impersonation & Tools | Software | Application | Attack Detection | Attack Addressing | Prevention”.

Εξετάστηκαν αρκετά άρθρα (περιοδικών ή συνεδρίων), καθώς και μελέτες που εμφανίστηκαν με τις παραπάνω λέξεις κλειδιά. Προκειμένου να διαχωριστούν και να φιλτραριστούν οι μελέτες, εφαρμόστηκαν κριτήρια επιλεξιμότητας (αποκλεισμού, συμπερίληψης και ποιότητας) της βιβλιογραφίας για να καθοριστεί αν θα συμπεριληφθεί ή θα αποκλειστεί το κάθε εντοπισμένο άρθρο από την μετέπειτα ανάλυση.

Τα κριτήρια αποκλεισμού ήταν τα εξής:

- Αποκλεισμός βιβλιογραφίας με ελλιπείς πληροφορίες, όπως άρθρα που στοχεύουν στον πολύ κόσμο (π.χ https://www.secnews.gr/180321/ti-einai-to-social-engineering-poies-oi-technikes-tou-kai-pos-na-prostatefteite/).

- Αποκλεισμός βιβλιογραφίας με διαθέσιμη μόνο περίληψη ή με περιορισμένη πρόσβαση για ανάγνωση του περιεχομένου.

- Αποκλεισμός βιβλιογραφίας δημοσιευμένης σε γλώσσα εκτός Ελληνικών ή Αγγλικών.

- Αποκλεισμός βιβλιογραφίας δημοσιευμένης πριν το 2008.

Τα κριτήρια συμπερίληψης ήταν τα εξής:

- Βιβλιογραφία που μπορεί να περιλαμβάνει ερευνητικές εργασίες και διδακτορικές διατριβές.

- Βιβλιογραφία που παρουσιάζει εργαλεία και μεθοδολογίες ως λύσεις που χρησιμοποιούνται γενικά από εταιρείες και οργανισμούς. (Δημοφιλείς και πολυχρησιμοποιημένες λύσεις).

- Βιβλιογραφία που επικεντρώνεται σε εργαλεία, μεθοδολογίες, στρατηγικές και λύσεις κοινωνικής μηχανικής.

Τα κριτήρια ποιότητας ήταν:

- Βιβλιογραφία με δύσκολη στην ανάγνωση περιεχόμενο.

- Βιβλιογραφία με μη παροχή ουσιαστικής λύσης.

- Βιβλιογραφία με κανένα είδος αποτίμησης/αξιολόγησης (review/assessment) της συνεισφοράς της.

- Βιβλιογραφία με μη επαληθεύσιμα αποτελέσματα.

Τα άρθρα που αρχικά εντοπίστηκαν ήταν 367 και τα άρθρα που διατηρήθηκαν και χρησιμοποιήθηκαν στην Διπλωματική εργασία 170.

4. Ανάλυση

Στο κεφάλαιο αυτό θα αναλύσουμε κάποιες μεθοδολογίες και τρόπους προστασίας από επιθέσεις κοινωνικής μηχανικής σε μεγαλύτερο βάθος. Επίσης, θα ταξινομήσουμε τρόπους αντιμετώπισης τέτοιων επιθέσεων με βάση κάποιες βασικές διαστάσεις.

4.1 Μεθοδολογίες και τρόποι προστασίας από επιθέσεις Κοινωνικής Μηχανικής

Η κατηγοριοποίηση των λύσεων προέρχεται από το [170]. Η κατηγοριοποίηση αυτή περιλαμβάνει τις εξής βασικές κατηγορίες.

- Πολιτική και έλεγχος διαδικασιών
- Εκπαίδευση
- Τεχνικοί τρόποι Προστασίας

Πολιτική και έλεγχος διαδικασιών

Η πολιτική και η διαδικασία παρέχουν ιεραρχικό έλεγχο μέσω κάποιων πλαισίων διαχείρισης και διαδικασιών, σε αντίθεση με τεχνικά συστήματα, όπως ένα λογισμικό ασφαλείας. Έχουν σχεδιαστεί για τη μείωση της έκθεσης σε επιθέσεις κοινωνικής μηχανικής και η καλά συντηρημένη πολιτική και οι οργανωτικές διαδικασίες βοηθούν στο μετριασμό της εμφάνισης μιας πιθανής εκμετάλλευσης χωρίς να στηρίζεται στις τεχνικές δυνατότητες των χρηστών του συστήματος.

Ο έλεγχος της πολιτικής και των διαδικασιών ορίζεται γύρω από το επιχειρησιακό περιβάλλον και το περιβάλλον των χρηστών. Ωστόσο, τα πλαίσια ασφαλείας που έχουν εισαχθεί για την αντιμετώπιση των επιθέσεων, έχουν προστεθεί ως πρόσθετα στοιχεία στην ευρύτερη τεχνική απειλή, και όχι στην στρατηγική ανάπτυξη της πολιτικής και του ελέγχου διαδικασιών. Η πολιτική πρέπει να είναι εγγενώς δομημένη με τη διαχείριση των ανθρώπων και να ενσωματωθεί στο πυρήνα όλων των πληροφοριακών συστημάτων.

Το βιβλίο που έχει δημοσιευθεί από τον Peltier [95], προσφέρει μία ολοκληρωμένη συγκέντρωση στρατηγικών για την εφαρμογή ασφαλούς πολιτικής και διαδικασίας για την διακυβέρνηση των πληροφοριακών συστημάτων, συνδυάζοντας πρότυπα και βέλτιστες πρακτικές κατευθυντήριες γραμμές για την ενσωμάτωση της ασφάλειας πληροφοριακών

συστημάτων σε όλους τους οργανωτικούς τομείς (από τους ρόλους και τις αρμοδιότητες των χρηστών έως τα ενσωματωμένα εργαλεία ασφαλείας).

Οι Frauenstein και Solms [96] έχουν αναπτύξει ένα ολιστικό πλαίσιο με επίκεντρο τον χρήστη, το οποίο στοχεύει ειδικά στις επιθέσεις phishing. Το πλαίσιο περιλαμβάνει ένα επίπεδο χρήστη, "ανθρώπινο τείχος προστασίας", στο πλαίσιο ασφάλειας πληροφοριών, επιτρέποντας τη διασταυρούμενη επικοινωνία μεταξύ των διαστάσεων της άμυνας (χρήστης, οργανωτική πολιτική και τεχνικοί έλεγχοι) και συνδυάζοντας μηχανισμούς όπως η ευαισθητοποίηση και η εκπαίδευση των χρηστών με την πολιτική και τα συνδεδεμένα τεχνολογικά εργαλεία.

Η πολιτική και οι διαδικασίες πρέπει να είναι ευέλικτες σε άγνωστες και απρόβλεπτες επιθέσεις και, ως εκ τούτου, κατάλληλες για το μεταβαλλόμενο τοπίο απειλών. Οι σταθερές κατευθυντήριες γραμμές μπορούν γρήγορα να ξεπεραστούν καθώς νέες μέθοδοι επιθέσεων αναπτύσσονται συνεχώς. Ως προσωπικό εργαλείο, η πολιτική και οι διαδικασία μπορεί να είναι αποτελεσματικά, για αυτό το λόγο οι κανόνες και οι κανονισμοί που έχουν σχεδιαστεί για την επιβολή και τη διατήρηση ασφάλειας είναι πολύ μεγάλοι σε έκταση, σχεδόν εντελώς άσχετοι με τον χρήστη και τη χρήση του σε συστήματα υπολογιστών και το Διαδίκτυο.

Εκπαίδευση

Μιας και οι επιθέσεις στοχεύουν τους ανθρώπους- χρήστες του συστήματος και όχι το σύστημα καθαυτό, οι επιθέσεις μπορούν να μειωθούν σε μεγάλο βαθμό με την κατάλληλη εκπαίδευση και ευαισθητοποίηση των χρηστών του συστήματος. Η εκπαίδευση αποτελεί βασικό στοιχείο του μοντέλου άμυνας και συγκεκριμένα σε επιθέσεις κοινωνικής μηχανικής αφορά διαδικασίες που οι τεχνικοί μηχανισμοί ασφαλείας έχουν αποδειχθεί ανεπαρκής.

Η διαδραστική εκπαίδευση όπως κουίζ, τεστ και παιχνίδια έχει [97-98] αποδειχθεί αρκετά αποτελεσματική. Επίσης τα δεδομένα τα οποία συλλέγονται από αυτές τις διαδικασίες είναι αρκετά χρήσιμα μιας και μπορεί να «μετρηθεί» η αποτελεσματικότητα μιας επίθεσης και να αλλάξει ανάλογα η πολιτική και ο έλεγχος διαδικασιών.

Η δουλειά του Lin [99] διεξήγαγε έρευνες για τη μέτρηση της αποτελεσματικότητας των συστημάτων ευαισθητοποίησης, που είναι ενσωματωμένα σε ένα σύστημα ασφαλείας, τα οποία προειδοποιούσαν τους χρήστες για ενδεχόμενα phishing attacks ή μη ασφαλείς συνδέσεις (HTTP και όχι HTTPS) χωρίς την απαραίτητη κρυπτογράφηση.

Παρομοίως μία πιο πρόσφατη έρευνα των Lee et al. [100], έδειξε ότι οι εικόνες ασφαλείας που προορίζονται να παρέχουν οπτική επιβεβαίωση της αυθεντικότητας του ιστότοπου στις τραπεζικές συναλλαγές μέσω διαδικτύου αποδείχθηκαν αναξιόπιστες, καθώς το 74% των χρηστών που συμμετείχε στην μελέτη εισήγαγε τον κωδικό πρόσβασης τους μετά την αφαίρεση των εικόνων ασφαλείας.

Άλλες έρευνες έχουν προτείνει συστήματα εκπαίδευσης «ενδιάμεσου λογισμικού» για την πρόσβαση στο διαδίκτυο [101]. Το ενδιάμεσο λογισμικό είναι μία πύλη εκπαίδευσης, υπεύθυνη για τη διαμεσολάβηση μιας σύνδεσης μεταξύ του χρήστη και του διαδικτύου. Μόνο μετά την επιτυχής ολοκλήρωση του εκπαιδευτικού κουίζ θα μπορούν οι χρήστες να έχουν πρόσβαση σε απομακρυσμένους πόρους του διαδικτύου.

Οι Stewart et al. [102] και πιο πρόσφατα οι Konak και Bartolacci [103] έχουν αναπτύξει εκπαιδευτικές πλατφόρμες με σκοπό την καλύτερη κατανόηση της δυναμικής των απειλών της κοινωνικής μηχανικής.

Μία μελέτη που διεξήχθη από τους Anderson et al. [104] χρησιμοποίησε νευρολογική σάρωση για να καθορίσει τον τρόπο που συμβαίνει μια αυτοματοποιημένη αντίδραση με βάση ένα επαναλαμβανόμενο θέμα στον ανθρώπινο εγκέφαλο, όταν παρουσιάζονται στους χρήστες κοινές στατικές προειδοποιήσεις ασφαλείας, όπως η προαναφερθέντα στην έρευνα των Lee et al. [105]. Η έρευνα αναλύει την συμπεριφορά του εγκεφάλου μέσω μαγνητικών απεικόνισης, όπου οι χρήστες αποδείχθηκε ότι παρουσίασαν πτώση στα τμήματα οπτικής επεξεργασίας του εγκεφάλου μετά από πολλαπλές εκθέσεις σε μια παρόμοια προειδοποίηση ασφαλείας. Ως λύση οι ερευνητές προτείνουν προειδοποιήσεις ασφαλείας που είναι πολυμορφικές, δηλαδή αλλάζουν συνεχώς εμφάνιση και συμπεριφορά κάθε φορά που δημιουργούνται (π.χ. διαφορετικό χρώμα, κείμενο, θέση, μέγεθος). Αυτό αποδείχθηκε ότι μειώνει τις αυτόματες αντιδράσεις των χρηστών και ήταν αποτελεσματικότερο.

Τεχνικοί Τρόποι Προστασίας

- **Μηχανισμοί Sandboxing:** Το Sandboxing είναι η διαδικασία δημιουργίας ενός απομονωμένου περιβάλλοντος υπολογιστή, συνήθως μέσω εικονικοποίησης, για τη δοκιμή μη αξιόπιστων λειτουργιών. Έχει εφαρμοστεί αποτελεσματικά ως λύση ασφάλειας σε διάφορους τομείς της πληροφορικής, από συγκεκριμένες πλατφόρμες κώδικα έως συστήματα περιήγησης [106], καθώς και στο τομέα της ασφάλειας των smartphone για τη βελτίωση της άμυνας έναντι κακόβουλου λογισμικού [107].

Ένα περιβάλλον sandboxing μπορεί να ελέγχει και να αναλύει τα χαρακτηριστικά που σχετίζονται με μία επίθεση, κατηγοριοποιώντας τη συμπεριφορά για την ανίχνευση ανωμαλιών [108]. Περαιτέρω μεθοδολογίες sandboxing μπορούν να παράσχουν έναν μηχανισμό ελέγχου της ακεραιότητας [109]. Αυτή η προσέγγιση βασίζεται σε μεγάλο βαθμό στην ακρίβεια των αλγορίθμων ανάλυσης συμπεριφοράς για την ανίχνευση ανωμαλιών και εισάγει το ζήτημα της αυξημένης ζήτησης πόρων. Μια τεχνική επίθεση που χρησιμοποιεί οπτικές μεθόδους εξαπάτησης αντί τεχνικών μηχανισμών για την εξαπάτηση των χρηστών θα χρειαζόταν να γίνει κατανοητή από το sandbox environment, αλλά αυτό είναι ένα πολύ σύνθετο έργο που απαιτεί δειγματοληψία.

Τα προγράμματα περιήγησης στον ιστό [110-111] έχουν εφαρμόσει τεχνολογίες sandbox για να αποτρέψουν τους ιστότοπους από το να χειραγωγούν τις δικές τους οπτικές και συμπεριφορικές ιδιότητες, όπως η γραμμή διευθύνσεων URL, οι καρτέλες του προγράμματος περιήγησης και οι δείκτες ασφαλείας. Ωστόσο, αυτοί οι μηχανισμοί εξακολουθούν να μην είναι σε θέση να αποτρέψουν την εκτέλεση των επιθέσεων παραποίησης, όπως αυτές που δημιουργούν μια ψεύτικη γραμμή διευθύνσεων χρησιμοποιώντας κώδικα που φαίνεται νόμιμος στο sandbox καθώς και είναι οπτικά παραπλανητικός για τον χρήστη. Αντ' αυτού, η περιορισμένη έρευνα που έχει διεξαχθεί στην εφαρμογή του sandboxing για επιθέσεις κοινωνικής μηχανικής επικεντρώνεται στην αξία της ως εργαλείο μάθησης.

Τα λειτουργικά συστήματα Windows και Linux εφαρμόζουν διάφορους μηχανισμούς sandbox για να αποτρέψουν την πραγματοποίηση μη εξουσιοδοτημένων αλλαγών σε ένα σύστημα- λειτουργώντας ως μια καλή γραμμή άμυνας ακόμη και όταν ένας χρήστης μπορεί να έχει εξαπατηθεί εν αγνοία του από μια επίθεση κοινωνικής μηχανικής. Ένα τέτοιο παράδειγμα είναι η secure-attention-sequence, ή secure-attention-key, όπως είναι πιο γνωστό, η οποία είναι μια ασφαλής μέθοδος για την αποτροπή των εφαρμογών που πλαστογραφούν ένα λειτουργικό σύστημα ή την διαδικασία σύνδεσης στο λειτουργικό σύστημα. Το σύστημα σύνδεσης προστατεύεται από πιθανές απειλές με τη χρήση του πυρήνα του λειτουργικού συστήματος για την αναστολή όλων των διεργασιών που εκτελούνται πριν από την έναρξή του. Το σύστημα ελέγχου πρόσβασης χρηστών (UAC) είναι

ένας άλλος μηχανισμός sandbox των Windows που χρησιμοποιεί την απομόνωση προνομίων διεπαφής χρήστη (UIPI) για να αποτρέπει τις εφαρμογές από μη εξουσιοδοτημένες αλλαγές προνομίων κατά τη διάρκεια της εκτέλεσης. Αυτό επιτυγχάνεται με την απομόνωση διεργασιών που επισημαίνονται με χαμηλότερο επίπεδο ακεραιότητας (π.χ. μη υπογεγραμμένη, μη αξιόπιστη εφαρμογή τρίτου μέρους) από την αποστολή μηνυμάτων σε διεργασίες υψηλότερου επιπέδου ακεραιότητας (π.χ. για να ζητήσουν μια ενέργεια που μπορεί να απαιτεί πρόσβαση στο root system), έως ότου το επιτρέψει ένας εξουσιοδοτημένος χρήστης στο UAC ή ένα έγκυρο πιστοποιητικό από εγκεκριμένη αρχή υπογραφής κώδικα [112].

Αυτή η προσέγγιση μπορεί σίγουρα να περιορίσει τον αντίκτυπο ορισμένων επιθέσεων κοινωνικής μηχανικής, αλλά είναι συγκεκριμένη στο λειτουργικό σύστημα και δεν μπορεί να παρατηρήσει τις αιτήσεις που γίνονται σε ένα απομακρυσμένο σύστημα στο οποίο δεν έχει κανέναν έλεγχο.

Ένα σύστημα με τίτλο "BLADE" (block all drive-by exploits) που αναπτύχθηκε από τους Lu et al. [113] παρέχει έναν μηχανισμό sandboxing για τον μετριασμό των εκμεταλλεύσεων drive-by download. Το σύστημα BLADE εισάγει μια ανεξάρτητη από το πρόγραμμα περιήγησης επέκταση του πυρήνα του λειτουργικού συστήματος που επιβάλλει έναν κανόνα ότι τα εκτελέσιμα αρχεία πρέπει να προέρχονται από μια ρητή ενέργεια του χρήστη που παρέχει τη συγκατάθεσή του. Στην περίπτωση μιας drive-by λήψης, όπου συνήθως δεν υπάρχει ρητή ενέργεια του χρήστη το σύστημα ανακατευθύνει σε μια ασφαλή τοποθεσία σε sandbox environment στο δίσκο του συστήματος του χρήστη, όπου αποτρέπεται η εκτέλεση. Αυτή η προσέγγιση δεν εξαρτάται από τον μηχανισμό εξαπάτησης ή απόκρυψης, καθώς ανταποκρίνεται μόνο όταν πραγματοποιείται μια λήψη χωρίς συγκεκριμένη συγκατάθεση του χρήστη. Σε δοκιμές, το BLADE μπόρεσε να αποκλείσει όλες τις drive-by επιθέσεις λήψης, σε ένα δείγμα πάνω από 1.900 ενεργών κακόβουλων διευθύνσεων URL με ελάχιστη επίδραση στην απόδοση του συστήματος.

Πρόσφατα, οι Bianchi et al. [114] ανέπτυξαν ένα εργαλείο για το λειτουργικό σύστημα Android, το οποίο έχει σχεδιαστεί για να αποτρέπει την εκτέλεση κακόβουλων εφαρμογών που εκτελούν αισθητικές (DV1), συμπεριφορικές (DV2) ή και τα δύο είδη παραπλανήσεων (DV3). Αυτό μπορεί να περιλαμβάνει την παραποίηση της εμφάνισης μιας εφαρμογής ή τη δημιουργία παραπλανητικών οπτικών ενδείξεων πάνω σε νόμιμες εφαρμογές με καταγραφή και ανάλυση κλήσεων διεπαφής προγράμματος εφαρμογής (API) στο Android GUI. Οι ερευνητές χρησιμοποιούν μια μέθοδο που ονομάζεται "στατική ανάλυση κώδικα", η οποία ανακρίνει τα bytecode της εφαρμογής, με βάση ένα σύστημα που προτείνεται από τους Egele et al. [115], κατά τη διάρκεια του χρόνου εκτέλεσης για να ανιχνεύσει τις αιτήσεις που γίνονται σε συγκεκριμένα API του Android, οι οποίες μπορούν να επιτρέψουν την εκτέλεση οπτικά παραπλανητικών στοιχείων στη γραφική διεπαφή ενός συστήματος Android. Ενώ η προσέγγιση αποδείχθηκε επιτυχής στον εντοπισμό εφαρμογών που χρησιμοποιούν αυτά τα API για την εκτέλεση οπτικής παραπλάνησης, οι ερευνητές διαπίστωσαν ότι υπήρχε ένα σημαντικό ποσοστό ψευδώς θετικών αποτελεσμάτων. Κρίσιμο είναι ότι, σε μια αξιολόγηση στην οποία συμμετείχαν 308 χρήστες, η προσέγγιση αυτή αποδείχθηκε ότι βελτιώνει σημαντικά την ικανότητα ενός χρήστη στον εντοπισμό κακόβουλων εφαρμογών.

- **Authorization, Authentication and Accounting (AAA):** είναι ένα πλαίσιο για την έξυπνη πρόσβαση σε υπολογιστές, την επιβολή πολιτικών, τον έλεγχο της χρήσης και την παροχή των απαραίτητων πληροφοριών για την τιμολόγηση υπηρεσιών. Εφαρμόζεται συνήθως σε ελεγχόμενα περιβάλλοντα, ιδίως όπου υπάρχει ένα ποικίλο τοπίο χρηστών που θέτει σε κίνδυνο τον έλεγχο και την προστασία των δεδομένων. Το πλαίσιο παρέχει ελέγχους

για την πρόσβαση σε πόρους (Authentication), την επιβολή οργανωτικής πολιτικής (Authorization) και τον έλεγχο της χρήσης των πόρων (δηλ. χρόνοι σύνδεσης, διάρκεια συνεδρίας, συσκευές στις οποίες έγινε πρόσβαση (Accounting)). Η χρήση του αποσκοπεί στη διασφάλιση ότι οι οργανισμοί έχουν ένα λεπτομερές επίπεδο διασφάλισης και ελέγχου σχετικά με το ποιος έχει πρόσβαση σε ένα σύστημα, με βάση δεδομένα σχετικά με ονόματα, ρόλους, σύνολα δεξιοτήτων κλπ. Σε μεγάλες εγκαταστάσεις, όπως τομείς και δίκτυα οργανισμών, η διαχείριση της AAA μπορεί να γίνει αποτελεσματικά μέσω εποπτείας, εξασφαλίζοντας εκτεταμένη ταυτοποίηση και εξουσιοδοτημένη πρόσβαση για τη λειτουργία εκτελέσιμων προγραμμάτων, πρόσβαση σε κοινόχρηστους φακέλους κ.ο.κ. Ωστόσο, σε ένα περιβάλλον οικιακού χρήστη, η AAA είναι λιγότερο πρακτική, επειδή η κεντρική εποπτική αρχή παραμένει στον χρήστη και όχι σε κάποιο ειδικό κυβερνοασφάλειας (cyber security specialist) [116].

Οι δυνατότητες του ατόμου απαιτούν τη σωστή χρήση των "ελάχιστων δικαιωμάτων χρήστη" και προνομίων διαχειριστή, όπου χρειάζεται, για την προστασία του συστήματος και την ασφάλεια των δεδομένων. Πολλά τεχνικά, αυτόνομα, βασισμένα σε συνεδρίες πρωτόκολλα, όπως το Kerberos [114], έχουν συνδυαστεί με επιτυχία με ασφαλείς κεντρικές πλατφόρμες ελέγχου ταυτότητας και εξουσιοδότησης [118]. Από την άλλη πλευρά, οι πλατφόρμες ασφαλείας, όπως το User Access Control της Microsoft, όχι μόνο υποδεικνύουν στους χρήστες ότι μια ενέργεια μπορεί να απαιτεί πιο προνομιακά δικαιώματα, αλλά εμφανίζεται στην πραγματικότητα όταν ένα πρόγραμμα προσπαθεί να αποκτήσει αυτά τα δικαιώματα για να εκτελεστεί και ζητά από τους χρήστες να ταυτοποιηθούν εκ των προτέρων [119]. Στη συνέχεια, μπορεί κανείς να λάβει τεκμηριωμένη απόφαση. Ωστόσο, το σύστημα εξακολουθεί να βασίζεται στη δέουσα επιμέλεια του χρήστη για τη σωστή επιλογή των επιλογών ελέγχου πρόσβασης χρήσης.

Η πρόοδος στη διαχείριση ταυτότητας, όπως το Forefront Identity Manager της Microsoft [120], έχει οδηγήσει στην ανάπτυξη ολοκληρωμένων λύσεων ασφάλειας AAA που έχουν ένα ευρύ εύρος εφαρμογών με αποτέλεσμα την αποτροπή επιθέσεων κοινωνικής μηχανικής. Ωστόσο, παρόλο που παρέχουν ευέλικτο έλεγχο των χρηστών και πρόσβαση σε πόρους, τα συστήματα διαχείρισης ταυτότητας δεν είναι δυναμικά ως προς τους μηχανισμούς μετριασμού των επιθέσεων αυτού του είδους.

- **Monitoring:** Σε αυτό το πλαίσιο, με τον όρο monitoring (παρακολούθηση), αναφερόμαστε συγκεκριμένα στην παρατήρηση της συμπεριφοράς του υπολογιστικού συστήματος, που δημιουργείται από ενέργειες του χρήστη/προγραμματιστή, μέσω συλλογής, μηχανισμούς συγκέντρωσης και ανάλυσης. Το Monitoring είναι ένας βασικός μηχανισμός ασφάλειας για επιθέσεις κοινωνικής μηχανικής, καθώς οι νέες επιθέσεις μπορούν να ταξινομηθούν μέσω της έκθεσης και του αποτελεσματικού ελέγχου ασφαλείας όπου εισάγονται μέσω της κατανόησης των ενεργειών του χρήστη που θέτουν σε κίνδυνο το σύστημα ή τον ίδιο τον χρήστη. Η παρακολούθηση ιστορικού απαιτούσε έντονη εποπτεία και εγκληματολογική έρευνα, ενώ οι νέες εξελίξεις επιτρέπουν τον εντοπισμό και την αντιμετώπιση επιθέσεων μέσω της χρήσης Honeypots [121]. Αυτά τα εικονικοποιημένα περιβάλλοντα επιτρέπουν στους επαγγελματίες της ασφάλειας να αναλύουν τις επιθέσεις, να ταξινομούν τα χαρακτηριστικά των επιθέσεων και να εντοπίζουν τις συμπεριφορές του χρήστη που μπορεί να οδηγήσουν σε παραβίαση. Η έρευνα που διεξάγεται στον τομέα της στατιστικής παρακολούθησης επικεντρώνεται στην ανάπτυξη σημασιολογικής αποφυγής μεταμφιεσμένων επιθέσεων με τη μοντελοποίηση της συμπεριφοράς αναζήτησης των χρηστών. Για παράδειγμα, οι Salem και Stolfo [122] εισήγαγαν μια νέα μέθοδο παρακολούθησης και μηχανικής

μάθησης προκειμένου να εντοπίσουν την ανώμαλη συμπεριφορά των χρηστών σε ένα σύστημα.

Ένα πολλά υποσχόμενο κομμάτι έρευνας σε αυτόν τον τομέα [123] αφορά την ανίχνευση κακόβουλων ιστότοπων με την παρακολούθηση της διαδρομής ανακατεύθυνσης που ακολουθείται για την επίτευξη ενός διαδικτυακού προορισμού, αντί της ανάλυσης του φυσικού ιστότοπου.

Οι Lu et al. [124] εισήγαγαν το "SURF", μια επέκταση του προγράμματος περιήγησης που έχει σχεδιαστεί για την ανίχνευση δηλητηρίασης μηχανών αναζήτησης με την παρακολούθηση των συνόδων χρήστη "αναζήτηση-μετά-επίσκεψη" και την ταξινόμηση των ανακατευθύνσεων σε νόμιμες ή κακόβουλες. Το συγκεκριμένο σύστημα ήταν σε θέση να επιτύχει με 99% ακρίβεια στην εμπειρική αξιολόγηση στον πραγματικό κόσμο.

Οι Li et al. [125] χαρτογράφησαν τοπολογικά ειδικούς κεντρικούς υπολογιστές σε κακόβουλες διαδικτυακές υποδομές, που εμπλέκονται στην ενορχήστρωση μονοπατιών ανακατεύθυνσης προς κακόβουλους ιστοτόπους. Η έρευνα χρησιμοποιεί έναν αλγόριθμο PageRank για την παρακολούθηση και τη σύλληψη υποδοχέων επίθεσης που είναι υπεύθυνοι για τη διαμεσολάβηση και την ανταλλαγή κίνησης για κακόβουλες δραστηριότητες. Η παρούσα μελέτη επικεντρώνεται στην επικάλυψη διαχείρισης που εφαρμόζουν οι εγκληματίες του κυβερνοχώρου για τη διεξαγωγή και τον έλεγχο της διανομής κακόβουλου λογισμικού μεγάλης κλίμακας. Στο πλαίσιο της μελέτης για τον εντοπισμό των αλυσίδων ανακατεύθυνσης, οι ερευνητές χρησιμοποίησαν ένα σύστημα που ονομάζεται "WarningBird" [126], το οποίο είναι ένα σύστημα URL ανίχνευσης για το Twitter, το οποίο έχει σχεδιαστεί για να προστατεύει από την υπό όρους ανακατεύθυνση, όπου ο επιτιθέμενος ανακατευθύνει τους ανιχνευτές ιστού (web crawler) σε μια νόμιμη τοποθεσία προορισμού αλλά τους χρήστες σε μία κακόβουλη τοποθεσία. Σε συνδυασμό με έναν ταξινομητή πραγματικού χρόνου, το WarningBird παρουσιάζει υψηλό βαθμό ακρίβειας έναντι μεγάλων δειγμάτων Tweets.

- **Integrity Checking:** Η ακεραιότητα των χρηστών και των εφαρμογών/δεδομένων είναι δύσκολο να διασφαλιστεί χωρίς απόδειξη ή ανάλυση, ιδίως εάν τα δεδομένα και τα επακόλουθα χαρακτηριστικά συμπεριφοράς προέρχονται από ένα εξωτερικό δίκτυο, συνάδελφο ή φίλο. Ο έλεγχος ακεραιότητας παρέχει στον χρήστη μια οπτική απόκριση και τεχνική διαβεβαίωση ως προς το αν το αρχείο, ο ιστότοπος, ή τα δεδομένα πρέπει να είναι αξιόπιστα [127-128]. Συστήματα, όπως το TripWire, υλοποιούν έλεγχο ακεραιότητας αρχείων που παρακολουθεί συνεχώς την ακεραιότητα των δεδομένων (π.χ. ενός αρχείου με βάση τις ιδιότητες και τη συμπεριφορά του), ενώ περαιτέρω εργασίες έχουν δείξει πώς αυτός ο αμυντικός μηχανισμός μπορεί να γίνει και δυναμικός [129].

Μια μελέτη που διεξήχθη από τους Dhanalakshmi και Chellappan [130] αξιολογεί μεθόδους για την ταξινόμηση των τύπων αρχείων, με σκοπό τον εντοπισμό και την ανίχνευση επιθέσεων μεταμφίεσης αρχείων. Η έρευνα αξιολογεί αλγορίθμους ανίχνευσης τύπου αρχείου, όπως η επικεφαλίδα του αρχείου (FHT), η ανάλυση πολλαπλών διακρίσεων (MDA) και η ανίχνευση σύνθετων αρχείων (CFD), για τον εντοπισμό του πραγματικού τύπου αρχείου. Χρησιμοποιώντας αυτή την προσέγγιση, οι ερευνητές είναι σε θέση να ταξινομήσουν θραύσματα ενός αρχείου ή ενσωματωμένα αρχεία για να προσδιορίσουν τον τύπο αρχείου, παρέχοντας έτσι ένα βασικό μηχανισμό μετριασμού για κοινές επιθέσεις μεταμφίεσης αρχείων.

Η έρευνα που διεξήχθη από τους Hara et al. 2009 [131] δείχνει πώς η χρήση απλών μεθόδων για τον προσδιορισμό της ακεραιότητας, με τον έλεγχο και τη σύγκριση οπτικής ομοιότητας με βάση το phishing, μπορεί να παρέχει ακριβή αποτελέσματα. Οι ερευνητές αποδεικνύουν ότι με την ανάλυση της ομοιότητας μεταξύ οπτικών συστατικών σε έναν ιστότοπο, χωρίς προηγούμενη γνώση ότι πρόκειται για πλατφόρμα επίθεσης, μπορούν να είναι σε θέση να προσδιορίσουν αυτόματα συγκεκριμένη διάταξη των ιστότοπων που έχουν παραποιηθεί από ιστοσελίδες phishing.

Πρόσφατες μελέτες έχουν προτείνει λύσεις για την ανίχνευση της ακεραιότητας των διευθύνσεων URL [132] μέσω αλγορίθμων βελτιστοποίησης και τεχνικών ανίχνευσης για τον μετριασμό των επιθέσεων. Τα εργαλεία ακεραιότητας έχουν επεκταθεί για την παροχή προστασίας από τα σημεία πρόσβασης WiFi "Evil Twin", αναλύοντας βασικά στοιχεία που σχετίζονται με την εμφάνιση αντιστοιχίσεων SSID-mac, καθώς και τους αλγόριθμους κατακερματισμού και κρυπτογράφησης που χρησιμοποιούνται [133].

- **Μηχανική Μάθηση (Machine Learning):** Η έρευνα έχει αποδείξει πώς η ανίχνευση κακόβουλου λογισμικού μέσω μηχανικής μάθησης μπορεί να είναι δυναμική, όπου κατάλληλοι αλγόριθμοι, όπως οι k-κοντινότεροι γείτονες (K-nearest neighbors), η μάθηση μέσω δέντρων απόφασης, οι μηχανές διανυσμάτων υποστήριξης καθώς και τα Μπεϋζιανά και νευρωνικά δίκτυα μπορούν να εφαρμοστούν για τη δημιουργία προφίλ αρχείων έναντι γνωστών και πιθανών εκμεταλλεύσεων και να διακρίνουν μεταξύ νόμιμων και παράνομων δεδομένων [134]. Οι αλγόριθμοι μηχανικής μάθησης έχουν εφαρμοστεί με επιτυχία για την ανίχνευση κακόβουλων ηλεκτρονικών μηνυμάτων, χρησιμοποιώντας τεχνικές ταξινόμησης ανωμαλιών, οπότε καταδεικνύουν τη δυνατότητα και την ικανότητα για περαιτέρω εφαρμογή σε άλλους συναφείς τομείς επιθέσεων κοινωνικής μηχανικής [135].

Η έρευνα που διεξάγεται για την αντιμετώπιση στις απειλές της κοινωνικής μηχανικής ειδικότερα, έχει δει την εφαρμογή νευρωνικών δικτύων για την πρόβλεψη της νομιμότητας και μη τηλεφωνικών κλήσεων. Χρησιμοποιώντας ένα κατασκευασμένο σύνολο δεδομένων, οι προσέγγιση έδειξε θετική ακρίβεια στις αποκρίσεις πρόβλεψης και, ως εκ τούτου, η μεθοδολογία ταξινόμησης μπορεί να είναι κατάλληλη για ένα ευρύτερο φάσμα τύπων επιθέσεων κοινωνικής μηχανικής. Ωστόσο, αξιόπιστα σύνολα δεδομένων που παράγονται από δεδομένα του πραγματικού κόσμου μπορεί να είναι απαραίτητα καθώς και η ικανότητα τυπικής ενσωμάτωσης των διαδικασιών απόφασης του αλγορίθμου σε κατάλληλα συστήματα ελέγχου πρόσβασης [136].

Μια προσπάθεια για τη δημιουργία αξιόπιστων ταξινομητών για τη διδασκαλία αλγορίθμων μηχανικής μάθησης έχει προταθεί στους Lee et al. [121], όπου ειδικά διαμορφωμένα honeypots παρέχουν χαρακτηριστικά που μπορούν να υποδείξουν και να βοηθήσουν στο φιλτράρισμα των spammer των κοινωνικών δικτύων. Η εφαρμογή της μηχανικής μάθησης για την άμυνα κατά των επιθέσεων κοινωνικής μηχανικής μπορεί να είναι στην εκπαίδευση μηχανών διανυσμάτων υποστήριξης με χρήση δυαδικών χαρακτηριστικών για την ανίχνευση πρώιμων μηνυμάτων spam. Ένας σημαντικός όγκος έρευνας που χρησιμοποιεί αλγορίθμους μηχανικής μάθησης έχει επικεντρωθεί στη δημιουργία αποτελεσματικών μηχανισμών ανίχνευσης για επιθέσεις phishing στο διαδίκτυο και στο ηλεκτρονικό ταχυδρομείο, όπως η εργασία των Basnet et al. [137] όπου δοκιμάζουν μια σειρά από τεχνικές ομαδοποίησης και φιλτραρίσματος για την ταξινόμηση δεδομένων για ακριβής πρόβλεψη της ανίχνευσης επιθέσεων.

Οι Bergholz et al. [138] ανέπτυξαν ένα σύστημα για την ανίχνευση ηλεκτρονικών μηνυμάτων phishing με βάση τα συστατικά χαρακτηριστικά του ηλεκτρονικού ταχυδρομείου, όπως το σώμα του κειμένου, η διεύθυνση του αποστολέα, ή ενσωματωμένες εικόνες, χρησιμοποιώντας συνδυασμούς μηχανικής μάθησης για ταξινόμηση καθώς και ταξινόμηση μηχανισμών μοντελοποίησης για φιλτράρισμα. Οι Bergholz et al. [139] συνέχισαν αυτή την έρευνα αποδεικνύοντας την αποτελεσματικότητα και τη λειτουργικότητα μιας σειράς προσεγγίσεων φιλτραρίσματος που, αν συνδυαστούν, μπορούν να παρέχουν αποτελεσματική ανίχνευση για αυτόν τον φορέα επίθεσης.

Η πρόοδος στην εφαρμογή της τεχνολογίας οντολογικής σημασιολογίας έχει δει εξελίξεις που αποδεικνύουν τη δυνατότητα εφαρμογής και σε επιθέσεις κοινωνικής μηχανικής, χρησιμοποιώντας ένα λογισμικό τεχνητής νοημοσύνης με ένα σύνολο δεδομένων που βασίζεται σε εκφράσεις ως μηχανισμούς πρόβλεψης για τη δημιουργία συμπερασμάτων σχετικά με δεδομένα κειμένου [140]. Αυτός ο τύπος μηχανισμού ανίχνευσης αποτελεί χρήσιμη εφαρμογή κατά των επιθέσεων phishing που παρατηρούνται σε μηνύματα ηλεκτρονικού ταχυδρομείου και ιστότοπους. Ενώ αυτή η τεχνολογία δεν ταξινομείται αυστηρά ως μηχανική μάθηση όσον αφορά τους αλγορίθμους που χρησιμοποιούνται, υπάρχει μια διαδικασία μάθησης που συντονίζει τις αποφάσεις για ένα κομμάτι δεδομένων, στην προκειμένη περίπτωση γλώσσα που βασίζεται σε κείμενο, για την εξαγωγή προγνωστικών συμπερασμάτων για κακόβουλη ή μη κακόβουλη επικοινωνία.

Έρευνα από τους Xiang et al. [141] εισήγαγε το CANTINA+, το οποίο είναι ένα πλαίσιο για την ανίχνευση δικτυακών τόπων phishing με χρήση αλγορίθμων μηχανικής μάθησης σε ένα εύρος συστημάτων, όπως οιμηχανές αναζήτησης και εφαρμογές τρίτων. Αυτή η έρευνα συγκεντρώνει πολλαπλά συστήματα, όπως μηχανές αναζήτησης και υπηρεσίες τρίτων με μηχανισμούς μηχανικής μάθησης για να την ακριβή ταξινόμηση ιστοσελίδων.

Στην εργασία των Cova et al. [142], χρησιμοποιήθηκαν τεχνικές μηχανικής μάθησης και εξομοίωσης για την ανίχνευση ανάρμοστης συμπεριφοράς JavaScript σε σχέση με καθιερωμένα προφίλ συμπεριφοράς για την αποτροπή drive-by downloads. Οι ερευνητές εφαρμόζουν 10 διαφορετικά χαρακτηριστικά μιας διαδικασίας επίθεσης για την κατηγοριοποίηση των δραστηριοτήτων που εμπλέκονται σε drive-by downloads, μιμούμενες συμπεριφορές ιστότοπων που αντιστοιχούν σε αυτά τα χαρακτηριστικά σε συνδυασμό με έναν ταξινομητή ανίχνευσης που υλοποιείται σε ένα εργαλείο ανοικτού κώδικα (open source) με την ονομασία JSAND. Οι ερευνητές παρουσιάζουν τα αποτελέσματα της εφαρμογής του JSAND σε ένα μεγάλης κλίμακας σύνολο δεδομένων κώδικα JavaScript. Η μελέτη είναι σε θέση να καταδείξει μέσω της χρήσης ενός εξομοιωτή, την ανίχνευση κακόβουλου κώδικα όταν συγκρίνεται με χαρακτηριστικά ενσωματωμένα σε ένα μαθημένο μοντέλο ταξινομητή (μηχανικής μάθησης). Η περαιτέρω έρευνα αποσκοπεί στην ανάπτυξη του συστήματος JSAND για την παροχή μιας επέκτασης του προγράμματος περιήγησης που παρέχει προληπτική ανίχνευση και μετριασμό των επιθέσεων drive-by download.

Τεχνικές μηχανικής μάθησης έχουν εφαρμοστεί για την ανίχνευση επιθέσεων phishing στο Twitter. Στο [143], ένας ανιχνευτής URL phishing για το Twitter επικεντρώνεται στο tweet περιεχόμενο, το μήκος, τα hashtags, τις αναφορές, την ηλικία του λογαριασμού και τον αριθμό των tweets σε συνδυασμό με χαρακτηριστικά URL (συντομευμένο κείμενο, μη τυποποιημένη δομή προθέματος και κατάληξης τομέα). Χρησιμοποιώντας αυτές τις μεταβλητές στο πλαίσιο τεχνικών ταξινόμησης μηχανικής μάθησης (naive Bayes, decision tree και random forest algorithms), ανιχνεύονται τεχνικές phishing με ακρίβεια 92,52%. Το σύστημα έχει αναπτυχθεί ως επέκταση του προγράμματος περιήγησης Chrome, λειτουργώντας σε πραγματικό χρόνο για να ταξινομήσει αν ένα tweet είναι μια επίθεση phishing ή μια νόμιμη επικοινωνία από έναν λογαριασμό Twitter. Επιπλέον, η μελέτη έχει

δοκιμαστεί σε μια κοινότητα χρηστών για να αξιολογηθεί η χρηστικότητα και η πρακτικότητά της επέκτασης στην εφαρμογή- με τους ερευνητές να υποστηρίζουν τη σχετική απλότητα και ευκολία χρήσης της. Η εργασία αντιπροσωπεύει έρευνα με αποτέλεσμα ένα προϊόν που μπορεί να χρησιμοποιηθεί από την κοινότητα του Διαδικτύου έναντι μιας γνωστής ευπάθειας μιας δημοφιλούς πλατφόρμας μέσων κοινωνικής δικτύωσης.

Οι Stringhini και Thonnard [144] ανέπτυξαν πρόσφατα ένα σύστημα ταξινόμησης και αποκλεισμού spear phishing μηνυμάτων ηλεκτρονικού ταχυδρομείου από παραβιασμένους λογαριασμούς ηλεκτρονικού ταχυδρομείου με την εκπαίδευση ενός συστήματος υποστήριξης διανυσματικής μηχανής με ένα μοντέλο συμπεριφοράς που βασίζεται στις συνήθειες ηλεκτρονικού ταχυδρομείου ενός χρήστη. Το σύστημα συλλέγει και διαμορφώνει προφίλ συμπεριφορικών χαρακτηριστικών που σχετίζονται με τη γραφή (π.χ. συχνότητα χαρακτήρων/λέξεων, στίξη, γραφίδα), τη σύνθεση και την αποστολή (π.χ. ώρα/ημερομηνία, χαρακτηρισμοί URL, περιεχόμενο αλυσίδας ηλεκτρονικού ταχυδρομείου), και αλληλεπίδραση (π.χ. επαφές ηλεκτρονικού ταχυδρομείου). Αυτά τα χαρακτηριστικά χρησιμοποιούνται για τη δημιουργία ενός προφίλ συμπεριφοράς ηλεκτρονικού ταχυδρομείου που συνδέεται με τον χρήστη, το οποίο ενημερώνεται συνεχώς κάθε φορά που ο χρήστης στέλνει ένα νέο μήνυμα ηλεκτρονικού ταχυδρομείου. Οι δοκιμές έδειξαν πως όσο μεγαλύτερο το ιστορικό ηλεκτρονικού ταχυδρομείου ενός χρήστη, τόσο χαμηλότερο είναι το ποσοστό των ψευδώς θετικών αποτελεσμάτων, το οποίο πέφτει κάτω από το 0,05%. για 7.000 μηνύματα ηλεκτρονικού ταχυδρομείου και άνω.

Οι αλγόριθμοι μηχανικής μάθησης παρέχουν μια αρχιτεκτονική για τη δημιουργία προφίλ σημασιολογικών επιθέσεων με προληπτικό τρόπο.

- **Βαθιά μάθηση (Deep Learning):** Πρόσφατες εξελίξεις στις προσεγγίσεις πρότειναν ότι η ταξινόμηση των δικτυακών τόπων phishing με τη χρήση νευρωνικών δικτύων θα πρέπει να υπερτερεί των παραδοσιακών αλγορίθμων μηχανικής μάθησης (Machine Learning). Ωστόσο, τα αποτελέσματα νευρωνικών δικτύων εξαρτώνται σε μεγάλο βαθμό από τη ρύθμιση των διαφορετικών μαθησιακών αλγορίθμων. Η παρούσα ενότητα περιγράφει τις προσεγγίσεις Deep Learning που βασίζονται σε εισβολές συστήματα ανίχνευσης εισβολών.

Οι συγγραφείς στο [145] ασχολήθηκαν με την εξέταση διαφορετικών τεχνικών και μίλησαν για διαφορετικές στρατηγικές για την ακριβή αναγνώριση επιθέσεων phishing. Από τις αξιολογημένες στρατηγικές, οι διαδικασίες Deep Learning που χρησιμοποίησαν για την εξαγωγή χαρακτηριστικών παρουσιάζουν καλές επιδόσεις λόγω της υψηλής ακρίβειας, ενώ είναι εύρωστες. Τα μοντέλα ταξινόμησης απεικονίζουν επίσης καλές επιδόσεις.

Οι συγγραφείς Maurya και Jain [146] πρότειναν μία anti-phishing δομή που εξαρτάται από τη χρήση μιας αναγνώρισης μοντέλων phishing (Μοντέλα που έχουν εκπαιδευτεί με phishing websites, τα οποίο κατηγοριοποιούν ιστοσελίδες ανάλογα με το αν είναι phishing ή όχι). Εξαρτάται από το Deep Learning, σε επίπεδο ISP για να εγγυηθεί την ασφάλεια. Αυτή η μεθοδολογία περιλαμβάνει ένα μεταβατικό επίπεδο ασφάλειας και τίθεται μεταξύ διαφόρων εργαζομένων και τελικών πελατών. Η επάρκεια της εκτέλεσης αυτής της δομής έγκειται στον τρόπο με τον οποίο ένα μοναδικός σκοπός αποκλεισμού μπορεί να εγγυηθεί ένα μεγάλο αριθμό των πελατών που προστατεύονται από μια συγκεκριμένη επίθεση phishing. Η επιβάρυνση υπολογισμού για τα μοντέλα ανακάλυψης phishing περιορίζεται διακριτά στους ISP ενώ στους τελικούς χρήστες παρέχεται ασφαλής βοήθεια ανεξάρτητα από τα σχέδια των πλαισίων τους χωρίς ιδιαίτερα αποδοτικές μηχανές επεξεργασίας.

Οι συγγραφείς Subasi et al. [147] πρότειναν μια σύγκριση του Adaboost και του multi boosting για την ανίχνευση phishing ιστότοπων. Χρησιμοποίησαν το αποθετήριο μηχανικής μάθησης UCI ως σύνολο δεδομένων με 11.055 περιπτώσεις και 30 χαρακτηριστικά. Τα AdaBoost και multi boost είναι τα προτεινόμενα σύνολα μάθησης σε αυτή την έρευνα για την αναβάθμιση της παρουσίασης των υπολογισμών των επιθέσεων phishing. Τα ensemble μοντέλα βελτιώνουν την έκθεση των ταξινομητών όσον αφορά την ακρίβεια, το μέτρο F και την περιοχή ROC. Τα πειραματικά αποτελέσματα αποκαλύπτουν ότι με τη χρήση μοντέλων ensemble, είναι δυνατή η αναγνώριση σελίδων phishing με ακρίβεια 97,61%.

Οι Abdelhamid et al. [148] πρότειναν μια σύγκριση με βάση το περιεχόμενο και τα χαρακτηριστικά του μοντέλου μάθησης. Χρησιμοποίησαν ένα σύνολο δεδομένων από το PhishTank, που περιέχει περίπου 11.000 παραδείγματα. Επίσης, χρησιμοποίησαν μια προσέγγιση που ονομάζεται ενισχυμένος δυναμικός κανόνας επαγωγής (eDRI) και ισχυρίστηκαν ότι η δυναμική επαγωγή κανόνων (eDRI) είναι ο πρώτος αλγόριθμος βαθιάς μηχανικής μάθησης που έχει ενσωματωθεί σε ένα εργαλείο κατά του phishing. Αυτός ο αλγόριθμος περνάει σύνολα δεδομένων με δύο κύριες συχνότητες και την ισχύ των κανόνων. Το σύνολο δεδομένων εκπαίδευσης αποθηκεύει μόνο "ισχυρά" χαρακτηριστικά και αυτά τα χαρακτηριστικά γίνονται μέρος του κανόνα ενώ άλλα αφαιρούνται.

Οι συγγραφείς στο Mao et al. [149] πρότειναν ένα σύστημα βασισμένο στη μάθηση για την επιλογή του σχεδιασμού της σελίδας που χρησιμοποιείται για τη διάκριση σελίδων επιθέσεων phishing για αποτελεσματικά χαρακτηριστικά διάταξης της σελίδας, χαρακτήρισαν τις κατευθυντήριες γραμμές και δημιούργησαν μία phishing ιστοσελίδα που χρησιμοποιεί δύο συμβατικούς αλγορίθμους μάθησης, SVM και DT. Δοκίμασαν τη μεθοδολογία σε δοκιμές πραγματικών σελίδων ιστοτόπων από το phishtank.com και το alexa.com.

Οι Jain και Gupta [150] πρότειναν τεχνικές και πραγματοποίησαν πειράματα σε περισσότερα από δύο σύνολα δεδομένων. Πρώτα στο Phishtank που περιέχει 1528 ιστότοπους phishing, έπειτα στο Openphish που περιέχει 613 ιστότοπους phishing, μετά στο Alexa που περιέχει 1600 νόμιμους ιστότοπους, μετά στην πύλη πληρωμών που περιέχει 66 νόμιμους ιστότοπους, και τέλος στον κορυφαίο τραπεζικό ιστότοπο που περιέχει 252 νόμιμους ιστότοπους. Με την εφαρμογή αλγορίθμων μηχανικής μάθησης, βελτίωσαν την ακρίβεια για την ανίχνευση phishing. Χρησιμοποίησαν RF, SVM, νευρωνικά δίκτυα (NN), LR και NB. Επίσης, χρησιμοποίησαν μια προσέγγιση εξαγωγής χαρακτηριστικών στην πλευρά του πελάτη.

- **Υβριδική Μάθηση:** Σε αυτή την ενότητα, παρουσιάζουμε τη σύγκριση των μοντέλων Υβριδικής Μάθησης τα οποία χρησιμοποιούνται από μελέτες τελευταίας τεχνολογίας.

Οι Kumar et al. [160] διαχώρισαν κάποια άσχετα χαρακτηριστικά από το περιεχόμενο και τις εικόνες και εφάρμοσαν SVM ως έναν δυαδικό ταξινομητή. Ομαδοποίησαν τα πραγματικά και τα ψεύτικα μηνύματα με στρατηγικές όπως η ανάλυση κειμένου, το tokenization λέξεων, και την απομάκρυνση των λέξεων στάσης (stop words). Οι Jain et al. [161] χρησιμοποίησαν την μετρική TF- IDF για να εντοπίσουν τα πιο σημαντικά χαρακτηριστικά των ιστότοπων που θα χρησιμοποιηθούν στην ερώτηση αναζήτησης. Η προτεινόμενη προσέγγιση είναι πιο ακριβής έναντι των υφιστάμενων τεχνικών που χρησιμοποιούν την παραδοσιακό TF-IDF μετρική.

Οι Adebowale et al. [162] πρότειναν μια υβριδική προσέγγιση που περιλαμβάνει αναζήτηση ευρετικού κανόνα και λογιστική παλινδρόμηση (SHLR) για την αποτελεσματική ανίχνευση επιθέσεων phishing. Οι συγγραφείς πρότειναν προσέγγιση τριών βημάτων: (1) οι περισσότεροι ιστότοποι που εμφανίζονται στο αποτέλεσμα ενός ερωτήματος αναζήτησης είναι νόμιμοι εάν ο τομέας της ιστοσελίδας ταιριάζει με το όνομα τομέα των ιστότοπων που

ανακτήθηκαν στα αποτελέσματα κατά το ερώτημα, (2) οι ευρετικοί κανόνες που ορίζονται από τα χαρακτηριστικά του χαρακτήρα και (3) ένα μοντέλο ML για την πρόβλεψη αν η ιστοσελίδα είναι είτε νόμιμη είτε αφορά επίθεση ηλεκτρονικού "ψαρέματος".

Οι Chiew et al. [163] χρησιμοποίησαν δύο σύνολα δεδομένων, το ένα από 5000 ιστοσελίδες phishing με βάση τις διευθύνσεις URL από το PhishTank και το δεύτερο από το OpenPhish. Άλλες 5000 νόμιμες ιστοσελίδες βασίστηκαν σε διευθύνσεις URL από την Alexa και το αρχείο Common Crawl5.

Οι Pandey et al. [164] χρησιμοποίησαν ένα σύνολο δεδομένων από το Website phishing dataset, το οποίο είναι διαθέσιμο σε απευθείας σύνδεση σε αποθετήριο του Πανεπιστημίου της Καλιφόρνια. Αυτό το σύνολο δεδομένων έχει 10 χαρακτηριστικά και 1353 περιπτώσεις. Εκπαίδευσαν ένα υβριδικό μοντέλο RF-SVM που πέτυχε ακρίβεια 94%.

Οι Niranjan et al. [165] πρότειναν μια τεχνική συνόλου μέσω της μεθόδου ψηφοφορίας και στοίβαξης (voting and stacking method). Επέλεξαν το σύνολο δεδομένων UCI ML phishing και πήραν μόνο 23 χαρακτηριστικά από τα 30 διαθέσιμα για την περαιτέρω ανίχνευση επιθέσεων. Από ένα σύνολο 11.055 περιπτώσεων, το σύνολο δεδομένων έχει 6157 νόμιμες και 4898 περιπτώσεις phishing. Χρησιμοποίησαν το μοντέλο EKRV για την πρόβλεψη των επιθέσεων.

Οι Patil et al. [166] πρότειναν μία υβριδική λύση που χρησιμοποιεί τρεις προσεγγίσεις: μαύρη λίστα και λευκή λίστα, ευρετικές μέθοδους και οπτική ομοιότητα. Η προτεινόμενη μεθοδολογία παρακολουθεί όλη την κυκλοφορία στο σύστημα του τελικού χρήστη και συγκρίνει κάθε διεύθυνση URL με τη λευκή λίστα αξιόπιστων τομέων. Ο ιστότοπος αναλύει διάφορες λεπτομέρειες για χαρακτηριστικά. Τα τρία αποτελέσματα είναι οι ύποπτοι ιστότοποι, οι ιστότοποι phishing και οι νόμιμοι δικτυακοί τόποι. Ο ταξινομητής ML χρησιμοποιείται για τη συλλογή δεδομένων και για τη δημιουργία ενός τελικού σκορ. Εάν η βαθμολογία είναι μεγαλύτερη από το κατώτατο όριο, τότε χαρακτηρίζει τη διεύθυνση URL ως επίθεση phishing και την αποκλείει αμέσως. Χρησιμοποίησαν τους LR(Logistic Regression), DT(Decision Trees)(Safavian and Langberde,1991) και RF(Random Forest) για να προβλέψουν την ακρίβεια των δοκιμαστικών τους ιστότοπων.

Οι Jagadeesan et al. [167] χρησιμοποίησαν RF και SVM [171] για την ανίχνευση επιθέσεων phishing. Χρησιμοποίησαν δύο τύπους συνόλων δεδομένων. Το πρώτο προέρχεται από το αποθετήριο μηχανικής μάθησης UCI, το οποίο έχει 30 χαρακτηριστικά. Αυτό το σύνολο δεδομένων αποτελείται από 2456 καταχωρίσεις phishing και μη phishing διευθύνσεων URL. Το δεύτερο σύνολο δεδομένων αποτελείται από από 1353 διευθύνσεις URL που έχει 10 χαρακτηριστικά και τρεις κατηγορίες: Phishing, non-Phishing και ύποπτες.

- **Ανίχνευση επιθέσεων βάση σεναρίων:** Σε αυτή την ενότητα, παρέχουμε μια σύγκριση των βασισμένων σε σενάρια μεθόδων ανίχνευσης επιθέσεων phishing που έχουν προταθεί από διάφορους ερευνητές. Οι μελέτες δείχνουν ότι τα διάφορα σενάρια λειτούργησαν με διάφορες μεθόδους και παρέχουν διαφορετικά αποτελέσματα.

Οι συγγραφείς Begum και Badugu [151] συζήτησαν ορισμένες προσεγγίσεις που είναι χρήσιμες για την ανίχνευση μιας επίθεσης phishing. Πραγματοποίησαν μια λεπτομερή έρευνα των υφιστάμενων τεχνικών όπως οι προσεγγίσεις που βασίζονται στη μηχανική μάθηση (ML), οι προσεγγίσεις που βασίζονται σε νευρωνικά δίκτυα και προσεγγίσεις ανίχνευσης με βάση τη συμπεριφορά για ανίχνευση επιθέσεων phishing.

Οι Yasin et al. [152] ενοποίησαν διάφορες μελέτες για να αποσαφηνίσουν διαφορετικές ασκήσεις των ειδικών κυβερνοασφάλειας. Επιπλέον, ισχυρίστηκαν ότι μια υψηλότερη κατανόηση των σεναρίων επιθέσεων της κοινωνικής μηχανικής θα ήταν δυνατή με τη χρήση επίκαιρων και τεχνικών διερεύνησης που βασίζονται σε παιχνίδια. Η προτεινόμενη στρατηγική για την ερμηνεία των σεναρίων επιθέσεων κοινωνικής μηχανικής αποτελεί μια τέτοια προσπάθεια για την ενδυνάμωση της κατανόησης των γενικών σεναρίων επιθέσεων.

Οι Fatima et al. [153] παρουσίασαν το PhishI ως έναν ακριβή τρόπο αντιμετώπισης της δομής γνήσιων παιγνίων (genuine games) για την εκπαίδευση στην ασφάλεια. Πρότειναν ένα σύστημα δομής παιγνίων που ενσωματώνει την ομάδα πληροφοριών για την κοινωνική δικτύωση, που χρειάζεται έγκυρους παίκτες. Στο παιχνίδι PhishI, τα μέλη απαιτείται να ανταλλάσσουν μηνύματα phishing και έχουν τη δυνατότητα να παρατηρήσουν τη βιωσιμότητα του σεναρίου επίθεσης. Τα αποτελέσματα έδειξαν ότι η προσοχή των μαθητών στις πιθανότητες spear-phishing βελτιώνεται και ότι η προστασία από την πρώτη πιθανή επίθεση αναβαθμίζεται. Επιπλέον, το παιχνίδι κατέδειξε μια ευεργετική αποτελεσματικότητα στην κατανόηση των μελών για τα ακραία διαδικτυακά δεδομένα και την αποκάλυψη πληροφοριών.

Οι Chiew et al. [154] συγκέντρωσαν τις επιθέσεις phishing λεπτομερώς μέσω των χαρακτηριστικών τους, του μέσου και του φορέα στον οποίο ζουν και των εξειδικευμένων μεθοδολογιών τους. Εκτός αυτού, αποδέχονται ότι οι πληροφορίες αυτές θα βοηθήσουν τον συνολικό πληθυσμό με τη λήψη προπαρασκευαστικών και προληπτικών δραστηριοτήτων ενάντια σε αυτές τις επιθέσεις phishing και τις πολιτικές για την εκτέλεση προσεγγίσεων για τον έλεγχο οποιασδήποτε περαιτέρω κατάχρησης από τους phishers. Η στήριξη μόνο στις οδηγίες ενός ειδικού κυβερνοασφάλειας ως προληπτικό μέτρο σε μια phishing επίθεση δεν αρκεί. Η έρευνά τους δείχνει ότι απαιτείται η βελτίωση έξυπνων πλαισίων για την αντιμετώπιση αυτών των εξειδικευμένων μεθοδολογιών, καθώς τα εν λόγω αντίμετρα θα έχουν τη δυνατότητα να αναγνωρίζουν και να απενεργοποιούν τόσο τις υπάρχουσες επιθέσεις όσο και τους νέους κινδύνους phishing.

Οι Yao et al. [155] χρησιμοποίησαν τη μέθοδο εξαγωγής λογότυπων χρησιμοποιώντας τη διαδικασία ανίχνευσης ταυτότητας για την ανίχνευση του phishing. Δύο μη επικαλυπτόμενα σύνολα δεδομένων δημιουργήθηκαν από ένα σύνολο 726 σελίδων. Οι σελίδες phishing προέρχονται από τον ιστότοπο PhishTank και οι σελίδες του νόμιμου δικτυακού τόπου προέρχονται από το διαδίκτυο. Είχαν περιορίσει το έργο τους και τις σελίδες που χρησιμοποίησαν, καθώς τα μέλη των τελικών χρηστών μπορεί να ήταν υπερβολικά προσεκτικοί σε σχέση με την πιθανή διπροσωπία και κατ' αυτόν τον τρόπο προοδευτικά προσεκτικοί στις αξιολογήσεις τους για κάθε μήνυμα ηλεκτρονικού ταχυδρομείου απ' ό,τι θα έκαναν αν ήταν στο κανονικό τους εργασιακό περιβάλλον.

Οι Parsons et al. [156] χρησιμοποίησαν τη μέθοδο ANOVA. Σε μια μελέτη που βασίστηκε σε σενάρια phishing, είχαν συνολικά 985 συμμετέχοντες. Η ανάλυση διακύμανσης δύο τρόπων επαναλαμβανόμενων μετρήσεων (ANOVA) χρησιμοποιήθηκε για την έρευνα του αντίκτυπου της αυθεντικότητας του ηλεκτρονικού ταχυδρομείου. Η έρευνα αυτή περιλάμβανε μόνο ένα phishing site και ένα πιστοποιημένο μήνυμα ηλεκτρονικού ταχυδρομείου με ένα από τα πρότυπα και δεν εξέτασε τον αντίκτυπο πολλών προτύπων μέσα σε ένα μήνυμα ηλεκτρονικού ταχυδρομείου. Ακολούθησε η σύγκριση του ταξινομητή γνωστού ως RF(Random Forest), ο οποίος είναι ο αλγόριθμος που χρησιμοποιείται περισσότερο από τους ερευνητές, σε σχέση με άλλους ταξινομητές.

Οι Tyagi et al. [157] χρησιμοποίησαν 30 χαρακτηριστικά για την ανίχνευση επιθέσεων μέσω του RF. Χρησιμοποίησαν και άλλους ταξινομητές, αλλά το αποτέλεσμά τους για το RF ήταν καλύτερο σε σχέση με αυτό των άλλων ταξινομητών.

Ομοίως, οι Mao et al. [158] συνέλεξαν ένα σύνολο δεδομένων 49 phishing websites από το PhishinTank.com. Χρησιμοποίησαν τέσσερις ταξινομητές μάθησης για την ανίχνευση επιθέσεων phishing και κατέληξαν στο συμπέρασμα ότι οι ταξινομητές RF είναι πολύ καλύτεροι από τους άλλους.

Οι συγγραφείς στο Sahingoz et al. [159] χρησιμοποίησαν το Ebbu2017 Phishing Dataset που περιέχει 73.575 διευθύνσεις URL στις οποίες 36.400 είναι νόμιμες διευθύνσεις URL και 37.175 είναι διευθύνσεις phishing. Πρότειναν επτά διαφορετικούς αλγόριθμους ταξινόμησης, συμπεριλαμβανομένων χαρακτηριστικών βασισμένων στην επεξεργασία φυσικής γλώσσας (NLP). Στην πραγματικότητα χρησιμοποίησαν ένα σύνολο δεδομένων το οποίο δεν χρησιμοποιείται συνήθως για την ανίχνευση επιθέσεων phishing.

4.2 Αξιολόγηση Τρόπων Αντιμετώπισης Επιθέσεων Κοινωνικής Μηχανικής

4.2.1 Κριτήρια

Σε αυτή την ενότητα θα ορίσουμε κριτήρια για την αξιολόγηση των τρόπων αντιμετώπισης επιθέσεων κοινωνικής μηχανικής και θα αποτιμήσουμε τους τρόπους αντιμετώπισης με βάση αυτά τα κριτήρια. Τα κριτήρια επινοήθηκαν στα πλαίσια της παρούσας διπλωματικής. Ορισμένα κριτήρια από αυτά που επινοήθηκαν εστιάζουν στην εφαρμοσιμότητα των μεθόδων προστασίας προς αξιολόγηση.

Μέθοδος Προστασίας: Ο τρόπος αντιμετώπισης μιας επίθεσης ή απειλής. Μπορεί να είναι η ανίχνευση, η πρόληψη ή η αντίδραση σε αυτή την επίθεση/απειλή.

Μέθοδος αποτίμησης: Με ποιόν τρόπο αποτιμάται η προτεινόμενη μέθοδος προστασίας (π.χ Εκπαίδευση, Πειραματική μελέτη, κλπ.)

Βαθμός Προστασίας: Πόσο αποδοτική (π.χ Καθόλου, Λίγο, Μέτρια, Αρκετά) είναι η μέθοδος προστασίας για τις επιθέσεις στις οποίες ειδικεύεται

Ευκολία Εφαρμογής: Πόσο εύκολα (π.χ Λίγο, Μέτρια, Αρκετά) εφαρμόζεται ο τρόπος αντιμετώπισης που προτείνεται σε ένα πληροφοριακό σύστημα ή σε ένα πλάνο ενός οργανισμού.

Μέρος εφαρμογής: Πού εφαρμόζεται η προσέγγιση που προτείνεται (π.χ. Πληροφοριακά Συστήματα, Οργανισμούς, Δίκτυα, Ανθρώπους,...)

Χρόνος Εφαρμογής: Πόσος χρόνος (π.χ σε ώρες, μέρες, μήνες,...) θα χρειαστεί ώστε να εφαρμοστεί πλήρως η προσέγγιση

Προσπάθεια Εφαρμογής: Πόση προσπάθεια (π.χ Λίγο, Μέτρια, Μεγάλη) απαιτείται για να εφαρμοστεί πλήρως η προσέγγιση

Κόστος Εφαρμογής: Πόσο στοιχίζει η εφαρμογή της μεθόδου προστασίας καθώς και η διατήρηση της (π.χ. Καθόλου, Λίγο, Μέτρια, Αρκετά)

4.2.2 Αξιολόγηση

Με βάση τις παραπάνω διαστάσεις και τρόπους αντιμετώπισης δημιουργήθηκαν οι παρακάτω πίνακες αξιολόγησης, ένας για κάθε είδος μεθόδων προστασίας.

Είδος Μεθόδου Προστασίας: Πολιτική - Διαδικασία

Πίνακας 3: Αξιολόγηση Μεθόδου Προστασίας: Πολιτική-Διαδικασία.

Μέθοδος Προστασίας	Είδος Μεθόδου Προστασίας	Μέθοδος Αποτίμησης	Τρόπος Αντιμετώπισης	Βαθμός Προστασίας	Ευκολία Εφαρμογής	Μέρος Εφαρμογής	Χρόνος εφαρμογής	Προσπάθεια εφαρμογής	Κόστος Εφαρμογής
[95]	Πολιτική-Διαδικασία	Θεωρητική Μελέτη	Ανίχνευση, Πρόληψη, Αντίδραση	Αρκετά	Μέτρια	Συστήματα, Δίκτυα, Άνθρωποι	Μήνες	Μέτρια	Μέτρια
[96]	Πολιτική-Διαδικασία	Θεωρητική Μελέτη	Πρόληψη	Μέτρια	Μέτρια	Συστήματα, Δίκτυα, Άνθρωποι	Μήνες	Μέτρια	Μέτρια

Είδος Μεθόδου Προστασίας: Εκπαίδευση

Πίνακας 4: Αξιολόγηση Μεθόδου Προστασίας: Εκπαίδευση.

Μέθοδος Προστασίας	Είδος Μεθόδου Προστασίας	Μέθοδος Αποτίμησης	Τρόπος Αντιμετώπισης	Βαθμός Προστασίας	Ευκολία Εφαρμογής	Μέρος Εφαρμογής	Χρόνος εφαρμογής	Προσπάθεια εφαρμογής	Κόστος Εφαρμογής
[97] [98]	Εκπαίδευση	Εκπαίδευση	Πρόληψη	Μέτρια	Αρκετά	Άνθρωποι	Μέρες	Λίγο	Λίγο
[99]	Εκπαίδευση	Εκπαίδευση	Πρόληψη, Αντίδραση	Λίγο	Αρκετά	Συστήματα	Ώρες	Λίγο	Καθόλου
[100]	Εκπαίδευση	Εκπαίδευση	Πρόληψη	Μέτρια	Αρκετά	Συστήματα	Ώρες	Λίγο	Καθόλου
[101]	Εκπαίδευση	Εκπαίδευση	Πρόληψη	Μέτρια	Μέτρια	Σύστημα	Ώρες	Μέτρια	Λίγο
[102] [103]	Εκπαίδευση	Εκπαίδευση	Πρόληψη	Μέτρια	Μέτρια	Σύστημα	Ώρες	Μέτρια	Λίγο
[104]	Εκπαίδευση	Εκπαίδευση	Ανίχνευση	Μεγάλη	Λίγο	Συστήματα	Μέρες	Μέτρια	Λίγο

Είδος Μεθόδου Προστασίας: Sanboxing

Πίνακας 5: Αξιολόγηση Μεθόδων Sandboxing.

Μέθοδος Προστασίας	Είδος Μεθόδου Προστασίας	Μέθοδος Αποτίμησης	Τρόπος Αντιμετώπισης	Βαθμός Προστασίας	Ευκολία Εφαρμογής	Μέρος Εφαρμογής	Χρόνος εφαρμογής	Προσπάθεια εφαρμογής	Κόστος Εφαρμογής
[106]	Τεχνικό-Sandboxing	Πειραματική Μελέτη	Αντίδραση	Μεγάλη	Μέτρια	Συστήματα	Μέρες	Μέτρια	Μέτρια
[108]	Τεχνικό-Sandboxing	Πειραματική Μελέτη	Ανίχνευση	Μέτριο	Μέτρια	Συστήματα	Μέρες	Μέτρια	Μέτρια
[110]	Τεχνικό-Sandboxing	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Λίγο	Συστήματα	Μήνες	Μεγάλο	Μεγάλο (Χρονικά)
[110][111]	Τεχνικό-Sandboxing	Πειραματική Μελέτη	Πρόληψη	Λίγο	Μέτρια	Συστήματα	Βδομάδες	Μέτρια	Μέτρια
[112]	Τεχνικό-Sandboxing	Πειραματική Μελέτη	Αντίδραση	Μέτρια	Αρκετά	Συστήματα	Μέρες	Μέτρια	Μέτρια
[113]	Τεχνικό-Sandboxing	Πειραματική Μελέτη	Αντίδραση	Μεγάλη	Αρκετά	Συστήματα	Βδομάδες	Λίγο	Μέτρια
[114]	Τεχνικό-Sandboxing	Πειραματική Μελέτη	Αντίδραση	Μεγάλη	Αρκετά	Συστήματα	Μέρες	Λίγο	Μεγάλο (Χρονικά)

Είδος Μεθόδου Προστασίας: AAA

Πίνακας 6: Αξιολόγηση Μεθόδων Προστασίας AAA.

Μέθοδος Προστασίας	Είδος Μεθόδου Προστασίας	Μέθοδος Αποτίμησης	Τρόπος Αντιμετώπισης	Βαθμός Προστασίας	Ευκολία Εφαρμογής	Μέρος Εφαρμογής	Χρόνος εφαρμογής	Προσπάθεια εφαρμογής	Κόστος Εφαρμογής
[118]	AAA	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Συστήματα	Μέρες	Μέτρια	Μέτρια
[119]	AAA	Πειραματική Μελέτη	Αντίδραση	Μέτρια	Μέτρια	Συστήματα	Μέρες	Μέτρια	Λίγο
[120]	AAA	Πειραματική Μελέτη	Αντίδραση	Μέτρια	Μέτρια	Συστήματα	Μέρες	Μέτρια	Μέτρια

Είδος Μεθόδου Προστασίας: Παρακολούθηση (Monitoring)

Πίνακας 7: Αξιολόγηση Μεθόδων Προστασίας Παρακολούθηση-Monitoring.

Μέθοδος Προστασίας	Είδος Μεθόδου Προστασίας	Μέθοδος Αποτίμησης	Τρόπος Αντιμετώπισης	Βαθμός Προστασίας	Ευκολία Εφαρμογής	Μέρος Εφαρμογής	Χρόνος εφαρμογής	Προσπάθεια εφαρμογής	Κόστος Εφαρμογής
[122]	Τεχνικό - Monitoring	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Συστήματα	Βδομάδες	Μέτρια	Μέτρια
[123]	Τεχνικό - Monitoring	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Μέθοδος	Ώρες	Μέτρια	Μέτρια
[124]	Τεχνικό - Monitoring	Πειραματική Μελέτη	Ανίχνευση	Μεγάλη	Αρκετά	Επέκταση προγράμματος περιήγησης	Ώρες	Λίγο	Καθόλου
[125]	Τεχνικό - Monitoring	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Αρκετά	Συστήματα διαχείρισης	Ώρες	Μέτρια	Λίγο
[126]	Τεχνικό - Monitoring	Πειραματική Μελέτη	Ανίχνευση	Μεγάλη	Αρκετά	Συστήματα	Ώρες	Λίγο	Καθόλου

Είδος Μεθόδου Προστασίας: Integrity Checking

Πίνακας 8: Αξιολόγηση Μεθόδων Προστασίας Integrity Checking.

Μέθοδος Προστασίας	Είδος Μεθόδου Προστασίας	Μέθοδος Αποτίμησης	Τρόπος Αντιμετώπισης	Βαθμός Προστασίας	Ευκολία Εφαρμογής	Μέρος Εφαρμογής	Χρόνος εφαρμογής	Προσπάθεια εφαρμογής	Κόστος Εφαρμογής
[127] [128] [129]	Τεχνικό - Integrity Checking	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Έλεγχος ακεραιότητας	Μήνες	Λίγο	Μέτρια
[130]	Τεχνικό - Integrity Checking	Πειραματική Μελέτη	Πρόληψη	Λίγο	Μέτρια	Έρευνα Αλγορίθμων	Μήνες	Μέτρια	Μέτρια

[131]	Τεχνικό - Integrity Checking	Πειραματική Μελέτη	Πρόληψη	Μέτρια	Αρκετά	Έρευνα μεθόδων	Βδομάδες	Μέτρια	Μέτρια
[132]	Τεχνικό - Integrity Checking	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Αλγόριθμοι Βελτιστοποίησης	Μήνες	Μέτρια	Μεγάλη

Είδος Μεθόδου Προστασίας: Machine Learning

Πίνακας 9: Αξιολόγηση Μεθόδων Προστασίας Machine Learning.

Μέθοδος Προστασίας	Είδος Μεθόδου Προστασίας	Μέθοδος Αποτίμησης	Τρόπος Αντιμετώπισης	Βαθμός Προστασίας	Ευκολία Εφαρμογής	Μέρος Εφαρμογής	Χρόνος εφαρμογής	Προσπάθεια εφαρμογής	Κόστος Εφαρμογής
[134]	Τεχνικό - Machine Learning	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Αρκετά	Έρευνα μεθόδων	Μήνες	Μεγάλη	Μέτρια
[136]	Τεχνικό - Machine Learning	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Συστήματα	Βδομάδες	Μέτρια	Μέτρια
[137]	Τεχνικό - Machine Learning	Πειραματική Μελέτη	Ανίχνευση	Αρκετά	Μέτρια	Τεχνικές Ομαδοποίησης	Μήνες	Μέτρια	Μέτρια
[138][139]	Τεχνικό - Machine Learning	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Συστήματα	Βδομάδες	Μέτρια	Μέτρια
[140]	Τεχνικό - Machine Learning	Πειραματική Μελέτη	Ανίχνευση	Λίγο	Μέτρια	Μηχανισμοί πρόβλεψης - εξαγωγή προγνωστικών	Μήνες	Μέτρια	Λίγο
[141]	Τεχνικό - Machine Learning	Πειραματική Μελέτη	Ανίχνευση	Καθόλου	Μέτρια	Πλαίσιο ανίχνευσης	Μήνες	Μέτρια	Μέτρια
[142]	Τεχνικό - Machine Learning	Πειραματική Μελέτη	Ανίχνευση	Μέτριο	Αρκετά	Επέκταση Προγράμματος περιήγησης	Ώρες	Μέτρια	Καθόλου
[143]	Τεχνικό - Machine Learning	Πειραματική Μελέτη	Ανίχνευση	Αρκετά	Αρκετά	Επέκταση Προγράμματος περιήγησης	Ώρες	Μέτρια	Καθόλου

[144]	Τεχνικό - Machine Learning	Πειραματική Μελέτη	Ανίχνευση	Αρκετά	Μέτρια	Συστήματα	Μήνες	Μέτρια	Μεγάλο (Χρονικά)

Είδος Μεθόδου Προστασίας: Deep Learning

Πίνακας 10: Αξιολόγηση Μεθόδων Προστασίας Deep Learning.

Μέθοδος Προστασίας	Είδος Μεθόδου Προστασίας	Μέθοδος Αποτίμησης	Τρόπος Αντιμετώπισης	Βαθμός Προστασίας	Ευκολία Εφαρμογής	Μέρος Εφαρμογής	Χρόνος εφαρμογής	Προσπάθεια εφαρμογής	Κόστος Εφαρμογής
[145]	Τεχνικό - Deep Learning	Πειραματική Μελέτη	Πρόληψη	Καθόλου	Λίγο	Μοντέλα Ταξινόμησης	Μήνες	Μέτρια	Μεγάλο (Χρονικά)
[146]	Τεχνικό - Deep Learning	Πειραματική Μελέτη	Ανίχνευση	Λίγο	Μέτρια	Μεθοδολογία	Μήνες	Μέτρια	Μέτρια
[147]	Τεχνικό - Deep Learning	Πειραματική Μελέτη	Ανίχνευση	Αρκετά	Λίγο	Συστήματα	Μήνες	Μέτρια	Μεγάλο (Χρονικά)
[148]	Τεχνικό - Deep Learning	Πειραματική Μελέτη	Πρόληψη	Λίγο	Μέτρια	Μεθοδολογία	Βδομάδες	Μέτρια	Μέτρια
[149]	Τεχνικό - Deep Learning	Πειραματική Μελέτη	Πρόληψη	Λίγο	Μέτρια	Μεθοδολογία	Βδομάδες	Μέτρια	Μέτρια(Χρονικά)
[150]	Τεχνικό - Deep Learning	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Συστήματα	Βδομάδες	Μέτρια	Μέτρια

Είδος Μεθόδου Προστασίας: Scenario-Based

Πίνακας 11: Αξιολόγηση Μεθόδων Προστασίας Scenario-Based.

Μέθοδος Προστασίας	Είδος Μεθόδου Προστασίας	Μέθοδος Αποτίμησης	Τρόπος Αντιμετώπισης	Βαθμός Προστασίας	Ευκολία Εφαρμογής	Μέρος Εφαρμογής	Χρόνος εφαρμογής	Προσπάθεια εφαρμογής	Κόστος Εφαρμογής
[151]	Τεχνικό-Scenario based	Πειραματική Μελέτη	Ανίχνευση	Λίγο	Μέτρια	Έρευνα	Βδομάδες	Μέτρια	Μέτρια
[152] [64]	Τεχνικό-Scenario based	Πειραματική Μελέτη	Πρόληψη	Λίγο	Μέτρια	Στρατηγική	Μήνες	Μεγάλο	Μέτριο
[153]	Τεχνικό-Scenario based	Πειραματική Μελέτη	Αντίδραση	Μέτρια	Μέτρια	Παίγνια	Βδομάδες	Λίγο	Μέτριο
[154]	Τεχνικό-Scenario based	Πειραματική Μελέτη	Πρόληψη	Λίγο	Μέτρια	Έρευνα	Βδομάδες	Μέτρια	Μέτριο
[155]	Τεχνικό-Scenario based	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Αρκετά	Μέθοδος	Μέρες	Λίγο	Λίγο
[156] [52]	Τεχνικό-Scenario based	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Μέθοδος	Βδομάδες	Μέτρια	Μέτριο
[158]	Τεχνικό-Scenario based	Πειραματική Μελέτη	Ανίχνευση	Λίγο	Λίγο	Έρευνα	Βδομάδες	Μέτρια	Λίγο
[159]	Τεχνικό-Scenario based	Πειραματική Μελέτη	Ανίχνευση	Λίγο	Μέτρια	Έρευνα	Βδομάδες	Λίγο	Μέτριο

Είδος Μεθόδου Προστασίας: Hybrid

Πίνακας 12: Αξιολόγηση Μεθόδων Προστασίας Hybrid.

Μέθοδος Προστασίας	Είδος Μεθόδου Προστασίας	Μέθοδος Αποτίμησης	Τρόπος Αντιμετώπισης	Βαθμός Προστασίας	Ευκολία Εφαρμογής	Μέρος Εφαρμογής	Χρόνος εφαρμογής	Προσπάθεια εφαρμογής	Κόστος Εφαρμογής
[162]	Τεχνικό - Hybrid	Πειραματική Μελέτη	Πρόληψη	Μέτρια	Μέτρια	Υβριδική προσέγγιση	Βδομάδες	Μέτρια	Λίγο
[164]	Τεχνικό - Hybrid	Πειραματική Μελέτη	Αντίδραση	Αρκετά	Μέτρια	Υβριδική προσέγγιση	Βδομάδες	Μέτρια	Μέτρια
[165]	Τεχνικό - Hybrid	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Υβριδική προσέγγιση	Μέρες	Μέτρια	Μέτρια
[166]	Τεχνικό - Hybrid	Πειραματική Μελέτη	Ανίχνευση	Μέτρια	Μέτρια	Υβριδική προσέγγιση - Μεθοδολογία	Μέρες	Μέτρια	Μέτριο
[167]	Τεχνικό - Hybrid	Πειραματική Μελέτη	Ανίχνευση	Αρκετά	Μέτρια	Υβριδική προσέγγιση - Μεθοδολογία	Μέρες	Μέτρια	Μέτριο

4.3 Κατηγοριοποίηση προσεγγίσεων με βάση την δεύτερη Κατηγοριοποίηση του θεωρητικού υπόβαθρου

Attack Vs Defence		Πολιτική και έλεγχος Διαδικασιών	Εκπαίδευση	Μηχανισμοί Sandboxing	Authorization, Authentication and Accounting (AAA)	Monitoring	Integrity Checking	Machine Learning	Deep Learning	Scenario-Based	Hybrid
ORCHESTRATION	TD1							[140](Raskin 2010) [144](Stringhini 2015)		[152](Yasin et al)	
	TD2										
	MD1-L			[107](Blasingel et al 2010) [108](Greamo and Gosh 2011)				[134](Singhal and Raul 2012)	[147](Subasi et al) [149](Mao) [150](Jain)	[151](Begum) [152](Yasin et al) [153] (Fatima) [159] (Sahingoz)	[161](Jain) [167] (Jagadeesan)
	MD1-R	[96](Fraunstein and Solms) [95](Peltier 2013)	[98](Arachilage 2012)	[106](Barth et al 2008) [107](Blasingel et al 2010) [113](Lu et al 2010)		[121](Lee et al 2012) [123](Stringhini et al 2011) [124](Lu et al 2011) [125](Li et al 2013) [126](Lee and kim 2012)	[127](FCS 2009) [128] (Webroot 2013) [132](Bhardwaj 2014)	[135](Dong-Her 2011) [121](Lee 2010) [137](Basnet 2008) [141](Xiang 2011) [142](Cova 2010) [143](Aggarwal 2012) [144](Stringhini 2015)	[146](Maurya and Jain) [148](Abdelhamid)	[155](Yao) [157](Tyagi)	[160](Kumar) [165](Niranjan) [166](Patil)
	MD2										
	MD3										
	MA1										
	MA2			[113](Lu et al 2010)		[121](Lee et al 2012) [123](Stringhini 2011) [124](Lu et al 2011) [125](Li et al 2013) [126](Lee 2012)	[128] (Webroot 2013)	[143](Aggarwal 2012)	[146](Maurya and Jain) [149](Mao) [150](Jain)	[151](Begum) [159] (Sahingoz)	[162](Adebowale) [164](Pandley) [165](Niranjan)
EXPLOITATION	DV1		[97](Kumaraguru 2009) [98](Arachilage 2012) [99](Lin 2011) [100](Lee 2015) [105](Lee 2015)	[114](Bianchi et al 2015)			[127](FCS 2009) [128] (Webroot 2013) [131](Hara et al 2009)	[135](Dong-Her 2011) [137](Basnet et al 2008) [138](Bergholz 2008) [141](Xiang 2011)	[146](Maurya and Jain) [147](Subasi et al) [148](Abdelhamid)	[151](Begum) [153] (Fatima) [154](Chiew) [155](Yao) [157](Tyagi)	[160](Kumar) [161](Jain) [163](Chew) [166](Patil) [167] (Jagadeesan)
	DV2		[100](Anderson 2013)	[114](Bianchi et al 2015)				[134](Singhal and Raul 2012) [121](Lee 2010) [142](Cova 2010)			
	DV3		[99](Lin2011) [100](Anderson 2013) [105](Lee 2015)	[114](Bianchi et al 2015)			[127](FCS 2009) [132](Bhardwaj 2014)	[143](Aggarwal 2012) [144](Stringhini 2015)			[161](Jain) [162](Adebowale) [163](Chew) [164](Pandley)
	IM1		[97](Kumaraguru 2009) [98](Arachilage 2012)					[135](Dong-Her 2011) [121](Lee 2010) [137](Basnet et al 2008) [138](Bergholz 2008) [141](Xiang 2011) [143](Aggarwal 2012)	[149](Mao) [150](Jain)		
	IM2			[106](Barth et al 2008) [112](Microsoft 2007) [114](Bianchi et al 2015)		[121](Lee et al 2012) [123](Stringhini et al 2011) [124](Lu et al 2011) [125](Li et al 2013) [126](Lee and kim 2012)	[127](FCS 2009) [129](Abdullah 2011) [130](Dhanalakshmi 2010) [132](Bhardwaj 2014)	[134](Singhal and Raul 2012) [137](Basnet 2008) [139](Bergholz 2010) [141](Xiang 2011) [142](Cova 2010) [143](Aggarwal 2012)	[147](Subasi et al)	[151](Begum) [154](Chiew) [155](Yao) [159] (Sahingoz)	[162](Adebowale) [165](Niranjan)
EXECUTION	AP1			[106](Barth et al 2008) [107](Blasingel et al 2010) [113](Lu et al 2010)	[116](Motiee et al 2010) [114](Neuman 1994) [118](Desmond et al 2008) [119](Motiee et al 2010) [120](Turner et al 2010)						
	AP2			[106](Barth et al 2008) [113](Lu et al 2010) [114](Bianchi et al 2015)							
	ES1			[106](Barth et al 2008) [113](Lu et al 2010)	[116](Motiee et al 2010) [114](Neuman 1994) [118](Desmond et al 2008) [119](Motiee et al 2010) [120](Turner et al 2010)			[144](Stringhini 2015)	[148](Abdelhamid) [149](Mao) [150](Jain)	[153] (Fatima) [157](Tyagi)	[160](Kumar) [161](Jain) [163](Chew) [167] (Jagadeesan)
	ES2	[96](Fraunstein and Solms)			[114](Neuman 1994) [118](Desmond et al 2008) [120](Turner et al 2010)				[150](Jain)		[164](Pandley) [166](Patil)

Εικόνα 5: Κατηγοριοποίηση προσεγγίσεων με βάση την δεύτερη κατηγοριοποίηση.

4.4 Ανάλυση και Σχολιασμός πινάκων

Σε αυτό το κεφάλαιο θα αναλύσουμε πόσο αποτελεσματικοί είναι οι τρόποι-προσεγγίσεις ώς προς την απόδοση τους στα κριτήρια που πρέπει να πληρούν. Παρακάτω γίνεται ανάλυση ανά κατηγορία μεθόδων προστασίας.

Πολιτική και έλεγχος διαδικασιών

Η πολιτική και ο έλεγχος διαδικασιών είναι απαραίτητες σε έναν οργανισμό, μιάς και παρέχουν μία συνολική προσέγγιση προστασίας. Κυρίως είναι διαδικασίες για την πρόληψη και την ανίχνευση πιθανών απειλών και επιθέσεων, αλλά και βήματα και διαδικασίες για την αντίδραση σε τέτοιες επιθέσεις.

Όπως βλέπουμε και στον Πίνακας 3 η ευκολία εφαρμογής είναι μέτρια και ο χρόνος εφαρμογής αντιστοιχεί σε μήνες, μιας και θέλει κάποιο χρόνο ώστε να εφαρμοστούν συνολικά τέτοιες προσεγγίσεις ασφάλειας. Υπάρχει αρκετά καλός βαθμός προστασίας αλλά κυρίως παρέχονται μόνο γενικές κατευθύνσεις για τις διαδικασίες ασφαλείας σε όλο το πληροφοριακό σύστημα και τους ανθρώπους που το χειρίζονται. Το κόστος εφαρμογής είναι μέτριο.

Η Πολιτική και ο έλεγχος διαδικασιών είναι από τους βασικούς πυλώνες στην γενική ασφάλεια ενός οργανισμού και τα αποτελέσματα είναι αρκετά ικανοποιητικά μιας και ασχολείται με όλα τα επίπεδα ασφάλειας (Τεχνικές επιθέσεις - Επιθέσεις Κοινωνικής Μηχανικής, κτλ.)

Εκπαίδευση

Η εκπαίδευση είναι ίσως η πιο αποδοτική και βασική αντιμετώπιση σε επιθέσεις κοινωνικής μηχανικής, μιας και ο αδύναμος κρίκος του συστήματος είναι ο άνθρωπος-χρήστης των πληροφοριακών συστημάτων.

Όπως βλέπουμε και στον Πίνακας 4 κυρίως ασχολείται με την πρόληψη και όχι τόσο με την ανίχνευση και την αντίδραση ενώ το βασικό μέρος εφαρμογής είναι αναμενόμενα οι άνθρωποι. Ο βαθμός προστασίας, η ευκολία εφαρμογής, ο χρόνος εφαρμογής και το κόστος εφαρμογής, μπορούν να ποικίλουν ανάλογα με το πρόγραμμα και τον τρόπο εκπαίδευσης που θα ακολουθήσει η κάθε εταιρεία-οργανισμός.

Υπάρχουν διάφοροι τρόποι εκπαίδευσης, όπου πέρα από μια απλή παρουσίαση ή σεμινάριο, μπορεί να πάρει τη μορφή διαδραστικών παιχνιδιών ή συστημάτων εκπαίδευσης, τα οποία κάνουν τη διαδικασία πιο ενδιαφέρον και ανταμείβουν τον χρήστη στην διαδικασία εκμάθησης.

Η εκπαίδευση έχει δείξει αρκετά καλά αποτελέσματα ως τρόπος προστασίας διότι και μειώνει σε αρκετά ικανοποιητικό βαθμό τις επιθέσεις κοινωνικής μηχανικής [105], αλλά πρέπει να γίνει εμπεριστατωμένα ώστε να έχει τέτοια αποτελέσματα ενώ πρέπει να είναι συνεχής ώστε να καλύψει καινούργιες επιθέσεις και τρόπους εκμετάλλευσης [100]. Το κόστος εφαρμογής ποικίλει ανάλογα με το πρόγραμμα που θα επιλεχθεί.

Τεχνικοί Τρόποι Προστασίας - Μηχανισμοί Sandboxing

Οι μηχανισμοί Sandboxing είναι ένας αρκετά καλός τρόπος προστασίας με χαμηλό κόστος σε σχέση με αυτά που προσφέρει. Όπως βλέπουμε και στο Πίνακας 5 η προστασία που

παρέχεται είναι Μέτρια προς Μεγάλη, εκτός από κάποιες συγκεκριμένες προσεγγίσεις οι οποίες βρίσκονται σε πιο πειραματικό στάδιο. Η ευκολία εφαρμογής είναι Μέτρια προς Μεγάλη μιας και το σύστημα υποστηρίζει τον χρήστη στην λήψη σωστών αποφάσεων ως προς την εκτέλεση εφαρμογών αμφίβολης προέλευσης σε ένα τέτοιο περιβάλλον μέσω UI χωρίς να είναι δυσνόητο και δύσκολο για τον απλό χρήστη. Ο χρόνος εφαρμογής κινείται από μέρες σε μήνες ανάλογα με την προσέγγιση που ακολουθείται και την επιλογή του περιβάλλοντος (διαδεδομένο και γρήγορα προσβάσιμο, είτε πειραματικό-ερευνητικό σε ανάπτυξης) και η προσπάθεια εφαρμογής κινείται σε Λίγο έως Μέτρια. Το κόστος ποικίλει πάλι ανάλογα με την επιλογή του περιβάλλοντος, αλλά κινείται προς το μέτριο.

Όπως βλέπουμε και στην Εικόνα 5 το εύρος των επιθέσεων που αντιμετωπίζονται είναι αρκετά μεγάλο διότι καλύπτονται πολλά βασικά χαρακτηριστικά των επιθέσεων. Αυτό οφείλεται στο γεγονός ύπαρξης ολόκληρου εικονικού συστήματος. Παρόλα αυτά υπάρχουν περιθώρια βελτίωσης στην λειτουργικότητα και τους πόρους που χρησιμοποιούνται στα εικονικά περιβάλλοντα για την αντιμετώπιση των εκάστοτε επιθέσεων.

Οι μηχανισμοί Sandboxing οι οποίοι χρησιμοποιούνται ήδη είναι αρκετά διαδεδομένοι σαν τρόπος ασφάλειας, με αρκετά καλό ποσοστό άμυνας σε επιθέσεις μεγάλου εύρους. Ερευνητικά δεν υπάρχει τόσο μεγάλο ενδιαφέρον σε σχέση με άλλους τρόπους αντιμετώπισης, όπως το Machine Learning ή το Deep Learning, μιας και η τεχνολογία είναι αρκετά συγκεκριμένη σε αυτό που κάνει, αλλά παραμένει ένας από τους πολύ καλούς τρόπους προστασίας και υπάρχουν κενά που πρέπει να καλυφθούν και να βελτιωθούν.

Τεχνικοί Τρόποι Προστασίας - Authorization, Authentication and Accounting (AAA)

Το AAA είναι ένα πλαίσιο που παρέχει αρκετά γενικούς μηχανισμούς ασφάλειας όπως η πολιτική και ο έλεγχος διαδικασιών. Όπως βλέπουμε και στο Πίνακας 6, ο τρόπος αντιμετώπισης είναι κυρίως στην ανίχνευση και στην αντίδραση με Μέτριο βαθμό προστασίας και ευκολία εφαρμογής. Εφαρμόζεται σε συστήματα όπου ο χρόνος εφαρμογής είναι συνήθως ώρες. Η προσπάθεια εφαρμογής είναι μέτρια και το κόστος εφαρμογής μέτριο προς μικρό αν χρησιμοποιούνται λύσεις ανοιχτού κώδικα και το κόστος ανεβαίνει αν χρησιμοποιούνται εμπορικές λύσεις.

Το AAA είναι ένας μέτριος τρόπος προστασίας και αρκετά συγκεκριμένος προς μεγάλους οργανισμούς και όχι για οικιακή χρήση, μιας και χρειάζεται μία κεντρική εποπτική αρχή (ειδικός κυβερνοασφάλειας), ώστε να λειτουργήσει σε ικανοποιητικό βαθμό με λίγο προς μέτριο κόστος.

Τεχνικοί Τρόποι Προστασίας - Monitoring

Το monitoring είναι ένας βασικός μηχανισμός ασφάλειας ο οποίος ανιχνεύει και ταξινομεί τις επιθέσεις στο σύστημα. Όπως βλέπουμε και στο Πίνακας 7 ο τρόπος αντιμετώπισης είναι η ανίχνευση και ο βαθμός προστασίας ενώ η ευκολία εφαρμογής μέτρια προς μεγάλη. Μιας και μιλάμε για κάτι τόσο ευρύ (που μπορεί να καλύπτει ένα ολόκληρο σύστημα) εφαρμόζεται με διάφορους τρόπους (Συστήματα, Μέθοδοι, Επεκτάσεις, κ.λ.π.). Ο χρόνος εφαρμογής εξαρτάται από την λύση και την επιλογή των εργαλείων (Ώρες - Μέρες- Μήνες) αλλά τα αποτελέσματα και η ταξινόμηση προκειμένου να έχουν νόημα καλό είναι να έχουν συλλεχθεί για μεγάλο χρονικό διάστημα. Το κόστος εφαρμογής είναι από μέτριο προς καθόλου μιας και υπάρχουν αρκετά τέτοια προγράμματα δωρεάν (SolarWinds ip Monitor free edition, nmap, κα.).

Το Monitoring είναι ένας πολύ καλός τρόπος αντιμετώπισης επιθέσεων, με αρκετά καλά αποτελέσματα [124], και πολυχρησιμοποιημένο ανά τίς εταιρείες και τους οργανισμούς όπου το χειρίζονται cyber security specialists και τεχνικοί IT. Υπάρχουν ήδη καταξιωμένα προγράμματα (π.χ Wireshark), και υπάρχει αρκετά μεγάλο ερευνητικό ενδιαφέρον.

Τεχνικοί Τρόποι Προστασίας - Integrity Checking

Όπως βλέπουμε και στο Πίνακας 8 οι κύριοι τρόποι αντιμετώπισης εδώ είναι η πρόληψη και η ανίχνευση, μιας και οι χρήστες προειδοποιούνται για τυχόν κακόβουλες ενέργειες ενώ δεν υπάρχει κάποια αντίδραση. Ο βαθμός προστασίας είναι μέτριος διότι η τελική απόφαση πέρα της προειδοποίησης βρίσκεται στην κρίση του χρήστη ενώ η ευκολία εφαρμογής μέτρια εκτός από κάποιες λίγες εξαιρέσεις που μπορεί να ποικίλουν ανάλογα σε ποιό στάδιο βρίσκονται και πόσο πειραματική είναι η εκάστοτε προσέγγιση. Η προσπάθεια εφαρμογής είναι μέτρια συνήθως με κάποιες εξαιρέσεις ανάλογα την έρευνα και τους αλγορίθμους που εξετάζονται ενώ ο χρόνος εφαρμογής μέτριος μιας και εξαρτάται από την υλοποίηση που ακολουθείται.

Τεχνικοί Τρόποι Προστασίας - Machine Learning

Το Machine Learning είναι από τους τρόπους αντιμετώπισης που γίνονται όλο και πιο συχνοί με μεγάλο ερευνητικό ενδιαφέρον. Όπως φαίνεται στο Πίνακας 9 στους τρόπους αντιμετώπισης είναι η ανίχνευση μιας και στις περισσότερες προσεγγίσεις εκπαιδεύονται αλγόριθμοι και έξυπνοι πράκτορες με βάση στατιστικά και υπαρκτές προσεγγίσεις. Ο βαθμός προστασίας, ο χρόνος εφαρμογής και το κόστος εφαρμογής ποικίλουν ανάλογα με την προσέγγιση και τον όγκο των δεδομένων που επιλέγονται κάθε φορά προς μάθηση. Η ευκολία εφαρμογής είναι μέτρια προς μεγάλη μιας και αφού ο αλγόριθμος έχει αποκτήσει την απαραίτητη “γνώση” μπορεί να συμπεριληφθεί σε συστήματα με σχετική ευκολία.

Χρησιμοποιώντας μεταβλητές χαρακτηρισμού με σύνολα δεδομένων εισόδου συμπεριφοράς (που συνήθως συλλέγονται μέσω παρακολούθησης), μπορούν να γίνουν ακριβείς προβλέψεις και μετρήσεις δεικτών ως προς τη σημασία της επίδρασης της συμπεριφοράς ενός αρχείου ή ενός χρήστη σε ένα σύστημα. Ενώ εργαλεία μηχανικής μάθησης έχουν κατασκευαστεί, δοκιμαστεί εκτενώς και αξιολογηθεί στην έρευνα, η εφαρμογή τους έχει επικεντρωθεί σε μεγάλο βαθμό στην αντιμετώπιση επιθέσεων phishing. Δεν υπάρχουν πολλές πραγματικές υλοποιήσεις στο πλαίσιο αναπτυγμένων και δημοφιλών υπηρεσιών ηλεκτρονικού ταχυδρομείου στον εμπορικό χώρο και, κατά συνέπεια, υπάρχει έλλειψη εμπειρικών αξιολογήσεων σε πραγματικά περιβάλλοντα που αντιμετωπίζουν χρήστες.

Το Machine Learning έχει μεγάλο ερευνητικό ενδιαφέρον, μιας και η χρήση και η επινόηση αλγορίθμων μηχανικής μάθησης είναι αρκετά διαδεδομένη τα τελευταία χρόνια. Τα αποτελέσματα είναι πολύ καλά, αλλά υπάρχει χώρος για να βελτιστοποιηθούν περαιτέρω αυτές οι λύσεις.

Τεχνικοί Τρόποι Προστασίας - Deep Learning

Το Deep Learning είναι και αυτό ένας τρόπος αντιμετώπισης με αρκετά μεγάλο ερευνητικό ενδιαφέρον. Με τα νευρωνικά δίκτυα να χρησιμοποιούνται όλο και περισσότερο και σιγά σιγά να αντικαθιστούν τους παραδοσιακούς αλγορίθμους μηχανικής μάθησης οι deep learning προσεγγίσεις γίνονται όλο και συχνότερες.

Όπως βλέπουμε και στο Πίνακας 10 στους τρόπους αντιμετώπισης είναι η ανίχνευση και η πρόληψη, μιας και η προσέγγιση είναι παρόμοια με το Machine Learning με τη διαφορά της χρήσης των νευρωνικών, αντί παραδοσιακών αλγορίθμων. Ο βαθμός προστασίας είναι από μέτριος προς καλός αλλά ακόμα αντιμετωπίζονται αρκετά προβλήματα με την χρήση των νευρωνικών δικτύων ενώ η ευκολία εφαρμογής είναι μέτρια, μιας και αυτή εξαρτάται από τον τρόπο προσέγγισης του προβλήματος και τον όγκο δεδομένων προς μάθηση. Το μέρος εφαρμογής συνήθως είναι συστήματα και μεθοδολογίες ενώ ο χρόνος εφαρμογής είναι από εβδομάδες έως μήνες, μιας και είναι αρκετά χρονοβόρο να εκπαιδευτεί το νευρωνικό δίκτυο με μεγάλο όγκο δεδομένων. Η προσπάθεια εφαρμογής είναι μέτρια και το κόστος μέτριο στα παραδείγματα πού είδαμε στη παρούσα διπλωματική. Οι τιμές αυτές μπορεί να ποικίλουν ανάλογα με το integration (ενσωμάτωση) της κάθε προσπάθειας εφαρμογής.

Το Deep Learning, όπως το Machine Learning, είναι αρκετά διαδεδομένο τα τελευταία χρόνια με μεγάλο ερευνητικό ενδιαφέρον. Τα αποτελέσματα και εδώ είναι μέτρια προς το παρόν, αλλά ενδέχεται να βελτιωθούν περισσότερο λόγω της έντασης της έρευνας που πραγματοποιείται σε αυτό το πεδίο..

Τεχνικοί Τρόποι Προστασίας - Ανίχνευση επιθέσεων βάση σεναρίων

Στις ανιχνεύσεις επιθέσεων βάση σεναρίων χρησιμοποιούνται σενάρια επιθέσεων από τους ερευνητές και ανάλογα με αυτά παράγονται τα αντίστοιχα αποτελέσματα. Από το Πίνακας 11 φαίνεται πως οι τρόποι αντιμετώπισης μπορεί να είναι ανίχνευση, πρόληψη αλλά και αντίδραση. Παρόμοια στον βαθμό προστασίας, στην ευκολία εφαρμογής, στον χρόνο εφαρμογής και στις προσπάθειες εφαρμογής ποικίλουν οι διαφορετικές αξιολογήσεις μιας και εξαρτώνται από το σενάριο που έχει επιλέξει ο ερευνητής.

Η ανίχνευση επιθέσεων βάση σεναρίων έχει μικτά αποτελέσματα, ανάλογα την υλοποίηση και το use case. Αλλά είναι αρκετά καλά τα αποτελέσματα εφόσον εφαρμοστούν σωστά οι αντίστοιχες μέθοδοι σε συγκεκριμένα σενάρια επιθέσεων. Μπορεί να χρησιμοποιηθεί για πιο ειδικές καταστάσεις και ανάγκες οργανισμών, ανάλογα με τα σενάρια που χρησιμοποιούνται ή είναι απαραίτητα προς χρήση. Παρόλα αυτά μπορούμε να χρησιμοποιήσουμε συνδυασμούς ανιχνεύσεων επιθέσεων βάση σεναρίων ώστε να έχουμε καλύτερα αποτελέσματα σε μεγαλύτερη κάλυψη.

Τεχνικοί Τρόποι Προστασίας - Υβριδική Μάθηση

Όπως φαίνεται και στο Πίνακας 12 οι τρόποι αντιμετώπισης μπορεί να είναι η Πρόληψη, η αντίδραση και η Ανίχνευση. Ο βαθμός προστασίας είναι μέτριος προς μεγάλος και η ευκολία εφαρμογής μέτρια. Πάλι εδώ οι τιμές αυτές είναι για τα συγκεκριμένα παραδείγματα και μπορούν να ποικίλουν ανάλογα με το μέρος ενσωμάτωση της λύσης. Το μέρος προσέγγισης είναι Υβριδική Προσέγγιση-Μεθοδολογία και ο χρόνος εφαρμογής κυμαίνεται από μέρες μέχρι εβδομάδες. Η προσπάθεια εφαρμογής είναι μέτρια και το κόστος μπορεί να ποικίλει από λίγο έως πολύ ανάλογα με την προσέγγιση και τον τρόπο που έχει επιλεχθεί από τους ερευνητές.

Η Υβριδική Μάθηση έχει πολύ καλά αποτελέσματα, με μεγάλο ερευνητικό ενδιαφέρον μιας και ενώνει διάφορους τρόπους προσέγγισης και στοχεύει σε αρκετά καλά αποτελέσματα.

Σχολιασμός Πινάκων

Πίνακας 13: Αξιολόγηση Μεθόδου Προστασίας: Πολιτική-Διαδικασία.

Η Πολιτική-Διαδικασία είναι μία γενική προσέγγιση της συνολικής ασφάλειας που ξεχωρίζει σε συνολικές προσεγγίσεις ασφαλείας εταιρειών και οργανισμών. Η απόδοση είναι καλή στις περισσότερες περιπτώσεις με μεγάλο βαθμό προστασίας και μέτρια ευκολία εφαρμογής. Το αρνητικό της συγκεκριμένης μεθόδου είναι ότι είναι αρκετά χρονοβόρα μιας και για να εφαρμοστεί πρέπει να συμμετέχουν όλοι όσοι συμμετέχουν στον οργανισμό.

Πίνακας 14: Αξιολόγηση Μεθόδου Προστασίας: Εκπαίδευση.

Η εκπαίδευση φαίνεται να είναι ο αποτελεσματικότερος τρόπος προστασίας από επιθέσεις κοινωνικής μηχανικής μιας και λόγω της φύσης αυτών των επιθέσεων οι τεχνικοί τρόποι προστασίας θα είναι πάντα λίγο περιορισμένοι. Αρκετές έρευνες έχουν δείξει πολλά υποσχόμενα αποτελέσματα, άμα αυτή η μέθοδος προστασίας εφαρμοστεί σωστά σε έναν οργανισμό. Ώς τρόπο αντιμετώπισης στοχεύει κυρίως στην πρόληψη και το Κόστος εφαρμογής είναι Μικρό. Η ευκολία εφαρμογής είναι μέτρια μιας και το “κόστος” είναι οι ώρες που θα αφιερωθούν για την εκπαίδευση και δεν υπάρχουν “τεχνικές” δυσκολίες.

Από τους τεχνικούς τρόπους ο καθένας έχει τα δυνατά και τα τρωτά του σημεία.

Πίνακας 15: Αξιολόγηση Μεθόδων Sandboxing.

Το SandBoxing είναι ένας διαδεδομένος τρόπος προστασίας που χρησιμοποιείται από πολλούς οργανισμούς και συστήματα. Για παράδειγμα τα Windows όταν ο χρήστης θέλει να τρέξει μία εφαρμογή ή ένα executable αμφιβόλου προέλευσης, παροτρύνουν τον χρήστη να τα τρέξει σε ένα περιβάλλον sandbox.

Σαν τρόπος αντιμετώπισης κυρίως μιλάμε για αντίδραση, με μέτριο προς μεγάλο βαθμό προστασίας, μεγάλη ευκολία εφαρμογής, μιας και έχουν αναπτυχθεί πολλές υποδομές και blueprints για sandboxing περιβάλλοντα. Το μέρος εφαρμογής είναι τα συστήματα και το κόστος εφαρμογής μέτριο.

Πίνακας 16: Αξιολόγηση Μεθόδων Προστασίας AAA.

Το AAA είναι και αυτό ένας αποτελεσματικός τρόπος αντιμετώπισης και οι περισσότεροι οργανισμοί τον έχουν υιοθετήσει μέσω κυρίως εφαρμογών τρίτων μερών (π.χ. VIP Access για είσοδο στο VPN του εκάστοτε οργανισμού).Η ευκολία εφαρμογής είναι μέτρια, μιας και αφού κατασκευαστεί η υποδομή, είναι πολύ εύκολη η ενοποίηση για πολλά άτομα στον οργανισμό. Το κόστος εφαρμογής είναι μέτριο προς λίγο.

Πίνακας 17: Αξιολόγηση Μεθόδων Προστασίας Παρακολούθηση-Monitoring.

Το Monitoring είναι από τους αποτελεσματικότερους τρόπους προστασίας, αν υπάρχει στον οργανισμό κάποιος cyber-security specialist ο οποίος διαχειρίζεται το δίκτυο και την κίνηση του και έχει τις κατάλληλες γνώσεις για να πάρει τις κατάλληλες αποφάσεις. Ο τρόπος αντιμετώπισης είναι η ανίχνευση και η ευκολία μέτρια προς αρκετά εύκολη, μιας και υπάρχουν εργαλεία ανοικτού κώδικα και αυτό που χρειάζεται είναι specialists που μπορούν να χρησιμοποιήσουν και να καταλάβουν αυτά τα εργαλεία. Ο χρόνος εφαρμογής είναι από ώρες μέχρι εβδομάδες, ανάλογα με το αν χρειάζεται κάποια ειδική λύση για τον εκάστοτε οργανισμό και το κόστος εφαρμογής λίγο έως μέτριο.

Πίνακας 18: Αξιολόγηση Μεθόδων Προστασίας Integrity Checking.

Το Integrity Checking ως τρόπος αντιμετώπισης εστιάζει στην ανίχνευση και την πρόληψη με μέτριο βαθμό προστασίας και μέτρια προς αρκετή ευκολία εφαρμογής. Ο χρόνος εφαρμογής είναι κυρίως μήνες, αλλά αυτό γίνεται γιατί στις λύσεις που αναλύσαμε στην παρούσα διπλωματική δεν υπήρχε κάποια υποδομή. Το κόστος εφαρμογής είναι Μέτριο.

Πίνακας 19: Αξιολόγηση Μεθόδων Προστασίας Deep Learning.

Το Deep Learning ως τρόπος αντιμετώπισης εστιάζει στην ανίχνευση και την πρόληψη. Ο βαθμός προστασίας ποικίλει μιας και εξαρτάται από το μοντέλο που έχει δημιουργηθεί και τα δεδομένα που έχουν χρησιμοποιηθεί για την δημιουργία του. Ο χρόνος εφαρμογής και εδώ ποικίλει μιας και ο χρόνος που χρειάζεται για να μαζευτούν όλα τα απαραίτητα δεδομένα δεν είναι σταθερός, αλλά αλλάζει. Στην πλειοψηφία η προσπάθεια εφαρμογής και το κόστος εφαρμογής είναι μέτρια.

Πίνακας 20: Αξιολόγηση Μεθόδων Προστασίας Scenario-Based.

Το Scenario-Based είναι πολύ αποτελεσματικό για συγκεκριμένες επιθέσεις εφόσον ο οργανισμός ξέρει από ποιά συγκεκριμένη επίθεση - σενάριο θέλει να προστατευτεί. Ο τρόπος αντιμετώπισης είναι η ανίχνευση, η πρόληψη και η αντίδραση. Ο βαθμός προστασίας είναι μικρός προς μέτριος, αλλά αυτό συμβαίνει γιατί οι ερευνητές χρησιμοποίησαν αυτόν τον τρόπο αντιμετώπισης σε πιο πειραματικές καταστάσεις και η ευκολία εφαρμογής και το κόστος εφαρμογής είναι μέτρια.

Πίνακας 21: Αξιολόγηση Μεθόδων Προστασίας Hybrid.

Οι Hybrid μέθοδοι προστασίας είναι αρκετά αποτελεσματικές επίσης αλλά με λίγο μεγαλύτερο κόστος εφαρμογής λόγω της τεχνογνωσίας και των υποδομών που χρειάζεται. Σαν τρόπο αντιμετώπισης έχουμε την πρόληψη, την αντίδραση και την ανίχνευση με μέτριο προς αρκετό βαθμό προστασίας και προσπάθεια εφαρμογής και κόστος εφαρμογής μέτριο.

<u>Παροχή Καθολικών αποτελεσμάτων</u>

Η Μηχανική Μάθησης, το Deep Learning και οι Hybrid μέθοδοι αντιμετώπισης είναι πολλά υποσχόμενες, μιας και εφόσον εφαρμοστούν σωστά μπορούν να ανιχνεύσουν και να αντιμετωπίσουν μεγάλο μέρος επιθέσεων, αλλά ακόμα υπάρχουν ερευνητικές δυσκολίες και κενά. Πιό “παραδοσιακές” μέθοδοι όπως το sandboxing, το AAA είναι αρκετά αποτελεσματικές και στα ανάλογα use cases. Η εκπαίδευση και η πολιτική και οι διαδικασίες είναι πολλά υποσχόμενες όσο περισσότεροι άνθρωποί εξοικειώνονται και κατανοούν βασικές αρχές της επιστήμης της πληροφορικής, στο να αποτρέψουν επιθέσεις κοινωνικής μηχανικής.

Σχολιασμός Κατηγοριοποίησης προσεγγίσεων με βάση την δεύτερη κατηγοριοποίηση

Εικόνα 5: Κατηγοριοποίηση προσεγγίσεων με βάση την δεύτερη κατηγοριοποίηση.

Στην φάση του orchestration οι καλύτερες λύσεις ως προς την κάλυψη των σχετικών κατηγοριών φαίνεται να είναι το Machine Learning και οι Scenario-Based με το Sandboxing, το Deep Learning και το Hybrid να ακολουθούν.

Στην φάση του exploitation οι καλύτερες λύσεις ως προς την κάλυψη των σχετικών κατηγοριών φαίνεται να είναι το Machine Learning με την Εκπαίδευση και το Sandboxing να ακολουθούν.

Στην φάση του execution οι καλύτερες λύσεις φαίνεται να είναι το AAA και το SandBoxing, με το Deep Learning και το Hybrid να ακολουθεί.

Το είδος μεθόδου προστασίας που καλύπτει καλύτερα τα περισσότερα χαρακτηριστικά των επιθέσεων κοινωνικής μηχανικής είναι το Sandboxing και το Machine Learning. Ακολουθούν το Deep Learning και οι Hybrid Προσεγγίσεις.

Το Hybrid περιλαμβάνει πολλούς τρόπους προστασίας που μπορεί να καλύπτουν διαφορετικά χαρακτηριστικά. Το Machine Learning λόγω της μεγάλης ερευνητικής του δράσης καλύπτει και αυτό μεγάλο όγκο κριτηρίων. Το sandboxing είναι ένας αρκετά "παραδοσιακός" τρόπος να απομονώσεις το περιβάλλον σου με σκοπό να "τρέξεις " με ασφάλεια ενδεχομένως κακόβουλες εφαρμογές και το deep learning είναι πολλά υποσχόμενο και με μεγάλη άνοδο τα τελευταία χρόνια.

Παρατηρούμε ότι στο orchestration δεν καλύπτονται καθόλου οι TD2, MD2, MD3 και MA1 ενώ καλύπτεται περισσότερο η MD1-R από την MD1. Στο Exploitation καλύπτεται λιγότερο η DV2 και περισσότερο οι DV1 και IM2. Στο execution καλύπτεται λιγότερο η AP2 και περισσότερο η ES1.

Το είδος μεθόδου προστασίας που έχει την μεγαλύτερη κάλυψη στη φάση της εκμετάλλευσης ως προς τις μεθόδους προστασίας που προτείνονται σε αυτό είναι το Machine Learning, το sandboxing με το AAA στην φάση της εκτέλεσης, ενώ στη φάση της ενορχήστρωσης ξεχωρίζουν ελάχιστα οι Machine Learning and scenario-based αλλά όπως αναφέρθηκε δεν μπορούμε να προστατευτούμε τόσο καλά σε αυτό το στάδιο όπου η επίθεση δεν έχει αρχίσει ενεργά. Επομένως, ανάλογα με το στάδιο της επίθεσης, κάποιοι τρόποι προστασίας είναι καλύτεροι από κάποιους άλλους. Συνεπώς, η καλύτερη ολική προσέγγιση είναι ένας συνδυασμός πολλών μεθόδων προστασίας από πολλά είδη. Σημειώνουμε, βέβαια, πως το Machine Learning φαίνεται πως υπερέχει σε 2 στάδια, οπότε λογικά η βάση μιας ολικής προσέγγισης θα πρέπει να είναι αυτό.

5. Προκλήσεις και Ερευνητικά Κενά

Οι οργανισμοί και οι επιχειρήσεις κατανοούν τη σημασία της ασφάλειας στον κυβερνοχώρο αλλά τείνουν να κατατάσσουν άλλους παράγοντες όπως την ευκολία χρήσης, την αισθητική και την απόδοση πολύ πιο μπροστά από τις ανησυχίες για την ασφάλεια. Η ευκολία χρήσης και η αισθητική είναι και αυτοί παράγοντες προσέλκυσης πελατών, κάτι το οποίο δεν μπορεί να αγνοηθεί. Το ζητούμενο όμως είναι να υπάρχει κάποιο trade-off μεταξύ των δύο (Ασφάλεια-Αισθητική/Ευκολία Χρήσης/Απόδοση) ώστε να δοθεί η απαραίτητη σημασία και στην ασφάλεια. Θα ήταν, επομένως, πολύ χρήσιμο να διεξαχθεί περαιτέρω έρευνα σχετικά με την κατανόηση της ψυχολογίας των διαδικτυακών καταναλωτών σχετικά με την αισθητική, την ευκολία χρήσης και απόδοση που προτιμούν, αλλά και ως προς το αν κατανοούν τη σημασία της ασφάλειας. Επίσης, η έρευνα πρέπει να στοχεύει στην αξιολόγηση του κατά πόσον ο χρήστης αφήνει τα χαρακτηριστικά ασφαλείας (Firewall, Πρόσθετα ασφάλειας για τον φυλλομετρητή, Email monitor, κ.λ.π) ενεργοποιημένα όταν του δίνεται η επιλογή να τα απενεργοποιήσει, μιας και η περαιτέρω πληροφορία και λογισμικό που

χρησιμοποιείται μπορεί να χαμηλώσει τις επιδόσεις του υπολογιστή του και να κουράσει τον χρήστη με “αχρείαστες” για αυτόν περαιτέρω πληροφορίες.

Ένα σημαντικό ερευνητικό κενό υπάρχει μεταξύ της έρευνας και της βιομηχανίας "όσον αφορά τα πραγματικά θετικά αποτελέσματα". Η βιομηχανία προσεγγίζει πιο ρεαλιστικές λύσεις, ενώ ερευνητικά χρησιμοποιούνται και δημιουργούνται αρκετοί μέθοδοι και εργαλεία που δεν έχουν απαραίτητα καλά αποτελέσματα αλλά ερευνητικό ενδιαφέρον. Η βιομηχανία και η έρευνα είναι αναγκαίες και οι δυο, μιας και στην βιομηχανία φαίνεται αν τα εργαλεία που χρησιμοποιούνται είναι όντως πρακτικά και ασφαλή για πραγματικές καταστάσεις ενώ η έρευνα προχωράει τον κλάδο και βρίσκει νέους και αποτελεσματικότερους τρόπους υπό προϋποθέσεις, για την αντιμετώπιση των επιθέσεων αυτών. Ενώ οι ακαδημαϊκές και βιβλιογραφικές έρευνες επικεντρώνεται ουσιαστικά στη μηχανική μάθηση και τις ευρετικές μεθόδους, υποθέτοντας θετικά αποτελέσματα, αυτά τα αληθινά θετικά αποτελέσματα είναι μερικές φορές υψηλά ψευδώς θετικά μιας και οι επιθέσεις και οι προϋποθέσεις που υπάρχουν για το εκάστοτε πείραμα δεν αντικατοπτρίζουν πάντα την πραγματικότητα. Ωστόσο, η βιομηχανία, λόγω της ύπαρξης αληθινών δεδομένων είναι δυνατόν να ελεγχθεί με τον καλύτερο δυνατό τρόπο η αποτελεσματικότητα και απόδοση μιας μεθόδου προστασίας. Ακόμα δεν υπάρχει μία ολοκληρωμένη λύση για το phishing σε κανένα στάδιο ελέγχου (Ενορχήστρωση, Εκμετάλλευση, Εκτέλεση). Οι ενδιαφερόμενοι σε κάθε στάδιο του κύκλου ζωής του phishing πρέπει να καταβάλουν συγκεκριμένες προσπάθειες για να αντιμετώπιση τέτοιων επιθέσεων ανάλογα με το στάδιο ελέγχου της επίθεσης και να δρομολογήσουν καινοτόμες νέες πρακτικές για να αποτρέψουν την απειλή.

Πολλές εταιρείες και οργανισμοί αναπτύσσουν σχέδια και στρατηγικές, καθώς και ξοδεύουν τεράστια ποσά για να εξαλείψουν τις επιθέσεις κοινωνικής μηχανικής, αλλά εξακολουθούν να υπάρχουν πολλοί περιορισμοί και μη αποτελεσματικά μέτρα που μπορεί να αντιμετωπίσουν κατά την εφαρμογή για την αντιμετώπιση των επιθέσεων κοινωνικής μηχανικής λόγω της εξάρτησης αυτών των επιθέσεων στη χειραγώγηση φυσικών καταστάσεων και την ψυχολογική εκμετάλλευση των ατόμων. Για παράδειγμα, οι περιορισμοί αυτοί μπορεί να αντιπροσωπεύονται στο διαφορετικό επίπεδο ευαισθητοποίησης, κουλτούρας και εκπαίδευσης των ατόμων και των εργαζομένων σε ιδρύματα σχετικά με τις επιθέσεις αυτές όταν εκτίθενται σε αυτές. Την εμπειρογνωμοσύνη των επιτιθέμενων και τη συνεχή εξέλιξη των τεχνικών και των μεθόδων που χρησιμοποιούν για να πραγματοποιήσουν τέτοιες επιθέσεις. Αύξηση των ανθρώπινων αδυναμιών και της ικανότητας ελέγχου και χειραγώγησής τους, μερικά από τα μειονεκτήματα των εργαλείων που χρησιμοποιούν οι οργανισμοί για να εντοπισμό και την αποτροπή επιθέσεων κοινωνικής μηχανικής, όπως το υψηλό κόστος, η πολυπλοκότητα και η πιθανότητα λάθους και τα περιορισμένα εργαλεία που χρησιμοποιούνται για την εξάλειψη αυτών των επιθέσεων, εκτός από τα ανθρώπινα λάθη που μπορεί να προκύψουν από την εφαρμογή τους από ορισμένους υπαλλήλους. Ως εκ τούτου, χρειαζόμαστε πιο προηγμένα, αποτελεσματικά μέτρα ως προς την ψυχολογία των εργαζομένων και τις αντιδράσεις τους σε τέτοιες επιθέσεις και τεχνολογίες ασφαλείας για τον περιορισμό αυτών των επιθέσεων οι οποίες στοχεύουν στις συχνότερες συμπεριφορές.

Όσον αφορά το Deep Learning και το Machine Learning είναι τεχνολογίες οι οποίες έχουν μεγάλη άνοδο τα τελευταία χρόνια και βρίσκονται ολοένα και αποτελεσματικότεροι τρόποι να εφαρμοστούν. Θα πρέπει να εξετάζονται συνεχώς νέοι αλγόριθμοι μηχανικής μάθησης και αποτελεσματικοί τρόποι να εφαρμόζονται αυτές οι τεχνολογίες στον τομέα της ασφάλειας και ειδικότερα για την αντιμετώπιση των επιθέσεων κοινωνικής μηχανικής.

6. Μελλοντικές Κατευθύνσεις

- Είναι σημαντικό να εκπαιδεύονται συνεχώς οι άνθρωποι σχετικά με τους τρόπους που μπορούν να ασφαλίζουν τους λογαριασμούς τους και να μειώσουν την πιθανότητα να γίνουν θύματα επιθέσεων κοινωνικής μηχανικής, μέσα από σεμινάρια.

- Να εκπαιδεύονται οι εργαζόμενοι, ειδικά σε μία περίοδο πανδημίας που η δουλειά εξ αποστάσεως γίνεται συχνότερη, ώστε να αναφέρουν περιστατικά όπου εκτίθενται σε τέτοιες επιθέσεις, στην εταιρία που εργάζονται.

- Οι εταιρίες και οι οργανισμοί να παρέχουν εκπαιδευτικά προγράμματα και σεμινάρια για το προσωπικό τους, όπου αναδεικνύονται διάφορες τεχνικές και επιθέσεις κοινωνικής μηχανικής και τρόποι αποφυγής τους, μέσα από την πλήρη κατανόηση τους και την λήψη κατάλληλων προφυλάξεων.

- Συνεχής ενημέρωση για πρόσφατες τεχνολογικές εξελίξεις και καινούργιες επιθέσεις ή καινούρια είδη επιθέσεων κοινωνικής μηχανικής.

- Συνεχής ενημέρωση λογισμικού και προγραμμάτων ασφαλείας.

- Εφαρμογή ολοκληρωμένης στρατηγικής και πολιτικής ασφάλειας σε μία εταιρία.

- Η χώρα να εφαρμόσει αυστηρούς νόμους για το έγκλημα στον κυβερνοχώρο.

- Να υπάρχει σε κάθε οργανισμό κάποιος αναλυτής κυβερνο-ασφάλειας (cyber security analyst), ο οποίος πέρα από την διενέργεια παρακολούθησης (monitoring) του συστήματος του οργανισμού και την υλοποίηση βασικών διαδικασιών ασφάλειας θα επιμορφώνει συνεχώς το προσωπικό με παρόμοιους τρόπους που προαναφέρθηκαν και στο Κεφάλαιο 4. Ανάλυση ([97], [98], [100]).

- Διενέργεια έρευνας για την παραγωγή νέων ή την πλήρη και κατάλληλη διαμόρφωση υπαρχόντων αλγορίθμων μηχανικής μάθησης, οι οποίοι θα είναι αποτελεσματικότεροι και στοχευμένοι σε προβλήματα ασφάλειας και ειδικότερα επιθέσεων κοινωνικής μηχανικής.

- Μελέτη & εφαρμογή υβριδικών λύσεων και μεθόδων ασφάλειας στοχευμένων στα συγκεκριμένα προβλήματα και κενά ασφαλείας του εκάστοτε οργανισμού.

- Η μελλοντική έρευνα που αφορά τις επιθέσεις κοινωνικής μηχανικής μπορεί να εξετάσει ψηφιακές υπογραφές ή πιστοποιητικά που προέρχονται από ένα κεντρικό σύστημα ελέγχου ταυτότητας για την απόδειξη της νομιμότητας των δεδομένων που παρουσιάζονται σε ένα αρχείο ή μια εφαρμογή (ιστότοπος, ηλεκτρονικό ταχυδρομείο, βοηθητικά προγράμματα), όπου τα δεδομένα μεταφορτώνονται, αναλύονται και εγκρίνονται, έτσι ώστε σε περίπτωση που τα δεδομένα λείπουν από αυτή την κεντρική αρχή ή χρησιμοποιούν άγνωστα πιστοποιητικά, να υπάρχει μειωμένη λειτουργικότητα έως ότου επαληθευτούν κατάλληλα για ασφαλή λειτουργία στο αντίστοιχο περιβάλλον(τομέας, κοινή πλατφόρμα/επιφάνεια εργασίας, φάκελος, αποθήκευση στο νέφος κ.λπ.).

7. Συμπεράσματα

Οι επιθέσεις κοινωνικής μηχανικής αυξάνονται ραγδαία και αποτελούν σοβαρή απειλή για την ασφάλεια των εταιρειών και οργανισμών συνεχώς ενώ προκαλούν σημαντική συναισθηματική και οικονομική ζημία με διάφορους τρόπους. Ως εκ τούτου, είναι σημαντικό να κατανοήσουμε τις μεθόδους που χρησιμοποιούν οι επιτιθέμενοι για την εφαρμογή των επιθέσεων αυτού του είδους, προκειμένου να χρησιμοποιήσουμε τα κατάλληλα μέτρα ασφάλειας για την προστασία των εταιρειών και των οργανισμών σε συνεχή βάση.

Στην παρούσα εργασία, παρέχουμε μια ολοκληρωμένη εξήγηση της έννοιας των επιθέσεων κοινωνικής μηχανικής, τα κύρια στάδια της υλοποίησής τους, καθώς και μια λεπτομερή ταξινόμησή τους σύμφωνα με τις διαστάσεις 2 διαφορετικών κατηγοριοποιήσεων, όπου και αυτές αναλύονται σε αυτή την αναφορά. Επιπλέον, παρέχουμε μια λεπτομερή εξήγηση των διαφορετικών και κοινών τύπων αυτών των επιθέσεων που χρησιμοποιούνται με στόχο την εξαπάτηση του θύματος για την ανάκτηση των απαιτούμενων πληροφοριών χωρίς τη γνώση του, όπως κωδικών πρόσβασης και αριθμών τραπεζικών λογαριασμών, γνωρίζοντας ότι μπορεί να υπάρχουν περισσότερες τεχνικές και άλλες μέθοδοι στις οποίες καταφεύγουν οι επιτιθέμενοι για την πραγματοποίηση των επιθέσεων τους. Όλα αυτά θα βοηθήσουν στη διάδοση της ευαισθητοποίησης σχετικά με αυτές τις επιθέσεις μεταξύ των ατόμων και των εργαζομένων σε ιδρύματα, καθώς και θα τους δώσουν τη δυνατότητα να αντιμετωπίσουν και να ανταποκριθούν κατάλληλα σε αυτές τις επιθέσεις, αν εκτεθούν σε αυτές, γεγονός που θα βοηθούσε στη μείωση των εγκλημάτων στον κυβερνοχώρο και την δημιουργία ενός πιο ασφαλούς περιβάλλοντος.

Επιπρόσθετα, αναλύσαμε διάφορες τεχνικές μετριασμού και μέτρα που απαιτούνται για την αποφυγή και την καταπολέμηση αυτών των επιθέσεων ως ένα βήμα αντιμετώπισή τους, διευκολύνοντας την ανάπτυξη αντιμέτρων και τη διεξαγωγή περαιτέρω έρευνας σε αυτόν τον τομέα. Τα τεχνικά αντίμετρα έχουν παραδοσιακά δυσκολευτεί να αντιμετωπίσουν ένα μεγάλο εύρος τύπων επιθέσεων κοινωνικής μηχανικής, ιδίως τις νέες και πιο δυναμικές. Δεν υπάρχει αμφιβολία ότι ο έλεγχος της πολιτικής, η εκπαίδευση των χρηστών μέσω εκστρατειών κατάρτισης και ευαισθητοποίησης καθώς και άλλα αντίμετρα διαχείρισης των χρηστών μπορούν να είναι αποτελεσματικά. Πιθανόν να είναι επωφελής ο συνδυασμός των τεχνικών αντιμέτρων και αντίμετρων διαχείρισης χρηστών σε ένα ενιαίο αυτόνομο σύστημα, στο οποίο οι χρήστες θα έχουν προληπτικό προφίλ όσον αφορά την ευαισθησία τους σε συγκεκριμένες επιθέσεις και θα τους χορηγούνται ανάλογα και δυναμικά δικαιώματα. Αυτό είναι ένα σημαντικό αποτέλεσμα της ανάλυσης που επιτελέσθηκε σε αυτή την αναφορά, η οποία βασίστηκε στην δεύτερη κατηγοριοποίηση των επιθέσεων κοινωνικής μηχανικής για την αποτίμηση της κάλυψής τους και σε ένα σύνολο από κριτήρια σύγκρισης των μεθόδων/αντιμέτρων προστασίας.

Προς αυτό το όραμα, είναι σημαντικό να έχουμε μια συνολική εικόνα του τοπίου απειλών από επιθέσεις κοινωνικής μηχανικής. Η δεύτερη κατηγοριοποίηση που παρέχεται σε αυτή την αναφορά και έχει προταθεί από τους Ryan Heartfield και George Loukas είναι ένα βήμα προς την αντιμετώπιση αυτής της πρόκλησης. Εισάγει μια δομημένη βάση για την ταξινόμηση των σημασιολογικών επιθέσεων, αναλύοντάς τις στα συστατικά τους στοιχεία και επιτρέπει έτσι τον εντοπισμό αντιμέτρων που εφαρμόζονται σε μια σειρά διαφορετικών επιθέσεων που μοιράζονται ένα υποσύνολο των χαρακτηριστικών τους.

Παρά ταύτα, καταλήξαμε στο συμπέρασμα ότι ακόμη και όταν χρησιμοποιείται το καλύτερο και πιο ακριβό λογισμικό ασφαλείας, εμείς ως άτομα, οργανισμοί ή εταιρείες είμαστε πάντα ευάλωτοι σε επιθέσεις κοινωνικής μηχανικής λόγω των ανθρωπίνων συναισθημάτων και τη ψυχολογία, ώστε να αντιμετωπίζουμε με ψυχραιμία και αποφασιστικότητα τις επιθέσεις αυτού του είδους που συχνά βασίζονται στην εκμετάλλευση και τη χειραγώγηση του ψυχολογικού παράγοντα του ανθρώπου στοιχείου για να κερδίσουν την εμπιστοσύνη του ή να τον αναγκάσουν να πραγματοποιήσει μη επιθυμητές ενέργειες, χωρίς την ανάγκη για τεχνική εμπειρογνωμοσύνη και εμπειρογνωμοσύνη σε θέματα ασφάλειας.

Ως εκ τούτου, εμείς πρέπει να διαδώσουμε την ευαισθητοποίηση και τη συνεχή κατάρτιση μεταξύ των μελών της κοινωνίας για να γνωρίζουν αυτές τις επιθέσεις με στόχο τη λήψη των απαραίτητων μέτρων ασφαλείας και προφυλάξεων για την αποτροπή αυτών των επιθέσεων αλλά και την ανίχνευσή/εντοπισμό τους όταν αυτές διενεργούνται.

8. Βιβλιογραφία

[1] K. Krombholz, H. Hobel, M. Huber, and E. Weippl, “Social engineering attacks on the knowledge worker,” SIN 2013 - Proc. 6th Int. Conf. Secur. Inf. Networks, no. November, pp. 28–35, 2013, doi: 10.1145/2523514.2523596.]

[2] J. D. Chenoweth, “Book Review: The Art of Deception: Controlling the Human Element of Security,” J. Inf. Priv. Secur., vol. 1, no. 2, pp. 69–70, 2005, doi: 10.1080/15536548.2005.10855769.

[3] F. L. Greitzer, J. R. Strozer, S. Cohen, A. P. Moore, D. Mundie, and J. Cowley, “Analysis of Unintentional Insider Threats Deriving from Social Engineering Exploits,” 2014, doi: 10.1109/SPW.2014.39.

[4] “Social Engineering Fundamentals, Part I: Hacker Tactics.” https://www.socialengineer.org/wiki/archives/PenetrationTesters/Pentest-HackerTactics.html (accessed Aug. 06, 2021).

[5] K. Krombholz, H. Hobel, M. Huber, and E. Weippl, “Social engineering attacks on the knowledge worker,” SIN 2013 - Proc. 6th Int. Conf. Secur. Inf. Networks, no. November, pp. 28–35, 2013, doi: 10.1145/2523514.2523596.

[6] F. Mouton, M. M. Malan, L. Leenen, and H. S. Venter, “Social engineering attack framework,” 2014 Inf. Secur. South Africa - Proc. ISSA 2014 Conf., no. August, 2014, doi: 10.1109/ISSA.2014.6950510.

[7] J. Van De Merwe and F. Mouton, “Mapping the Anatomy of Social Engineering Attacks to the Systems Engineering Life Cycle,” Proc. Elev. Int. Symp. Hum. Asp. Inf. Secur. Assur., no. HAISA, pp. 24–40, 2017.

[8] A. Sood and R. Enbody. 2014. Targeted Cyber Attacks: Multi-staged Attacks Driven by Exploits and Malware. Syngress

[9] I. Rouf, R. Miller, H. Mustafa, T. Taylor, S. Oh, W. Xu, M. Gruteser, W. Trappe, and I. Seskar. 2010. Security and privacy vulnerabilities of in-car wireless networks: A tire pressure monitoring system case study. In Proceedings of the 19th USENIX Security Symposium.

[10] CESG. 2015. Common Cyber Attacks: Reducing the Impact. Retrieved from https://www.gov.uk/government/ uploads/system/uploads/attachment_data/file/400106/Common_Cyber_Attacks-Reducing_The_Impact. pdf.

[11] H. Orman. 2009. The compleat story of phish. IEEE Internet Computing 17, 1, 87–91.

[12] B. Chaffin. 2014. Someone Targets Hong Kong Protesters Using Jailbroken iPhones with Malware. Retrieved from http://www.macobserver.com/tmo/article/someone-targets-hong-kong-protesters-using-jailbrokeniphones-with-malware.

[13] M. T. Banday, J. A. Qadri, and N. A. Shah. 2009. Study of Botnets and Their Threats to Internet Security. Retrieved from http://sprouts.aisnet.org/594/1/Botnet_Sprotus.pdf

[14] V. Sharma. 2011. An analytical survey of recent worm attacks. In IJCSNS(11), Vol. 11, 99–103.

[15] H. Hwang, G. Jung, K. Sohn, and S. Park. 2008. A study on MITM (man in the middle) vulnerability in wireless network using 802.1 X and EAP. In Information Science and Security (ICISS). IEEE, 164–170.

[16] J. M. Briones, M. A. Coronel, and P. Chavez-Burbano. 2013. Case of study: Identity theft in a university WLAN evil twin and cloned authentication web interface. In Proceedings of the 2013 World Congress on Computer and Information Technology (WCCIT'13). IEEE, 1–4.

[17] D. Fisher. 2015. Massive, Decades-Long Cyber Espionage Framework Uncovered. Retrieved from http:// threatpost.com/massive-decades-long-cyberespionage-framework-uncovered/111080d.

[18] T. M. Chen. 2003. Trends in viruses and worms. The Internet Protocol Journal 6, 3, 23–33.

[19] P. Ducklin. 2014. Anatomy of an Android SMS Virus—Watch Out for Text Messages, Even from Your Friends! Retrieved from https://nakedsecurity.sophos.com/2014/06/29/anatomy-of-an-android-sms-virus-watchout-for-text-messages-even-from-your-friends/.

[20] N. P. P. Mavromatis and M. A. R. F. Monrose. 2008. All your iframes point to us. In USENIX Security Symposium. USENIX, 1–16.

[21] D. Emm. 2005. The changing face of malware. In Proceedings of the IWWST.

[22] L. Corrons. 2010. The business of rogueware. In Web Application Security, vol. 72. 7.

[23] G. S. Bindra. 2011. Masquerading as a trustworthy entity through portable document file (PDF) format. In Privacy, Security, Risk and Trust (PASSAT). IEEE, 784–789.

[24] N. Leavitt. 2005. Instant messaging: A new target for hackers. Computer 38, 7, 20–23.

[25] T. Lauinger, V. Pankakoski, D. alzarotti, and E. Kirda. 2010. Honeybot, your man in the middle for automated social engineering. In Proceedings of the 3rd USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET'10).

[26] R. Heartfield and G. Loukas. 2013. On the feasibility of automated semantic attacks in the cloud. In Computer and Information Sciences III. Springer, London, 343–351.

[27] Z. Li, K. Zhang, Y. Xie, F. Yu, and X. Wang. 2012. Knowing your enemy: Understanding and detecting malicious web advertising. In Proceedings of the 2012 ACM Conference on Computer and Communications Security. ACM.

[28] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu. 2011. The socialbot network: When bots socialize for fame and money. In Proceedings of the 27th Annual Computer Security Applications Conference. ACM, 93–102.

[29] Z. Coburn and G. Marra. 2008. Realboy Believable Twitter Bots. Retrieved from http://ca.olin.edu/ 2008/realboy/.

[30] KeeLog. 2015. KeeLog Key Grabber Internal Module PS2 2GB. Retrieved from https://www.keelog.com/.

[31] B. Anderson and B. Anderson. 2010. Seven Deadliest USB Attacks. Syngress

[32] M. E. Johnson, D. McGuire, and N. D. Willey. 2008. The evolution of the peer-to-peer file sharing industry and the security risks for users. In Proceedings of the 41st Annual Hawaii International Conference on System Sciences. IEEE, 383–383.

[33] S. Shin, J. Jung, and H. Balakrishnan. 2006. Malware prevalence in the KaZaA file-sharing network. In Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement. ACM.

[34] R. Heartfield and G. Loukas. 2013. On the feasibility of automated semantic attacks in the cloud. In Computer and Information Sciences III. Springer, London, 343–351.

[35] K. P. Yee. 2005. Guidelines and Strategies for Secure Interaction Design. Chapter 13, 247–273. Retrieved from http://sid.toolness.org/ch13yee.pdf.

[36] C. E. Drake, J. O. Jonathan, and J. K. Eugene. 2004. Anatomy of a phishing email. In CEAS.

[37] R. Dhamija, D. J. Tygar, and M. Hearst. 2006. Why phishing works. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM.

[38] TrendMicro. 2014. Malaysia Airlines Flight 370 News Used To Spread Online Threats. Retrieved from http://blog.trendmicro.com/trendlabs-security-intelligence/malaysia-airlines-flight-370-news-used -to-spread-online-threats/.

[39] Y. Song, C. Yang, and G. Gu. 2010. Who is peeping at your passwords at Starbucks? To catch an evil twin access point. In Proceedings of the 2010 IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'10). IEEE, 323–332.

[40] A. P. Felt and D. Wagner. 2011. Phishing on Mobile Devices. In W2SP

[41] N. Leavitt. 2005. Instant messaging: A new target for hackers. Computer 38, 7, 20–23.

[42] H. Xu, N. Wang, and J. Grossklags. 2011. Third-party apps on Facebook: Privacy and the illusion of control. In Proceedings of the 5th ACM Symposium on Computer Human Interaction for Management of Information Technology. ACM.

[43] L. Corrons. 2010. The business of rogueware. In Web Application Security, vol. 72. 7.

[44] J. Giles. 2010. Scareware the inside story. New Scientist , Article 205, 2753, 38–41.

[45] Pierluigi Paganini. 2014. Phishing goes mobile with cloned banking app into Google Play. Retrieved from http://securityaffairs.co/wordpress/26134/cyber-crime/phishing-goes-mobile-cloned-banking-app-googleplay.html.

[46] K. Nohl and J. Lehl. 2014. BadUSBOn accessories that turn evil. In Black Hat USA.

[47] D. Sullivan. 2008. What Is Search Engine Spam? The Video Edition, url =. (2008).

[48] M. Cova, C. Kruegel, and G. Vigna. 2010. Detection and analysis of drive-by-download attacks and malicious JavaScript code. In Proceedings of the 19th International Conference on World Wide Web. ACM, 281–290.

[49] B. Krishna. 2011. Malicious emails masquerade as office printer messages. Symantec Connect Blog - Symantec Intelligence.ONLINE. Retrieved from http://www.symantec.com/connect/ blogs/malicious-emails-masquerade-office-printer-messages-0.

[50] K. Selvaraj and N. F. Gutierrez. 2010. The rise of PDF malware. Symantec Security Response. (2010).

[51] G. S. Bindra. 2011. Masquerading as a trustworthy entity through portable document file (PDF) format. In Privacy, Security, Risk and Trust (PASSAT). IEEE, 784–789.

[52] C. E. Drake, J. O. Jonathan, and J. K. Eugene. 2004. Anatomy of a phishing email. In CEAS

[53] Y. Song, C. Yang, and G. Gu. 2010. Who is peeping at your passwords at Starbucks? To catch an evil twin access point. In Proceedings of the 2010 IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'10). IEEE, 323–332.

[54] J. Szurdi, B. Kocso, G. Cseh, J. Spring, M. Felegyhazi, and C. Kanich. 2014. The long tail of typosquatting domain names. In Proceedings of the 23rd USENIX Security Symposium (USENIX Security 14). USENIX, 191–206.

[55] P. Agten, W. Joosen, F. Piessens, and N. Nikiforakis. 2015. Seven months' worth of mistakes: A longitudinal study of typosquatting abuse. In Proceedings of the 22nd Network and Distributed System Security Symposium (NDSS'15).

[56] J. Giles. 2010. Scareware the inside story. New Scientist , Article 205, 2753, 38–41.

[57] C. Dhinakaran, J. K. Lee, and D. Nagamalai. 2009. "Reminder: Please update your details": Phishing trends. In Proceedings of the 1st International Conference on Networks and Communications (NETCOM'09). IEEE, 295–300

[58] A. Kumar, M. Chaudhary, and N. Kumar, "Social Engineering Threats and Awareness: A Survey," undefined, 2015.

[59] "What Is Shoulder Surfing?" https://www.lifelock.com/learn-identity-theftresources-what-is-shoulder-surfing.html (accessed Aug. 06, 2021).

[60] S. Ahmad, "Social Engineering Techniques Contrast Study," Int. J. Eng. Stud., vol. 9, no. 1, pp. 105–110, 2017, Accessed: Aug. 06, 2021. [Online]. Available: http://www.ripublication.com.

[61] "What Is Social Engineering and How Does It Work? | Synopsys." https://www.synopsys.com/glossary/what-issocial-engineering.html (accessed Aug. 06, 2021).

[62] J. A. Chaudhry, S. A. Chaudhry, and R. G.Rittenhouse, “Phishing attacks and defenses,”Int. J. Secur. its Appl., vol. 10, no. 1, pp. 247–256, 2016, doi: 10.14257/ijsia.2016.10.1.23.

[63] J. Janulevi˘,“applied sciences E-mail-Based Phishing Attack Taxonomy,” pp. 1–15, 2020

[64] G. Ho, A. Sharma, M. Javed, V. Paxson, and D. Wagner, “Detecting Credential Spearphishing Attacks in Enterprise Settings,” Accessed: Aug.06, 2021. [Online]. Available:https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/ho

[65] V. Bhavsar, A. Kadlak, and S. Sharma, “Study on Phishing Attacks,” Artic. Int. J. Comput.Appl., vol. 182, no. 33, pp. 975–8887, 2018, doi:10.5120/ijca2018918286.

[66] V. Bhavsar, A. Kadlak, and S. Sharma, “Study on Phishing Attacks,” Artic. Int. J. Comput.Appl., vol. 182, no. 33, pp. 975–8887, 2018, doi:10.5120/ijca2018918286.

[67] “What Is Angler Phishing And How Do I Avoid Becoming A Victim?” https://blog.knowbe4.com/what-is-anglerphishing-and-how-do-i-avoid-becoming-avictim (accessed Aug. 06, 2021).

[68] [Abu-Nimeh and Nair 2006; Laurie and Laurie 2003]

[69] https://www.myalignedit.com/it-services/cybersecurity/what-are-baiting-attacks-and-how-can-you-prevent-them/#:~:text=Baiting%20attacks%20are%20social%20engineering%20attacks%20that%20use,social%20engineering%20techniques%2C%20baiting%20relies%20on%20psychological%20manipulation

[70] A. Jain, H. Tailang, H. Goswami, S. Dutta, M. S. Sankhla, and R. Kumar, “Social Engineering: Hacking a Human Being through Technology,” IOSR J. Comput. Eng., vol. 18, no. 5, pp. 94–100, 2016, doi: 10.9790/0661-18050594100.

[71] F. L. Greitzer, J. R. Strozer, S. Cohen, A. P. Moore, D. Mundie, and J. Cowley, “Analysis of Unintentional Insider Threats Deriving from Social Engineering Exploits,” 2014, doi: 10.1109/SPW.2014.39.

[72] D. Airehrour, N. V. Nair, and S. Madanian, “Social engineering attacks and countermeasures in the New Zealand Banking System: Advancing a user-reflective mitigation model,” Inf., vol. 9, no. 5, 2018, doi: 10.3390/info9050110.

[73] https://www.imperva.com/learn/application-security/pretexting/

[74] A. Jain, H. Tailang, H. Goswami, S. Dutta, M. S. Sankhla, and R. Kumar, “Social Engineering: Hacking a Human Being through Technology,” IOSR J. Comput. Eng., vol. 18, no. 5, pp. 94–100, 2016, doi: 10.9790/0661-18050594100.

[75] D. Airehrour, N. V. Nair, and S. Madanian, “Social engineering attacks and countermeasures in the New Zealand Banking System: Advancing a user-reflective mitigation model,” Inf., vol. 9, no. 5, 2018, doi: 10.3390/info9050110.

[76] F. Breda, H. Barbosa, and T. Morais, “Social Engineering and Cyber Security,” INTED2017 Proc., vol. 1, no. March, pp. 4204–4211, 2017, doi: 10.21125/inted.2017.1008.

[77] P. L. Gallegos-Segovia, P. E. Vintimilla-Tapia, J. F. Bravo-Torres, I. F. Yuquilima-Albarado, V. M. Larios-Rosillo, and J. D. Jara-Saltos, “Social engineering as an attack

vector for ransomware," 2017 Chil. Conf. Electr. Electron. Eng. Inf. Commun. Technol. CHILECON 2017 - Proc., vol. 2017-Janua, no. June 2020, pp. 1–6, 2017, doi: 10.1109/CHILECON.2017.8229528.

[78] A. Kumar, M. Chaudhary, and N. Kumar, "Social Engineering Threats and Awareness: A Survey," undefined, 2015.

[79] H. J. Chittooparambil, B. Shanmugam, S. Azam, K. Kannoorpatti, M. Jonkman, and G. N. Samy, "A review of ransomware families and detection methods," Adv. Intell. Syst. Comput., vol. 843, pp. 588–597, 2019, doi: 10.1007/978-3-319- 99007-1_55.

[80] J. Bullée and M. Junger, "The Palgrave Handbook of International Cybercrime and Cyberdeviance," Palgrave Handb. Int. Cybercrime Cyberdeviance, no. March, 2020, doi: 10.1007/978-3-319-90307-1.

[81] A. A. Alsufyani and S. Alzahrani, "Social Engineering Attack," Int. J. Adv. Res. Eng. Technol., vol. 11, no. 11, pp. 965–975, 2020, doi: 10.34218/IJARET.11.11.2020.089.

[82] "Consumer Information Blog - for 2018- December | FTC Consumer Information." https://www.consumer.ftc.gov/blog/2018/12/wh at-social-security-scam-sounds&lang=en (accessed Aug. 06, 2021).

[83] https://blog.mailfence.com/quid-pro-quo-attacks/

[84] A. P. Felt and D. Wagner. 2011. Phishing on Mobile Devices. In W2SP

[85] F. Callegati, W. Cerroni, and M. Ramilli. 2009. Man-in-the-middle attack to the HTTPS protocol. IEEE Security and Privacy 7, 1, 78–81.

[86] S. Ford, M. Cova, C. Kruegel, and G. Vigna. 2009. Analyzing and detecting malicious flash advertisements. In Proceedings of the Annual Computer Security Applications Conference (ACSAC'09). IEEE, 363– 372

[87] J. R. Jacobs. 2011. Measuring the Effectiveness of the USB Flash Drive as a Vector for Social Engineering Attacks on Commercial and Residential Computer Systems. Master's thesis. Embry-Riddle Aeronautical University

[88] F. Howard and O. Komili. 2010. Poisoned search results: How hackers have automated search engine poisoning attacks to distribute malware. Sophos Technical Papers (2010). https://www.sophos. com/medialibrary/PDFs/technical%20papers/sophosseoinsights.pdf.

[89] A. Patel, Q. Qassim, and C. Wills, "A survey of intrusion detection and prevention systems," Inf. Manag. Comput. Secur., vol. 18, no. 4, pp. 277– 290, 2010, doi: 10.1108/09685221011079199.

[90] J. D. Bustard, J. N. Carter, and M. S. Nixon, "Targeted biometric impersonation," 2013 Int. Work. Biometrics Forensics, IWBF 2013, 2013, doi: 10.1109/IWBF.2013.6547323.

[91] R. Islam and J. Abawajy, "A multi-tier phishing detection and filtering approach," J. Netw. Comput. Appl., vol. 36, no. 1, pp. 324–335, Jan. 2013, doi: 10.1016/J.JNCA.2012.05.009.

[92] T. M. Chen and V. Wang, "Web filtering and censoring," Computer (Long. Beach. Calif)., vol. 43, no. 3, pp. 94–97, Mar. 2010, doi: 10.1109/MC.2010.84.

[93] S. Sinha, “Building an Nmap Network Scanner,” Begin. Ethical Hacking with Python, pp. 165– 168, 2017, doi: 10.1007/978-1-4842-2541-7_24.

[94] https://www.holmsecurity.com/blog/what-is-nmap

[95] R. T. Peltier. 2013. Information Security Policies, Procedures, and Standards: Guidelines for Effective Information Security Management. CRC Press

[96] E. D. Frauenstein and R. von Solms. 2013. An enterprise anti-phishing framework. In Information Assurance and Security Education and Training. Springer Berlin Heidelberg, 196–203

[97] P. Kumaraguru. 2009. PhishGuru: A System for Educating Users About Semantic Attacks. Ph.D. Dissertation. Carnegie Mellon University

[98] G. N. A. Arachchilage, S. Love, and M. Scott. 2012. Designing a mobile game to teach conceptual knowledge of avoiding phishing attacks. International Journal for e-Learning Security 2, 2, 127–132

[99] E. Lin, S. Greenberg, E. Trotter, D. Ma, and J. Aycock. 2011. Does domain highlighting help people identify phishing sites? In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2075–2084

[100] J. Lee, L. Bauer, and M. L. Mazurek. 2015. The effectiveness of security images in Internet banking. IEEE Internet Computing 19, 1, 54–62

[101] E. Kritzinger and S. H. von Solms. 2010. Cyber security for home users: A new way of protection through awareness enforcement. Computer and Security 29, 8, 840–847.

[102] K. E. Stewart, J. W. Humphries, and T. R. Andel. 2009. Developing a virtualization platform for courses in networking, systems administration and cyber security education. In Proceedings of the 2009 Spring Simulation Multiconference. Society for Computer Simulation International.

[103] A. Konak and M. Bartolacci. 2012. Broadening E-commerce information security education using virtual computing technologies. In Proceedings of the 2012 Networking and Electronic Commerce Research Conference

[104] B. B. Anderson, C. B. Kirwan, J. L. Jenkins, D. Eargle, S. Howard, and A. Vance. 2013. How polymorphic warnings reduce habituation in the braininsights from an fMRI study. In Proceedings of of CHI15.

[105] J. Lee, L. Bauer, and M. L. Mazurek. 2015. The effectiveness of security images in Internet banking. IEEE Internet Computing 19, 1, 54–62

[106] A. Barth, C. Jackso, C. Reis, and TGC Team. 2008. The Security Architecture of the Chromium Browser. Retrieved from http://seclah.stanford.edu/websec/chromium

[107] T. Blasing, L. Batyuk, A. D. Schmidt, S. A. Camtepe, and S. Albayrak. 2010. An Android application sandbox system for suspicious software detection. In Proceedings of the 5th International Conference on Malicious and Unwanted Software (MALWARE). IEEE, 55–62.

[108] C. Greamo and A.Ghosh. 2011. Sandboxing and virtualisation: Modern tools for combating malware. In Security and Privacy,9, 2, 79–82.

[109] E. F. Brickell, J. F. Cihula, C. D. Hall, and R. Uhlig. 2011. Method of improving computer security through sandboxing. US Patent No. 7,908,653. (2011).

[110] Mozilla Firefox. 2015. Mozilla Wiki—Security/Sandbox. Retrieved from https://wiki.mozilla.org/Security/ Sandbox

[111] Chromium. 2015. The Chromium Projects—Sandbox. Retrieved from http://www.chromium.org/developers/ design-documents/sandbox.

[112] Microsoft. 2007. The Windows Vista and Windows Server 2008 Developer Story: Windows Vista Application Development Requirements for User Account Control. Retrieved from https://msdn. microsoft.com/en-us/library/aa905330.aspx

[113] L. Lu, V. Yegneswaran, P. Porras, and W. Lee. 2010. Blade: An attack-agnostic approach for preventing driveby malware infections. In Proceedings of the 17th ACM Conference on Computer and Communications Security. ACM, 440–450.

[114] A. Bianchi, J. Corbetta, L. Invernizzi, Y. Fratantonio, C. Kruegel, and G. Vigna. 2015. What the app is that? Deception and countermeasures in the Android user interface. In Proceedings of the 36th IEEE Symposium on Security and Privacy. IEEE.

[115] M. Egele, D. Brumley Y. Fratantonio, and C. Kruegel. 2013. An empirical study of cryptographic misuse in Android applications. In Proceedings of the 2013 ACM SIGSAC Conference on Computer and Communications Security. ACM, 73–84.

[116] S. Motiee, K. Hawkey, and K. Beznosov. 2010. Do windows users follow the principle of least privilege?: investigating user account control practices. In Proceedings of the 6th Symposium on Usable Privacy and Security. ACM.

[117] B. C. Neuman and T. Ts'o. 1994. Kerberos: An authentication service for computer networks. Communications Magazine 32, 9, 33–38.

[118] B. Desmond, J. Richards, R. Allen, and A. G. Lowe-Norris. 2008. Active Directory: Designing, Deploying, and Running Active Directory. O'Reilly Media.

[119] S. Motiee, K. Hawkey, and K. Beznosov. 2010. Do windows users follow the principle of least privilege?: investigating user account control practices. In Proceedings of the 6th Symposium on Usable Privacy and Security. ACM.

[120] B. Turner, D. Lundell, J. Zamora, and C. Calderon. 2010. Microsoft Forefront Identity Manager 2010 Technical Overview. Technical Report. Retrieved from http://download.microsoft.com/ download/0/8/4/0846D14C-B2D5-4BEA-9061-311BBF5BB76B/FIM%202010%20Technical%20Overview. docx.

[121] K. Lee, J. Caverlee, and S. Webb. 2010. The social honeypot project: Protecting online communities from spammers. In Proceedings of the 19th International Conference on World Wide Web. ACM

[122] M. B. Salem and S. J. Stolfo. 2011. Modeling user search behavior for masquerade detection. In Recent Advances in Intrusion Detection. Springer Berlin Heidelberg

[123] G. Stringhini, C. Kruegel, and G. Vigna. 2013. Shady paths: Leveraging surfing crowds to detect malicious web pages. In Proceedings of the 2013 ACM SIGSAC conference on Computer and communications security. ACM, 133–144.

[124] L. Lu, R. Perdisci, and W. Lee. 2011. Surf: Detecting and measuring search poisoning. In Proceedings of the 18th ACM Conference on Computer and Communications Security. ACM, 467–476.

[125] Z. Li, S. Alrwais, Y. Xie, F. Yu, and X. Wang. 2013. Finding the linchpins of the dark web: A study on topologically dedicated hosts on malicious web infrastructures. In Proceedings of the 2013 IEEE Symposium on Security and Privacy (SP'13). IEEE, 112–126.

[126] S. Lee and J. Kim. 2012. WarningBird: Detecting suspicious URLs in Twitter stream. In NDSS

[127] FirstCyberSecurity. 2009. Protecting Your Brand Online and Creating Customer Confidence. Retrieved from http://www.firstcybersecurity.com/main/IPRiskMReview.pdf.

[128] Webroot. 2013. Webroot Real-Time Anti-Phishing Service. Retrieved from http://www.webroot.com/ shared/pdf/WAP-Anti-Phishing-102013.pdf.

[129] Z. H. Abdullah, N. I. Udzir, R. Mahmod, and K. Samsudin. 2011. Towards a dynamic file integrity monitor through a security classification. Internal Journal of New Computer Architectures and Their Applications (IJNCAA) 1, 3, 766–779.\

[130] R. Dhanalakshmi and C. Chellappan. 2010. Detection and recognition of file masquerading for e-mail and data security. In Recent Trends in Network Security and Applications. Springer, Berlin, 253–262.

[131] M. Hara, A. Yamada, and Y. Miyake. 2009. Visual similarity-based phishing detection without victim site information. In Proceedings of the IEEE Symposium on Computational Intelligence in Cyber Security(CICS'09). IEEE, 30–36.

[132] T. Bhardwaj, K. T. Sharma, and M. R. Pandit. 2014. Social engineering prevention by detecting malicious URLs using artificial bee colony algorithm. In Proceedings of the 3rd International Conference on Soft Computing for Problem Solving. Springer, 355–363.

[133] Darknet. 2015. EvilAP Defender Detect Evil Twin Attacks. Retrieved from http://www.darknet.org.uk/ 2015/04/evilap-defender-detect-evil-twin-attacks/.

[134] P. Singhal and N. Raul. 2012. Malware detection module using machine learning algorithms to assist in centralized security in enterprise networks. International Journal of Network Security Its Applications 4, 1, 6 pages.

[135] S. Dong-Her, C. Hsiu-Sen, C. Chun-Yuan, and B. Lin. 2011. Internet security: Malicious e-mails detection and protection. Industrial Management and Data Systems 104, 7, 613–623.

[136] H. Sandouka, A. J. Cullen, and I. Mann. 2009. Social engineering detection using neural networks. In Proceedings of the International Conference on CyberWorlds (CW'09). IEEE, 273–278.

[137] R. Basnet, S. Mukkamala, and A. H. Sung. 2008. Detection of phishing attacks: A machine learning approach. In Soft Computing Applications in Industry. Springer, Berlin, 373–383.

[138] A. Bergholz, J. H. Chang, G. Paa, F. Reichartz, and S. Strobel. 2008. Improved phishing detection using model-based features. In CEAS

[139] A. Bergholz, J. De Beer, S. Glahn, M. F. Moens, G. Paa, and S. Strobel. 2010. New filtering approaches for phishing email. Journal of Computer Security 18, 1, 7–35.

[140] V. Raskin, J. M. Taylor, and C. F. Hempelmann. 2010. Ontological semantic technology for detecting insider threat and social engineering. In Proceedings of the 2010 Workshop on New Security Paradigms. ACM.

[141] G. Xiang, J. Hong, C. P. Rose, and L. Cranor. 2011. CANTINA+: A feature-rich machine learning framework for detecting phishing web sites. ACM Transactions on Information and System Security (TISSEC) 14, 2, Article 21

[142] M. Cova, C. Kruegel, and G. Vigna. 2010. Detection and analysis of drive-by-download attacks and malicious JavaScript code. In Proceedings of the 19th International Conference on World Wide Web. ACM, 281–290.

[143] A. Aggarwal, A. Rajadesingan, and P. Kumaraguru. 2012. PhishAri: Automatic realtime phishing detection on twitter. In eCrime Researchers Summit (eCrime). IEEE, 1–12.

[144] G. Stringhini and O. Thonnard. 2015. That Aint You: Blocking spearphishing through behavioral modelling. In Detection of Intrusions and Malware, and Vulnerability Assessment. Springer, 78–97.

[145] Shie, E. W. S. (2020). Critical analysis of current research aimed at improving detection of phishing attacks. Selected computing research papers, p. 45.

[146] Maurya, S., & Jain, A. (2020). Deep learning to combat phishing. Journal of Statistics and Management Systems, pp. 1–13.

[147] Subasi, A., Molah, E., Almkallawi, F., & Chaudhery, T. J. (2017). Intelligent phishing website detection using random forest classifier. In 2017 International conference on electrical and computing technologies and applications (ICECTA) (pp. 1–5). IEEE

[148] Abdelhamid, N., Thabtah, F., Abdel-jaber, H. (2017). Phishing detection: A recent intelligent machine learning comparison based on models content and features. In 2017 IEEE international conference on intelligence and security informatics (ISI) (pp. 72–77). IEEE.

[149] Mao, J., Bian, J., Tian, W., Zhu, S., Wei, T., Li, A., et al. (2018). Detecting phishing websites via aggregation analysis of page layouts. Procedia Computer Science, 129, 224–230.

[150] Jain, A. K., & Gupta, B. B. (2018). Towards detection of phishing websites on client-side using machine learning based approach. Telecommunication Systems, 68(4), 687–700.

[151] Begum, A., & Badugu, S. (2020). A study of malicious url detection using machine learning and heuristic approaches. In Advances in decision sciences, security and computer vision, image processing (pp. 587–597). Berlin: Springer.

[152] Yasin, A., Fatima, R., Liu, L., Yasin, A., & Wang, J. (2019). Contemplating social engineering studies and attack scenarios: A review study. Security and Privacy, 2(4), e73.

[153] Fatima, R., Yasin, A., Liu, L., & Wang, J. (2019). How persuasive is a phishing email? A phishing game for phishing awareness. Journal of Computer Security, 27(6), 581–612.

[154] Chiew, K. L., Yong, K. S. C., & Tan, C. L. (2018). A survey of phishing attacks: Their types, vectors and technical approaches. Expert Systems with Applications, 106, 1–20

[155] Yao, W., Ding Y., & Li, X. (2018). Logophish: A new twodimensional code phishing attack detection method. In 2018 IEEE international conference on parallel and distributed processing with applications, ubiquitous computing and communications, big data and cloud computing, social computing and networking, sustainable computing and communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom) (pp. 231–236). IEEE.

[156] Parsons, K., Butavicius, M., Delfabbro, P., & Lillie, M. (2019). Predicting susceptibility to social influence in phishing emails. International Journal of Human-Computer Studies, 128, 17–26.

[157] Tyagi, I., Shad, J., Sharma, S., Gaur, S., & Kaur, G. (2018). A novel machine learning approach to detect phishing websites. In 2018 5th International conference on signal processing and integrated networks (SPIN) (pp. 425–430). IEEE.

[158] Mao, J., Bian, J., Tian, W., Zhu, S., Wei, T., Li, A., et al. (2019). Phishing page detection via learning classifiers from page layout feature. EURASIP Journal on Wireless Communications and Networking, 2019(1), 43.

[159] Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). Machine learning based phishing detection from urls. Expert Systems with Applications, 117, 345–357.

[160] Kumar, A., Chatterjee, J. M., & Díaz, V. G. (2020). A novel hybrid approach of svm combined with nlp and probabilistic neural network for email phishing. International Journal of Electrical and Computer Engineering, 10(1), 486.

[161] Jain, A. K., Parashar, S., Katare, P., & Sharma, I. (2020). Phishskape: A content based approach to escape phishing attacks. Procedia Computer Science, 171, 1102–1109.

[162] Adebowale, M. A., Lwin, K. T., Sanchez, E., & Hossain, M. A. (2019). Intelligent web-phishing detection and protection scheme using integrated features of images, frames and text.Expert Systems with Applications, 115, 300–313.

[163] Chiew, K. L., Tan, C. L., Wong, K., Yong, K. S., & Tiong, W. K. (2019). A new hybrid ensemble feature selection framework for machine learning-based phishing detection system.Information Sciences, 484, 153–166.

[164] Pandey, A., Gill, N., Nadendla, K. S. P., & Thaseen, I. S. (2018). Identification of phishing attack in websites using random forestsvm hybrid model. In International conference on intelligent systems design and applications (pp. 120–128). Springer.

[165] Niranjan, A., Haripriya, D., Pooja, R., Sarah, S., Shenoy, P. D., & Venugopal, K. (2019). Ekrv: Ensemble of knn and random committee using voting for efficient classification of phishing. In Progress in advanced computing and intelligent engineering (pp. 403–414). Springer.

[166] Patil, V., Thakkar, P., Shah, C., Bhat, T., & Godse, S. (2018). Detection and prevention of phishing websites using machine learning approach. In 2018 Fourth international conference on computing communication control and automation (ICCUBEA) (pp. 1–5). IEEE.

[167] Jagadeesan, S., Chaturvedi, A., & Kumar, S. (2018). Url phishing analysis using random forest. International Journal of Pure and Applied Mathematics, 118(20), 4159–4163.

[168] https://www.audax.gr/kyvernoegklima/koinwniki-mixaniki/

[169] A SURVEY OF SOCIAL ENGINEERING ATTACKS: DETECTION AND PREVENTION TOOLS 1 NOOR AMMAR ODEH, 2 DERAR ELEYAN, 3 AMNA ELEYAN

[170] A Taxonomy of Attacks and a Survey of Defence Mechanisms for Semantic Social Engineering Attacks RYAN HEARTFIELD and GEORGE LOUKAS

[171] Saiping Xu, etc 2019 - Combining random forest and support vector machines for object-based rurand-land-cover classification using high spatial resolution imagery