**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**

**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

# ΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΑΝΙΧΝΕΥΣΗΣ ΑΠΑΤΗΣ

Πολυξένη Π. Φραγκοπούλου

Διπλωματική Εργασία
που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των απαιτήσεων για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στην Εφαρμοσμένη Στατιστική.

Πειραιάς
2013

**UNIVERSITY OF PEIRAIUS**

**DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE**

**POSTGRADUATE PROGRAMME IN APPLIED STATISTICS**

# STATISTICAL FRAUD DETECTION

By

Polyxeni P. Fragopoulou

MSc Dissertation

submitted to the Department of Statistics and Insurance Science of the University of Piraeus in partial fulfillment of the requirements for the degree of Master of Science in Applied Statistics.

Piraeus, Greece
2013

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίσθηκε από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. …….. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών.

Τα μέλη της Επιτροπής ήσαν:

- Καθηγητής Κούτρας Μάρκος (Επιβλέπων)
- Επίκουρη Καθηγήτρια Βερροπούλου Γεωργία
- Επίκουρος Καθηγητής Τζαβελάς Γεώργιος

Η έγκριση της Διπλωματική Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνωμών του συγγραφέα.

*Στους Γονείς Μου*

# Ευχαριστίες

Στο σημείο αυτό θα ήθελα να ευχαριστήσω τα άτομα εκείνα που συνέβαλλαν στην ολοκλήρωση της διπλωματικής μου εργασίας. Ιδιαίτερα θα ήθελα να εκφράσω τις ευχαριστίες μου στον επιβλέποντα καθηγητή μου κ. Κούτρα Μάρκο, για τις σημαντικές του παρατηρήσεις, οι οποίες βοήθησαν στην ολοκλήρωση της διπλωματικής εργασίας αλλά και για όλη τη διάρκεια των μεταπτυχιακών μου σπουδών. Επίσης θα ήθελα να ευχαριστήσω τα μέλη της τριμελής επιτροπής την Επίκουρη Καθηγήτρια κ. Βερροπούλου Γεωργία και Επίκουρο Καθηγητή Τζαβελά Γεώργιο, για την συμμετοχή τους στην εξεταστική επιτροπή.

Τέλος θα ήθελα να ευχαριστήσω την οικογένεια μου για τη συμπαράσταση και την κατανόηση που μου προσέφερε καθώς με στήριξε όλη αυτή τη διάρκεια της ακαδημαϊκής μου πορείας. Χωρίς αυτή θα ήταν αδύνατο να τις πραγματοποιήσω.

# Περίληψη

Στη σημερινή εποχή η πρόοδος της τεχνολογίας και της εξέλιξης της παγκόσμιας επικοινωνίας συνέβαλλαν στην αύξηση της απάτης, με συνέπεια την απώλεια δισεκατομμυρίων δολαρίων το χρόνο παγκοσμίως. Αν και η πρόληψη με χρήση της τεχνολογίας είναι ο καλύτερος τρόπος να μειώσουμε την απάτη, οι απατεώνες θα συνεχίζουν να βρίσκουν τρόπους ώστε να μπορούν να μην γίνονται αντιληπτοί παρά τα μέτρα πρόληψης. Για το λόγο αυτό η επιστήμη προσπαθεί να αναπτύξει μηχανισμούς που εμποδίζουν ή ανακαλύπτουν τέτοιου είδους συμπεριφορές. Η ανάπτυξη συστημάτων μεθοδολογιών που έχουν ως στόχο την παραπλάνηση ή εξαπάτηση για την ανίχνευση της απάτης είναι απαραίτητες αφού δίνουν τη δυνατότητα να εντοπιστούν και να καταδικαστούν οι απατεώνες. Η ενεργοποίησή τους αρχίζει μόλις αποτύχει η πρόληψη της απάτης.

Η ανίχνευση απάτης είναι ένας κλάδος της επιστήμης που στοχεύει στην ανακάλυψή της. Η πρόοδος της επιστήμης των υπολογιστών έδωσε την ευκαιρία στους επιστήμονες να αναπτύξουν νέες τεχνικές και μεθόδους μετατρέποντας έτσι τους υπολογιστικά απαιτητικούς αλγορίθμους σε εφικτές λύσεις. Ταυτόχρονα όμως, η ευκολία με την οποία μπορούμε να συγκεντρώσουμε και να αποθηκεύσουμε δεδομένα με ψηφιακό τρόπο δημιούργησε ογκώδη και συνεχώς αυξανόμενα σύνολα δεδομένων. Στατιστικές μέθοδοι ανίχνευσης της απάτης εφαρμόστηκαν με επιτυχία σε διάφορους τομείς οι οποίοι περιλαμβάνουν μεγάλα σύνολα δεδομένων, ορισμένοι από τους οποίους είναι η υγεία, οι τηλεπικοινωνίες, η οικονομία και η τεχνολογία. Δεν είναι λοιπόν παράξενο το ενδιαφέρον της επιστημονικής κοινότητας για ανάπτυξη τεχνικών και μεθόδων εξαγωγής πληροφορίας από αυτά τα ποικιλόμορφα και τεράστια σύνολα δεδομένων.

Στην παρούσα διπλωματική παρουσιάζονται ορισμένες από τις στατιστικές μεθόδους που χρησιμοποιούνται για την ανίχνευση της απάτης σε κάποιους τομείς και δίνονται ορισμένα αναλυτικά παραδείγματα για να γίνει πιο κατανοητός ο μηχανισμός με τον οποίο οδηγούμαστε στην ανακάλυψη της απάτης με βάση διαθέσιμα στατιστικά δεδομένα.

# Abstract

Nowadays the technological progress and the evolution of global communication contributed to the increase in fraud, resulting in the loss of billions of dollars worldwide each year. Although prevention by the aid of technology is the best way to reduce fraud, fraudsters will continue to invent ways allowing circumvention of such measures. Science will always move on developing mechanisms that prevent or discover human activity aiming at manipulation or deceive. Methodologies for the detection of fraud are necessary if we are enabling to identify and convict fraudsters. Their activation starts once fraud prevention has failed.

Fraud detection is a sector of science that has undertaken the responsibility of discovering fraud. The progress in computer science equipped the scientists with the opportunity to develop new techniques and methods, transforming thus the computationally intensive algorithms into feasible solutions. At the same time, the ease with which we can collect and store data in a digital way has created bulky and continuously growing data sets. Statistical methods for fraud detection have been successfully applied in various sectors of society which include large data sets, some of them are health, telecommunications, economy and technology. It is not surprising the interest of the scientific community in finding techniques and methods for extracting information from these diverse and huge data sets.

In this Msc thesis, we shall present some of the statistical methods used to detect fraud in a variety of some areas, and is given some detailed examples in order to understand more the mechanism by which we are led to the discovery of fraud based on available statistical data

# Περιεχόμενα

# Κατάλογος Πινάκων

# Κατάλογος Σχημάτων

# ΚΕΦΑΛΑΙΟ 1

# Ανίχνευση Απάτης



## 1.1 Εισαγωγή

Ο συνοπτικός ορισμός της **απάτης** (*fraud*) που δίνεται με βάση το *Oxford Dictionary* είναι ο ακόλουθος: "παράνομη ή εγκληματική εξαπάτηση η οποία έχει ως σκοπό να προκαλέσει οικονομικό ή προσωπικό κέρδος." Σύμφωνα με τον ορισμό αυτό γίνεται σαφές πως η απάτη είναι μια εσκεμμένη μη νομική πράξη είτε για προσωπικό μας όφελος είτε για να βλάψουμε κάποιο άλλο άτομο. Επιπροσθέτως θα μπορούσε να θεωρηθεί ακόμη ως ένα αδίκημα όταν συνοδεύεται από παραβίαση του αστικού δικαίου που σκοπός της είναι η εξαπάτηση φυσικών και νομικών προσώπων ώστε να υποστούν ζημιές, συχνά με τη μορφή κάποιας νομισματικής απώλειας. Υπάρχουν πολλά είδη απάτης σε ορισμένα εκ των οποίων θα πραγματοποιηθεί εκτενής αναφορά στην Ενότητα 1.4, ωστόσο το κίνητρο για κάθε μία δεν είναι πάντα το ίδιο όσον αφορά την υλοποίησή της. Κατά κύριο λόγο σκοπό της απάτης αποτελεί η εύκολη και γρήγορη απόκτηση χρήματος, ωστόσο υπάρχουν κάποιοι τομείς, όπως για παράδειγμα στην επιστήμη, όπου η απάτη χρησιμοποιείται για την απόκτηση κύρους και όχι για την κατάκτηση κάποιου χρηματικού επάθλου.

Τα τελευταία χρόνια, λόγω της ραγδαίας εξέλιξης της σύγχρονης τεχνολογίας, και της εξέλιξης της παγκόσμιας επικοινωνίας, παρέχονται ακόμη περισσότεροι τρόποι ώστε να μπορέσει κάποιος να διαπράξει απάτη με αποτέλεσμα αυτή να αυξάνεται δραματικά και να έχει ως συνέπεια την απώλεια δισεκατομμυρίων δολαρίων παγκοσμίως το χρόνο. Ο κίνδυνος της απάτης είναι καθοριστικός για όλες τις επιχειρήσεις. Οι διοικήσεις των σύγχρονων επιχειρήσεων καλούνται να προβλέψουν τον κίνδυνο απάτης σε κάθε μορφή, να λάβουν μέτρα αποτρεπτικά πριν αυτός προκύψει και τέλος να τον αντιμετωπίσουν αποτελεσματικά όταν εμφανιστεί. Γι' αυτόν το λόγο θα αναφερθούμε στη συνέχεια στη διάκριση που αφορά την πρόληψη και την ανίχνευση της απάτης.

Ο όρος **πρόληψη της απάτης** (*fraud prevention*) αναφέρεται σε μέτρα που πρέπει να λαμβάνονται ώστε να την εμποδίσουμε. Τα μέτρα αυτά περιλαμβάνουν:

**α. περίτεχνα σχέδια** (*elaborate designs*), **ίνες φθορισμού** (*fluorescent Fibers*), **πολυτονικά σχέδια** *(multitone drawings)*, **υδατογραφήματα** (*watermarks*), **ελάσματα μετάλλου** (*laminated metal strips*) και **ολογραφήματα** (*holographs*) για τα τραπεζογραμμάτια,

**β. προσωπικούς αριθμούς αναγνώρισης** (*personal identification numbers*) τραπεζικών καρτών, **συστήματα ασφαλείας Διαδικτύου** (*Internet security systems*) όσον αφορά τις συναλλαγές μέσω πιστωτικών καρτών,

**γ. σύστημα αναγνώρισης συνδρομητών καρτών κινητών τηλέφωνων** (*Subscriber Identity Module (SIM) cards for mobile phones*) όσον αφορά την ασφάλεια χρήσης καρτών SIM,

**δ. κωδικούς πρόσβασης σε συστήματα ηλεκτρονικών υπολογιστών** (*passwords on computer systems*) για την αποτροπή εισβολής στα συστήματά τους.

Αν και η πρόληψη με χρήση της τεχνολογίας είναι ο καλύτερος τρόπος να μειώσουμε την απάτη, οι απατεώνες θα συνεχίζουν να βρίσκουν τρόπους ώστε να μπορούν να παρακάμπτουν τα μέτρα αυτά. Επομένως θα πρέπει να βρεθεί ένας συμβιβασμός μεταξύ του κόστους και της δυσχέρειας που έχει η κάθε μέθοδος πρόληψης της απάτης σε σχέση με την αποτελεσματικότητά της.

Ο όρος **ανίχνευση της απάτης** (*fraud detection*) αναφέτεραι στον εντοπισμό της το συντομότερο δυνατό, αφού προηγουμένως έχει διαπραχθεί. Η ανίχνευση απάτης αφορά την παρακολούθηση της συμπεριφοράς του πληθυσμού των χρηστών, προκειμένου να εκτιμηθεί, να ανιχνευθεί, και να εντοπίσει πιθανή ανεπιθύμητη συμπεριφορά. Ανεπιθύμητη συμπεριφορά είναι ένας ευρύς όρος συμπεριλαμβανομένης της εγκληματικότητας, της απάτης, της εισβολής, και της αθέτησης του λογαριασμού. Μεθοδολογίες για την ανίχνευση της απάτης είναι ουσιαστικές αν δίνουν τη δυνατότητα να εντοπιστούν και να καταδικαστούν οι δράστες και η δραστηριοποίησή της αρχίζει όταν αποτύχει η πρόληψη της απάτης.

Ένα απλό παράδειγμα πρόληψης της απάτης πιστωτικών καρτών βασίζεται στην καλή φύλαξή τους καθώς επίσης και στην κατάλληλη μέριμνά τους. Αν παρ'όλα αυτά τα στοιχεία της κάρτας μας έχουν κλαπεί, τότε θα πρέπει να είμαστε σε θέση να εντοπίσουμε, το συντομότερο δυνατό, ότι η απάτη έχει διαπραχτεί και να κινηθούμε κατάλληλα ώστε να διασφαλίσουμε την προσωπική μας ασφάλεια.

Μόλις μία μέθοδος ανίχνευσης γίνει γνωστό ότι έχει τεθεί σε εφαρμογή, οι απατεώνες θα προσπαθήσουν να εφαρμόσουν νέες στρατηγικές. Καθώς όλο και περισσότεροι άνθρωποι εισβάλλουν στον τομέα της απάτης πολλοί από αυτούς δεν γνωρίζουν τις τεχνικές ανίχνευσης που είχαν χρησιμοποιηθεί στο παρελθόν και ήταν επιτυχείς γι'αυτό και οι στρατηγικές που

εφαρμόζουν τους οδηγούν σε απάτες που μπορούν εύκολα να ανιχνευθούν. Ωστόσο η ανάπτυξη νέων μεθόδων ανίχνευσης γίνεται όλο και δυσκολότερη από το γεγονός ότι η ανταλλαγή ιδεών μεταξύ των ενδιαφερομένων για τον εντοπισμό της απάτης είναι πολύ περιορισμένη. Συνήθως τα σύνολα δεδομένων δεν είναι διαθέσιμα και τα αποτελέσματα συχνά λογοκρίνονται, καθιστώντας έτσι δύσκολη την αξιολόγηση και την υλοποίηση των τεχνικών που έχουν ήδη αναπτυχθεί.

Πολλά προβλήματα ανίχνευσης απάτης περιλαμβάνουν τεράστια σύνολα δεδομένων που εξελίσσονται συνεχώς. Για παράδειγμα, η επιχείρηση πιστωτικών καρτών Barclaycard εισάγει στις βάσεις δεδομένων της περίπου 350 εκατομμύρια συναλλαγές ετησίως στο Ηνωμένο Βασίλειο μόνο (Hand, Blunt, Kelly and Adams (2000)), η βασιλική τράπεζα της Σκωτίας, που έχει τη μεγαλύτερη εμπορική επιχείρηση πιστωτικών καρτών στην Ευρώπη, καταχωρεί δισεκατομμύρια συναλλαγές ετησίως και το AT&T καταγράφει περίπου 275 εκατομμύρια κλήσεις καθημερινώς (Cortes and Pregibon (1998)). Η επεξεργασία αυτών των συνόλων δεδομένων σε μια αναζήτηση για δόλιες συναλλαγές ή κλήσεις απαιτεί πολύ περισσότερα από μία απλή καινοτομία του στατιστικού προτύπου, και χρειάζεται γρήγορους και αποδοτικούς αλγορίθμους. Οι τεχνικές εξόρυξης δεδομένων μπορούν να βοηθήσουν σημαντικά προς αυτή την κατεύθυνση. Οι προαναφερθέντες αριθμοί δείχνουν επίσης την πιθανή αξία της ανίχνευσης απάτης: εάν το 0.1% των συναλλαγών σε 100 εκατομμυρία πρόκειται για απάτη, και η μέση απώλεια της εταιρείας ανά απάτη είναι μόλις £10, η επιχείρηση χάνει συνολικά £1 εκατομμύρια λίρες.

Τα στατιστικά εργαλεία που χρησιμοποιούνται για την ανίχνευση απάτης είναι πολλά και ποικίλουν. Συνήθως βασίζονται στη σύγκριση των παρατηρηθέντων δεδομένων με τις αναμενόμενες τιμές οι οποίες μπορούν να προκύψουν με διάφορους τρόπους σύμφωνα με το περιβάλλον στο οποίο ανήκουν. Όσον αφορά τις στατιστικές μεθόδους ανίχνευσης απάτης μπορούν να διακριθούν σε εποπτευόμενες και μη.

Οι **εποπτευόμενες μέθοδοι** (*supervised methods*) χρησιμοποιούν μία βάση δεδομένων από δόλια ή νόμιμα αρχεία για την κατασκευή ενός πρότυπου μοντέλου εκχωρώντας τις νέες παρατηρήσεις σε μία από αυτές τις δύο κατηγορίες. Επιπλέον οι μέθοδοι αυτές μπορούν να χρησιμοποιηθούν για την ανίχνευση περιπτώσεων απάτης οι οποίες έχουν εμφανιστεί στο παρελθόν.

Οι **μη εποπτευόμενες μέθοδοι** (*unsupervised methods*) αναζητούν τους λογαριασμούς ή τους πελάτες, οι οποίοι συμπεριφέρονται «ασυνήθιστα». Σε αυτές τις μεθόδους οι ακραίες τιμές αποτελούν τη βασική μορφή παρατήρησης ώστε να προβούμε σε έλεγχο αν κάτι δεν έχει πάει καλά και χρειάζεται επανεξέταση. Ωστόσο είτε η χρησιμοποίηση των εποπτευόμενων μεθόδων

είτε των μη μας οδηγεί στο συμπέρασμα ότι το αποτέλεσμα που θα πάρουμε θα μας δώσει απλά μια ένδειξη της πιθανότητας απάτης. Η στατιστική ανάλυση από μόνη της δεν μπορεί να διαβεβαιώσει ότι ένα συγκεκριμένο αρχείο, ή μια εγγραφή ή μια κίνηση αποτελεί απάτη. Αυτό το οποίο μπορεί να επισημάνει είναι το κατά πόσο το αρχείο αυτό ενδέχεται να είναι περισσότερο ύποπτο σε σχέση με κάποια άλλα.

Η ανίχνευση της απάτης είναι απλά η ικανότητα να ανακαλύψουμε ότι διεπράχθη ένα (οικονομικό συνήθως) αδίκημα. Ένα σύστημα ανίχνευσης θα προσπαθήσει να εντοπίσει σχέδια και τάσεις ύποπτης συμπεριφοράς. Συνήθως, το σύστημα θα δημιουργήσει ένα **βαθμό (σκορ) υποψίας** (*suspicion score*) ο οποίος θα δείχνει πόσο πιθανό είναι μια περίπτωση να αποτελεί απάτη. Αυτά τα αποτελέσματα μπορούν στη συνέχεια να διαταχτούν στη σειρά οπότε ιδιαίτερη προσοχή θα πρέπει να δοθεί σε εκείνα με τα υψηλότερα αποτελέσματα ή σε εκείνα που εμφανίζουν μια ξαφνική αύξηση. Όσο υψηλότερο είναι το score, τόσο πιο ασυνήθιστη φαίνεται να είναι η αντίστοιχη κίνηση. Επομένως περιπτώσεις που υπερβαίνουν το κατώτατο όριο ενός ορισμένου σκορ υποψίας χρειάζεται να διερευνηθούν περισσότερο. Η αποτελεσματικότητα του συστήματος ανίχνευσης θα εξαρτηθεί τελικά από την ταχύτητα με την οποία ανιχνεύεται το έγκλημα από το εύρος των εγκλημάτων που μπορεί να ανιχνεύσει, και από τον αριθμό των ψευδών συναγερμών που παράγονται. Το γεγονός ότι υπάρχουν πολλοί διαφορετικοί τρόποι με τους οποίους μπορεί να διαπραχτεί η απάτη και πολλά διαφορετικά σενάρια στα οποία μπορεί να εμφανιστεί, σημαίνει ότι υπάρχουν πολλοί διαφορετικοί τρόποι για να υπολογιστούν τα ύποπτα αυτά αποτελέσματα. Βασικό πρόβλημα όμως αποτελεί το γεγονός ότι είναι υπερβολικά δαπανηρό να αναλάβουμε μια λεπτομερή έρευνα για όλα τα αρχεία, γι αυτό ακριβώς το λόγο θα πρέπει να εστιάσουμε την έρευνα στις περιπτώσεις εκείνες που πιστεύουμε ενδεχομένως πως είναι πιθανότερο απατηλές.

Δυστυχώς το να μπορέσουμε να ανιχνεύσουμε την απάτη είναι κάτι αρκετά δύσκολο και απαιτεί μεγάλο κόστος. Ακόμη και αν μπορεί να βρεθεί μια μέθοδος ανίχνευσης η οποία να αναγνωρίζει σωστά το 99% των νόμιμων αρχείων ως νόμιμα και 99% των δόλιων αρχείων ως περιπτώσεις απάτης τότε θα μπορούσε να θεωρηθεί ως ένα εξαιρετικά αποτελεσματικό σύστημα. Ωστόσο ακόμη και κάτι τέτοιο θα είχε πολλές αρνητικές επιπτώσεις αφού με ένα τέτοιο σύστημα ένας μικρός αριθμός αρχείων θα ήταν δόλια. Επομένως εμείς για να μπορέσουμε στο τέλος τα εντοπίσουμε τα αρχεία αυτά θα πρέπει να τα εξετάσουμε όλα, κάτι το οποίο θα έχει και ιδιαίτερο κόστος. Συνεπώς για να ελαττώσουμε την απάτη θα πρέπει να βρεθεί μία λύση ανάμεσα στο κόστος και στην προσπάθεια που θα καταβάλλουμε για την ανίχνευσή της. Με λίγα λόγια θα πρέπει να υπάρξει ένας οικονομικός συμβιβασμός μεταξύ του κόστους έτσι ώστε να μην

ξοδεύουμε πολλά χρήματα για την ανίχνευση της απάτης και της εξοικονόμησης που θα γίνει από τον εντοπισμό της. Παρόλα αυτά πολλές φορές τα πράγματα είναι πιο πολύπλοκα, για παράδειγμα στην περίπτωση μιας τράπεζας που έχει διαπιστώσει την ύπαρξη απάτης αλλά έχει κάνει λίγα πράγματα για την ανίχνευσή της, μπορεί να προκληθεί έλλειψη εμπιστοσύνης μεταξύ του πελάτη και της τράπεζας αλλά και να χάσει η ίδια το κύρος της.

## 1.2 Ιστορική Αναδρομή

Στην Ενότητα 1.1 δώσαμε τον ορισμό της απάτης με βάση τον οποίο η απάτη υπό μία ευρύτερη έννοια είναι μια εξαπάτηση που γίνεται για το προσωπικό κέρδος. Ωστόσο έχει και μια δεύτερη πιο συγκεκριμένη έννοια η οποία έχει να κάνει με το έγκλημα ή διαφορετικά με μία σκόπιμη εξαπάτηση άλλων ανθρώπων προκειμένου αυτοί να βλαφθούν - συνήθως, για να λάβουν οι δράστες της απάτης την ιδιοκτησία ή τις υπηρεσίες των άλλων αδίκως. Η απάτη δεν αποτελεί ένα νέο πρόβλημα το οποίο αντιμετωπίζουμε μόνο στη σύγχρονη εποχή αλλά είναι τόσο παλιά όσο η ίδια η ανθρωπότητα. Πράγματι, θα μπορούσε κάποιος να σκεφτεί τι ακριβώς γινόταν στο παρελθόν και τότε θα μπορούσε να ισχυριστεί ότι η απάτη είναι μεγαλύτερης ηλικίας από όσο θα μπορούσε να φανταστεί. Αρκεί να σκεφτούμε ότι ακόμη και τα ζώα μπορεί να συμπεριφέρονται με δόλιο τρόπο ώστε να εξαπατήσουν τα υπόλοιπα για προσωπικό τους όφελος καθώς επίσης και για να μπορέσουν να προστατευθούν από τους εχθρούς τους. Αποτελούσε ένα είδος "καμουφλαζ" γι'αυτά. Ωστόσο η έννοια της «εγκληματικής» συμπεριφοράς του ανθρώπου είναι αποκλειστική.

Ένα περιστατικό απάτης προέκυψε το 1981 όταν ο Dr. John Darsee του Harvard Medical School, ο οποίος κοινοποίησε τα αποτελέσματά του από πειράματα που πραγματοποιήσε πάνω σε σκυλιά, τα οποία είχαν υποστεί έμφραγμα του μυοκαρδίου, λέγεται ότι είχαν εμβολιαστεί με πειραματικά φάρμακα καρδιακών παθήσεων. Το 1982, ωστόσο, δύο επιτροπές διερεύνησης αναφέρουν ότι τα πειράματα αυτά δεν πραγματοποιήθηκαν ποτέ.

Μία άλλη περίπτωση απάτης αναφέρεται σχετικά με την παραποίηση στοιχείων που συμμετέχουν στην Ιατρική έρευνα η οποία βασίζεται στην μεταμόσχευση δέρματος, μια διαδικασία που μπορεί να ενισχύσει την ασφάλεια της μεταμόσχευσης οργάνων. Κατά την περίοδο 1967 – 1974 ο Dr. William A. Summerlin του διάσημου Ινστιτούτου Sloan Kettering της Νέας Υόρκης ανέφερε ότι είχε μεταμοσχευθεί με επιτυχία δέρμα από μαύρα ποντίκια σε λευκά. Το 1974, ωστόσο, υπό την πίεση των συναδέλφων του, ο Dr. Summerlin ομολόγησε ότι είχε χρησιμοποιήσει ένα μαύρο μαρκαδόρο για να σκουρύνει μια περιοχή του μοσχεύματος που είχε μεταμοσχευθεί στην πραγματικότητα από ένα λευκό ποντίκι σε κάποιο άλλο. Στη συνέχεια, η επιτροπή ερευνητών της Sloan Kettering διαπίστωσε ότι ο Summerlin είχε διαστρεβλώσει επίσης

τα αποτελέσματα της προηγούμενης έρευνάς του που αφορούσε την μεταμόσχευση κερατοειδούς από ανθρώπινα πτώματα σε κουνέλια. Και επομένως αυτό είχε ως αποτέλεσμα το συνολικό έργο του Summerlin να αποσυρθεί εξαιτίς ότι τα αποτελέσματά του θεωρήθηκαν πλαστά.

Ένα ακόμη παράδειγμα απάτης αποτελεί η περίπτωση των Jaffer and Cameron (2006). Πρόκειται ίσως και για την πιο διάσημη περίπτωση στην Αγγλία η οποία αναφέρεται στον Malcolm Pearce, ο οποίος ήταν λέκτορας στο St George's Hospital Medical School του Λονδίνου και στον Geoffrey Chamberlain, καθηγητής και επικεφαλής του τμήματος. Το 1994 δημοσίευσαν ένα έγγραφο στην Βρετανική εφημερίδα σχετικά με τη Μαιευτική και τη Γυναικολογία, στην οποία ο Pearce ήταν βοηθός αρχισυντάκτη και ο Chamberlain ο εκδότης. Το έγγραφο αυτό ισχυριζόταν την επιτυχία της μεθόδου σχετικά με την εκ νέου εμφύτευση μιας εξωμήτριας κύησης. Στο ίδιο τεύχος, ο Pearce είχε δημοσιεύσει μια τυχαιοποιημένη ελεγχόμενη κλινική δοκιμή. Λίγους μήνες αργότερα, ένας κατώτερος ερευνητής στο τμήμα τους ειδοποίησε τις αρχές ότι η περίπτωση της εκ νέου εμφύτευσης ήταν ένα έργο μυθοπλασίας και ότι η τυχαιοποίηση των ασθενών στην μελέτη δεν υφίστατο. Η υπόθεση οδήγησε στην παραίτηση του Chamberlain και ο Pearce διαγράφηκε από το ιατρικό μητρώο.

## 1.3 Μέθοδοι Απάτης

Η δραματική αύξηση των περιπτώσεων απάτης που έχουν ως αποτέλεσμα την απώλεια δισεκατομμυρίων δολαρίων ετησίως σε όλο τον κόσμο, οδήγησε στην ανάπτυξη διάφορων σύγχρονων τεχνικών για την ανίχνευση της. Οι τεχνικές αυτές συνεχώς αναπτύσσονται και εφαρμόζονται σε πολλούς επιχειρηματικούς τομείς. Στην παρούσα ενότητα θα αναφέρουμε τις μεθόδους και τα εργαλεία που χρησιμοποιούνται για τη στατιστική ανίχνευση της απάτης. Η **Στατιστική** (*Statistic*) και η **Μηχανική μάθηση** (*Machine learning*) είναι δύο από τις επιστημονικές περιοχές που χρησιμοποιούνται για τον εντοπισμό της, και έχουν εφαρμοστεί επιτυχώς για την ανίχνευση των δραστηριοτήτων όπως το ξέπλυμα χρήματος, η απάτη πιστωτικών καρτών, η απάτη τηλεπικοινωνιών και η παράνομη εισβολή σε υπολογιστικά συστήματα καθώς και η ιατρική και επιστημονική απάτη. Αυτά είναι λίγα από τα παραδείγματα τα οποία θα αναφέρουμε στην Ενότητα 1.4.

Με βάση τη διάκριση που έγινε στην ανίχνευση της απάτης, των εποπτευόμενων μεθόδων και μη, θα παρουσιάσουμε τις τεχνικές που χρησιμοποίεί η κάθε μέθοδος ξεχωριστά. Οι εποπτευόμενες μέθοδοι, όπως αναφέραμε στην Ενότητα 1.1, χρησιμοποιούν δείγματα από δόλια ή νόμιμα αρχεία για την κατασκευή μοντέλου το οποίο θα αποδίδει ένα βαθμό υποψίας για τις νέες περιπτώσεις. Ωστόσο για να μπορέσουμε να κάνουμε τον διαχωρισμό αυτό και να

αποδώσουμε βαθμό υποψίας, η απάτη προηγουμένως θα πρέπει να έχει διαπραχθεί. Για αυτές τις μεθόδους, κατά τους Hand (1981) και McLachlan (1992), παραδοσιακές τεχνικές στατιστικής ταξινόμησης όπως η **γραμμική διακριτή ανάλυση** (*linear discriminant analysis*) και η **λογιστική διάκριση** (*logistic discrimination*), αποδείχθηκαν αποτελεσματικά εργαλεία για πολλές εφαρμογές. Ωστόσο ακόμη πιο ισχυρά εργαλεία είναι τα **νευρωνικά δίκτυα** (*neural networks*) σύμφωνα με αναφορές των Ripley (1996), Hand (1997) και Webb (1999) των οποίων η χρήση είναι ευρέως διαδεδομένη όσον αφορά την ανίχνευση απάτης τηλεπικοινωνιών και την απάτη πιστωτικών καρτών. Επιπλέον για την ανίχνευση της απάτης πιστωτικών καρτών και τηλεπικοινωνιών χρησιμοποιήθηκαν ***rule-based methods,*** τις οποίες θα αναφέρουμε αναλυτικά σε επόμενη ενότητα. Για την ανίχνευση ιατρικής απάτης οι Major and Riedinger (1992) δημιούργησαν ένα **στατιστικό σύστημα γνώσης** (*knowledge/statistical-based system*) όπου οι εξειδικευμένες γνώσεις ενσωματώνονται με στις στατιστικές τεχνικές. Επιπλέον η **Ανάλυση Συνδέσμων** (*Link analysis*) είναι μία μέθοδος ανίχνευσης της απάτης νομιμοποίησης παράνομου χρήματος και της απάτης τηλεπικοινωνιών.

Κατά τον σχηματισμό ενός εποπτευόμενου εργαλείου για την ανίχνευση της απάτης θα πρέπει να λάβουμε υπόψη μας το άνισο μέγεθος που θα υπάρχει στα σύνολα δεδομένων των δύο κατηγοριών καθώς επίσης και το κόστος που θα προκύψει από τα διαφορετικά είδη της εσφαλμένης ταξινόμησης. Ένα παράδειγμα το οποίο μπορούμε να αναφέρουμε είναι στις περιπτώσεις των πιστωτικών συναλλαγών όπου ίσως κάποιες να επισημανθούν εσφαλμένες, π.χ., μια απατηλή συναλλαγή μπορεί να μείνει απόρρητη και να μην παρατηρηθεί και στο τέλος να επισημανθεί ως νόμιμη αλλά μπορεί ακόμη να συμβεί και το αντίθετο και κάποια νόμιμη συναλλαγή να αναφερθεί εσφαλμένα ως δόλια. Γενικά θα πρέπει να σκεφτούμε σοβαρά το κόστος των εξεταζόμενων παρατηρήσεων και αν τελικά μας ωφελεί το να ανιχνεύσουμε την απάτη. Ίσως να είναι προτιμότερο να μην ασχοληθούμε καν αν πρόκειται για μία αρκετά χρονοβόρα και δαπανηρή διαδικασία ενώ αυτό το οποίο θα ανιχνεύσουμε δεν είναι τόσο σημαντικό. Επιπλέον είναι γνωστό ότι η ένταξη μίας παρατήρησης σε μία κατηγορία είναι αβέβαιη.

Στις μη εποπτευόμενες μεθόδους, στις οποίες δεν υπάρχουν προηγούμενα δεδομένα από νόμιμες και δόλιες παρατηρήσεις, οι τεχνικές οι οποίες χρησιμοποιούνται βασίζονται στην ανίχνευση των ακραίων τιμών. Στις τεχνικές αυτές κατασκευάζουμε μια βασική κατανομή η οποία θα παρουσιάζει την κανονική συμπεριφορά έτσι ώστε να μπορέσουμε να εντοπίσουμε τις παρατηρήσεις εκείνες οι οποίες παρουσιάζουν τη μέγιστη απόκλιση από τον τύπο αυτό. Ένα παράδειγμα μιας τέτοιας μεθόδου είναι η χρήση αριθμητικής ανάλυσης χρησιμοποιώντας τον

νόμο του Benford ή διαφορετικά ο νόμος του πρώτου ψηφίου. Ο νόμος του Benford ισχυρίζεται ότι όταν έχουμε αριθμητικά δεδομένα, ενώ θα περιμέναμε για το πρώτο ψηφίο των αριθμών, όλα τα ψηφία από το 1 ως το 9 να έχουν την ίδια συχνότητα εμφάνισης, δεν συμβαίνει κάτι τέτοιο. Σύμφωνα με το νόμο αυτό το πρώτο ψηφίο είναι το 1 το οποίο εμφανίζεται συχνότερα από τα υπόλοιπα (συγκεκριμένα περίπου 30%). Επιπλέον, όσο μεγαλύτερο το ψηφίο, τόσο λιγότερο πιθανό είναι να εμφανιστεί ως κύριο ψηφίο ενός αριθμού, για παράδειγμα το 9 εμφανίζεται λιγότερο από όλα τα υπόλοιπα ψηφία με ποσοστό περίπου 5%. Μέχρι πρόσφατα, ο νόμος αυτός θεωρήθηκε ως μια απλή μαθηματική περιέργεια με καμία εμφανή χρήσιμη εφαρμογή. Ο νόμος του πρώτου ψηφίου φαίνεται να δουλεύει καλύτερα σε μεγάλα δείγματα αριθμητικών μεγεθών καθώς επίσης σε αριθμούς επιλεγμένους τυχαία από διαφορετικές πηγές δεδομένων. Ωστόσο, οι Nigrini και Mittermaier (1997) και Nigrini (1999) έδειξαν ότι ο νόμος Benford μπορεί να χρησιμοποιηθεί για την ανίχνευση της απάτης.

Είναι γνωστό ότι οι απατεώνες συνεχώς θα ενημερώνονται για τα νέα μέτρα πρόληψης και ανίχνευσης της απάτης γι' αυτό το λόγο, προκειμένου να μπορέσουμε να την ελαττώσουμε όσο το δυνατό περισσότερο, θα πρέπει οι τεχνικές ανίχνευσης της απάτης να ενημερώνονται και να εξελίσσονται συνεχώς. (Βλέπε Burge Shawe-Taylor (1997), Fawcett and Provost (1997a), Cortes, Pregibon and Volinsky (2001) και Senator (2000)).

Παρά το γεγονός ότι τα βασικά στατιστικά μοντέλα για την ανίχνευση της απάτης μπορεί να χαρακτηριστούν ως επιβλέψιμα ή και μη, οι περιοχές εφαρμογής της ανίχνευσης δεν μπορεί να περιγραφούν με ευκολία λόγω της ποικιλομορφίας των διαφορετικών λειτουργιών που παρουσιάζουν αλλά και της ποικιλίας και της ποσότητας των διαθέσιμών τους στοιχείων. Και τα δύο αυτά χαρακτηριστικά οδηγούν στην επιλογή του κατάλληλου εργαλείου ανίχνευσης της απάτης.

## 1.4 Ταξινόμηση Ειδών Απάτης

Καθώς είναι πολύ δύσκολο να καλύψουμε όλες τις περιοχές στις οποίες μπορούν να εφαρμοστούν οι στατιστικές μέθοδοι ανίχνευσης της απάτης, επιλέξαμε να αναφερθούμε σε εκείνες τις περιοχές για τις οποίες υπάρχει ένας μεγάλος όγκος δεδομένων και αρχείων καθώς επίσης και αρκετή βιβλιογραφία με χρήσιμο υλικό. Η ταξινόμηση των ειδών της απάτης είναι η ακόλουθη:

- **Απάτη Πιστωτικών Καρτών** (*Credit Card Fraud*)
- **Νομιμοποίηση εσόδων από παράνομες δραστηριότητες** (*Money Laundering*)
- **Απάτη Τηλεπικοινωνιών** (*Telecommunications Fraud*)
- **Εισβολή σε Υπολογιστικά Συστήματα** (*Computer Intrusion*)
- **Ιατρική και Επιστημονική Απάτη** (*Medical and Scientific Fraud*)

Στη συνέχεια θα περιγράψουμε εν συντομία το κάθε είδος απάτης ξεχωριστά.

*α. Απάτη Πιστωτικών Καρτών*

Είναι γνωστό ότι οι απάτες μέσω πιστωτικών καρτών αποτελούν ένα σοβαρό και εντεινόμενο πρόβλημα. Δυστυχώς δεν είναι εύκολο να ποσοτικοποιήσουμε την έκταση αυτού του είδους απάτης καθώς τα νούμερα συνεχώς αλλάζουν. Ένας ακόμη σημαντικό πρόβλημα αποτελεί το γεγονός ότι οι επιχειρήσεις είναι απρόθυμες να δημοσιεύσουν τα νούμερα αυτά γιατί δε θέλουν να φοβίσουν το κοινό, γι'αυτό και γίνονται πολλές εκτιμήσεις εκ των οποίων θα αναφερθούν αντιπροσωπευτικά κάποιες. Για παράδειγμα, ο Leonard (1993) δήλωσε ότι το κόστος της απάτης με βάση την κάρτα Visa/MasterCard στον Καναδά τις χρονολογίες 1989, 1990 και 1991 ήταν $19, 29 και 46 εκατομμύρια (Καναδικά), αντίστοιχα. Οι Ghosh και Reilly (1994) δήλωσαν την απώλεια γύρω στα $850 εκατομμύρια το χρόνο στις Ηνωμένες Πολιτείες για όλους τους τύπους απατών των πιστωτικών καρτών. Επιπλέον οι Aleskerov, Freisleben και Rao (1997) ανέφεραν τις εκτιμήσεις $700 εκατομμυρίων για την απάτη της Visa/MasterCard στις Ηνωμένες Πολιτείες για κάθε έτος και $10 δισεκατομμύρια παγκοσμίως το 1996. Είναι σημαντικό να αναφέρουμε ότι από το 1996 μέχρι το 2000 στο Ηνωμένο Βασίλειο υπήρξε μία σημαντική αύξηση στις συνολικές απώλειες που σημειώθηκαν μέσω της απάτης των πιστωτικών καρτών. Συγκεκριμένα τα έτη 1997, 1998, 1999 και 2000 σημειώθηκαν απώλειες γύρω στα £122 εκατομμύρια, £135 εκατομμύρια, £188 εκατομμύρια και £293 εκατομμύρια αντίστοιχα. Ωστόσο οι απώλειες που σημειώθηκαν λόγω απάτης στις Βρετανικές κάρτες μειώθηκαν κατά 7% από £365.4 εκατομμύρια το 2010 σε £341.0 εκατομμύρια το 2011, με αποτέλεσμα την τριετή μείωση κατά 45% περίπου. Οι απώλειες αυτές είναι στα χαμηλότερα επίπεδα από το 2000. Αυτή η υγιής τάση είναι αποτέλεσμα των προσπαθειών του κλάδου για την αποτροπή, την ανίχνευση και την δίωξη των απατεώνων.

Το κύριο βάρος για την ανίχνευση απάτης που σχετίζεται με πιστωτικές κάρτες πέφτει στην ίδια την επιχείρηση και στον εκδότη καρτών. Είναι πολύ σημαντικό το ενδιαφέρον που θα δείξουν όσον αφορά την εμπόδιση της απάτης και την ανίχνευσή της όσο το δυνατόν συντομότερα όταν έχει αποτύχει η πρόληψη, γιατί έτσι θα υπάρχει εμπιστοσύνη μεταξύ της εταιρείας και του πελάτη.

Οι πιο συνηθισμένοι τύποι απάτης μέσω πιστωτικών καρτών είναι οι εξής:

- Συναλλαγές μέσω χαμένων/κλεμμένων καρτών.
- Έκδοσης καρτών με χρήση ψευδών προσωπικών ή και οικονομικών στοιχείων.

- Πλαστογραφία, δηλαδή δημιουργία κάρτας κλώνου που θα μπορεί να χρησιμοποιηθεί σε συναλλαγές σε επιχειρήσεις ενώ η κανονική κάρτα βρίσκεται στα χέρια του κατόχου της.
- Υποκλοπή και παράνομη χρήση του αριθμού της κάρτας σε συναλλαγές από απόσταση ή στο Internet.
- ATM Skimming, δηλαδή η υποκλοπή των στοιχείων της μαγνητικής ταινίας της κάρτας και του PIN κατά τη διάρκεια συναλλαγής σε ATM.
- Υποκλοπή προσωπικών στοιχείων πελατών και στοιχείων καρτών μέσα από αρχεία που τηρούν οι επιχειρήσεις.

Ο πλέον συνηθισμένος τρόπος απάτης πιστωτικών καρτών από αυτούς που αναφέραμε προηγουμένως είναι η χρήση μιας κλεμμένης κάρτας. Σε αυτήν την περίπτωση, ο απατεώνας ξοδεύει όσο το δυνατόν περισσότερα και όσο πιο σύντομα μπορεί, προτού να ανιχνευθεί η κλοπή της κάρτας και την ακυρώσουν. Όσο πιο γρήγορα γίνει η ανίχνευση της κλοπής της πιστωτικής κάρτας τόσο το καλύτερο καθώς θα μπορέσουμε έτσι να αποτρέψουμε μεγάλες απώλειες χρημάτων.

Ο τύπος απάτης που μπορεί να προκαλέσει τις μεγαλύτερες απώλειες είναι με αγορές που γίνονται από απόσταση καθώς σε τέτοιου είδους περιπτώσεις δε χρειάζεται η παρουσία του χρήστη-αγοραστή, αυτοπροσώπως την ώρα της συναλλαγής. Το μόνο που χρειάζεται για τέτοιου είδους συναλλαγές είναι η δήλωση των στοιχείων της κάρτας. Τέτοιου είδους συναλλαγές γίνονται συνήθως μέσω τηλεφωνικών αγορών ή αγορών μέσω του Διαδικτύου. Για να μπορέσουμε να επιτύχουμε κάτι τέτοιο χρειάζεται να αποκτήσουμε τα στοιχεία της κάρτας εν αγνοία του κατόχου. Αυτό μπορεί να συμβεί με διάφορους τρόπους, ένας απο τους οποίους είναι το «Skimming». Σύμφωνα με τη μέθοδο αυτή, οι απατεώνες αντιγράφουν παρανόμως τη μαγνητική λωρίδα της πιστωτικής κάρτας μέσω μιας μικρής φορητής συσκευής ανάγνωσης καρτών η οποία συνήθως τοποθετείται στην υποδοχή της κάρτας του ATM. Μόλις η κάρτα εισάγεται μέσα στην υποδοχή καρτών η συσκευή αυτή αντιγράφει τα στοιχεία της.

Οι βάσεις δεδομένων πιστωτικών καρτών περιέχουν πληροφορίες για κάθε συναλλαγή. Αυτές οι πληροφορίες περιλαμβάνουν τη συμπλήρωση πολλών στοιχείων, π.χ τον αριθμό του λογαριασμού, το είδος της αγοράς, τον τύπο της πιστωτικής κάρτας, το μέγεθος της συναλλαγής και άλλα πολλά. Κάποια από τα στοιχεία αυτά αφορούν αριθμητικά δεδομένα όπως για παράδειγμα το μέγεθος της συναλλαγής, ενώ κάποια άλλα είναι στοιχεία για κατηγορικές μεταβλητές όπως για παράδειγμα ο τύπος της πιστωτικής κάρτας. Τέτοιου είδους δεδομένα έχουν

οδηγήσει στην εφαρμογή της στατιστικής, της μηχανικής μάθησης και των εργαλείων της εξόρυξης δεδομένων.

Το να μπορέσουμε να ανιχνεύσουμε ότι ένας λογαριασμός βρίσκεται σε κίνδυνο μπορεί να γίνει, για μεμονωμένους πελάτες, χρησιμοποιώντας την συνολική τους εικόνα σε προηγούμενες συναλλαγές. Στην αρχή της έκδοσης μίας κάρτας οι χρήστες είναι αρκετά επιφυλακτικοί στις αγορές τους, επομένως ξαφνικές αγορές πολλών μικρών ηλεκτρικών συσκευών ή κοσμημάτων καθώς επίσης η άμεση χρήση μίας νέας κάρτας σε πολλά διαφορετικά σημεία είναι ένα αρκετά ύποπτο στοιχείο.

*β. Νομιμοποίηση εσόδων από παράνομες δραστηριότητες*

Με την έννοια «ξέπλυμα χρήματος» ή «ξέπλυμα μαύρου χρήματος» όπως αλλιώς αναφέρεται ώστε να δοθεί η δέουσα έμφαση, χαρακτηρίζεται οποιαδήποτε οικονομική συναλλαγή που γίνεται συνήθως μέσω χρημάτων ώστε παράνομα ποσά που λαμβάνονται από κάποιον να "μεταμφιέζονται" με τέτοιο τρόπο ώστε τελικά να εμφανίζονται ως νόμιμα. Η διαδικασία της νομιμοποίησης περιλαμβάνει οποιοδήποτε είδος αξιόποινης πράξης οι οποίες μπορεί να κυμαίνονται από φοροδιαφυγή και πλαστογραφία μέχρι τα ναρκωτικά και την εμπορία ανθρώπων. Αυτό που θέλουν να πετύχουν οι εγκληματίες αυτής της κατηγορίας είναι να αποκρύψουν την παράνομη προέλευση των χρημάτων, προσπαθώντας να μεγιστοποιήσουν το κέρδος.

Το Γραφείο Αξιολόγησης Τεχνολογίας (*Office of Technology Assessment* (*OTA*)) αναφέρει για το 1995 ότι: «Ομοσπονδιακές υπηρεσίες εκτιμούν γύρω στα $ 300 δισεκατομμύρια δολάρια νομιμοποίησης ετησίως, σε παγκόσμιο επίπεδο. Από αυτά περίπου τα 40 με 80 δισεκατομμύρια δολάρια μπορεί να είναι κέρδη από ναρκωτικά που έγιναν στις Ηνωμένες Πολιτείες». Το 1996 το Διεθνές Νομισματικό Ταμείο εκτιμά ότι 2-5% της παγκόσμιας οικονομίας σε όλο τον κόσμο συμμετέχουν στη νομιμοποίηση παράνομου χρήματος. Ωστόσο, η Ομάδα Χρηματοοικονομικής Δράσης για το Ξέπλυμα Χρήματος (*Financial Action Task Force on Money Laundering* (*FATF*)), ένα διακυβερνητικό όργανο που συγκροτήθηκε για την καταπολέμηση της νομιμοποίησης παράνομου χρήματος, παραδέχτηκε ότι «συνολικά είναι απολύτως αδύνατο να παραγάγει μια αξιόπιστη εκτίμηση του ποσού που νομιμοποιείται και ως εκ τούτου η FATF δεν δημοσιεύει οποιαδήποτε ποσά».

Τα εμβάσματα αποτελούν ένα φυσικό τρόπο νομιμοποίησης παράνομου χρήματος. Σύμφωνα με αναφορές της OTA, το 1995 πραγματοποιούνται μεταφορές χρημάτων ημερησίως που φτάνουν γύρω στο μισό εκατομμύριο μέσω των συστημάτων Fedwire και CHIPS, με σχεδόν

το ένα τέταρτο του ενός εκατομμύριου μεταφορών να το χρησιμοποιεί το σύστημα SWIFT. Από αυτά εκτιμάται περίπου ότι 0,05 έως 0,1% των συναλλαγών συμμετέχουν σε εσόδα από παράνομες δραστηριότητες. Η ανίχνευση της νομιμοποίησης παράνομου χρήματος είναι μία αρκετά δύσκολη διαδικασία σε σύγκριση με την ανίχνευση της απάτης σε άλλους τομείς όπως π.χ στην περίπτωση των πιστωτικών καρτών. Η ανίχνευση της απάτης όσον αφορά το ξέπλυμα χρήματος μπορεί να πάρει χρόνια μέχρι ένας προσωπικός λογαριασμός να θεωρηθεί μέρος μίας παράνομης διαδικασίας. Αυτό οφείλεται στο ότι οι διαθέσιμες πληροφορίες που έχουν για τους κατόχους λογαριασμών οι επενδυτικές τράπεζες είναι αρκετά λιγότερες από εκείνες που έχουν οι τράπεζες λιανικών επενδύσεων. Επομένως ένα καλύτερο σύστημα καταγραφής των πελατών θα βοηθούσε αρκετά.

Στις Η.Π.Α το 1970 έγιναν προσπάθειες για την καταπολέμηση του φαινομένου της νομιμοποίησης εσόδων από παράνομες δραστηριότητες υποχρεώνοντας όλες τις τράπεζες να αναφέρουν στις αρχές όσες συναλλαγές ανέρχονταν σε ποσό μεγαλύτερο των £10.000. Για το λόγο αυτό οι δράστες έπρεπε να βρουν μεθόδους ώστε να συμβαδίσουν με τις νέες τακτικές της νομοθεσίας. Μία από αυτές τις μεθόδους, η οποία αποτελεί πλέον και την πιο διαδεδομένη μέθοδο νομιμοποίησης εσόδων από παράνομες ενέργειες, είναι η **διάρθρωση** (*structuring*) γνωστή και ως «μέθοδος του μυρμηγκιού» (*smurfing*), η διαδικασία της οποίας είναι σχετικά απλή. Η μέθοδος αυτή βασίζεται στη διάσπαση του "βρώμικου" χρήματος σε μικρότερα όπου το καθένα δε θα ξεπερνάει κάποιο όριο αναφοράς (π.χ τα £10.000 που όριζαν οι Η.Π.Α). Για τη μεταφορά αυτή θα πρέπει να υπάρχουν κάποιοι μεσολαβητές οι οποίοι είναι υπεράνω υποψίας, αλλά και για να μπορέσουν να αναλάβουν να κάνουν τυχόν καταθέσεις σε τραπεζικούς λογαριασμούς. Άλλες μέθοδοι που χρησιμοποιούνται από τους δράστες για τη νομιμοποίηση παράνομων εσόδων είναι οι εξής:

- Η δημιουργία κάλυψης καταστημάτων: πρόκειται για ιδρύματα και εταιρείες κέλυφος οι οποίες αποκρύπτουν τον πραγματικό ιδιοκτήτη των χρημάτων.
- Ο παλαιότερος και απλούστερος τρόπος, είναι η φυσική μεταφορά χρήματος η οποία γίνεται με απλή τοποθέτηση των χρημάτων σε διάφορες συσκευασίες, βαλίτσες, εμπορεύματα τα οποία αποστέλλονται με απλό ταχυδρομείο στον προορισμό τους.
- Η μέθοδος που βασίζεται στο εμπόριο εσόδων από παράνομες δραστηριότητες υπό ή-υπερτιμώντας τα τιμολόγια , προκειμένου να αποκρύψουν τη μεταφορά χρημάτων
- Μία τεχνική η οποία χρησιμοποιείται σε συναλλαγές που αφορούν μεγάλα ποσά είναι η δόλια συνεργασία μεταξύ τραπεζικών υπαλλήλων ή και τραπεζικού ιδρύματος με τους

απατεώνες από τους οποίους μπορούν να αντλήσουν πληροφορίες που αφορούν τις μεθόδους ανίχνευσης που έχουν εφαρμοστεί.

Η διαδικασία νομιμοποίησης εσόδων από παράνομες ενέργειες περιλαμβάνει 3 στάδια τα οποία είναι:

**i. Τοποθέτηση** (*Placement*): πρόκειται για τη κατάθεση χρημάτων σε κάποιο τραπεζικό σύστημα ή σε κάποια νόμιμη επιχείρηση. Το στάδιο αυτό περιλαμβάνει ακόμη και τη «φυσική» διασυνοριακή μεταφορά χαρτονομισμάτων. Κλασσικά παραδείγματα είναι η αγορά αγαθών μεγάλης αξίας, όπως έργα τέχνης, αεροπλάνα, πολύτιμα μέταλλα και πετρώματα ή ακόμη χρηματιστηριακών τίτλων, επιταγών και άλλων τραπεζογραμμάτιων.

**ii. Διαστρωμάτωση** (*Layering*): πρόκειται για τη διεξαγωγή σύνθετων χρηματοοικονομικών συναλλαγών, προκειμένου να καταστεί η ανίχνευση της προέλευσής του αδύνατη.

**iii. Ολοκλήρωση** (*Integration*): συνεπάγεται στην απόκτηση του πλούτου που προέρχεται από τις συναλλαγές παράνομων κεφαλαίων με την νομιμοποίησή τους.

Ωστόσο είναι δύσκολο ως σχεδόν ακατόρθωτο να περιγράψουμε πότε μία συναλλαγή μπορεί να θεωρηθεί ως απάτη. Αυτό που θα μπορούσε να βοηθήσει είναι οι κινήσεις που γίνονται πάνω σε έναν λογαριασμό, π.χ. αν δούμε κάποιον να καταθέτει το ποσό των £10.000 δε θα θεωρηθεί ύποπτο ενώ αν γίνουν πολλές τέτοιου είδους καταθέσεις και σε διαφορετικές τράπεζες τότε αυτό θεωρείται ύποπτο. Ακόμη και μία κατάθεση ενός μεγάλου ποσού δε θα κριθεί ύποπτη αλλά μια τέτοια κατάθεση για την οποία ακολουθεί ανάληψή της σε σύντομο χρονικό διάστημα θεωρείται αρκετά ύποπτη.

*γ. Απάτη Τηλεπικοινωνιών*

Λόγω της ραγδαίας εξέλιξης της τεχνολογίας καθώς επίσης και του ιδιαίτερου χαμηλού κόστους που έχει η κινητή τηλεφωνία τα τελευταία χρόνια έχει οδηγήσει στην σημαντική αύξηση ιδιοκτησίας κινητών τηλεφώνων μιας και έχουν γίνει αρκετά προσιτά για όλο τον πληθυσμό. Αυτό έχει καταλήξει στο αποτέλεσμα η απάτη παγκόσμιας κινητής τηλεφωνίας να έχει αυξηθεί σημαντικά κάνοντας πολλές εκτιμήσεις για το κόστος που έχει αυτό το είδος απάτης. Παραδείγματα τέτοιων εκτιμήσεων είναι τα εξής: οι Cox *et al.* (1997) έδωσαν μία εκτίμηση του κόστους γύρω στο £1 εκατομμύριο το χρόνο. Επιπλέον ο Hoath (1998) και η εταιρεία Fraud Management Limited (FML (2003)) αναφέρουν απώλειες σύμφωνα με παγκόσμια απάτη τηλεπικοινωνιών οι οποίες εκτιμώνται σε δεκάδες δισεκατομμύρια δολάρια κάθε έτος. Ο οργανισμός ελέγχου σύνδεσης απάτης (*Communications Fraud Control Association* (*cfca.org*)) εκτιμά σε διαστήματα την έκταση της απάτης των τηλεπικοινωνιών παγκοσμίως. Το 1999 αυτή η εκτίμηση ήταν $ 12 δισεκατομμύρια, το

2003 ήταν μεταξύ $ 35 και $ 40 δισεκατομμύρια, το 2006 ήταν μεταξύ $ 55 και $ 60 δισεκατομμύρια, και το 2009 ήταν μεταξύ $ 70 και 78 δισεκατομμυρίων δολαρίων. Παρόλο που μπορεί να έχουμε πολλά παραδείγματα τα οποία να μας δείχνουν τα ποσά της απάτης αυτής τα οποία είναι και αρκετά μεγάλα, δε πρέπει ωστόσο να ξεχνάμε ότι πρόκειται για εκτιμήσεις και όχι για ακριβή τιμές.

Υπάρχουν πολλά διαφορετικά είδη απάτης τηλεπικοινωνιών όπως αναφέρουν οι Shawe-Taylor *et al.* (2000), τα οποία μπορούν να ταξινομηθούν σε δύο κατηγορίες.
Αυτές είναι:

- Η **Απάτη Συνδρομής** (*subscription fraud*) η οποία εμφανίζεται όταν ο απατεώνας εγγράφεται σε μία υπηρεσία δηλώνοντας ψευδή στοιχεία χωρίς να έχει την πρόθεση πληρωμής. Επομένως όσες συναλλαγές γίνονται μέσω αυτού του λογαριασμού θα είναι δόλιες. Οι λογαριασμοί αυτοί συνήθως χρησιμοποιούνται για αγορές μέσω κλήσεων είτε για εντατική προσωπική χρήση. Ο Hoath (1998) χαρακτηρίζει ότι ίσως αυτή η μορφή αποτελεί την πιο σημαντική και διαδεδομένη απάτη τηλεπικοινωνιών παγκοσμίως. Η απάτη συνδρομής μπορεί να διαιρεθεί σε δύο κατηγορίες : (α) για το κέρδος, (β) για προσωπικό όφελος. Μία τεχνική ανίχνευσης της απάτης συνδρομής είναι μέσω της εξέτασης του λογαριασμού πληρωμής και την παρακολούθηση της συμπεριφοράς των πελατών. Η απάτη αποκαλύπτεται σε κάποιο σημείο από τους ανεξόφλητους λογαριασμούς που συσσωρεύονται και οι οποίοι είναι μεγάλοι.
- Η **Υπερτιθέμενη Απάτη** (*Superimposed fraud*) στην οποία οι απατεώνες προβαίνουν σε παράνομη χρήση ενός νόμιμου λογαριασμού χρησιμοποιώντας διάφορες τεχνικές. Η απάτη αυτή μπορεί να εντοπιστεί συνήθως με την εμφάνιση άγνωστων κλήσεων σε έναν λογαριασμό. Υπάρχουν διάφοροι τρόποι πραγματοποίησης της απάτης αυτής μερικοί απο τους οποίους είναι:

  (α) η **κλωνοποίηση κινητών τηλεφώνων** (mobile phone cloning) δηλαδή η απάτη μέσω της αντιγραφής των στοιχείων λογαριασμών νομότυπων συνδρομητών,

  (β) τα **είδωλα** (*ghosting*) όπου πρόκειται για την απόκτηση δωρεάν ή φθηνής τιμής με τεχνικά μέσα της παραπλάνησης του δίκτυου,

  (γ) το **ανακάτεμα** (*tumbling)*, που εκμεταλλευόμενο αδυναμίες στα συστήματα αναγνώρισης συνδρομητών επέτρεπε στον απατεώνα να αλλάζει συνεχώς αριθμούς αναγνώρισης συνδρομητών.

Η υπερτιθέμενη απάτη είναι συνήθως πιο δύσκολη ως προς την ανίχνευσή της καθώς μπορεί να μείνει για ένα μεγάλο χρονικό διάστημα χωρίς να έχει εντοπιστεί. Οι Davis και Goyal. (1993) χαρακτήρισαν αυτές τις απάτες σημαντικές απειλές για τα έσοδα των πάροχων κινητής

τηλεφωνίας. Τα πρώτα συστήματα ανίχνευσης απάτης στις τηλεπικοινωνίες αναζητούσαν τις περιπτώσεις που ο ίδιος αριθμός χρησιμοποιούταν ταυτόχρονα από περισσότερους από ένα χρήστες (μηχανισμοί αναγνώρισης επικαλυπτόμενων τηλεφωνημάτων) και εκείνες που γινόντουσαν σε σύντομο χρονικό διάστημα από τοποθεσίες που απείχαν πολύ μεταξύ τους (παγίδες ταχύτητας).

Η αγορά τηλεπικοινωνιών με την πάροδο του χρόνου θα γίνει ακόμη πιο πολύπλοκη με αποτέλεσμα να υπάρχουν πιο πολλές πιθανότητες να συμβεί απάτη στον κλάδο αυτό. Μέχρι στιγμής η έκταση της απάτης υπολογίζεται λαμβάνοντας υπόψη παράγοντες όπως για παράδειγμα το μήκος κλήσεων και τα τιμολόγια. Γι'αυτό το λόγο η νέα γενιά των κινητών τηλεφώνων θα πρέπει να λάβει υπόψην της περισσότερο το περιεχόμενο των κλήσεων και την προτεραιότητα της ίδιας της κλήσης.

*δ. Εισβολή σε Υπολογιστικά Συστήματα*

Η ταχεία επέκταση του δικτύου υπολογιστών σε ολόκληρο τον κόσμο έχει κάνει την ασφάλεια τους ένα κρίσιμο ζήτημα. Η απάτη εισβολής σε συστήματα υπολογιστών έχει σε αρκετές περιπτώσεις μεγάλες επιπτώσεις και η ανίχνευση της εισβολής τους είναι ένα αρκετά ευαίσθητο θέμα που απαιτεί εκτεταμένη έρευνα. Οι εισβολείς υπολογιστών ή «Hackers» όπως αλλιώς τους επικαλούμαστε, μπορούν να θεωρηθούν ως μοντέρνοι διαρρήκτες καθώς πολλοί από αυτούς προσπαθούν να βρουν κωδικους, να κλέψουν πληροφορίες, να διαβάσουν τα e-mail ενώ κάποιοι άλλοι απλά θέλουν να αποδείξουν ότι απλά κατάφεραν να εισβάλλουν στον υπολογιστή κάποιου και να εισέλθουν στα συστήματά του. Θα πρέπει όμως να αναφέρουμε ότι μπορούν να δεχτούν εισβολή όχι μόνο απλοί άνθρωποι στον υπολογιστή τους και στα συστήματά τους αλλά ακόμη και για πολύ καλά προστατευμένες περιοχές μπορεί να βρίσκονται σε κίδυνο. Στη συνέχεια θα αναφέρουμε δύο παραδείγματα που αφορούν την εισβολή σε τέτοιου είδους συστήματα σε πολύ καλά προστατευόμενες περιοχές στις οποίες μπόρεσαν να εισβάλλουν.

Το πρώτο παράδειγμα είναι "η περίπτωση ενός 16-χρονου αγοριού, του Jonathon James από το Μαϊάμι που ήταν γνωστός στο διαδίκτυο ως "c0mrade". Ήταν ο πρώτος ανηλίκας ο οποίος καταδικάστηκε σε φυλάκιση 6 μηνών λόγω κατασκοπείας του Πενταγώνου και των συστημάτων ηλεκτρονικών υπολογιστών της NASA. Η NASA ισχυρίστηκε ότι ο 16-χρονος έκλεψε πάνω από 1,7 εκατομμύρια δολάρια αξίας του λογισμικού, και χρειάστηκε να κλείσει τα συστήματα υπολογιστών της κάτι το οποίο της κόστισε 41.000 δολάρια. Πως όμως ένα 16-χρονο αγόρι μέσα από ένα εκατομμύριο ανθρώπων μπόρεσε να κατασκοπεύσει τα συστήματα της NASA της ασφάλειας των συστημάτων της? Ο ίδιος ο μικρός απάντησε ότι ο κωδικός τους ήταν

χάλια και ότι σίγουρα το λογισμικό της δεν άξιζε 1,7 εκατομμύρια δολάρια όπως ισχυρίστηκαν οι ίδιοι."

Το δεύτερο παράδειγμα εισβολής σε υπολογιστικά συστήματα είναι ότι το Φεβρουάριο του 2000, σε ηλικία 15 χρονών, ο Michael Calse, γνωστός με το ψευδώνυμο Malfiaboy, κατάφερε να ρίξει τις σελίδες των Yahoo!, Amazon, eBay, CNN και Dell. Κατάφεραν να τον εντοπίσουν μετά από σχόλια που έκανε σε chatroom για τα κατορθώματά του. Μέχρι και ο Bill Clinton είχε μπλεχτεί στην υπόθεση. Ο Michael Calse μπήκε για 8 μήνες στη φυλακή όταν ήταν 16 χρονών. Μελετητές υπολόγισαν ότι οι επιθέσεις του προκάλεσαν ζημιές στην παγκόσμια οικονομία της τάξης των 1.6 δισεκατομμυρίων δολαρίων.

Αν καταφέρουμε να εμποδίσουμε την εισβολή των hacker από την διείσδυσή τους στα συστήματα υπολογιστών ή αν μπορέσουμε να την ανιχνεύσουμε εγκαίρως τότε το εν λόγω έγκλημα θα μπορέσει να εξαλειφθεί οριστικά. Όπως όμως συμβαίνει και με την απάτη γενικότερα έτσι και εδώ μόλις κάποια μέθοδος εισβολής ανιχνευθεί και γίνει γνωστή η διαδικασία αντιμετώπισής της, ο εισβολέας θα αναζητήσει έναν καινούριο τρόπο εξαπάτησης. Λόγω της ραγδαίας εξέλιξης των δικτύων ηλεκτρονικών υπολογιστών η ασφάλεια τους αποτελεί ένα σημαντικό ζήτημα. Η εισβολή ορίζεται ως ένα σύνολο δράσεων που λαμβάνει ένα υπολογιστικό σύστημα από μια κανονική κατάσταση σε μια κατάσταση με κίνδυνο. Και εδώ όπως και στις άλλες απάτες χρησιμοποιούμε εποπτευόμενες και μη εποπτευόμενες μεθόδους μόνο που στα πλαίσια της εισβολής Η/Υ τις ονομάζουμε **ανίχνευση κακής χρήσης** (*misuse detection*) και **ανίχνευση ανωμαλιών** (*anomaly detection*) αντίστοιχα.

Η ανίχνευση κακής χρήσης πρόκειται για μία προσέγγιση ανίχνευσης επιθέσεων στα αδύνατα σημεία ενός συστήματος και γίνεται προσπάθεια για να περιγράψουν την συμπεριφορά των εισβολέων προκειμένου να αναγνωρίσουν μία πιθανή περίπτωση τέτοιας επίθεσης. Αυτές οι μέθοδοι μπορούν να εντοπίσουν μόνο τις εισβολές εκείνες οι οποίες έχουν εμφανιστεί στο παρελθόν.

Όσον αφορά τώρα την ανίχνευση ανωμαλιών βασίζεται στην περιγραφή συμπεριφορών έξω από τα κανονικά μοτίβα χρήσης του συστήματος. Έχουν ως στόχο δηλαδή να εντοπίσουν ότι είναι ασυνήθιστο, γι'αυτό και αυτού του είδους οι μέθοδοι αναφέρονται σε πρόβλημα προτύπων σε δεδομένα που δεν συμμορφώνονται σύμφωνα με την αναμενόμενη συμπεριφορά. Αυτά τα μη συμμορφούμενα πρότυπα συχνά αναφέρονται σε ακραίες τιμές, ανωμαλίες, εξαιρέσεις, εκπλήξεις, ιδιομορφίες κ.α. Από αυτές οι ακραίες τιμές και οι ανωμαλίες αποτελουν τις συχνότερες μορφές όσον αφορά την ανίχνευση ανωμαλιών. Ωστόσο η ίδια μπορεί να οδηγήσει

σε εσφαλμένους συναγερμούς λόγω της αλλαγής της συμπεριφοράς των χρηστών και επομένως ίσως οδηγηθούμε σε ένα μεγάλο ποσοτό λαθών.

Πολλοί όταν σκέφτονται την προστασία του υπολογιστή τους και την ασφάλεια του δικτύου τους έχουν στο μυαλό τους την εγκατάσταση ενός **τείχου προστασίας** (*Firewall*). Τα τείχη προστασίας είναι ευρέως διαδεδομένα ως ένα πρώτο επίπεδο προστασίας σε έναν πολύπλοκο σχεδιασμό ασφαλείας. Κατά κύριο λόγο ενεργούν ως συσκευή ελέγχου πρόσβασης ρυθμισμένη ούτως ώστε να επιτρέπει ή να απορρίπτει πακέτα δεδομένων που περνούν από ένα δίκτυο υπολογιστών σε ένα άλλο. Μολονότι το τείχος προστασίας είναι απαραίτητο για την ασφάλεια του δικτύου του υπολογιστή, δεν επαρκεί μονάχα αυτό. Υπάρχουν κατάλληλα εμπορικά προϊόντα τα οποία βοηθούν στην ανίχνευση εισβολής του υπολογιστή όπως για παράδειγμα τα ***next generation intrusion detection expert system (NIDES)*** και τα ***Cisco secure intrusion detection system (CSIDS)***. Μια λεπτομερειακή περιγραφή τους μπορεί κανείς να βρει στην τεχνική αναφορά των Anderson *et al.* (1995).

*ε. Ιατρική και Επιστημονική Απάτη*

Όπως είδαμε, ένας από τους τομείς που μπορεί να συμβεί απάτη είναι η Ιατρική. Η εμφάνισή της στον κλάδο αυτό μπορεί να γίνει με διάφορες μορφές. Η απάτη και η εξαπάτηση στην ιατρική έρευνα είναι κάτι το οποίο προκαλεί μεγάλη ανησυχία σε πολλές χώρες, καθώς μπορεί να οδηγήσει σε λανθασμένα συμπεράσματα τα οποία θα έχουν ως συνέπεια την απώλεια της εμπιστοσύνης του κοινού τόσο στην ίδια την ιατρική έρευνα όσο και στους γιατρούς. Το να "κατασκευάζουν" τα δεδομένα σε διάφορες κλινικές δοκιμές δεν είναι κάτι το οποίο θα πρέπει να μας εντυπωσιάσει, καθώς η προϊστορία μας έχει δείξει ότι κάτι τέτοιο συμβαίνει αρκετά συχνά (Buyse *et al.*, (1999)).

Ένα παράδειγμα που αναφέρουν οι Wu and Carlsson, (2010) αφορά τον Cyril Burt ο οποίος μέχρι και πριν τον θάνατό του το 1971 ήταν ένας αγαπητός ψυχολόγος ο οποίος συνέβαλε στη ψυχολογία και στη στατιστική. Ο Burt ήταν γνωστός για τις μελέτες του σχετικά με την κληρονομικότητα της ευφυΐας. Λίγο πριν το θάνατό του όμως η έρευνά του τέθηκε σε ανυποληψία μετά από στοιχεία που βρέθηκαν τα οποία δείχνουν ότι είχε πλαστογραφήσει τα δεδομένα της έρευνας. Κάποιοι μελετητές ωστόσο υποστήριξαν ότι ο Burt δεν διέπραξε απάτη. Αν και η υπόθεση του Burt είναι ακόμα συζητήσιμη, προκάλεσε σημαντικές αρνητικές επιδράσεις στη φήμη του και αυτό δίνει ένα μάθημα στους ερευνητές για το πόσο σημαντικό είναι να είναι ειλικρινείς.

Η Ιατρική απάτη σχετίζεται πολλές φορές με την **απάτη ασφάλισης** (*Insurance Fraud)*. Τα είδη της ασφαλιστικής απάτης είναι πολύ διαφορετικά, και εμφανίζονται σε όλους τους τομείς της ασφάλισης, ενώ παράλληλα ποικίλουν σημαντικά και ως προς την σοβαρότητά τους. Μπορούν να κυμανθούν από ελαφρώς υπερβολικούς ισχυρισμούς σε ισχυρισμούς που γίνονται σκοπίμως ώστε να προκαλέσουν ατυχήματα ή ζημιές. Η Ασφαλιστική απάτη αποτελεί ένα πολύ σημαντικό πρόβλημα, και οι κυβερνήσεις και οι άλλοι οργανισμοί καταβάλλουν προσπάθειες για να αποτρέψουν τέτοιες δραστηριότητες. Ο Allen (2000), ένας στατιστικός, μαζί με το Utah Bureau of Medicaid Fraud, αναφέρουν ότι πάνω από 10% των 800 εκατομμυρίων δολαρίων ετησίως μπορεί να κλαπεί.

Θα πρέπει σε αυτό το σημείο να αναφέρουμε ότι η Ιατρική δεν είναι η μοναδική επιστημονική περιοχή στην οποία κάποιοι έχουν παραποιήσει τα δεδομένα τους ώστε να μπορέσουν να υποστηρίξουν κάποιο ισχυρισμό. Τέτοιου είδους προβλήματα απάτης υπάρχουν και στον τομέα της επιστήμης τα οποία θέλουν εξίσου μεγάλη προσοχή. Ένα τέτοιο παράδειγμα αποτελεί η περίπτωση του καθηγητή Hwang Woo-suk του Εθνικού Πανεπιστημίου της Σεούλ, ο οποίος βρέθηκε ένοχος εκτενούς επιστημονικής απάτης. Μέχρι τον Ιανουάριο του 2006 θεωρούνταν ως ένας από τους ειδήμονες παγκοσμίως στην κλωνοποίηση και την έρευνα των εμβρυονικών κυττάρων. Ωστόσο το ίδιο το Πανεπιστήμιό του ανακάλυψε ότι είχε υπονομεύσει όλες τις σειρές κυττάρων που υποστήριζε σε άρθρα του στο περιοδικό Science το 2004 και το 2005 ότι είχε πάρει από κλωνοποιημένα ανθρώπινα έμβρυα. Η περίπτωση αυτή του Hwang δεν θα πρέπει να παραμεριστεί αλλά θα πρέπει να ληφθεί σοβαρά υπόψην και να γίνουν περισσότερες έρευνες στον τομέα της επιστήμης. Σε μία έρευνα που έγινε από 3247 επιστήμονες, περισσότεροι από το ένα τρίτο (1/3) εξ'αυτών ομολόγησαν ότι είχαν υποπέσει σε ακαδημαϊκό παράπτωμα – από παραποίηση ή λογοκλοπή (1,5%), μέχρι απόκρυψη στοιχείων από προηγούμενη έρευνά τους (6%), παράβλεψη χρήσης λανθασμένων δεδομένων απο άλλους (12,5%) και τέλος αλλαγή του σχεδιασμού, της μεθοδολογίας ή των αποτελεσμάτων ως αποτέλεσμα πίεσης από πηγή χρηματοδότησής τους (15,5%). Οι κοινωνιολόγοι που έκαναν την έρευνα αυτή προειδοποίησαν ότι η υπερβολική σημασία που δινόταν στις διάσημες υποθέσεις είχε ως αποτέλεσμα πολλά μικρότερης σημασίας παραπτώματα να αγνοούνται. Σημείωσαν μάλιστα ότι η έκταση των παραπτωμάτων που βρέθηκαν ήταν «εντυπωσιακή ως προς το εύρος και τη διάδοσή τους».

Δυστυχώς, ελάχιστες έρευνες μεγάλου βεληνεκούς έχουν γίνει πάνω στη διάδοση της επιστημονικής απάτης και έτσι είναι πολύ δύσκολο να γνωρίζουμε επακριβώς το μέγεθος του προβλήματος αυτού. Υπάρχουν δύο σημαντικά βιβλία, τα οποία αναφέρουν πολυάριθμες τέτοιες περιπτώσεις και τα οποία αποδεικνύουν ότι η απάτη στην επιστήμη έχει ξεκινήσει πολλά χρόνια πριν. Αυτά είναι:

- η καινοτόμος μελέτη "*Betrayers of the truth: Fraud and Deceit in the Halls of Science*" (Προδότες της αλήθειας: Απάτη και Εξαπάτηση στους Χώρους της Επιστήμης) των William Broad και Nicolas Wade το 1982.
- Το βιβλίο "*The Great Betrayal: Fraud in Science*" (η Μεγάλη Προδοσία: η Απάτη στην Επιστήμη) του Horace Judson το 2004.

# ΚΕΦΑΛΑΙΟ 2

# Ανίχνευση Απάτης Πιστωτικών Καρτών



## 2.1 Εισαγωγή

Οι **πιστωτικές κάρτες** (*Credit Cards*) αφορούν ένα σύγχρονο και διαδεδομένο τρόπο συναλλαγών, μία μορφή του λεγόμενου "πλαστικού χρήματος", παρέχοντας στους χρήστες τους τη δυνατότητα να πραγματοποιεί αγορές αγαθών και υπηρεσιών χωρίς άμεση καταβολή της αξίας τους, εντός βέβαια των πιστωτικών ορίων τους, από επιχειρήσεις που είναι συμβεβλημένες με τον τραπεζικό οργανισμό που τις εξέδωσε. Η δυνατότητα αυτή για ορισμένες κάρτες επεκτείνεται και στο εξωτερικό, ενώ άλλες μπορούν να χρησιμοποιηθούν για απεριόριστο όριο αγορών, με μόνη προϋπόθεση τη μηνιαία εξόφληση του λογαριασμού. Οι πιστωτικές κάρτες εκδίδονται κυρίως από πιστωτικά ιδρύματα (π.χ. τράπεζες) και, μεταξύ άλλων, η χρήση τους παρέχει τα ακόλουθα πλεονεκτήματα:

- **i.** ευκολία στις συναλλαγές σε όσες περιπτώσεις ο κάτοχος της κάρτας δεν έχει ή δεν θέλει να έχει μαζί του μετρητά καθώς παρέχεται η δυνατότητα ανάληψης μετρητών 24 ώρες το 24ωρο, ανάλογα βέβαια με το ύψος του πιστωτικού ορίου.
- **ii.** ασφάλεια στις συναλλαγές, γιατί ο κάτοχος της κάρτας δεν χρειάζεται να έχει μαζί του μετρητά διακινδυνεύοντας έτσι να τα χάσει.
- **iii.** εξασφάλιση περιόδου χάριτος αρκετών ημερών (π.χ. 25 ή 40 ημέρες) χωρίς τόκο από την ημερομηνία έκδοσης του λογαριασμού έως την ημερομηνία πληρωμής
- **iv.** λειτουργώντας ως κάρτες ηλεκτρονικών συναλλαγών παρέχουν τη δυνατότητα στους κατόχους τους να διενεργούν τραπεζικές πράξεις μέσω των Αυτόματων Ταμειολογιστικών Μηχανών (ΑΤΜ), όπως αναλήψεις, καταθέσεις, μεταφορά ποσών από λογαριασμό σε λογαριασμό, εξόφληση της δόσης ή ακόμη να έχει ενημέρωση για το υπόλοιπο των λογαριασμών του ή της πιστωτικής του κάρτας.

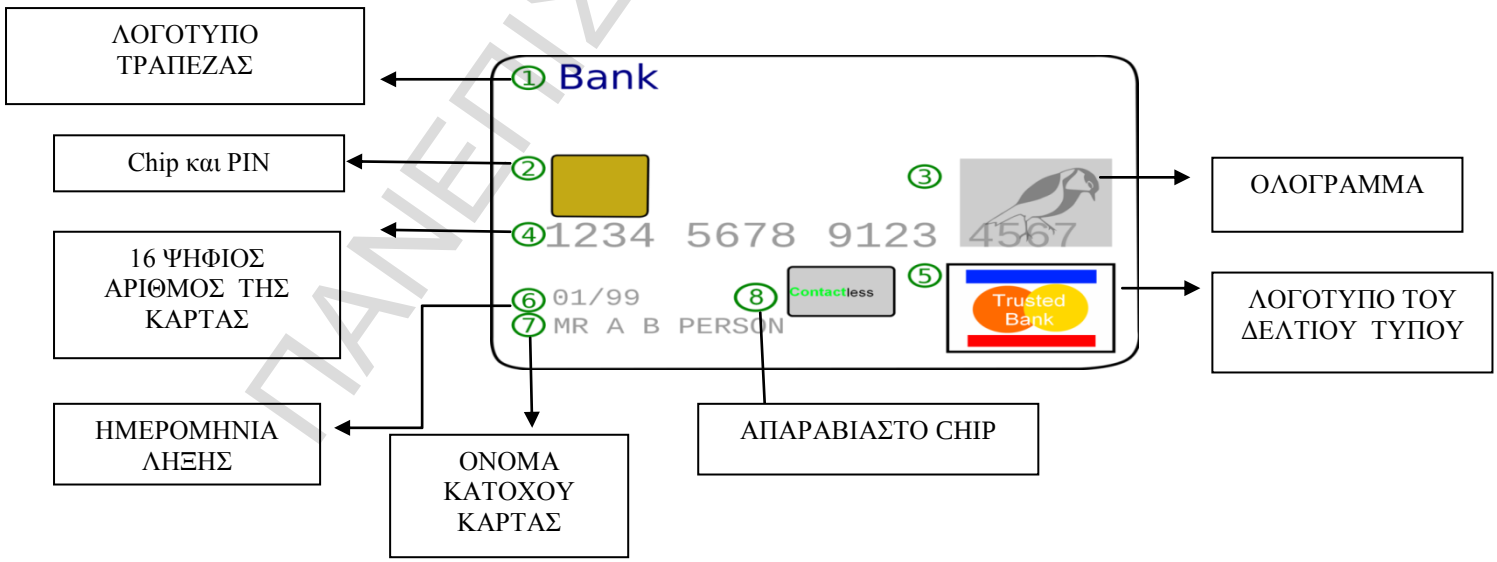
Οι κάρτες πληρωμής είναι ηλεκτρονικές κάρτες που περιέχουν πληροφορίες, οι οποίες μπορούν να χρησιμοποιηθούν για πληρωμές. Υπάρχουν τρία είδη καρτών πληρωμής: το ένα από αυτά και το πιο διαδεδομένο είναι οι πιστωτικές κάρτες τις οποίες αναφέραμε παραπάνω και τα

άλλα δύο είναι οι **χρεωστικές κάρτες** (*Debit Cards*) και οι **χρεωστικές κάρτες άμεσης πληρωμής** (*Debit Cards Direct Payment*). Στις χρεωστικές κάρτες το υπόλοιπο της κάρτας πληρώνεται ολόκληρο μόλις γίνει η λήψη μηνιαίας δήλωσης. Συνήθως, ο κάτοχος μιας χρεωστικής κάρτας λαμβάνει ένα δάνειο για 30 εως 45 μέρες το οποίο είναι ίσο με το υπόλοιπο της δήλωσής του. Όσον αφορά τώρα τις χρεωστικές κάρτες άμεσης πληρωμής, τα χρήματα για ένα αγοραζόμενο είδος αφαιρούνται αμέσως από τον τραπεζικό λογαριασμό του κατόχου. Η μεταφορά των χρημάτων από το λογαριασμό του κατόσου στο λογαριασμό του εμπόρου γίνεται σε 1 με 2 μέρες.

Η πιστωτική κάρτα έχει τη μορφή μιας πλαστικής κάρτας η οποία φέρει στη μια πλευρά της με ανάγλυφα στοιχεία έναν 16ψήφιο αριθμό της κάρτας, το ονοματεπώνυμο του κατόχου της, τη λήξη ισχύος της, καθώς και το πιστωτικό κατάστημα το οποίο τη χορήγησε. Στην άλλη πλευρά συνήθως υπάρχει η μαγνητική ταινία, θέση για την υπογραφή του κατόχου της και ο λογότυπος του οργανισμού που την εξέδωσε. Σε κάθε κάρτα υπάρχει ένας προσωπικός αριθμός αναγνώρισης ή αλλιώς P.I.N. (Personal Identification Number) ο οποίος είναι ένας απόρρητος κωδικός αριθμός που ισοδυναμεί με την υπογραφή του κατόχου της. Αυτός είναι απαραίτητος, σε συνδυασμό με την κάρτα, για την πραγματοποίηση συναλλαγών και πρέπει να παραμένει αυστηρά προσωπικός. Παρακάτω ακολουθεί μία απεικόνιση των στοιχείων που υπάρχουν επάνω σε κάθε πιστωτική κάρτα.

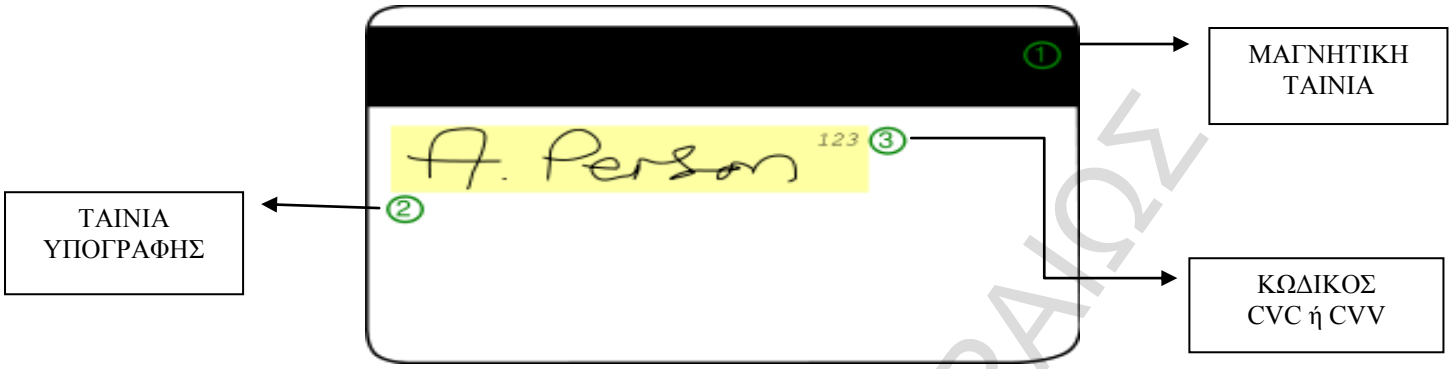
**Σχεδιάγραμμα 2.1.α.** *Απεικόνιση της μπροστινής όψης μιας πιστωτικής κάρτας*

Πηγή : *http:///www..Wikipedia..org*



**Σχεδιάγραμμα 2.1.β.** *Απεικόνιση της πίσω όψης μιας πιστωτικής κάρτας*

Πηγή : [http:///www..Wikipedia..org](http:///www..Wikipedia..org)




Σε γενικές γραμμές οι πιστωτικές κάρτες μπορούν να χωριστούν σε τρεις κατηγορίες: στην πρώτη ανήκουν όσες μπορούν να χρησιμοποιηθούν μόνο στο εσωτερικό της χώρας, στη δεύτερη οι κάρτες που η ισχύ τους επεκτείνεται και στο εξωτερικό, ενώ στην τρίτη κατηγορία ανήκουν οι κάρτες που χαρακτηρίζονται ως "χρυσές", "prestige", κτλ οι οποίες παρέχουν υψηλά πιστωτικά όρια και συνοδεύονται συνήθως από προνόμια και παροχές όπως για παράδειγμα ισχυρά ασφαλιστικά πακέτα και νομική προστασία.

Το λεγόμενο "πλαστικό χρήμα" εκτιμάται από τους ειδικούς ότι θα αποτελέσει τομέα ιδιαίτερης ανάπτυξης κατά τη μετάβαση στον 21ο αιώνα και αναμένεται ότι ένα σημαντικό μερίδιο του τραπεζικού μάρκετινγκ θα αφορά τη διάδοση και τη γενίκευση της χρήσης των πιστωτικών καρτών και γενικότερα των διάφορων μορφών πλαστικού χρήματος. Ήδη, όλο και περισσότερα μεγάλα καταστήματα στο εξωτερικό εκδίδουν πιστωτικές κάρτες, προσπαθώντας με τον τρόπο αυτό να διατηρήσουν έναν σημαντικό αριθμό πελατών.

Πρόσφατη καινοτομία είναι η έκδοση πιστωτικών καρτών που είναι συνδεδεμένες με αθλητικά σωματεία (π.χ. Παναθηναϊκός FC-Visa), πολιτιστικούς οργανισμούς (π.χ. Artion Visa σε συνεργασία με τον Οργανισμό Μεγάρου Μουσικής Αθηνών) κ.ά. Στόχος της είναι η διεύρυνση της πελατείας του τραπεζικού φορέα, η προβολή του συνεργαζόμενου φορέα και η εξυπηρέτηση του πελάτη (π.χ. εκπτώσεις, εξασφάλιση εισιτηρίων, θέσεων σε εκδηλώσεις του φορέα κ.ά.). Σημαντική αύξηση παρουσιάζουν τα τελευταία χρόνια και οι χρεωστικές τραπεζικές κάρτες, με τις οποίες μπορεί κανείς να πραγματοποιεί αγορές με απευθείας χρέωση του λογαριασμού του χωρίς κανένα όριο ή επιβάρυνση με τόκους. Τέλος, τα επόμενα χρόνια εκτιμάται ότι θα κυκλοφορήσουν και στην Ελλάδα οι λεγόμενες "έξυπνες κάρτες" (*smart cards*), που θα αποτελέσουν έναν σημαντικό νέο τρόπο συναλλαγών λειτουργώντας ως ηλεκτρονικά πορτοφόλια.

Πολλοί θεωρούν ότι οι πιστωτικές κάρτες και η χρήση τους αποτελεί μία πρόσφατη τάση η οποία χρονολογείται κατά τα τέλη του 20ου αιώνα, ωστόσο η αλήθεια είναι ότι οι πιστωτικές κάρτες υπήρχαν πολύ παλαιότερα, σε αντίθεση με όσα πιστεύουν οι περισσότεροι άνθρωποι. Σύμφωνα με την εγκυκλοπαίδεια Britannica, «η χρήση των πιστωτικών καρτών ξεκίνησε από τις Ηνωμένες Πολιτείες κατά τη διάρκεια της δεκαετίας του 1920, όταν μεμονωμένες επιχειρήσεις, όπως οι εταιρείες πετρελαίου και αλυσίδες ξενοδοχείων, άρχισαν να εκδίδουν πιστωτικές κάρτες στους πελάτες τους». Ωστόσο, κάποιοι λένε ότι η χρήση πιστωτικών καρτών άρχισε ήδη από τη δεκαετία του 1890 στην Ευρώπη. Οι ερευνητές λένε επίσης ότι οι πιστωτικές κάρτες εκείνες τις ημέρες δεν ήταν φτιαγμένες από πλαστικό, αλλά πιθανότατα από μεταλλικά νομίσματα, μεταλλικές πλάκες, ζελατίνη, μέταλλο, ίνες ή χαρτί.

Το 1946, ο διευθυντής John Biggins της Εθνικής Τράπεζας Flatbush του Μπρούκλιν στη Νέα Υόρκη επινόησε την πρώτη πιστωτική κάρτα τραπέζης. Η πρώτη παγκόσμια πιστωτική κάρτα, η οποία μπορούσε να χρησιμοποιηθεί σε διάφορα μέρη, εισήχθη από την εταιρεία Diners Club το 1950. Η ιστορία ξεκίνησε το 1949 όταν στο τέλος ενός επαγγελματικού δείπνου, ο Frank McNamara διαπίστωσε ότι δεν είχε μετρητά μαζί του για να πληρώσει το λογαριασμό και έτσι αναγκάστηκε να τηλεφωνήσει στη γυναίκα του. Επηρεασμένος από το συγκεκριμένο γεγονός, ο McNamara αποφάσισε να δώσει μία μόνιμη λύση στο πρόβλημα αυτό δίνοντας στον κόσμο την ιδέα χρησιμοποίησης μίας πιστωτικής κάρτας σε πολλά διαφορετικά καταστήματα, όπως εστιατόρια και βενζινάδικα, με σκοπό να πληρώνονται οι λογαριασμοί των εστιατορίων χωρίς ο πελάτης να χρειάζεται να έχει μετρητά μαζί του. Επομένως το Φεβρουάριο του 1950 εφευρέθηκε η πρώτη πιστωτική κάρτα Diners με ιδρυτή της τον Frank McNamara. Στη συνέχεια η American Express έκδοσε την πρώτη της πιστωτική κάρτα το 1958 και η Τράπεζα της Αμερικής έκδοσε την Visa αργότερα.

Οι 5 πιο γνωστές πιστωτικές κάρτες είναι οι ακόλουθες:

- Visa International,
- MasterCard,
- American Express,
- Discover,
- Diners Club,

Ωστόσο υπάρχουν και άλλες επιχειρήσεις που δραστηριοποιούνται στο χώρο των καρτών οι οποίες προσπαθούν να διεισδύσουν στην αγορά. Δύο σημαντικές επιχειρήσεις καρτών, που συνεργάζονται με μεγάλο αριθμό τραπεζών μελών, έχουν έρθει να εξουσιάσουν το χώρο των καρτών παγκοσμίως. Αυτοί είναι η Visa International και η MasterCard, με τη Visa να έχει

ηγετική θέση στον τομέα της καινοτομίας πιστωτικών καρτών. Αυτό είχε ως αποτέλεσμα την αναγνώριση της ως κορυφαία ένωση πιστωτικών καρτών στον κόσμο, έχοντας εκδώσει πάνω από ένα δισεκατομμύριο κάρτες, και έχοντας απορροφήσει πάνω από 50% του συνόλου των συναλλαγών με πιστωτικές κάρτες που πραγματοποιούνται σε όλο τον κόσμο.

Όσον αφορά τώρα την **απάτη πιστωτικών καρτών**, πρόκειται για έναν ευρύ όρο ο οποίος αφορά την κλοπή και την απάτη που διαπράττεται μέσω πιστωτικών καρτών ή κάθε άλλου παρόμοιου μηχανισμού πληρωμών. Ο σκοπός της μπορεί να είναι η απόκτηση αγαθών χωρίς την καταβολή χρημάτων ή την απόκτηση παράνομων χρηματικών ποσών από ένα λογαριασμό. Οι απάτες με πιστωτικές κάρτες ποικίλουν ωστόσο οι πιο σημαντικές αφορούν συναλλαγές τις οποίες αναφέραμε και στην Ενότητα 1.4 και οι οποίες είναι οι ακόλουθες:

- Χαμένες ή κλεμμένες κάρτες: πρόκειται για κάρτες οι οποίες έχουν κλαπεί από τους κατόχους τους ή είτε έχουν χάσει οι οποίες αντιπροσωπεύουν το 23% όλων των περιπτώσεων απάτης καρτών. Συχνά, οι κάρτες αυτές έχουν κλαπεί από το χώρο εργασίας, γυμναστήριο, και αφύλακτα οχήματα.
- Πλαστές πιστωτικές κάρτες: υπερβαίνουν το 37% του συνόλου των χρημάτων που χάνονται μέσα από τις απάτες πιστωτικών καρτών. Για να φτιάξουν πλαστές κάρτες οι δράστες χρησιμοποιούν την τελευταία τεχνολογία για να "ξαφρίσουν" πληροφορίες που περιέχονται στις μαγνητικές λωρίδες των καρτών και να περάσουν τα χαρακτηριστικά ασφαλείας, όπως τα ολογράμματα.
- Μη χρήση κάρτας: η απάτη αυτή περιλαμβάνει το 10% του συνόλου των ζημιών. Πρόκειται για την υποκλοπή και την παράνομη χρήση του αριθμού της κάρτας σε συναλλαγές από απόσταση ή στο Internet. Αυτό μπορεί να συμβεί, δίνοντας τα στοιχεία της πιστωτικής κάρτας στο τηλέφωνο ή σε ιστοσελίδες στο Internet.
- Υποκλοπή προσωπικών στοιχείων πελατών και στοιχείων καρτών μέσα από αρχεία που διαθέτουν οι επιχειρήσεις. Αυτού του είδους η απάτη αποτελεί το 4% του συνόλου των ζημιών, και εμφανίζεται όταν εγκληματίες κάνουν αίτηση για μια κάρτα χρησιμοποιώντας τα στοιχεία της ταυτότητας ή πληροφορίες κάποιου άλλου.
- Συναλλαγές μέσω έκδοσης καρτών με χρήση ψευδών προσωπικών ή και οικονομικών στοιχείων. Πρόκειται για μία μορφή απάτης στην οποία ένας εγκληματίας χρησιμοποιεί κλεμμένα ή πλαστά έγγραφα για να ανοίξει ένα λογαριασμό στο όνομα κάποιου άλλου.
- ATM Skimming: αναφέρεται σε μια διαδικασία κατά την οποία μια ειδική συσκευή χρησιμοποιείται για την αντιγραφή των στοιχείων της μαγνητικής ταινίας της κάρτας.

Αυτή η συσκευή τοποθετείται συνήθως κρυφά σε ένα μηχάνημα ATM ως μία συσκευή ανάγνωσης καρτών.

Επομένως, σύμφωνα με τα παραπάνω η απάτη μέσω πιστωτικών καρτών μπορεί να διαιρεθεί σε δύο είδη: στις ***offline*** απάτες και στις ***online*** απάτες. Offline απάτες έχουμε στις περιπτώσεις που χρησιμοποιείται μια κλεμμένη πιστωτική κάρτα με φυσικό τρόπο, έχει να κάνει δηλαδή με οτιδήποτε εκτός από τη χρήση της κάρτας για πληρωμές μέσω διαδυκτίου. Είναι ο πιο ευθύς τρόπος απάτης, με τον απατεώνα να προσπαθεί να ξοδέψει όσο περισσότερα χρήματα μπορεί σε ένα σύντομο χρονικό διάστημα. Σε αυτές τις απάτες είναι κρίσιμη για τον πραγματικό ιδιοκτήτη η άμεση δήλωση της απώλειας ή κλοπής της κάρτας ώστε η τράπεζα να προχωρήσει στην ακύρωσή της. Στις offline απάτες συγκαταλέγονται οι απάτες με πιστωτικές κάρτες που εκδίδονται με πλαστά στοιχεία. Σε αυτές τις περιπτώσεις ο κάτοχος-απατεώνας δεν βιάζεται και η απάτη αποκαλύπτεται συνήθως αφού λήξουν αρκετοί απλήρωτοι λογαριασμοί.

Όμως ο σοβαρότερος τρόπος απάτης μέσω πιστωτικών καρτών αφορά στην κλοπή των λεπτομερειών της πιστωτικής κάρτας, όπως η αντιγραφή των στοιχείων της (κλωνοποίηση της κάρτας μέσω ειδικών μηχανημάτων) ή η υποκλοπή τους κατά τη διάρκεια τηλεφωνικών ή διαδικτυακών αγορών. Αυτές είναι οι online απάτες, στις οποίες ο απατεώνας έχει μεγαλύτερο χρονικό διάστημα για να δράσει οπότε και είναι οι απάτες με τις μεγαλύτερες απώλειες από την πλευρά των τραπεζών. Η αποκάλυψη της απάτης συμβαίνει μετά την έκδοση ενός λογαριασμού, αφού αυτός πρώτα φτάσει στον πελάτη και εκείνος αναφέρει τη διαφωνία του στην τράπεζα.

Στα τέλη του 2005, η MasterCard και η Visa έφτασε πωλήσεις άνω των $ 190,6 δισεκατομμυρίων δολαρίων, μέσα από $ 56,4 εκατομμύρια πιστωτικές κάρτες σε ολόκληρο τον Καναδά. Στατιστικές μελέτες έδειξαν ότι περίπου $ 2,8 εκατομμύρια δολάρια χάθηκαν εξαιτίας απάτες πιστωτικών καρτών, από δόλια χρήση της MasterCard και της Visa. Το συνολικό κόστος της απάτης των κατόχων και των εκδοτών πιστωτικών καρτών ανέρχεται γύρω στα $ 500 εκατομμύρια δολάρια το χρόνο. Το 2006 στο Ηνωμένο Βασίλειο υπήρξε μια αύξηση γύρω στο 25% σχετικά με την παράνομη χρήση τους, με απώλειες που έφταναν το ύψος των £ 535 εκατομμυρίων δολαρίων, σύμφωνα με τον τραπεζικό κλάδο. Ο **Οργανισμός Υπηρεσιών Εκκαθάρισης Πληρωμών** (*Association of Payment Clearing Services* (*Apacs*)) αναφέρει ότι η πρώτη αύξηση μέσα σε ένα διάστημα τριών ετών οφειλόταν κυρίως σε κλεμμένες και πλαστές πιστωτικές κάρτες οι οποίες χρησιμοποιούνταν στο εξωτερικό. Η απάτη καρτών στο εξωτερικό αυξήθηκε κατά 77% το 2005 σε £ 208 εκατομμύρια δολάρια, δηλαδή περίπου το 39% του συνόλου. Παρ'όλα αυτά οι απώλειες αυξήθηκαν καθώς οι κάρτες χρησιμοποιούνταν με αθέμιτο

τρόπο ώστε να αγοράσουν περισσότερα προϊόντα μέσω τηλεφώνου, Διαδικτύου ή μέσω ταχυδρομείου.

Όπως μπορούμε να δούμε από τα παραδείγματα τα οποία αναφέραμε παραπάνω, η απάτη πιστωτικών καρτών αποτελεί ένα σοβαρό και εντεινόμενο πρόβλημα. Ωστόσο, η ανίχνευση της πιστωτικής απάτης είναι εξαιρετικά δύσκολη ταυτόχρονα αποτελεί ένα δημοφιλές πρόβλημα ως προς την επίλυσή του. Για την ανακάλυψη της απάτης πιστωτικών καρτών, μπορούν να χρησιμοποιηθούν μέθοδοι που αναζητούν αλλαγές στα μοτίβα συναλλαγών καθώς και μέθοδοι που αναζητούν συγκεκριμένα μοτίβα αγορών τα οποία θεωρούνται ύποπτα. Ένας τρόπος είναι να μοντελοποιήσουμε ατομικά την προηγούμενη συμπεριφορά κάθε πελάτη ή ακόμη και να την συγκρίνουμε με μια αναμενόμενη συμπεριφορά. Επιπλέον ένας ακόμη τρόπος ανακάλυψης της απάτης πιστωτικών καρτών είναι με την αναζήτηση ορισμένων συμπεριφορών που σχετίζονται με τέτοιου είδους περιπτώσεις. Ενώ υπάρχουν μοντέλα πρόβλεψης για την ανίχνευση της πιστωτικής απάτης τα οποία είναι σε ενεργή χρήση στην πράξη, είναι σχετικά λίγες οι δημοσιευμένες μελέτες για την ανίχνευση της απάτης πιστωτικών καρτών με χρήση τεχνικών εξόρυξης δεδομένων, πιθανότατα λόγω έλλειψης των διαθέσιμων δεδομένων για την πραγματοποίηση σχετικής επιστημονικής έρευνας. Μεταξύ αυτών, οι περισσότερες εργασίες έχουν εξετάσει τη χρήση **νευρωνικών δικτύων** (*neural networks*), **μηχανών στήριξης διανυσμάτων** (*support vector machines*), **τυχαίων δασών** (*random forests*), **λογιστικής παλινδρόμησης** (*logistic regression*), της **μεθόδου κοντινότερου γείτονα** (*nearest neighbor*), και άλλες μεθόδους ανίχνευσης τις οποίες αναφέραμε στην Ενότητα 1.3. Άλλες τεχνικές που έχουν προταθεί για την ανίχνευση της απάτης πιστωτικών καρτών περιλαμβάνουν μεθόδους που βασίζονται στην **λογική** (*reasoning*), και στα **κρυμμένα μοντέλα Markov** (*hidden Markov models*).

Ο Joseph King-Fung Pun το 2011 πραγματοποίησε μία ανάλυση 11 μηνών σε δεδομένα που αφορούσαν συναλλαγές με πιστωτικές κάρτες μίας τράπεζας του Καναδά. Η έρευνα αυτή βασιζόταν στην ανίχνευση απάτης πιστωτικών καρτών. Οι μέθοδοι που χρησιμοποίησε στην ανάλυσή του ήταν: η λογιστική παλινδρόμηση, τα δέντρα ταξινόμησης, τα νευρωνικά δίκτυα και ο K-κοντινότερος γείτονας. Σκοπός της έρευνας αυτής ήταν η εξοικονόμηση χρημάτων που θα μπορούσε να επιτευχθεί από τον εντοπισμό παράνομων συναλλαγών. Το σύνολο των μεταβλητών που χρησιμοποίησε για την ανάλυσή του αποτελούταν από 29 μεταβλητές, ωστόσο λόγω περιορισμού του χώρου, θα αναφέρουμε ορισμένες από αυτές οι οποίες παρουσιάζονται στον Πίνακα 2.1.

| **Όνομα μεταβλητής** | **Μορφή μεταβλητής** | **Αριθμητική/Κατηγορική μεταβλητή** | **Περιγραφή της μεταβλητής** |
|---|---|---|---|
| Card Number | 16 ψήφιος αριθμός (1111111111111111) | Αριθμητική | Αριθμός πιστωτικής κάρτας που συνδέεται με τη συναλλαγή |
| Card Type | 5 ψήφιος αλφαριθμητικός χαρακτήρας (1A1A1) | Κατηγορική | Κατηγορία της πιστωτικής κάρτας (για παράδειγμα Rewards card, travel card, κλπ) |
| Fraud | 1 αλφαβητικό ψηφίο (A) | Κατηγορική | Ετικέτα απάτης για τη συναλλαγή με τιμές Ναι ή Οχι. |
| POS Entry | 1, 2, 3 ή 4 αριθμητικοί χαρακτήρες (1, 11, 111 ή 1111) | Αριθμητική | Περιγράφει τη χορήγηση αίτηση άδειας που τέθηκε στο Σημείο Πώλησης (POS) |
| Transaction Amount | 1 ως 8 αριθμητικοί χαρακτήρες (±1111.1111) | Αριθμητική | Το ποσό σε δολάρια που συνδέεται με κάθε συναλλαγή |
| User Country | 2 ως 5 αλφαβιτικοί χαρακτήρες (AA ή AAAAA) | Κατηγορική | Η χώρα που βρίσκεται ο κάτοχος της κάρτας |
| Times Difference (minutes) | 1 ως 8 αριθμητικοί χαρακτήρες (±1111.1111) | Αριθμητική | Ο χρόνος σε λεπτά μεταξύ της τρέχουσας συναλλαγής και της προηγούμενης συναλλαγής |

**Πίνακα 2.1** Πίνακας μεταβλητών παραδείγματος

Τα αποτελέσματα στα οποία κατέληξε η έρευνα έδειξαν ότι καλύτερος αλγόριθμος στην ανίχνευση απάτης αυναλλαγών με πιστωτικές κάρτες ήταν τα δέντρα ταξινόμησης. Ο αλγόριθμος αυτός οδηγούσε σε μία γρηγορότερη ανίχνευση της απάτης. Επιπλέον η απόδοση για την ανίχνευση βελτιώθηκε από 24% σε 34% με αποτέλεσμα την εξοικονόμης απο $1,8 σε $2,6 εκατομμύρια το χρόνο. Στη συνέχεια θα περιγράψουμε εκτενέστερα κάποιες από τις μεθόδους που χρησιμοποίησε στην ανάλυσή του. Επίσης θα αναφέρουμε και ορισμένα παραδείγματα για την εφαρμογή των μεθόδων αυτών σε ένα μικρό αριθμό του συνόλου δεδομένων απο το οποίο έγινε η ανάλυση, ώστε να γίνει κατανοητή η εύρεση της απάτης με τη χρήση των αλγορίθμων που αναφέραμε και παραπάνω.

## 2.2 Λογιστική Παλινδρόμηση

Η λογιστική παλινδρόμηση είναι μία τεχνική σχεδιασμένη για την πραγματοποίηση ανάλυσης δεδομένων που αφορούν την μελέτη και την πρόβλεψη τιμών κάποιας κατηγορικής εξαρτημένης μεταβλητής και κατηγορικών ή ποσοτικών ανεξάρτητων μεταβλητών.

Η μελέτη της σχέσης της κατηγορικής εξαρτημένης μεταβλητής δεν μπορεί να πραγματοποιηθεί μέσω του αλγορίθμου της γραμμικής παλινδρόμησης για δύο βασικούς λόγους:

- Πρώτον, όταν προβλέπουμε τις τιμές μία κατηγορικής εξαρτημένης μεταβλητής, στην ουσία υπολογίζουμε την πιθανότητα με την οποία η εξαρτημένη μεταβλητή θα λάβει κάποια συγκεκριμένη τιμή. Η τιμή της πιθανότητας αυτής θα πρέπει, εξ ορισμού, να παίρνει τιμές μεταξύ του 0 και του 1. Ωστόσο με τη χρήση της γραμμικής πολλαπλής παλινδρόμησης η εκτιμώμενη τιμή της μεταβλητής απόκρισης μπορεί να μην ανήκει στο διάστημα [0,1] όπως και θα έπρεπε και γι' αυτόν το λόγο δε μπορούμε να τη χρησιμοποιήσουμε.
- Δεύτερον, στη γραμμική παλινδρόμηση θα πρέπει να ικανοποιείται η υπόθεση της ισότητας των διακυμάνσεων (ομοσκεδαστικότητα). Ωστόσο, στην περίπτωση που η εξαρτημένη μεταβλητή είναι διχοτομική, έχει τυπική απόκλιση $\sqrt{p(1-p)}$, όπου $p$ είναι η μέση τιμή της μεταβλητής. Λόγω της συναρτησιακής σχέσης της τυπικής απόκλισης με την μέση τιμή, η ομοιογένεια της διακύμανσης των τιμών της εξαρτημένης μεταβλητής δεν είναι δυνατόν να ικανοποιείται.

Σε πολλές περιπτώσεις η μεταβλητή απόκρισης σε ένα μοντέλο που μας ενδιαφέρει είναι διακριτή. Στη μελέτη μας αυτή η εξαρτημένη μεταβλητή η «απάτη» είναι κατηγορηματική και έχει δύο τιμές τις οποίες μπορεί να πάρει, επομένως πρόκειται για δίτιμη μεταβλητή. Η λογιστική παλινδρόμηση είναι μία τεχνική η οποία χρησιμοποιείται ευρέως σε τέτοιου είδους προβλήματα.

Ας συμβολίσουμε με $Y_i$, για $\mathrm{i} = 1, \dots, n$ μία τυχαία μεταβλητή η οποία παίρνει τιμές $y_i$=0 ή 1. Στην τιμή 1 αντιστοιχούμε το ενδεχόμενο της "επιτυχίας", δηλαδή το να είναι μία συναλλαγή «δόλια», και στην τιμή 0 το ενδεχόμενο της "αποτυχίας". Η πιθανότητα επιτυχίας, δηλαδή η πιθανότητα κάποια συναλλαγή να χαρακτηριστεί ως «δόλια», θα συμβολίζεται με $p_i$.

Θεωρούμε λοιπόν ότι η τυχαία μεταβλητή $Y_i \sim Bernoulli\ (1, p_i)$ και η συνάρτηση πιθανότητας των $Y_i$ είναι:

$$P(Y_i = y_i) = p_i^{y_i}(1-p_i)^{1-y_i} \quad , \ y_i\text{=0 ή 1.}$$

Σκοπός μας είναι να εκτιμήσουμε την πιθανότητα επιτυχίας χρησιμοποιώντας ένα σύνολο ερμηνευτικών μεταβλητών. Η γενική μορφή του μοντέλου της λογιστικής παλινδρόμησης είναι:

$$ln(odds) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} \ ,$$

όπου $\beta_i$ είναι οι παράμετροι και $\boldsymbol{X} = (X_1, X_2, \dots, X_k)$ είναι οι ερμηνευτικές ή επεξηγηματικές τυχαίες μεταβλητές.

Το δεξί μέρος της εξίσωσης δημιουργείται από ένα γραμμικό συνδυασμό των ανεξάρτητων μεταβλητών που συμμετέχουν στο μοντέλο της παλινδρόμησης. Το αριστερό μέρος περιέχει τις

τιμές της εξαρτημένης μεταβλητής με την μορφή του λογαρίθμου των odds δηλαδή, του λογαρίθμου της ποσότητας:

$$odds = p_i/(1-p_i),$$

όπου $p_i$ είναι η πιθανότητα επιτυχίας, δηλαδή συμβολίζει την πιθανότητα κάποια συναλλαγή να χαρακτηριστεί ως «δόλια».

Η **σχετική πιθανότητα** (*odds ratio*) εκφράζει την πιθανότητα να συμβεί το γεγονός που έχει ορισθεί σαν επιτυχία του πειράματος. Η έννοια της σχετικής πιθανότητας είναι σημαντική για την ερμηνεία των παραμέτρων σε ένα μοντέλο λογιστικής παλινδρόμησης. Τιμή της σχετικής πιθανότητας μεγαλύτερη του 1 δηλώνει ότι το ενδεχόμενο στον αριθμητή είναι πιο πιθανό να συμβεί από αυτό στον παρονομαστή.

Θυμίζουμε επίσης ότι:

$$logit(odds) = \log\left(\frac{p_i}{1-p_i}\right),$$

Άρα η συνάρτηση $logit$ αναφέρεται στο λογάριθμο της σχετικής πιθανότητας του ενδεχομένου που μας ενδιαφέρει.

Οι συντελεστές των ανεξάρτητων μεταβλητών στην εξίσωση της παλινδρόμησης εκτιμούνται βάση της μεθόδου μέγιστης πιθανοφάνειας ή της μεθόδου ελαχίστων τετραγώνων.

Σε ένα μοντέλο λογιστικής παλινδρόμησης ένας αρχικός έλεγχος είναι η στατιστική σημαντικότητα του, εξετάζοντας την **απόκλισή** (*deviance*) του από το κορεσμένο μοντέλο[1]. Διαισθητικά, όσο πιο μικρή είναι η απόκλιση ενός μοντέλου, τόσο πιο κοντά είναι στο κορεσμένο μοντέλο, και αυτό παρέχει ένδειξη καλής προσαρμογής.

Επομένως :

- έστω ότι έχουμε δύο **εμφωλευμένα μοντέλα**[2] (*nested models*), $M_1$ και $M_2$
- έστω ότι οι συναρτήσεις πιθανοφάνειας των δύο μοντέλων είναι $L(M_1)$ και $L(M_2)$ αντίστοιχα
- έστω ότι το μοντέλο $M_1$ έχει απόκλιση $D_1$ και βαθμούς ελευθερίας $df_1$, ενώ το μοντέλο $M_2$ έχει απόκλιση $D_2$ και βαθμούς ελευθερίας $df_2$.

Αν θέσουμε:

$$l(M_1) = logL(M_1) \ , \ l(M_2) = logL(M_2)$$

τότε η στατιστική συνάρτηση για τον έλεγχο του λόγου πιθανοφανειών θα δίνεται από τη σχέση:

[1]Κορεσμένο ονομάζεται το μοντέλο το οποίο έχει τόσες παραμέτρους όσα και τα δεδομένα του.
[2]Εμφωλευμένα μοντέλα $M_1$ και $M_2$ ονομάζουμε τα μοντέλα εκείνα για τα οποία το σύνολο των επεξηγηματικών μεταβλητών του $M_1$ είναι υποσύνολο αυτών του $M_2$

$$D_{1,2} = -2log\frac{L(M_1)}{L(M_2)} = -2(l(M_1) - l(M_2)).$$

Η ποσότητα $D_{1,2}$ ακολουθεί την κατανομή $x_k^2$ , όπου ο ακέραιος *k* είναι η διαφορά των βαθμών ελευθερίας για τα κατάλοιπα ανάμεσα στα δύο μοντέλα, δηλαδή $k = df_1 - df_2$.

Πολλές φορές, σε ένα μοντέλο παλινδρόμησης, εκτός από τη στατιστική σημαντικότητα του μοντέλου μας ενδιαφέρει να εξετάσουμε και τη στατιστική σημαντικότητα κάθε ερμηνευτικής μεταβλητής ξεχωριστά. Σε αυτή την περίπτωση ο έλεγχος που χρησιμοποιούμε είναι ο εξής:

$$H_0: \beta_i = 0 \text{ έναντι } H_1: \beta_i \neq 0.$$

Μια εναλλακτική μέθοδος που μπορούμε να χρησιμοποιήσουμε για να ελέγξουμε τη σημαντικότητα μιας μεταβλητής σε ένα μοντέλο λογιστικής παλινδρόμησης είναι ο έλεγχος του Wald. Η στατιστική που χρησιμοποιείται για τον έλεγχο αυτό είναι η εξής:

$$z = \frac{\hat{\beta}}{s.e.(\widehat{\beta})},$$

όπου $\hat{\beta}$ είναι ο εκτιμητής της μέγιστης πιθανοφάνειας και το $s.e.(\widehat{\beta})$ είναι το τυπικό σφάλμα της εκτίμησης.

Συνήθως ο έλεγχος του Wald και ο έλεγχος με το λόγο πιθανοφανειών δίνουν το ίδιο αποτέλεσμα, ωστόσο για μικρά δείγματα μπορεί τα αποτελέσματα να διαφέρουν. Επομένως σύμφωνα με τα παραπάνω, εκτιμώντας τις παραμέτρους των τυχαίων μεταβλητών μπορούμε να κατασκευάσουμε ένα 100(1–α)% διάστημα εμπιστοσύνης για κάθε μία από αυτές, η γενική μορφή του οποίου είναι:

$$\hat{\beta} \pm z_{a/2}\, s.e.(\widehat{\beta}),$$

Σύμφωνα λοιπόν με όσα αναφέραμε, μπορούμε να κατατάξουμε το σύνολο των συναλλαγών με τη χρήση πιστωτικών καρτών σε «δόλιες» και «νόμιμες» ανάλογα με τις εκτιμήσεις των πιθανοτήτων που προκύπτουν από το μοντέλο λογιστικής παλινδρόμησης. Στη συνέχεια θα περιγράψουμε ένα παράδειγμα εφαρμογής της μεθόδου σε ένα μικρό σύνολο των δεδομένων που χρησιμοποίησε στην ανάλυσή του ο Joseph και θα παρουσιάσουμε συνοπτικά κάποια αποτελέσματα. Τα χαρακτηριστικά που περιλαμβάνει το σύνολο δεδομένων με το οποίο ασχοληθήκαμε είναι: το ποσό της συναλλαγής που πραγματοποιήθηκε, ο χρόνος της συναλλαγής και τον τύπο της κάθε συναλλαγής (δόλια ή νόμιμη).

| Instance | Transaction amount = x1 (thousands $) | Timestamp = x2 (minutes) | Fraud Classification |
|---|---|---|---|
| 1 | 0.35 | 0.9 | NO |
| 2 | 0.12 | 0.3 | YES |
| 3 | 0.47 | 0.6 | YES |

**Πίνακας 2.2.α.** Δεδομένα συναλλαγής πιστωτικών καρτών για λογιστική παλινδρόμηση

Τα δεδομένα του πίνακα θα χρησιμοποιηθούν για να κατασκευάσουμε ένα μοντέλο λογιστικής παλινδρόμησης. Η συνάρτηση του μοντέλου μας θα έχει την εξής μορφή:

$$Log\left(\frac{\widehat{p_i}}{1-\widehat{p_i}}\right) = b_0 + b_1X_1 + \cdots + b_iX_i$$

Και η πιθανότητα που μας ζητείται να εκτιμήσουμε είναι:

$$\widehat{p_i} = \frac{e^{b_0+b_1X_1+\cdots+b_iX_i}}{1+e^{b_0+b_1X_1+\cdots+b_iX_i}},$$

όπου $X_i$ για $i = 1,2$ είναι οι ανεξάρτητες μεταβλητές, $b_0$ είναι μία σταθερά, $b_i$ με $i = 1,2$ είναι οι παράμετροι των τυχαίων μεταβλητών και $p_i$ είναι η πιθανότητα επιτυχίας, δηλαδή η πιθανότητα να έχουμε «απάτη». Η μεταβλητή απόκρισής μας είναι η «Fraud» και παίρνει δύο τιμές. Την τιμή 1 την παίρνει όταν "Fraud=Yes", δηλαδή έχουμε πραγματοποίηση του γεγονότος, και την τιμή 0 την παίρνει όταν "Fraud=No".

Η μηδενική υπόθεση που εξετάζουμε στο πρόβλημά μας είναι η εξής:

**$H_0$:** $\hat{p}$ >0,5 η συναλλαγή είναι δόλια

**$H_1$:** δεν ισχύει η $H_0$

Χρησιμοποιώντας το στατιστικό πακέτο Spss βρέθηκαν οι εκτιμήσεις των παραμέτρων για το μοντέλο μας με τη βοήθεια του εκτιμητή μέγιστης πιθανοφάνειας και τα αποτελέσματα που μας έδωσε ήταν τα ακόλουθα:

- το $b_0 = 34.402$
- το $b_1 = 74.085$
- το $b_2 = -86.433$

Επομένως με βάση τα παραπάνω αποτελέσματα η εξίσωση του μοντέλου μας θα είναι:

$$Log\left(\frac{\widehat{p_i}}{1-\widehat{p_i}}\right) = 34.402 + 74.085X_1 - 86.433X_2$$

και η πιθανότητα που μας ζητείται να εκτιμήσουμε είναι:

$$\widehat{p_i} = \frac{e^{34.402+74.085X_1-86.433X_2}}{1+e^{34.402+74.085X_1-86.433X_2}},$$

Αυτή η εξίσωση μπορεί να χρησιμοποιηθεί για να προβλέψουμε αν μία μελλοντική περίπτωση θεωρηθεί ως «δόλια» ή μη. Για παράδειγμα, ας υποθέσουμε ότι έχουμε μια νέα συναλλαγή με πιστωτική κάρτα που θέλουμε να προβλέψουμε αν πρόκειται για «απάτη» ή μη. Έστω ότι το ποσό της συναλλαγής είναι 500$ και το timestamp είναι 1 λεπτό. Για να προσδιορίσουμε την πιθανότητα απάτης για τη συναλλαγή αυτή θα εισάγουμε τις νέες τιμές στο μοντέλο που βρήκαμε. Τα αποτελέσματα που θα έχουμε θα είναι τα ακόλουθα:

$$\hat{p} = \frac{e^{34.402+74.085\,(0.5)-86.433\,(1)}}{1+e^{34.402+74.085\,(0.5)-86.433\,(1)}} = 3{,}094 \text{ x } 10^{-7}$$

Σύμφωνα με το παραπάνω μοντέλο βλέπουμε ότι η νέα μας συναλλαγή θα είναι νόμιμη, εφόσον η εκτιμόμενη πιθανότητα είναι μικρότερη από 0.5 και είναι κοντά στο 0. Επομένως απορρίπτουμε τη μηδενική μας υπόθεση την οποία είχαμε ορίσει.

Επομένως έχοντας ένα δείγμα με τιμές $x_1$, $x_2$,...,$x_m$ τότε ως p-value του συγκεκριμένου δείγματος ορίζεται η πιθανότητα

$$p\text{-value} = \Pr(T(X_1, X_2, ..., X_m) > T(x_1, x_2, ..., x_m)/ H_0) = \Pr(T(X) > T(x)/ H_0),$$

η οποία μπορεί να θεωρηθεί ως η πιθανότητα να εμφανιστεί ένα τόσο ή ακόμη και πιο «ακραίο» δείγμα από αυτό που εμφανίστηκε, δεδομένου ότι ισχύει η $H_0$. Συνεπώς, αν το p-value ενός τ.δ. $X_1$, $X_2$, ..., $X_n$, βρεθεί «κοντά» στο 0, τότε μπορούμε να πούμε ότι η πιθανότητα να εμφανιστεί ένα «τέτοιο» δείγμα (ενώ ισχύει η $H_0$) είναι πολύ μικρή και σε αυτή την περίπτωση φυσιολογικά συμπεραίνουμε ότι δεν πρέπει να ισχύει η $H_0$. Εξάλλου αποδεικνύεται ότι αν

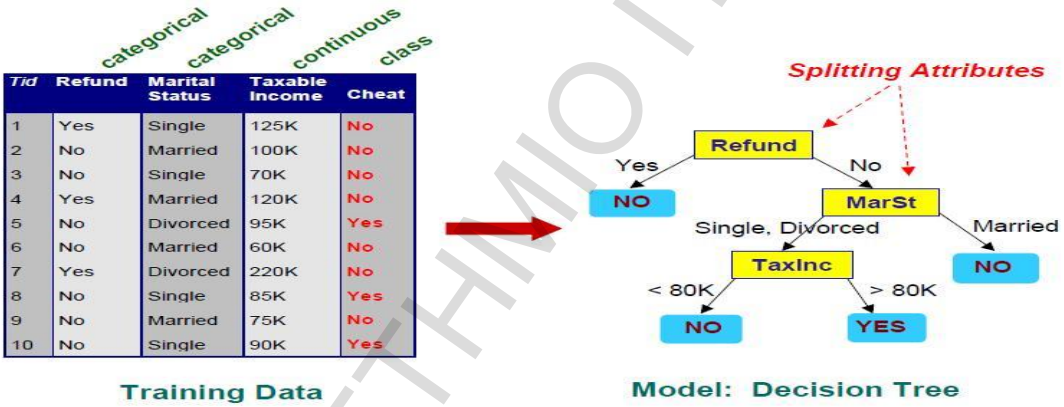
- το $p$-value < $a$ τότε απορρίπτουμε την $H_0$ σε ε.σ. α
- το $p$-value > $a$ τότε δεχόμαστε την $H_0$ σε ε.σ. α

## 2.3 Δέντρα Απόφασης

Τα δέντρα απόφασης (*Decision Trees*) αποτελούν μία από τις πιο διαδεδομένες μεθόδους αυτόματης εξαγωγής λογικών συμπερασμάτων. Η μορφή ενός δέντρου απόφασης φέρει τα χαρακτηριστικά μίας δενδρικής δομής, το οποίο όμως είναι κατασκευασμένο από πάνω προς τα κάτω. Ο πρώτος κόμβος ονομάζεται ρίζα και ακολουθούν οι κόμβοι ελέγχου (ή ενδιάμεσοι κόμβοι) και οι τερματικοί κόμβοι. Οι ενδιάμεσοι κόμβοι περιέχουν έναν έλεγχο και δημιουργούν δύο ή περισσότερους απογόνους. Οι τερματικοί κόμβοι είναι αυτοί που οδηγούν στην ταξινόμηση της εξεταζόμενης περίπτωσης σε μία από τις προεπιλεγμένες κλάσεις.

Τα δέντρα απόφασης συνιστούν μοντέλα τα οποία δίνουν τη δυνατότητα της ανάπτυξης συστημάτων κατηγοριοποίησης τα οποία μπορούν να προβλέπουν και να κατηγοριοποιούν

μελλοντικές καταστάσεις βασιζόμενα σε ένα σύνολο κανόνων απόφασης (*decision rules*). Η βασική ιδέα των δέντρων απόφασης είναι ο διαχωρισμός των δεδομένων αναδρομικά σε υποσύνολα ώστε κάθε ένα από αυτά να περιέχει ομοειδής καταστάσεις της μεταβλητής της οποίας την τιμή θέλουμε να προβλέψουμε. Σε κάθε σημείο στο οποίο το δέντρο διαχωρίζεται σε κλάδους, όλα τα χαρακτηριστικά εισόδου εκτιμώνται για να βρεθεί η επίδρασή τους στην μεταβλητή εξόδου, έτσι κάθε ένα μονοπάτι του δέντρου συνιστά και έναν κανόνα απόφασης. Το παρακάτω σχήμα δείχνει ένα παράδειγμα δέντρου απόφασης για να προβλέψει αν το άτομο είναι «απατεώνας». Στο δέντρο της απόφασης, η ρίζα και οι εσωτερικοί κόμβοι ελέγχου περιέχουν όρους που αποδίδουν σε χωριστές υπο-ομάδες που έχουν διαφορετικά χαρακτηριστικά. Η διαδικασία ταξινόμησης σταματά όταν στους τερματικούς κόμβους έχει επιτευχθεί η κατηγοριοποίηση. Συγκεκριμένα στο παράδειγμα της εικόνας που φαίνεται παρακάτω ο τερματικός κόμβος του δέντρου της εικόνας αποδίδεται σε μία κατηγορία (μεταβλητή εξόδου) που έχει την ένδειξη Ναι ή Όχι, δηλαδή τον χαρακτηρισμό του πελάτη σε «απατεώνα» ή μη.




Ένας ορισμός για το δέντρο απόφασης που χρησιμοποιείται στην κατηγοριοποίηση είναι ο ακόλουθος:

Έστω μία βάση δεδομένων $D=\{t_1,..., t_n\}$, όπου $t_i=t_{i1}, \ldots, t_{ih}$ και έστω ότι το σχήμα της βάσης δεδομένων περιέχει τα ακόλουθα γνωρίσματα: $\{A_1, A_2, ..., A_h\}$. Επίσης, έστω ότι έχουμε ένα σύνολο κατηγοριών $C=\{C_1, \ldots, C_m\}$. Ένα δέντρο απόφασης είναι μία δενδρική δομή που σχετίζεται με τη D και έχει τις ακόλουθες ιδιότητες:

- Κάθε εσωτερικός κόμβος έχει ως ετικέτα ένα γνώρισμα, $A_i$
- Κάθε τόξο έχει ως ετικέτα ένα κατηγόρημα που μπορεί να εφαρμοστεί στο γνώρισμα του κόμβου-γονέα
- Κάθε φύλλο (τερματικός κόμβος ) έχει ως όνομα μία κλάση, $C_j$

Ένα δέντρο απόφασης συνήθως κατασκευάζεται ακολουθώντας την εξής διαδικασία:

1. Στην αρχή ξεκινάμε με έναν κόμβο που περιέχει όλες τις εγγραφές. Ωστόσο θα πρέπει να κάνουμε μία αρχική επιλογή των γνωρισμάτων διάσπασης καθώς κάποια από τα γνωρίσματα της βάσης δεδομένων πρέπει να παραλειφθούν μιας και δεν εξυπηρετούν στην κατηγοριοποίηση.
2. Στη συνέχεια γίνεται διάσπαση του κόμβου, δηλαδή μοίρασμα των εγγραφών με βάση μια συνθήκη-διαχωρισμού σε κάποιο από τα γνωρίσματα. Ένα κριτήριο για την επιλογή του κατάλληλου γνωρίσματος διάσπασης είναι εκείνο το οποίο θα οδηγήσει στο μικρότερο δέντρο. Ένα άλλο κριτήριο βασίζεται στην επιλογή του γνωρίσματος που παράγει τους πιο ομοιογενείς (pure) κόμβους. Για το σκοπό αυτό χρησιμοποιείται μία συνάρτηση καταλληλότητας (fitness function). Χαρακτηριστικές συναρτήσεις καταλληλότητας είναι: α) Κέρδος πληροφορίας-Gain (ID3), β) Λόγος κέρδους πληροφορίας-Gain Ratio (C4.5), γ) Δείκτης Gini. (Περισσότερες πληροφορίες για τις συναρτήσεις αυτές αναφέρει στις σημειώσεις του ο Dunham M. (2004))
3. Έπειτα ακολουθούμε μία αναδρομική κλήση του βήματος 2 σε κάθε κόμβο.
4. Ωστόσο αφού κατασκευαστεί το δέντρο, θα κάνουμε κάποιες βελτιστοποιήσεις, δηλαδή θα γίνει το λεγόμενο **κλάδεμα του δέντρου** (*tree pruning*).

Για τα ίδια δεδομένα εκπαίδευσης θα μπορούσε να δημιουργηθεί ένα άλλο δέντρο απόφασης με διαφορετική κατηγοριοποίηση. Τα διαφορετικά δένδρα προκύπτουν από τις επιλογές των χαρακτηριστικών που θα χρησιμοποιηθούν ως ρίζα και ως κόμβοι–γονείς. Αυτό εξαρτάται από τον πιο αλγόριθμο θα χρησιμοποιήσουμε για το πρόβλημά μας. Οι αλγόριθμοι κατασκευής δένδρων απόφασης διαφέρουν στο πώς καθορίζουν το «καλύτερο» χαρακτηριστικό και τα αντίστοιχα «καλύτερα» κατηγορήματα. Μόλις καθοριστεί αυτό, ο κόμβος με τα τόξα (κλαδιά) του τοποθετούνται στο δένδρο. Στη συνέχεια ο αλγόριθμος συνεχίζει αναδρομικά, κάνοντας την ίδια διαδικασία για τους κόμβους των υποδένδρων έως ότου ο αλγόριθμος να φτάσει στον κόμβο τερματισμού.

Παρακάτω θα περιγράψουμε ένα πρόβλημα ανίχνευσης απάτης πιστωτικών καρτών με τη βοήθεια δέντρων απόφασης. Στο παράδειγμά μας ο αλγόριθμος που θα χρησιμοποιηθεί για την κατηγοριοποίηση των δεδομένων μας είναι ο ID3. Η βασική ιδέα του αλγορίθμου ID3 είναι η επιλογή χαρακτηριστικών διάσπασης που περιέχουν το μεγαλύτερο κέρδος πληροφορίας. Η έννοια που χρησιμοποιείται για να μετρηθεί η πληροφορία καλείται **εντροπία** (*Entropy*). Χρησιμοποιούμε το μέτρο της εντροπίας ώστε να μετρήσουμε το πόσο ανομοιογενές είναι ένα σύνολο δεδομένων. Το μέτρο αυτό παίρνει τιμές στο διάστημα [0,1] και ορίζεται ως εξής:

$$H(p_1, p_2, \dots, p_s) = \sum_{i=1}^{s} (p_i \log(\frac{1}{p_i}))\ ,$$

όπου $p_1, p_2, \dots, p_s$ είναι πιθανότητες των οποίων το άθροισμα ισούται με 1.

Στη συνέχεια ο ID3 θα επιλέξει το χαρακτηριστικό διάσπασης που θα έχει το μεγαλύτερο **κέρδος** (*Gain*) πληροφορίας[3]. Καταλήγοντας, ο ID3 αλγόριθμος υπολογίζει το κέρδος μιας διάσπασης χρησιμοποιώντας των εξής τύπο:

$$Gain(D,S) = H(D) - \sum_{i=1}^{s} P(D_i)H(D_i)\ ,$$

όπου H($D$) η εντροπία του $D$ πριν το διαχωρισμό και $H(D_i)$ η εντροπία του $D_i$ μετά το διαχωρισμό. Όσο μεγαλύτερη είναι η μείωση (το «άλμα» προς το 0), τόσο μεγαλύτερο είναι το κέρδος Gain($D$,$S$).

Στη συνέχεια στους πίνακες που ακολουθούν παρουσιάζονται ένα μικρό δείγμα του συνόλου των δεδομένων πιστωτικών καρτών τα οποίο προέρχενται από μία τράπεζα στον Καναδά (Πίνακας 2.3.α). Τα χαρακτηριστικά που χρησιμοποιούνται στο παράδειγμά μας είναι: η Χώρα στην οποία έγινε η συναλλαγή (*Country*), το είδος της κάρτας που χρησιμοποιήθηκε (*Card Type*), *POS Entry* και ο κωδικός ασφαλείας της πιστωτικής κάρτας (*Security Code*). Επιπλέον στον Πίνακα 2.3.β παρουσιάζονται οι μετρήσεις που έγιναν για κάθε χαρακτηριστικό ξεχωριστά. Για παράδειγμα για το χαρακτηριστικό *Country* έχουμε 3 μεταβλητές (*Canada, USA, Mexico*) όπου για την κάθε μία έχουμε μετρήσει πόσες φορές εμφανίστηκε απάτη ή όχι σε κάποια συναλλαγή που έγινε με πιστωτική κάρτα. Επομένως όπως φαίνεται στον πίνακα των μετρήσεων, για τη χώρα Καναδά είχαμε 2 φορές εμφάνιση απάτης συναλλαγής με πιστωτική κάρτα και 3 φορές εμφάνισης νόμιμης συναλλαγής. Με τον ίδιο τρόπο επεξηγούνται και οι υπόλοιπες μετρήσεις στον πίνακα 2.3.β και για τα υπόλοιπα χαρακτηριστικά του παραδείγματός μας. Στο τέλος του Πίνακα 2.3.β παρουσιάζεται και το συνολικό άθροισμα των περιπτώσεων εμφάνισης νόμιμης ή παράνομης χρήσης πιστωτικών καρτών για το σύνολο των χαρακτηριστικών. Με βάση αυτά τα δεδομένα θα προσπαθήσουμε να κατασκευάσουμε ένα δέντρο απόφασης για την ανίχνευση απάτης σε μία συναλλαγή.

[3]Το κέρδος πληροφορίας μετρά τη μείωση της εντροπίας που θα προκληθεί αν χωριστεί το σύνολο των δεδομένων με βάση κάποιο χαρακτηριστικό.

| Instance | Country | Card Type | POS Entry | Security Code | Legit or Fraud |
|---|---|---|---|---|---|
| 1 | Canada | Gold | Swiped | False | Legit |
| 2 | Canada | Gold | Swiped | True | Legit |
| 3 | Mexico | Gold | Swiped | False | Fraud |
| 4 | USA | Classic | Swiped | False | Fraud |
| 5 | USA | Platinum | Keyed | False | Fraud |
| 6 | USA | Platinum | Keyed | True | Legit |
| 7 | Mexico | Platinum | Keyed | True | Fraud |
| 8 | Canada | Classic | Swiped | False | Legit |
| 9 | Canada | Platinum | Keyed | False | Fraud |
| 10 | USA | Classic | Keyed | False | Fraud |
| 11 | Canada | Classic | Keyed | True | Fraud |
| 12 | Mexico | Classic | Swiped | True | Fraud |
| 13 | Mexico | Gold | Keyed | False | Fraud |
| 14 | USA | Classic | Swiped | True | Legit |

**Πίνακας 2.3.α.** Πίνακας συνόλου δεδομένων πιστωτικών καρτών με τέσσερα χαρακτηριστικά

| Country | | | Card Type | | | POS Entry | | | Security Code | | | Fraud | Legit |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Fraud | Legit | | Fraud | Legit | | Fraud | Legit | | Fraud | Legit | 9 | 5 |
| **Canada** | 2 | 3 | **Gold** | 2 | 2 | **Swiped** | 3 | 4 | **True** | 3 | 3 | | |
| **USA** | 3 | 2 | **Classic** | 4 | 2 | **Keyed** | 6 | 1 | **False** | 6 | 2 | | |
| **Mexico** | 4 | 0 | **Platinum** | 3 | 1 | | | | | | | | |

**Πίνακας 2.3.β.** Μετρήσεις για το σύνολο δεδομένων πιστωτικών καρτών

Καταρχάς για την κατασκευή του δέντρου μας θα πρέπει να καθορίσουμε τη ρίζα του δέντρου αποφάσεων. Προκειμένου να βρούμε τη ρίζα του δέντρου θα πρέπει να βρούμε ποιο γνώρισμα παράγει τους πιο ομοιογενείς κόμβους. Αρχικά σύμφωνα με τα δεδομένα του Πίνακα 2.3.α υπολογίζουμε την εντροπία σχετικά με το αν μία συναλλαγή είναι δόλια ή νόμιμη. Η εντροπία αυτή υπολογίζεται με βάση τον τύπο που αναφέραμε και παραπάνω και θα είναι ως εξής:

Εντροπία (*Fraud*)=Εντροπία(*9,5*)

$= -p \log_2 p - q \log_2 q$

$= -\left[\left(\frac{9}{14}\right)\log_2\left(\frac{9}{14}\right)\right] - \left[\left(\frac{5}{14}\right)\log_2\left(\frac{5}{14}\right)\right]$

$= -(-0.4098) - (0.5305)$

$= 0.94$

Στη συνέχεια υπολογίζουμε την εντροπία της απάτης για κάθε ένα από τα τέσσερα χαρακτηριστικά. Παραδείγματος χάριν θα περιγράψουμε πιο αναλυτικά πως υπολογίζουμε την

εντροπία για το χαρακτηριστικό *Country* το οποίο περιλαμβάνει 3 μεταβλητές. Ο υπολογισμός της εντροπίας για το χαρακτηριστικό αυτό είναι το άθροισμα της κάθε μίας για την κάθε μεταβλητή που περιέχεται. Τα αποτελέσματα που θα έχουμε βασίζονται στον Πίνακα 2.3.γ ο οποίος περιλαμβάνει τις μετρήσεις του συνόλου δεδομένων που έχουμε για την εμφάνιση νόμιμης ή παράνομης συναλλαγής μέσω πιστωτικών καρτών για τις 3 χώρες που έχουμε. Με τον ίδιο τρόπο υπολογίζονται και για τα υπόλοιπα χαρακτηριστικά. Τα αποτελέσματα που θα πάρουμε είναι τα ακόλουθα:

| **Country** | | | |
|---|---|---|---|
| | **Fraud** | **Legit** | **Total** |
| **Canada** | 2 | 3 | 5 |
| **USA** | 3 | 2 | 5 |
| **Mexico** | 4 | 0 | 4 |
| | | **Total** | 14 |

**Πίνακας 2.3.γ.** Μέτρηση δεδομένων για το χαρακτηριστικό "Country"

$$\text{Εντροπία}(\textit{Fraud, Country}) = \left[\left(\frac{5}{14}\right) * \text{Εντροπία}(2,3)\right] + \left[\left(\frac{5}{14}\right) * \text{Εντροπία}(3,2)\right] + \left[\left(\frac{4}{14}\right) * \text{Εντροπία}(4,0)\right] = \left\{\left(\frac{5}{14}\right) * \left[\left(\frac{2}{5}\right)\log_2\left(\frac{2}{5}\right)) - (\frac{3}{5})\log_2\left(\frac{3}{5}\right)\right]\right\} + \left\{\left(\frac{5}{14}\right) * \textit{35}\log\textit{235)}-\textit{(25)}\log\textit{225}+\textit{414}*\textit{44}\log\textit{244)}-\textit{(04)}\log\textit{204}=\textit{514}*\textit{0.97}+\textit{514}*\textit{0.97}+\textit{0} = 0.6929$$

| **Card Type** | | | |
|---|---|---|---|
| | **Fraud** | **Legit** | **Total** |
| **Gold** | 2 | 2 | 4 |
| **Classic** | 4 | 2 | 6 |
| **Platinum** | 3 | 1 | 4 |
| | | **Total** | 14 |

**Πίνακας 2.3.δ.** Μέτρηση δεδομένων για το χαρακτηριστικό "Card Type"

$$\text{Εντροπία}(\textit{Fraud, Card Type}) = \left[\left(\frac{4}{14}\right) * \text{Εντροπία}(2,2)\right] + \left[\left(\frac{6}{14}\right) * \text{Εντροπία}(4,2)\right] + \left[\left(\frac{4}{14}\right) * \text{Εντροπία}(3,1)\right] = \left[\left(\frac{4}{14}\right) * 1\right] + [\left(\frac{6}{14}\right) * 0.91 + [\left(\frac{4}{14}\right) * 0.81] = 0.9071$$

| **POS Entry** | | | |
|---|---|---|---|
| | **Fraud** | **Legit** | **Total** |
| **Swiped** | 3 | 4 | 7 |
| **keyed** | 6 | 1 | 7 |
| | | **Total** | 14 |

**Πίνακας 2.3.ε.** Μέτρηση δεδομένων για το χαρακτηριστικό "POS Entry"

$$\text{Εντροπία}(\textit{Fraud, POS Entry}) = \left[\left(\frac{7}{14}\right) * \text{Εντροπία}(3,4)\right] + \left[\left(\frac{7}{14}\right) * \text{Εντροπία}(6,1)\right] = \left[\left(\frac{7}{14}\right) * 0.98\right] + [\left(\frac{7}{14}\right) * 0.59] = 0.785$$

| **Security Code** | | | |
|---|---|---|---|
| | **Fraud** | **Legit** | **Total** |
| **True** | 3 | 3 | 6 |
| **False** | 6 | 2 | 8 |
| | | **Total** | 14 |

**Πίνακας 2.3.ζ.** Μέτρηση δεδομένων για το χαρακτηριστικό "Security Code"

$$\text{Εντροπία}(\textit{Fraud, Security Code}) = \left[\left(\frac{6}{14}\right) * \text{Εντροπία}(3,3)\right] + \left[\left(\frac{8}{14}\right) * \text{Εντροπία}(6,2)\right] = \left[\left(\frac{6}{14}\right) * 1\right] + \left[\left(\frac{8}{14}\right) * 0.81 = 0.8914\right.$$

Για να επιλέξουμε τώρα ποιο από τα 4 χαρακτηριστικά αυτά (*Country, Card Type, POS Entry, Security Code*) θα αποτελέσει τη ρίζα του δέντρου μας θα πρέπει να βρούμε το γνώρισμα εκείνο το οποίο έχει το μεγαλύτερο κέρδος πληροφορίας. Το κέρδος για κάθε ένα από τα 4 χαρακτηριστικά υπολογίζεται με βάση τον τύπο για το Gain που αναφέραμε και τα αποτελέσματα που θα έχουμε είναι:

- Gain (*Fraud,Country*) = 0.94 – 0.6929 = 0.2471,
- Gain (*Fraud,Card Type*) = 0.94 – 0.9071 = 0.0329,
- Gain (*Fraud,POS Entry*) = 0.94 – 0.785 = 0.155,
- Gain (*Fraud,Security Code*) = 0.94 – 0.8914 = 0.0486.

Επομένως σύμφωνα με τα παραπάνω αποτελέσματα σχετικά με το κέρδος πληροφορίας κάθε γνωρίσματος, βλέπουμε ότι ως ρίζα του δέντρου μας θα χρησιμοποιήσουμε το γνώρισμα "Country" καθώς έχει το μεγαλύτερο κέρδος σε σχέση με τα υπόλοιπα γνωρίσματα. Στη συνέχεια θα υπολογίσουμε τις εντροπίες για τη χώρα "Canada" σε σχέση με κάθε ένα από τα υπόλοιπα γνωρίσματά μας. Την ίδια διαδικασία θα ακολουθήσουμε και για τις άλλες δύο χώρες ώστε να δούμε ποιο γνώρισμα θα χρησιμοποιηθεί στη συνέχεια για την διαμόρφωση του δέντρου μας.

| **Fraud (Country=Canada)** | | | |
|---|---|---|---|
| **Card Type** | **Fraud** | **Legit** | |
| **Gold** | 0 | 2 | 2 |
| **Classic** | 1 | 1 | 2 |
| **Platinum** | 0 | 1 | 1 |
| | | **Total** | 4 |

**Πίνακας 2.3.η.** Μέτρηση δεδομένων για τα χαρακτηριστικά "Country=Canada" και "Card Type"

$$\text{Εντροπία}(\textit{Country=Canada}) = \text{Εντροπία}(2,3) = 0{,}97$$

Εντροπία(*Country=Canada, Card Type*) = $\left[\left(\frac{2}{5}\right) * \text{Εντροπία}(0,2)\right] + \left[\left(\frac{2}{5}\right) *$ Εντροπία$1,1+25*$Εντροπία$0,1=0+25*1+0=0.4$

| **Fraud (Country=Canada)** | | | |
|---|---|---|---|
| **POS Entry** | **Fraud** | **Legit** | |
| **Swiped** | 0 | 3 | 3 |
| **Keyed** | 2 | 0 | 2 |
| | | **Total** | 5 |

**Πίνακας 2.3.θ.** Μέτρηση δεδομένων για τα χαρακτηριστικά "Country=Canada" και "POS Entry"

Εντροπία(*Country=Canada, POS Entry*) = $\left[\left(\frac{3}{5}\right) * \text{Εντροπία}(0,3)\right] + \left[\left(\frac{2}{5}\right) *$ Εντροπία$2,0=0$

| **Fraud (Country=Canada)** | | | |
|---|---|---|---|
| **Security Code** | **Fraud** | **Legit** | |
| **False** | 1 | 2 | 3 |
| **True** | 1 | 1 | 2 |
| | | **Total** | 5 |

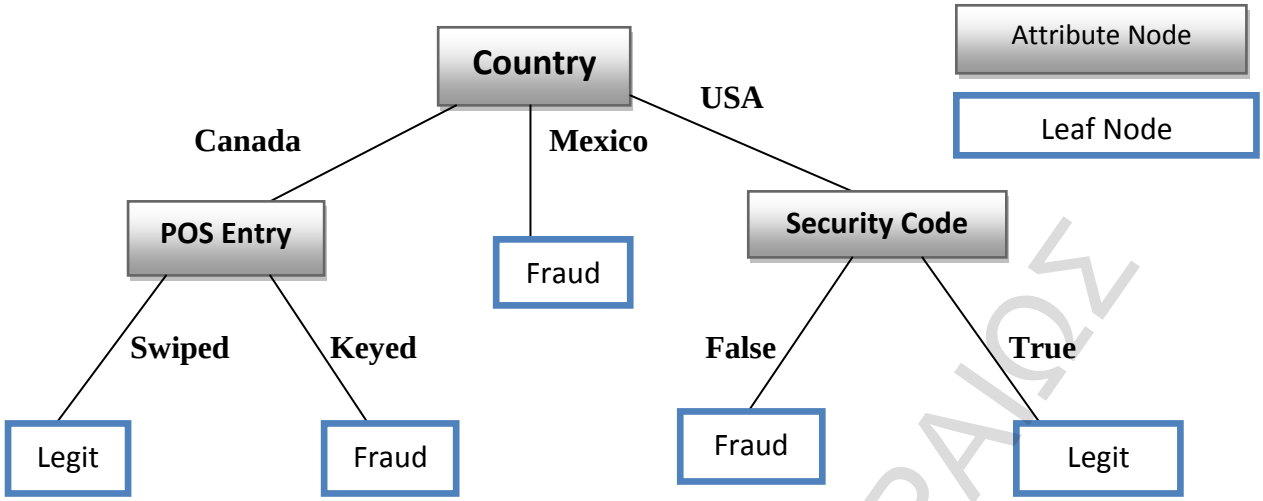
**Πίνακας 2.3.ι.** Μέτρηση δεδομένων για τα χαρακτηριστικά "Country=Canada" και "Security Code"

Εντροπία(*Country=Canada, Security Code*) = $\left[\left(\frac{3}{5}\right) * \text{Εντροπία}(1,2)\right] + \left[\left(\frac{2}{5}\right) * \text{Εντροπία}(1,1)\right] = \left[\left(\frac{3}{5}\right) * 0.91\right] + \left[\left(\frac{2}{5}\right) * 1\right] = 0.946$

Υπολογίζοντας πάλι το κέρδος πληροφορίας για τη χώρα "Canada" σε σχέση με τα υπόλοιπα γνωρίσματα θα πάρουμε τα ακόλουθα αποτελέσματα:

- Gain(*Country=Canada, Card Type*) = 0.97 – 0.4 = 0.57
- Gain(*Country=Canada, POS Entry*) = 0.97 – 0 = 0.97
- Gain(*Country=Canada, Security Code*) = 0.97 – 0.946 = 0.024

Από τα αποτελέσματα αυτά βλέπουμε ότι μεγαλύτερο κέρδος πληροφορίας έχει το γνώρισμα "POS Entry", επομένως ο κόμβος που ενώνεται στη συνέχεια με τη χώρα "Canada" είναι το γνώρισμα "POS Entry". Την ίδια διαδικασία ακολουθούμε και για τις άλλες δύο χώρες και το τελικό δέντρο απόφασης που θα πάρουμε είναι το επόμενο:



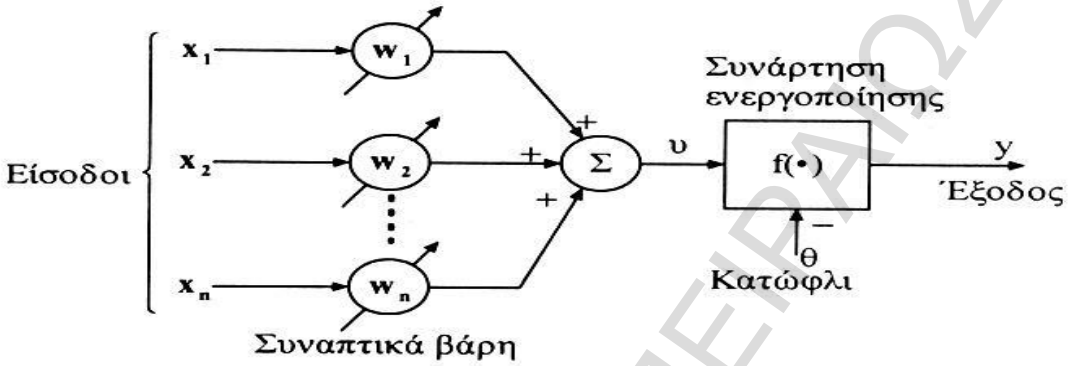

**Εικόνα 2.3.α.** Δέντρο απόφασης για τα δεδομένα του Πίνακα 2.3.α

Επομένως σύμφωνα με τον αλγόριθμο που χρησιμοποιήσαμε για την κατασκευή του δέντρου απόφασης επέλεξε ως σημαντική μεταβλητή το χαρακτηριστικό ***Country***. Με βάση το χαρακτηριστικό αυτό παρατηρούμε ότι αν κάποιο άτομο προέρχεται από την χώρα *Mexico*, το δέντρο τότε κατευθείαν καταλήγει σε κάποιο τερματικό κόμβο. Με λίγα λόγια μπορούμε να πούμε ότι οποιαδήποτε συναλλαγή μέσω πιστωτικής κάρτας γίνει στη χώρα αυτή τότε είναι απάτη. Όσον αφορά τώρα τα άτομα εκείνα που προέρχονται από την χώρα *Canada* και έχουν ***POS Entry Keyed*** τότε η συναλλαγή πρόκειται για απάτη. Αντίθετα αν είναι από τον Καναδά και έχει ***POS Entry*** *Swiped* η συναλλαγή είναι νόμιμη. Με τον ίδιο τρόπο ερμηνεύται και το τρίτο χαρακτηριστικό που βρέθηκε σημαντικό για την κατασκευή του δέντρου απόφασης. Επομένως όταν θα έρθει ένα νέο άτομο στην τράπεζα αυτή και δώσει τα στοιχεία του με βάση αυτά θα μπορέσουν να τον κατατάξουν σε ποια ομάδα θα ανήκει.

## 2.4 Νευρωνικά Δίκτυα

Το νευρωνικό δίκτυο είναι ένα δίκτυο από απλούς υπολογιστικούς **κόμβους ή νευρώνες** (*Processing Units-PUs*), που συνδέονται μεταξύ τους. Οι νευρώνες είναι τα δομικά στοιχεία του δικτύου, και σε αυτούς συντελείται όλη η επεξεργασία της πληροφορίας. Κάθε κόμβος λαμβάνει ένα σήμα εισόδου που είναι οι συνολικές πληροφορίες από άλλους κόμβους ή εξωτερικά ερεθίσματα, το επεξεργάζεται τοπικά μέσω μιας **συνάρτησης ενεργοποίησης** (*activation function*) και παράγει μία τιμή εξόδου. Η εν λόγω έξοδος είτε κατευθύνεται στο περιβάλλον, είτε τροφοδοτείται ως είσοδος σε άλλους νευρώνες του δικτύου. Οι νευρώνες είναι συνδεδεμένοι μεταξύ τους με τις λεγόμενες **συνάψεις** (*synapses*). Ο βαθμός αλληλεπίδρασης είναι διαφορετικός για κάθε ζεύγος νευρώνων και καθορίζεται από τα λεγόμενα **συναπτικά βάρη** (*synaptic weights*).

Υπάρχουν διάφοροι τύποι νευρώνων. Το είδος νευρώνα που θα επιλεχθεί για να δομηθεί ένα συγκεκριμένο τεχνητό νευρωνικό δίκτυο, εξαρτάται από τη φύση του εκάστοτε προβλήματος που εξετάζουμε. Σε πολλές περιπτώσεις χρησιμοποιείται συνδυασμός διαφορετικών ειδών νευρώνα. Στο σχήμα που ακολουθεί βλέπουμε το βασικό μοντέλο ενός νευρώνα.




**Εικόνα 1.4.α** *Σχηματική αναπαράσταση ενός μη-γραμμικού νευρώνα*

Με βάση την Εικόνας 2.4.α μπορούμε να διακρίνουμε τρεις φάσεις λειτουργίας του μοντέλου αυτού οι οποίες περιγράφονται παρακάτω:

Στην πρώτη φάση κάθε είσοδος πολλαπλασιάζεται με το αντίστοιχο συναπτικό βάρος που της αντιστοιχεί. Στη συνέχεια οι σταθμισμένες πλέον είσοδοι υπολογίζουν το ολικό άθροισμα των γινομένων. Τέλος, στην Τρίτη φάση, εφαρμόζεται η συνάρτηση ενεργοποίησης το αποτέλεσμα της οποίας δίνει την έξοδο του νευρώνα.

Ας συμβολίσουμε με $x_{ki}$ την $i$-οστή είσοδο του $k$ νευρώνα, με $w_{ki}$ το $i$-οστό συναπτικό βάρος του $k$ νευρώνα και με $\Phi(\cdot)$ τη συνάρτηση ενεργοποίησης του νευρωνικού δικτύου. Τότε η έξοδος $y_k$ του $k$ νευρώνα δίνεται από τον τύπο:

$$y_k = \Phi(\sum_{i=0}^{N} x_{ki}\, w_{ki})$$

Στον $k$-οστό νευρώνα υπάρχει ένα συναπτικό βάρος $w_{k0}$ με ιδιαίτερη σημασία, το οποίο καλείται **πόλωση** ή **κατώφλι** (*bias, threshold*). Η τιμή της εισόδου του είναι πάντα η μονάδα, $x_{k0} = 1$. Εάν το συνολικό άθροισμα από τις υπόλοιπες εισόδους του νευρώνα είναι μεγαλύτερο από την τιμή αυτή, τότε ο νευρώνας ενεργοποιείται. Εάν είναι μικρότερο, τότε ο νευρώνας παραμένει ανενεργός.

Έχουμε 3 βασικούς τύπους ενεργοποίησης:

**a.** ***Βηματική*** (*step transfer function*)

Η βηματική συνάρτηση ενεργοποίησης έχει τη μορφή:

$$\Phi(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

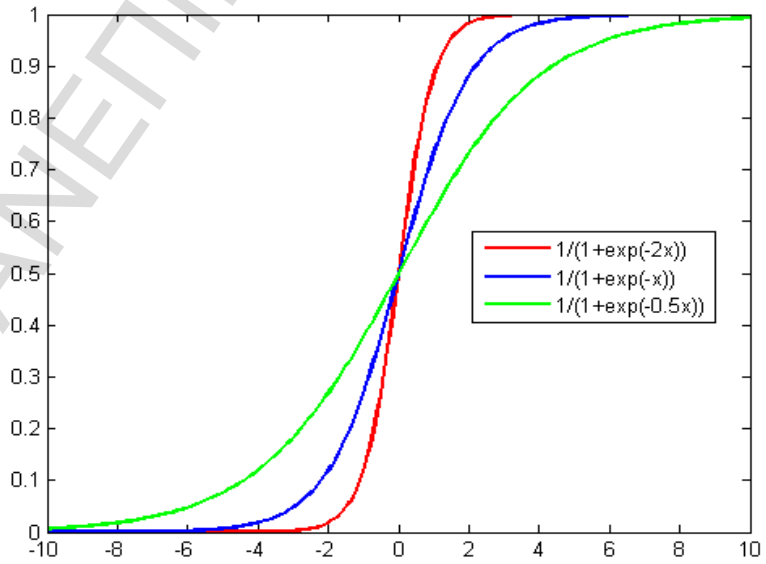
Η βηματική συνάρτηση δεν θεωρείται χρήσιμη ως συνάρτηση ενεργοποίησης στα τεχνητά νευρωνικά δίκτυα, γιατί είναι προτιμότερη η χρησιμοποίηση μιας παραγωγίσιμης συνάρτησης, ενώ η βηματική συνάρτηση έχει παράγωγο ίση με το μηδέν και δεν παραγωγίζεται στο $x$=0. Γι'αυτό το λόγο προέκυψε η ανάγκη συναρτήσεων ενεργοποίησης οι οποίες είναι συνεχείς και παραγωγίσιμες σε όλο το πεδίο ορισμού τους.

**b.** ***Μη γραμμική*** (*non-linear transfer function*)

Η μη γραμμική συνάρτηση ενεργοποίησης που χρησιμοποιείται συνήθως στα νευρωνικά δίκτυα καλείται σιγμοειδής συνάρτηση. Ένα από τα πιο γνωστά παραδείγματα σιγμοειδούς συνάρτησης είναι η **λογιστική συνάρτηση** (*logistic function*), η οποία δίνεται από τον ακόλουθο τύπο:

$$\Phi(x) = \frac{1}{1+e^{-ax}}, \; -\infty<x<\infty$$

Όπου $a$ είναι η παράμετρος κλίσης. Μεταβάλλοντας την παράμετρο $a$ παίρνουμε διαφορετικές συναρτήσεις με διάφορες κλίσεις. Στο σχήμα που ακολουθεί βλέπουμε την γραφική παράσταση της λογιστικής συνάρτησης για διάφορες τιμές του $a$, όπως για την τιμή $a$=0.5 (πράσινη γραμμή), $a$=1 (μπλε γραμμή) και α=2 (κόκκινη γραμμή).




**Εικόνα 2.4.β** *Γραφική παράσταση της λογιστικής συνάρτησης για α=0.5, α=1 και α=2*

Μια άλλη μορφή σιγμοειδούς συνάρτησης είναι η **συνάρτηση τόξου εφαπτομένης** (*arctangent function*), δηλαδή:

$$\Phi(x) = tanh\, x$$

**c.** ***Γραμμική*** (*linear transfer function*)

Με την εισαγωγή μιας μη γραμμικής συνάρτησης ενεργοποίησης (όπως π.χ. η λογιστική ή το τόξο εφαπτομένης), ο νευρώνας γίνεται μη γραμμικός. Αντίστοιχα, ένα τεχνητό νευρωνικό δίκτυο που αποτελείται από τέτοιους νευρώνες θα είναι μη γραμμικό. Αυτή η μη γραμμικότητα των νευρωνικών δικτύων αποτελεί ένα σημαντικό πλεονέκτημα έναντι άλλων γνωστών μεθόδων αντιμετώπισης πολλών προβλημάτων. Για παράδειγμα, όταν σε ένα πρόβλημα πρόβλεψης το σύστημα που μελετάμε είναι μη γραμμικό και ιδιαίτερα όταν παρουσιάζει μια χαοτική συμπεριφορά, τότε τα γνωστά γραμμικά μοντέλα πρόβλεψης αδυνατούν να δώσουν σωστά αποτελέσματα. Στις περιπτώσεις αυτές, τα μη γραμμικά τεχνητά νευρωνικά δίκτυα είναι προτιμότερα.

Ωστόσο επειδή πολλές φορές στην πράξη χρειάζεται κάποιοι από τους νευρώνες ενός μη γραμμικού νευρωνικού δικτύου να είναι γραμμικοί, μία ακόμη συνάρτηση ενεργοποίησης που χρησιμοποιείται είναι η γραμμική, της οποίας ο τύπος είναι:

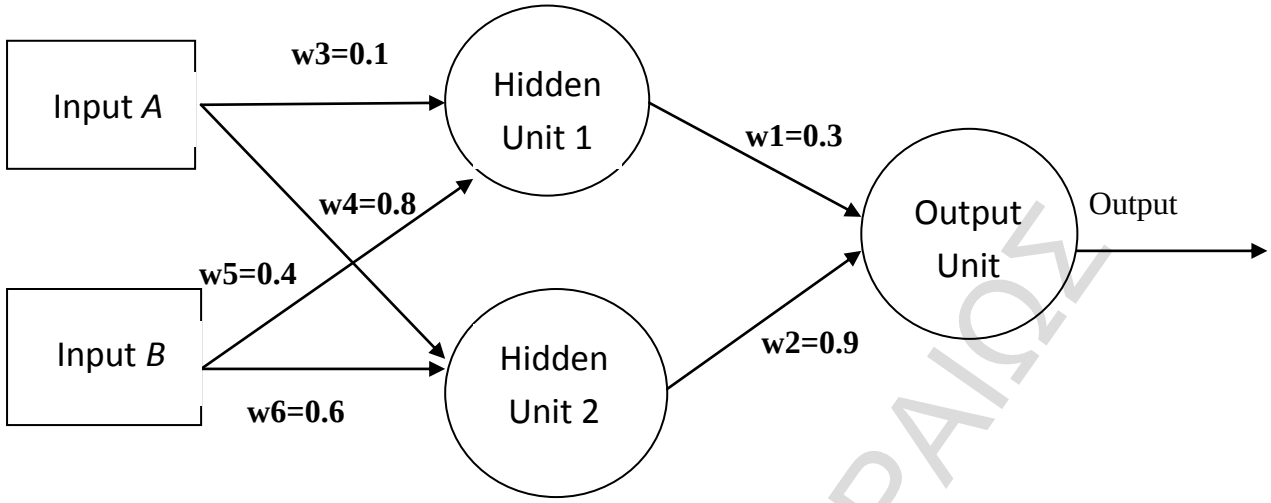
$$\Phi(x) = x$$

Στη συνέχεια θα περιγράψουμε την εφαρμογή της μεθόδου ενός νευρωνικού δικτύου στο σύνολο δεδομένων του πίνακα που ακολουθεί (Πίνακας 2.4.α).

| **Instance** | **Transaction amount (thousands $)** | **Timestamp (minutes)** | **Classification (0.5=legit, 1=fraud)** |
|---|---|---|---|
| 1 | 0.35 | 0.9 | 0.5 |
| 2 | 0.12 | 0.3 | 1 |
| 3 | 0.47 | 0.6 | 1 |

**Πίνακας 2.4.α.** Δεδομένα συναλλαγής πιστωτικών καρτών για χρήση της μεθόδου νευρωνικού δικτύου

Αρχικά επιλέγουμε ένα απλό δίκτυο όπως αυτό που φαίνεται παρακάτω (Εικόνα 2.4.γ σχηματική απεικόνιση του νευρωνικού μας δικτύου) του οποίου τα αρχικά βάρη είναι τυχαίοι αριθμοί. Οι νευρώνες σε αυτό το δίκτυο χρησιμοποιούν σιγμοειδή συνάρτηση ενεργοποίησης.




**Εικόνα 2.4.γ.** Δομή ενός απλού νευρωνικού δικτύου με τυχαία βάρη

Στο πιο πάνω δίκτυο δίνεται ως είσοδος ένα διάνυσμα **x** το οποίο αντιστοιχεί σε δύο χαρακτηριστικά με δεδομένα προηγούμενων συναλλαγών με τη χρήση πιστωτικών καρτών. Αφού εκπαιδευτεί το δίκτυο με προηγούμενα δεδομένα, θα είναι σε θέση να παράξει στην έξοδο μία τιμή που θα αντιπροσωπεύει την τιμή που «εκτιμήθηκε» για μία συναλλαγή που δεν είχε «ξαναδεί». Για να δώσουμε ένα παράδειγμα λειτουργίας στην ανίχνευση απάτης πιστωτικών καρτών με τη μέθοδο των νευρωνικών δικτύων, ας τρέξουμε αρχικά το δίκτυο με τα δεδομένα του πρώτου γεγονότος από τον Πίνακα 2.4.α τα οποία θα αποτελέσουν τις εισόδους *A* και *B* αντίστοιχα. Οι έξοδοι από κάθε νευρώνα υπολογίζονται ως εξής:

Είσοδος *A*= 0.35 , Είσοδος *B*=0.9

| | |
|---|---|
| Εισαγωγή των δεδομένων στην "Hidden Unit 1" | (0.35x0.1) + (0.9x0.8) = 0.755 |
| Εξαγωγή των δεδομένων από την "Hidden Unit 1" | $\frac{1}{1+e^{-0.755}} = 0.68$ |
| Είσοδος των δεδομένων στην "Hidden Unit 2" | (0.9x0.6) + (0.35x0.4) = 0.68 |
| Έξοδος των δεδομένων από την "Hidden Unit 2" | $\frac{1}{1+e^{-0.68}} = 0.6637$ |
| Είσοδος των δεδομένων στο "Output Unit" | (0.3x0.68) + (0.9x0.6637) = 0.8013 |
| Έξοδος των δεδομένων από το "Output Unit" | $\frac{1}{1+e^{-0.8013}} = 0.69$ |

**Πίνακας 2.4.β.** Πίνακας αποτελεσμάτων από κάθε νευρώνα

Στη συνέχεια υπολογίζουμε το σφάλμα εξόδου, το οποίο υπολογίζεται με βάση τη διαφορά ανάμεσα στην **τιμή αναφοράς** (*target value*), η οποία στο παράδειγμά μας για το πρώτο γεγονός ισούται με 0.5, και στην **τιμή εξόδου** (*output value*), την οποία και υπολογίσαμε παραπάνω και

την βρήκαμε 0.69. Η τιμή 0.5 που χρησιμοποιείται για την ταξινόμηση μίας συναλλαγής με τη χρήση πιστωτικής κάρτας ως δόλια οφείλεται στη συνάρτηση ενεργοποίησης που χρησιμοποιείται στο παράδειγμα αυτό. Ως συνάρτηση ενεργοποίησης χρησιμοποιείται η λογιστική συνάρτηση η οποία πρόκειται για μία από τις πιο γνωστές σιγμοειδούς συναρτήσεις. Το σφάλμα εξόδου σε κάθε νευρώνα συμβολίζεται με τον όρο $\delta$ το οποίο από τον ακόλουθο τύπο:

Σφάλμα εξόδου ($\delta$) = (στόχος - έξοδος) x (1 - έξοδος) x έξοδος

= (0.5 – 0.69) x (1 – 0.69) x 0.69

= -0.0406

Το σφάλμα στο στρώμα εξόδου συναρτάται με την παράγωγο της συνάρτησης ενεργοποίησης. Στο παράδειγμά μας αυτό χρησιμοποιείται η λογιστική συνάρτηση της οποίας τον τύπο έχουμε αναφέρει πιο πάνω στην παράγραφο 2.4. Μία χρήσιμη ιδιότητα της σιγμοειδούς αυτής συνάρτησης είναι η εξής:

$$\Phi'^{(x)} = \frac{d}{dx}\frac{1}{1+e^{-x}} = \frac{1}{(1+e^{-x})^2}(e^{-x})$$

$$= \frac{1}{1+e^{-x}}\left(1 - \frac{1}{1+e^{-x}}\right) = \Phi^{(x)}(1-\Phi^{(x)})$$

Επομένως ο όρος (1 - έξοδος) x έξοδος οφείλεται στην ιδιότητα της σιγμοειδούς συνάρτησης που χρησιμοποιείται στο νευρώνα καθώς για τον υπολογισμό της εξόδου χρησιμοποιούμε τη συνάρτηση ενεργοποίησης. Τα βάρη που συνδέουν το **κρυφό στρώμα** (*hidden layer*) με τις **μονάδες εξόδου** (*output unit*) ενημερώνονται ως εξής:

- $w_1^*$ = $w_1$ + ($\delta$ x είσοδος από την "hidden unit 1" ως την "output unit") = 0.3 + (-0.0406 x 0.68) = 0.272392
- $w_2^*$ = $w_2$ + ($\delta$ x είσοδος από την "hidden unit 2" ως την "output unit") = 0.9 + (-0.0406 x 0.6637) = 0.87305

Σε αντίθεση με το επίπεδο εξόδου, τα σφάλματα του κρυφού επιπέδου δεν μπορούν να υπολογιστούν άμεσα καθώς δεν υπάρχει μία τιμή αναφοράς στο επίπεδο αυτό και ως εκ τούτου τα σφάλματα αυτά δίνονται μέσα από τις μονάδες εξόδου. Επομένως ο υπολογισμός των συγκεκριμένων σφαλμάτων γίνεται με τη χρήση των σφαλμάτων από την μονάδα εξόδου πολλαπλασιασμένα με τα νέα βάρη των μονάδων εξόδου που υπολογίσαμε προηγουμένως. Τα αποτελέσματα που θα έχουμε θα είναι τα εξής:

- $\delta_1$ = $\delta$ x $w_1^*$ = -0.0406 x 0.272392 x [(1 – έξοδος "hidden unit 1") x έξοδος "hidden unit 1"] = -0.0406 x 0.272392 x [(1 – 0.68) x 0.68] = -2.406 x $10^{-3}$

- $\delta_2 = \delta$ x $w_2^*$ = -0.0406 x 0.87305 x [(1 – έξοδος "hidden unit 2") x έξοδος "hidden unit 2"] = -0.0406 x 0.87305 x [(1 – 0.6637) x 0.6637] = $-7.916 \times 10^{-3}$

Χρησιμοποιώντας τώρα τα αποτελέσματα των σφαλμάτων που βρήκαμε παραπάνω, τα νέα βάρη κρυφού επιπέδου μπορούν να υπολογιστούν ως εξής:

- $w_3^* = w_3 + (\delta_1$ x Είσοδος $A) = 0.1 + (-2.406 \times 10^{-3} \times 0.35) = 0.0992$
- $w_4^* = w_4 + (\delta_1$ x Είσοδος $B) = 0.8 + (-2.406 \times 10^{-3} \times 0.9) = 0.7978$
- $w_5^* = w_5 + (\delta_2$ x Είσοδος $A) = 0.4 + (-7.916 \times 10^{-3} \times 0.35) = 0.3972$
- $w_6^* = w_6 + (\delta_2$ x Είσοδος $B) = 0.6 + (-7.916 \times 10^{-3} \times 0.9) = 0.5928$

Έτσι ολοκληρώνεται η πρώτη επανάληψη στην οποία ενημερώνονται όλα τα βάρη του νευρωνικού δικτύου με τα δεδομένα του πρώτου γεγονότος. Δουλεύοντας μέσα στο δίκτυο με τα νέα βάρη η νέα τιμή της τελικής εξόδου υπολογίζεται να είναι 0.683. Αυτό οδηγεί σε ένα νέο μειωμένο σφάλμα, το οποίο ισούται με -0.183. Το αποτέλεσμα που πήραμε για την τελική έξοδο δηλώνει ότι η συναλλαγή μας είναι δόλια. Τιμές μεγαλύτερες του 0.5 δηλώνουν ότι η συναλλαγή είναι δόλια ενώ αντίθετα τιμές μικρότερες του 0.5 δηλώνουν ότι η συναλλαγή μέσω μία πιστωτικής κάρτας είναι νόμιμη. Την ίδια διαδικασία που αναλύσαμε παραπάνω ακολουθούμε για όλες τις περιπτώσεις του συνόλου δεδομένων.

## 2.5 Μέθοδος Κοντινότερου Γείτονα

Μια πολύ γνωστή και ευρέως χρησιμοποιούμενη τεχνική κατηγοριοποίησης που βασίζεται στη χρήση μέτρων βασισμένων στην απόσταση είναι αυτή των ***k*-κοντινότερων γειτόνων** (*K-Nearest Neighbor* (*KNN*)). Η τιμή *k* (ο αριθμός των κοντινότερων γειτόνων), ο οποίος χρησιμοποιείται για την επίτευξη της κατηγοριοποίησης με τη μεγαλύτερη δυνατή ακρίβεια, είναι σταθερός και γνωστός εκ των προτέρων. Όταν το μέγεθος της βάσης δεδομένων όπου αναζητούνται οι κοντινότεροι γείτονες είναι μεγάλο, οι αλγόριθμοι της σειριακής και της δυαδικής αναζήτησης δε μπορούν να εφαρμοστούν εξαιτίας του χρόνου που απαιτούν. Η μέθοδος κοντινότερου γείτονα είναι μια γενική μέθοδος με εφαρμογές στην κατασκευή μοντέλων πρόβλεψης νέων τιμών που μπορεί να χρησιμοποιηθεί και για την κατάταξη παρατηρήσεων. Η βασική ιδέα της μεθόδου είναι πως έχουμε ένα δείγμα και θέλουμε για μια νέα παρατήρηση με γνωστές τιμές και για ένα διάνυσμα μεταβλητών *x*, να προβλέψουμε την τιμή μια μεταβλητής *y*. Τότε χρησιμοποιούμε για την πρόβλεψή μας την πληροφορία που περιέχουν οι τιμές του δείγματος που μοιάζουν περισσότερο με την νέα παρατήρηση για την οποία θέλουμε να κάνουμε πρόβλεψη.

Ένα πρώτο βήμα στη μέθοδο αυτή είναι ο καθορισμός των παραμέτρων που απαιτούνται, δηλαδή η τιμή *k*, η οποία είναι γνωστή εκ των προτέρων όπως αναφέραμε και παραπάνω, καθώς

επίσης και το μέτρο απόστασης που θα χρησιμοποιηθεί. Υπάρχουν διάφοροι τρόποι επιλογής κατάλληλου μέτρου απόστασης ανάλογα με τη φύση του προβλήματος που εξετάζουμε κάθε φορά. Ένα από τα μέτρα απόστασης το οποίο είναι γνωστότερο για τη μέθοδο αυτή είναι η **ευκλείδια απόσταση** (*Euclidean distance*). Η ευκλείδια απόσταση οποία μεταξύ δύο παρατηρήσεων $p = (p_1, p_2, \dots, p_n)$ και $q = (q_1, q_2, \dots, q_n)$ δίνεται από τον ακόλουθη τύπο :

$$d(p,q) = d(q,p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}.$$

Παρακάτω θα εξετάσουμε ένα πρόβλημα ανίχνευσης απάτης πιστωτικών καρτών με τη χρήση της μεθόδου *k*-κοντινότερου γείτονα. Στον παρακάτω πίνακα που ακολουθεί παρουσιάζονται τα δεδομένα που θα χρησιμοποιηθούν για τη μέθοδο KNN. Ως μέτρο απόστασης για παράδειγμα αυτό θα χρησιμοποιήσουμε το **τετράγωνο της ευκλείδιας απόστασης** (*Squared Euclidean distance*) και *k*=3. Στην πρώτη στήλη του Πίνακα 2.5.α παρουσιάζονται το σύνολο των περιστατικών που θα εξετάσουμε, στη δεύτερη μας δίνεται το ποσό της συναλλαγής, στην τρίτη στήλη έχουμε τον χρόνο που χρειάστηκε για τη συναλλαγή και στην τελευταία στήλη έχουμε το χαρακτηρισμό της κάθε συναλλαγής ως «νόμιμη» ή μη.

| **Instance** | **Transaction amount (thousands $)** | **Timestamp (minutes)** | **Classification** |
|---|---|---|---|
| 1 | 25 | 25 | Fraud |
| 2 | 25 | 15 | Fraud |
| 3 | 12 | 15 | Legit |
| 4 | 7 | 15 | Legit |

**Πίνακας 2.5.α.** Σύνολο Δεδομένων για το παράδειγμα KNN

Ας υποθέσουμε ότι θέλουμε να προβλέψουμε μια συναλλαγή της οποίας το ποσό συναλλαγής ισούται με 12 δολάρια και έχει timestamp 25 λεπτά. Αρχικά θα πρέπει να υπολογίσουμε την απόσταση ανάμεσα στα δεδομένα του Πίνακα 2.5.α και τη νέα παρατήρηση μας που θέλουμε να κατατάξουμε (12$, 25 λεπτά). Ο πίνακας με τα αποτελέσματα που θα πάρουμε δίνεται παρακάτω:

| **Instance** | **Transaction amount (thousands $)** | **Timestamp (minutes)** | **Square Distance** |
|---|---|---|---|
| 1 | 25 | 25 | $(25\text{-}12)^2+(25\text{-}25)^2=169$ |
| 2 | 25 | 15 | $(25\text{-}12)^2+(15\text{-}25)^2=269$ |
| 3 | 12 | 15 | $(12\text{-}12)^2+(15\text{-}25)^2=100$ |
| 4 | 7 | 15 | $(7\text{-}12)^2+(15\text{-}25)^2=125$ |

**Πίνακας 2.5.β.** Αποστάσεις τετραγώνου ανάμεσα στα δεδομένα και τις νέες παρατηρήσεις

Στη συνέχεια θα πρέπει να ταξινομήσουμε τις αποστάσεις που υπολογίσαμε στον Πίνακα 2.5.β (*Square Distance*) από τη μικρότερη προς τη μεγαλύτερη και να προσδιορίσουμε που η νέα περίπτωση που εξετάζουμε (12$, 25 λεπτά) βρίσκεται εντός των 3-πλησιέστερων γειτόνων. Τα αποτελέσματα που θα έχουμε θα είναι τα ακόλουθα:

| **Instance** | **Transaction amount (thousands $)** | **Timestamp (minutes)** | **Square Distance** | **Lies within K-Nearest Neighbors? (k=3)** | **Classification of nearest neighbor** |
|---|---|---|---|---|---|
| 3 | 12 | 15 | $(12-12)^2+(15-25)^2=100$ | Yes | Legit |
| 4 | 7 | 15 | $(7-12)^2+(15-25)^2=125$ | Yes | Legit |
| 1 | 25 | 25 | $(25-12)^2+(25-25)^2=169$ | Yes | Fraud |
| 2 | 25 | 15 | $(25-12)^2+(15-25)^2=269$ | No | --- |

**Πίνακας 2.5.γ.** Ταξινόμηση των κοντινότερων γειτόνων

Σύμφωνα λοιπόν με τα αποτελέσματα που πήραμε από τον Πίνακα 2.5.γ βλέπουμε ότι η νέα περίπτωση που θέλουμε να εξετάσουμε (12$, 25 λεπτά) θα χαρακτηριζόταν ως νόμιμη συναλλαγή.

# ΚΕΦΑΛΑΙΟ 3

# Ανίχνευση Διαδυκτιακής Απάτης



## 3.1 Εισαγωγή

Στο κεφάλαιο αυτό θα αναφερθούμε στην ανίχνευση διαδυκτιακής απάτης η οποία περιλαμβάνει την ανίχνευση απάτης στον κλάδο των τηλεπικοινωνιών καθώς επίσης και τις εισβολές σε ηλεκτρονικά συστήματα. Στις επόμενες ενότητες αυτού του κεφαλαίου (Ενότητα 3.2 και Ενότητα 3.3) θα αναφέρουμε ορισμένα συνοπτικά στοιχεία για την κάθε μία μορφή απάτης και στη συνέχεια θα περιγράψουμε για την κάθε μία ξεχωριστά ορισμένες στατιστικές μεθόδους οι οποίες χρησιμοποιούνται για την ανίχνευσή τους.

Πολλές τεχνικές ανίχνευσης χρησιμοποιούνται στα συστήματα τηλεπικοινωνιών, από τις οποίες οι πιο διάσημες είναι αυτές των νευρωνικών δικτύων και των δέντρων αποφάσεων και ταξινόμησης. Άλλες τεχνικές που χρησιμοποιούνται εξίσου είναι η ανάλυση συστάδων, τα μοντέλα Markov ή οι κανόνες Bayesian. Στην Ενότητα 3.2 θα περιγράψουμε κάποιες από τις τεχνικές αυτές από τις οποίες θα αναφέρουμε και ορισμένα παραδείγματα εφαρμογής τους.

Όσον αφορά τώρα οι τεχνικές που χρησιμοποιούνται για την ανίχνευση απάτης σε ηλεκτρονικά συστήματα είναι πολλές. Συγκεκριμένα οι Ju και Vardi (2001) περιγράφουν την εφαρμογή των μοντέλων Markov ενώ οι Ryan *et al.* (1997) εφάρμοσαν τεχνικές νευρωνικών δικτύων. Ο Marchette (2001) αναφέρει διάφορες τεχνικές ανίχνευσης απάτης ηλεκτρονικών υπολογιστών, πολλές από τις οποίες αναφέρονται στο βιβλίο «Δίκτυα Υπολογιστών». Το βιβλίο αυτό εκδόθηκε τον Οκτώβρη του 2000 και περιγράφει αρκετά παραδείγματα νέων προσεγγίσεων στην ανίχνευση εισβολής ηλεκτρονικών υπολογιστών. Ένα σύστημα ανίχνευσης θα πρέπει να είναι σε θέση να προσαρμόζεται με τις αλλαγές που πραγματοποιούνται στα προφιλ των χρηστών καθώς επίσης με τα δίκτυα που αλλάζουν συνεχώς συμπεριφορά με την πάροδο του χρόνου. Στην Ενότητα 3.3 θα αναφέρουμε κάποιες από τις τεχνικές ανίχνευσης και θα παρουσιάσουμε και ορισμένα παραδείγματά τους.

## 3.2 Απάτες Τηλεπικοινωνιών

Τα τελευταία χρόνια παρατηρήσαμε όλοι την άνθιση της βιομηχανίας των τηλεπικοινωνιών. Αυτή η γιγαντιαία εξάπλωση οφείλεται κυρίως στην πρόοδο της τεχνολογίας των κινητών τηλεφώνων που έχει επιτρέψει την παραγωγή τους με μικρό κόστος. Η ανάπτυξη αυτή έχει μεγαλώσει τον αριθμό των χρηστών τηλεπικοινωνίας αλλά μαζί τους αυξήθηκε και το ποσό της απάτης που σχετίζεται με αυτήν.

Αρκετά άρθρα δίνουν διαφορετικά ποσά ή ποσοστά που περιγράφουν την έκταση της απάτης. Το 3% με 5% των εσόδων από τις τηλεπικοινωνίες παγκοσμίως χάνεται εξαιτίας της απάτης, κάτι το οποίο αναλογεί σε 55 δισεκατομμύρια δολάρια κάθε χρόνο. Μάλιστα οι νέες εταιρίες τηλεπικοινωνιών έχουν μεγαλύτερες απώλειες που σε ποσοστά μπορεί να ανέρχονται και σε 20% των εσόδων τους (Cahill et al. (2002)). Σύμφωνα με μία παγκόσμια έρευνα που διεξήχθη το 2009 ο οργανισμός απάτης τηλεπικοινωνιών ανακοίνωσε πως η απώλεια των εταιρειών που αφορούσαν απάτες υπολογίζεται ότι έφτανε το ποσό των 72-80 δισεκατομμυρίων δολαρίων. Τα αποτελέσματα της έρευνας αυτής έδειξαν ότι το ποσοστό απώλειας των εταιρειών ήταν πάνω από 34% των εσόδων τους σύμφωνα από αυτό που είχε δείξει η έρευνα το 2005.

Σύμφωνα με τα παραπάνω η ανίχνευση της απάτης στη βιομηχανία τηλεπικοινωνιών είναι σημαντική καθώς έχει ως αποτέλεσμα οι εταιρείες και οι προμηθευτές να χάνουν ένα σημαντικό μέρος των εσόδων τους. Για την ανίχνευση της απάτης χρησιμοποιούνται κατάλληλα μοντέλα συμπεριφοράς των χρηστών και στη συνέχεια γίνεται εφαρμογή αυτοματοποιημένων μεθόδων, προκειμένου να γίνει διάκριση της κανονικής από τη δόλια χρήση. Μια δυσκολία που αντιμετωπίζουν οι υπηρεσίες ανίχνευσης της απάτης είναι ότι συνήθως χρειάζεται να επεξεργαστούν τεράστια σύνολα δεδομένων, τα οποία εξελίσσονται διαρκώς. Τα δεδομένα που παράγονται από τα τηλεπικοινωνιακά δίκτυα είναι τεράστια και μπορούν να είναι της τάξεως των gigabyte ανά μέρα, επομένως οι τεχνικές **εξόρυξης δεδομένων** (*data mining*) είναι ιδιαίτερα δημοφιλείς. Ένας επιπλέον λόγος είναι ότι συναντούμε και εδώ ποικιλία στο είδος των δεδομένων: ημερομηνία και ώρα, διάρκεια κλήσης, είδος κλήσης, γεωγραφική προέλευση, γεωγραφικός προορισμός, κτλ.

Η εξόρυξη δεδομένων πρόκειται για μία σειρά από τεχνικές που βασίζονται σε ανάπτυξη αλγορίθμων που είναι χρήσιμες σε πολλούς και ετερόκλητους κλάδους όπως οι: οικονομία, βιοστατιστική, δημογραφία, μετεωρολογία και γεωλογία. Σήμερα οι βάσεις δεδομένων αποθηκεύουν στοιχεία μεγέθους terabytes. Μέσα σε αυτό το χαοτικό πλήθος δεδομένων κρύβεται μία σημαντική πληροφορία. Ο κλάδος που ασχολείται με αυτό το έργο, την ανάλυση μεγάλου συνόλου δεδομένων και την εξαγωγή πληροφοριών από αυτα, είναι γνωστός ως εξόρυξη δεδομένων. Συνεπώς η εξόρυξη

δεδομένων είναι μία περιοχή που χρειάζεται να χρησιμοποιήσει πληθώρα από εργαλεία ανάλυσης δεδομένων για να ανακαλύψει πρότυπα και σχέσεις σε δεδομένα που θα μπορέσουν να χρησιμοποιηθούν για να κάνουν προβλέψεις. Οι λειτουργίες της εξόρυξης δεδομένων χωρίζονται σε δύο βασικές κατηγορίες: την περιγραφή ή την πρόβλεψη. Η πρόβλεψη στοχεύει στον υπολογισμό της μελλοντικής αξίας ή στην πρόβλεψη της συμπεριφοράς κάποιων μεταβλητών που παρουσιάζουν ενδιαφέρον και οι οποίες βασίζονται στην συμπεριφορά άλλων μεταβλητών. Η περιγραφή επικεντρώνεται στην ανακάλυψη προτύπων και αναπαριστά τα δεδομένα μιας πολύπλοκης βάσης δεδομένων με ένα κατανοητό και αξιοποιήσιμο τρόπο. Χαρακτηριστικό παράδειγμα της μεθόδου αυτής αποτελεί η βάση δεδομένων AT&T. Πρόκειται για τη μεγαλύτερη τηλεφωνική εταιρεία παγκοσμίως η οποία περιέχει 350 εκατομμύρια προφίλ χρηστών και επεξεργάζεται 275 εκατομμύρια κλήσεις την εβδομάδα[4].

Στην διεθνή βιβλιογραφία έχουν αναφερθεί αρκετές μορφές απάτης στον κλάδο των τηλεπικοινωνιών. Όλες οι περιπτώσεις απάτης τηλεπικοινωνιών μπορεί πραγματικά να θεωρηθούν ως σενάρια απάτης που σχετίζονται με τον τρόπο που αποκτήθηκε η πρόσβαση στο δίκτυο. Ωστόσο, με δεδομένη την πληθώρα των τηλεπικοινωνιακών υπηρεσιών και την εφευρετικότητα των απατεώνων μπορεί κανείς να βρεθεί αντιμέτωπος με ποικίλες τεχνικές απάτης. Σε γενικές γραμμές, η ανίχνευση της απάτης επικεντρώνεται στην ανάλυση της δραστηριότητας των χρηστών και οι σχετικές προσεγγίσεις χωρίζονται σε δύο βασικές υποκατηγορίες: η μία ερευνά τα όρια μεταξύ της νόμιμης και της παράνομης συμπεριφοράς και η άλλη προσπαθεί να ανιχνεύσει ακραίες αλλαγές στη συμπεριφορά των χρηστών. Και στις δύο περιπτώσεις, η ανάλυση επιτυγχάνεται με τη βοήθεια στατιστικών και πιθανοτικών μεθόδων, νευρωνικών δικτύων και συστημάτων που βασίζονται στην εύρεση κανόνων συσχέτισης. Παρ 'όλα αυτά, θα πρέπει κανείς να έχει κατά νου ότι η υπερβολική χρήση ανίχνευσης της απάτης μπορεί να μην οδηγήσει στον εντοπισμό κάποιου απατεώνα αλλά στον καλύτερο πελάτη ενός παρόχου.

Στις επόμενες παραγράφους της ενότητας αυτής θα περιγράψουμε τον τρόπο με τον οποίο οι τεχνικές εξόρυξης δεδομένων μπορούν να χρησιμοποιηθούν ώστε να εξάγουμε σημαντικές πληροφορίες οι οποίες είναι "θαμμένες" στο σύνολο δεδομένων που μας δίνεται. Όπως αναφέραμε και σε προηγούμενο κεφάλαιο (Κεφάλαιο 2) το ίδιο και εδώ η εξόρυξη δεδομένων μπορεί να χρησιμοποιηθεί για τον προσδιορισμό της τηλεπικοινωνιακής απάτης, τη βελτίωση της αποτελεσματικότητας της εμπορίας, καθώς και τον εντοπισμό βλαβών του δικτύου. Εμείς θα αναφέρουμε δύο συγκεκριμένες εφαρμογές ανίχνευσης τηλεπικοινωνιακής απάτης από τις οποίες η μία βασίζεται στην εφαρμογή δέντρων απόφασης και η άλλη στη χρήση νευρωνικών δικτύων.

[4] Περισσότερες πληροφορίες σχετικά με το παράδειγμα αυτό αναφέρουν στην εργασία τους οι Cortes και Pregibon (1998)

Στα τηλεπικοινωνιακά συστήματα οι συναλλαγές των χρηστών καθώς επίσης και η συμπεριφορά τους περιέχονται στις **λεπτομερείς εγγραφές κλήσεων** (*Call Detail Record* (CDR)) του κάθε **ιδιωτικού υποκαταστήματος συναλλαγής** (*Private Branch Exchange* (PBX)). Το CDR περιέχει στοιχεία όπως: η αναγνώριση κλήσης, η διάρκεια της κλήσης, η ημερομηνία, ώρα της κλήσης, κλπ. Στον τομέα των τηλεπικοινωνιών ο στόχος των προφίλ συμπεριφοράς του χρήστη είναι να διακρίνει έναν κανονικό χρήστη από έναν απατεώνα. Τα προφίλ συμπεριφοράς του χρήστη κατασκευάστηκαν από τα ανεπεξέργαστα δεδομένα που περιέχονται στις CDR. Για τις αναλύσεις που έγιναν στο παράδειγμα που θα εξετάσουμε παρακάτω προτάθηκαν και μελετήθηκαν δύο διαφορετικά είδη προφίλ, στα οποία έγιναν διάφορες δοκιμές ώστε να αξιολογηθεί η ικανότητα ανίχνευσης της απάτης κάθε διαφορετικής δομής προφίλ. Η έρευνα αυτή πραγματοποιήθηκε από τους Constantinos S. Hilas και John N. Sahalos το 2007.

Το πρώτο προφίλ (Profile 1) έχει δημιουργηθεί από τη συσσωρευμένη εβδομαδιαία συμπεριφορά του χρήστη. Το προφίλ αυτό αποτελείται από επτά πεδία, τα οποία είναι η μέση τιμή και η τυπική απόκλιση του αριθμού των κλήσεων ανά εβδομάδα (calls), η μέση τιμή και η τυπική απόκλιση της διάρκειας (dur) των κλήσεων ανά εβδομάδα, ο μέγιστος αριθμός των κλήσεων, η μέγιστη διάρκεια μιας κλήσης και το μέγιστο κόστος μιας κλήσης (Πίνακας 3.2.α.) όπου και τα τρία αυτά πεδία υπολογίζονται σε περίοδο μιας εβδομάδας.

| Mean (calls) | Std (calls) | Mean (dur) | Std (dur) | Max (calls) | Max (dur) | Max (cost) |
|---|---|---|---|---|---|---|

**Πίνακας 3.2.α.** Το βασικό διάνυσμα για την παρουσίαση της εβδομαδιαίας συμπεριφοράς των χρηστών

Στη συνέχεια στον Πίνακα 3.2.β χρησιμοποιούνται τα ημερήσια δεδομένα για την κατασκευή ενός δεύτερου προφίλ (Profile 2) για να ελένξουν τα καθημερινά χαρακτηριστικά απάτης. Τα ημερήσια στοιχεία που χρησιμοποιούνται για την κατασκευή του είναι ο αριθμός κλήσεων ανά ημέρα (Calls), η διάρκεια των κλήσεων αυτών (Dur), η αντίστοιχη χρέωση μονάδων (Units), η μέγιστη διάρκεια μίας κλήσης (MaxDur), και οι μέγιστες μονάδες για μία κλήση σε αυτήν την ημέρα (MaxUnits).

| Calls | Dur | Units | MaxDur | MaxUnits |
|---|---|---|---|---|

**Πίνακας 3.2.β.** Το βασικό διάνυσμα για την παρουσίαση της καθημερινής συμπεριφοράς των χρηστών

Στην ενότητα αυτή θα παρουσιάσουμε την εφαρμογή ενός δέντρου απόφασης για την ανίχνευση απάτης σε πραγματικά δεδομένα που προέρχονται από μια βάση δεδομένων που διαθέτουν τα αρχεία της CDR από έναν οργανισμό PBX για μία περίοδο οκτώ ετών. Η βάση δεδομένων που χρησιμοποιήθηκε περιλαμβάνει 22.000 τηλεφωνικές κλήσεις οι οποίες

αντιστοιχούσαν σε 5.541 ημέρες (2702 ημέρες νόμιμης συμπεριφοράς και 2.839 με μία τουλάχιστον δόλια συμπεριφορά). Η εβδομαδιαία συνάθροιση αυτών των κλήσεων για κάθε χρήστη που εκπροσωπείται μέσω του προφίλ χρήστη (Πίνακας 3.2.α) απέδωσε 2.014 διανύσματα με 7 μεταβλητές το καθένα. Οι κανονικές ή νόμιμες χρήσεις αποτελούν το 51,1% του συνόλου των δεδομένων (class 1), ενώ η απάτη είναι το υπόλοιπο 48,9% (class 2). Η πολιτική χρέωσης του οργανισμού αυτού είναι ότι χρεώνει μόνο εθνικές, διεθνής και προς κινητά κλήσεις. Οι αστικές κλήσεις δεν χρεώνονται και επομένως δεν περιλαμβάνονται στο παράδειγμά μας. Στο παράδειγμά μας ο αλγόριθμος που χρησιμοποιήθηκε για την ταξινόμηση της απάτης σε δύο κατηγορίες είναι ο C4.5 (μία πλήρης περιγραφή του αλγορίθμου αυτού παρουσιάζεται στο βιβλίο του Quinlan J. (1993)), στόχος του οποίου ήταν να προσδιορίσει τις τιμές των κατάλληλων μεταβλητών που διαχωρίζουν καλύτερα τις τάξεις. Ο αλγόριθμος αυτός αποτελεί επέκταση του γνωστού αλγόριθμου ID3 τον οποίο περιγράψαμε στο Κεφάλαιο 2.

Τα δέντρα απόφασης χρησιμοποιήθηκαν με τη βοήθεια κανόνων προκειμένου να γίνει διάκριση μεταξύ της φυσιολογικής και της δόλιας συμπεριφοράς στο τηλεπικοινωνιακό δίκτυο. Οι λεπτομερείς λογαριασμοί εξετάστηκαν από έναν εμπειρογνώμονα στο συγκεκριμένο τομέα και κάθε τηλεφώνημα χαρακτηρίστηκε είτε ως απάτη είτε ως κανονική κλήση. Κάθε προφίλ από κάθε χρήστη σημάνθηκε σύμφωνα με δύο διαφορετικούς τρόπους. Ο πρώτος τρόπος ήταν να εντοπίσει την πρώτη ημέρα δραστηριοποίησης του απατεώνα. Στη συνέχεια ο λογαριασμός του κάθε χρήστη χωρίστηκε σε δύο σύνολα, ένα **πριν την απάτη** (*pre-fraud*) και ένα **μετά την απάτη** (*post-fraud*). Τα pre- και post- σχετίζονται με την πρώτη μέρα εμφάνισης της δόλιας δραστηριότητας. Η άλλη σήμανση της προσέγγισης ήταν πιο λεπτομερής. Αν κατά τη διάρκεια μιας ημέρας δεν είχε παρουσιαστεί καμία δόλια δραστηριότητα, τότε ολόκληρη η μέρα σημαδευόταν ως κανονική. Αν εμφανιζόταν τουλάχιστον μία δόλια δραστηριότητα κατα τη διάρκεια της ημέρας τότε ολόκληρη η μέρα χαρακτηριζόταν ως απάτη.

Για να εργαστούμε με δέντρα αποφάσεων κατασκευάζεται ένα προφίλ των χαρακτηριστικών συμπεριφοράς απάτης βάση του οποίου εξάγουμε συμπεράσματα για το αν η περίπτωση που εξετάζουμε πρόκειται για απάτη σύμφωνα με ένα σύνολο περιπτώσεων συμπεριφοράς απάτης που έχουμε ήδη συναντήσει.

Γενικά, ο αντικειμενικός στόχος ενός συστήματος ανίχνευσης απάτης είναι η μεγιστοποίηση του αριθμού των σωστών προβλέψεων αλλά και η διατήρηση των εσφαλμένων προβλέψεων σε αποδεκτά επίπεδα. Τα περισσότερα συστήματα ανίχνευσης απάτης προσπαθούν να ποσοτικοποιήσουν την απάτη και να της αποδώσουν κάποια χρηματική αξία ή όφελος. Με την ελαχιστοποίηση της πιθανότητας μη ανίχνευσης της απάτης ή της εσφαλμένης σήμανσης ως

απάτης μπορεί να επιτευχθεί με αρκετά υψηλή πιθανότητα μία σωστή διάγνωση. Παραθέτουμε στη συνέχεια ορισμένους τεχνικούς όρους που λαμβάνονται υπ' όψιν κατά την αξιολόγηση ενός συστήματος ανίχνευσης απάτης.

- **Ρυθμός εσφαλμένων σημάνσεων απάτης** (*false alarm ή false positive rate,* (FP)): Ονομάζεται το ποσοστό των νόμιμων επιθέσεων που εσφαλμένα χαρακτηρίστηκαν ως απατηλές.
- **Ρυθμός εσφαλμένων αρνητικών** χαρακτηρισμών (*false negative rate,* (FN)): Ονομάζεται το ποσοστό των απατών που λανθασμένα χαρακτηρίστηκαν ως νόμιμες.
- **Ρυθμός θετικών σημάνσεων** (*true positive rate, (TP)*): Ονομάζεται το ποσοστό των νόμιμων επιθέσεων που σωστά χαρακτηρίστηκαν ως απατηλές

Δύο ακόμη παράμετροι που χρησιμοποιούνται συνήθως είναι η **ανάκληση** (*recall*) και η **ακρίβεια** (*precision*). Η ανάκληση είναι στην πραγματικότητα η ίδια όπως το ποσοστό TP ενώ η ακρίβεια είναι ο αριθμός των αντικειμένων που έχουν ταξινομηθεί σωστά με βάση τον συνολικό αριθμό των αντικειμένων που κατατάσσονται στην ίδια κατηγορία. Η ακρίβεια υπολογίζεται από τον τύπο

Precision= $\frac{TP}{TP+FP}$ και εκφράζει την ορθότητα της μεθόδου που χρησιμοποιούμε

Πολύ συχνά χρησιμοποιείται και ένα στατιστικό μέτρο F το οποίο πρόκειται για ένα μέτρο της ακρίβειας ενός τεστ και υπολογίζεται ως εξής:

$$F - measure = \frac{2 \times recall \times precision}{recall + precision} = \frac{2TP}{2TP + FP + FN}$$

Γενικά, τα δέντρα απόφασης θεωρούνται μια ταξινομημένη δομή δέντρου που χρησιμοποιείται για να ταξινομήσει τάξεις με βάση κάποιους **κανόνες** (*rules*) τύπου IF ... THEN για τα **χαρακτηριστικά** (*attributes*) της **τάξης** (*class*). Οι κανόνες αυτοί είναι εύκολοι στην κατανόηση και στην επεξήγηση. Ένα άλλο χαρακτηριστικό που τα καθιστά δημοφιλή στην αναγνώριση προτύπων είναι ότι μπορούν να αντιμετωπίσουν αριθμητικές ή μη αριθμητικές μεταβλητές εξίσου καλά.

Θεωρητικά, ο αριθμός των μεταβλητών που χρησιμοποιούνται κατά την περιγραφή ενός προβλήματος μπορεί να βελτιώσει τη διακριτική ισχύ των αλγορίθμων μηχανικής μάθησης. Ωστόσο, ορισμένες μεταβλητές μπορεί να συσχετίζονται ή να είναι άσχετες με το πρόβλημα. Η χρήση τέτοιων

μεταβλητών "μπερδεύει" αλγόριθμους μηχανικής μάθησης. Γι' αυτό το λόγο είναι σύνηθες να προηγείται η διαδικασία επιλογής ενός χαρακτηριστικού γνωρίσματος που προσπαθεί να κρατήσει τα σημαντικότερα γνωρίσματα. Υπάρχουν αρκετές αυτοματοποιημένες μέθοδοι υπάρχουν που υποβοηθούν τη διαδικασία επιλογής χαρακτηριστικού γνωρίσματος. Η μείωση των διαστάσεων του προβλήματος αποδίδει μία πιο εύκολα ερμηνεύσιμη αναπαράσταση του προβλήματος. Επομένως όσο λιγότερες είναι οι μεταβλητές τοσο πιο κατανοητό και εύκολα ερμηνεύσιμο είναι το πρόβλημά μας.

Οι περισσότερες μέθοδοι που χρησιμοποιούνται για την διαδικασία επιλογής των χαρακτηριστικών γνωρισμάτων έχουν ως στόχο την αναζήτηση του χαρακτηριστικού εκείνου που συσχετίζεται έντονα με τις κλάσεις του προβλήματος και θα είναι λιγότερα συσχετισμένο με τα άλλα γνωρίσματα. Στο συγκεκριμένο παράδειγμα εφαρμόστηκαν τέσσερις διαφορετικές μέθοδοι επιλογής χαρακτηριστικών: the Best First, the Exhaustive Search, the Greedy Backward Elimination, and the Ranker method (Witten and Frank (2000)). Η τελευταία εφάρμοσε ως μέτρα σχετικότητας τόσο το **Κέρδος Πληροφορίας** (*Information Gain*) όσο και τον **δείκτη Κέρδους** (*Gain Ratio*). Όλες οι μέθοδοι έδωσαν παρόμοια αποτελέσματα, τα οποία παρουσιάζονται στον Πίνακα 3.2.γ. Σύμφωνα με τον πίνακα αυτό βλέπουμε ότι τα χαρακτηριστικά 1-4, δηλαδή η μέση τιμή και η τυπική απόκλιση του αριθμού των κλήσεων ανά εβδομάδα (calls) καθώς επίσης η μέση τιμή και η τυπική απόκλιση της διάρκειας (dur) των κλήσεων ανά εβδομάδα είναι ιδιαίτερα σημαντικά με βάση τις περισσότερες μεθόδους. Αντιθέτως τα χαρακτηριστικά 6 και 7, δηλαδή η μέγιστη διάρκεια μιας κλήσης ανά εβδομάδα και το μέγιστο κόστος μιας κλήσης ανά εβδομάδα χαρακτηρίζονται ως άσχετα ενώ το χαρακτηριστικό 5 που είναι ο μέγιστος αριθμός κλήσεων ανά εβδομάδα εξαρτάται από τη μέθοδο που χρησιμοποιείται για το αν θεωρείται σημαντικό.

| Attribute | Feature Selection Method | | | | |
|---|---|---|---|---|---|
| | Best First | Exhaustive Search | Greedy Backward Elimination | Ranker (Information Gain) | Ranker (Gain Ratio) |
| | Number of folds (%) | | | Average Rank (rank) | Average Rank (rank) |
| 1 MeanCalls | 10 (100%) | 10 (100%) | 10 (100%) | 1,6 ± 0,49 (1) | 1,7 ± 0,78 (2) |
| 2StdCalls | 8 (80%) | 9 (100%) | 8 (80%) | 3,8 ± 0,4 (4) | 2,9 ± 1,04 (3) |
| 3 MeanDur | 10 (100%) | 10 (100%) | 10 (100%) | 5 ± 0 (5) | 3,9 ± 0,54 (4) |
| 4 StdDur | 10 (100%) | 10 (100%) | 10 (100%) | 1,9 ± 1,22 (2) | 1,7 ± 0,64 (1) |
| 5 MaxCalls | 2 (20%) | 0 (0%) | 2 (20%) | 2,7 ± 0,46 (3) | 4,8 ± 0,4 (5) |

| | | | | | |
|---|---|---|---|---|---|
| 6 MaxDur | 0 (0%) | 0 (0%) | 0 (0%) | 7 ± 0 (7) | 6,2 ± 0,4 (6) |
| 7 MaxUnits | 0 (0%) | 0 (0%) | 0 (0%) | 6 ± 0 (6) | 6,8 ± 0 (7) |

**Πίνακας 3.2.γ.** Επιλογή χαρακτηριστικού για την εβδομαδιαία παρουσιάση της συμπεριφοράς των χρηστών

Για τα εβδομαδιαία προφίλ των χρηστών κατασκευάστηκαν τέσσερα διαφορετικά δέντρα. Για να μπορέσουμε να έχουμε μία βάση συγκρίσεων ανάμεσα στα δέντρα αυτά, το πρώτο δέντρο (Tree1) χτίστηκε χωρίς κλάδεμα ή επιλογή χαρακτηριστικού γνωρίσματος. Το δεύτερο δέντρο (Tree2) αλλάχθηκε μόνο ο αριθμός των αντικειμένων στα φύλλα ώστε να είναι 20. Στο τρίτο δέντρο (Tree3) εφαρμόστηκαν και οι τεχνικές κλαδέματος και η διαδικασία επιλογής χαρακτηριστικού γνωρίσματος. Τέλος, στο τέταρτο δέντρο (Tree4) εφαρμόστηκε το **μειωμένο σφάλμα κλαδέματος** (*reduced error pruning*) (Εικόνα 3.2.α.). Λόγω περιορισμών χώρου παρουσιάζεται μόνο το τελευταίο δέντρο.

Σε κάθε φύλλο εμφανίζονται δύο αριθμοί που χωρίζονται με μία κάθετο (/). Ο πρώτος είναι ο συνολικός αριθμός των αντικειμένων στο φύλλο, και ο δεύτερος είναι ο αριθμός των αντικειμένων που ταξινομείται εσφαλμένα στο φύλλο (Εικόνα 3.2.α). Ο αριθμός των φύλλων για κάθε δέντρο και το αντίστοιχο μέγεθός του παρουσιάζονται στον ακόλουθο πίνακα (Πίνακα 3.2.δ). Στον ίδιο πίνακα εικονίζονται πέντε στατιστικά μέτρα ακρίβειας για τα προαναφερθέντα δέντρα τα οποία είναι το ποσοστό TP, το ποσοστό FP, η ανάκληση, η ακρίβεια και το στατιστικό μέτρο F. Σύμφωνα με τον παρακάτω πίνακα βλέπουμε ότι το ποσοστό των αντικειμένων που ταξινομούνται σωστά είναι μεγαλύτερο από 85% για όλα τα δέντρα ενώ βλέπουμε ότι η απόδοση χειροτερεύει ελαφρώς καθώς προχωράμε προς τα μικρότερα δέντρα μετά την εφαρμογή του κλαδέματος και της διαδικασίας επιλογής χαρακτηριστικών. Ωστόσο, υπάρχει μία σημαντική βελτίωση σχετικά με τη δυνατότητα περιγραφής ενός μικρότερου δέντρου καθώς είναι πιο εύκολα ερμηνεύσιμο.

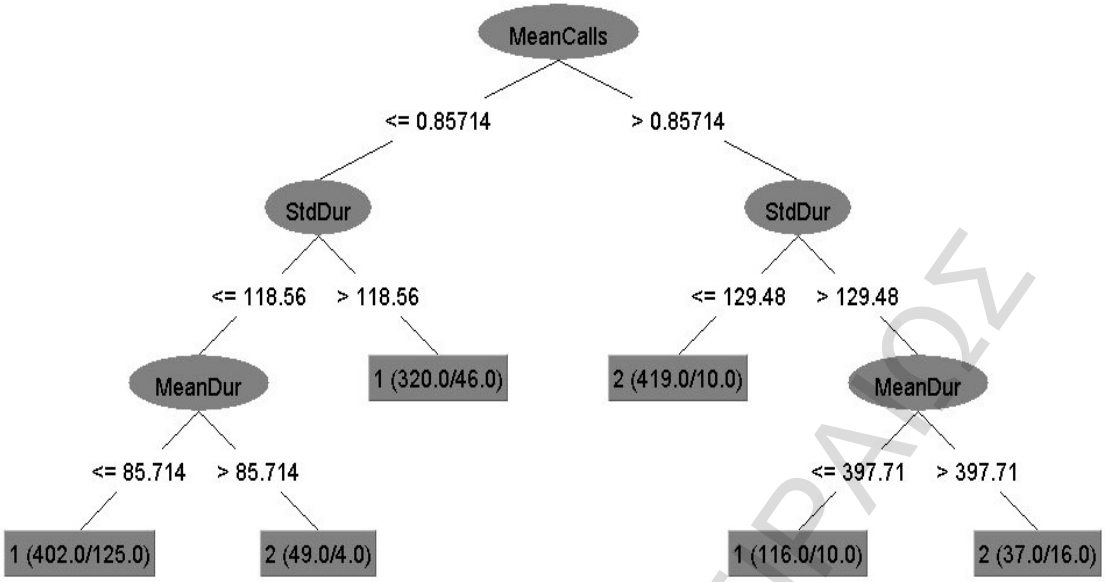
| **Statistics for Tree1 (no. of leaves: 63, size of tree: 125)** | | | | | |
|---|---|---|---|---|---|
| TP Rate | FP Rate | Precision | Recall | F-Measure | Class |
| 0,937 | 0,153 | 0,865 | 0,937 | 0,899 | 1 |
| 0,847 | 0,063 | 0,928 | 0,847 | 0,885 | 2 |
| **Statistics for Tree2 (no. of leaves: 20, size of tree: 39)** | | | | | |
| TP Rate | FP Rate | Precision | Recall | F-Measure | Class |
| 0,918 | 0,168 | 0,851 | 0,918 | 0,884 | 1 |
| 0,832 | 0,082 | 0,907 | 0,832 | 0,868 | 2 |

| **Statistics for Tree3 (no. of leaves: 13, size of tree: 25)** | | | | | |
|---|---|---|---|---|---|
| TP Rate | FP Rate | Precision | Recall | F-Measure | Class |
| 0,946 | 0,202 | 0,83 | 0,946 | 0,884 | 1 |
| 0,798 | 0,054 | 0,933 | 0,798 | 0,86 | 2 |
| **Statistics for Tree4 (no. of leaves: 6, size of tree: 11)** | | | | | |
| TP Rate | FP Rate | Precision | Recall | F-Measure | Class |
| 0,927 | 0,199 | 0,83 | 0,927 | 0,876 | 1 |
| 0,801 | 0,073 | 0,913 | 0,801 | 0,853 | 2 |

**Πίνακας 3.2.δ.** Στατιστικά των δέντρων τα οποία αντιστοιχούν στην εβδομαδιαία συμπεριφορά των χρηστών

Στο σχήμα (Εικόνα 3.2.α.) που ακολουθεί παρουσιάζεται το δέντρο απόφασης για το τελευταίο δέντρο (Tree4) η παρακολούθηση του οποίου μας παρέχει ενδείξεις για την κατασκευή κανόνων για τη διάκριση μεταξύ κανονικής και δόλιας χρήσης. Ο δεξιός κόμβος της ρίζας του δέντρου μας (MeanCalls> 0857) καλύπτει το 77% των περιπτώσεων απάτης. Ένα ενδιαφέρον αποτέλεσμα που φαίνεται από το δέντρο απόφασης που σχηματίστηκε είναι ότι η τυπική απόκλιση της διάρκειας των κλήσεων παίζει σημαντικό ρόλο στον προσδιορισμό της απάτης. Πιο συγκεκριμένα βλέπουμε ότι στις περιπτώσεις που έχουμε απάτη η τιμή της τυπικής απόκλισης είναι χαμηλότερη από ότι στις περιπτώσεις που έχουμε κανονικές συνθήκες χρήσης μίας κλήσης. Μια αναμενόμενη πτυχή της συμπεριφοράς των απατεώνων είναι εμφανής στα χαμηλότερα φύλλα της Εικόνας 3.2.α. Όπως φαίνεται οι απατεώνες τείνουν να κάνουν κλήσεις μεγάλης διάρκειας. Οι πιο ισχυροί κανόνες που προκύπτουν είναι οι εξής:

- IF MeanCalls <0,86 THEN class = 1 (εμπιστοσύνη: 71,98%, κάλυψη: 70,48%)
- IF MeanCalls> 0,86 και StdDur <129,5 THEN class = 2 (εμπιστοσύνη.: 97,5%, κάλυψη:. 41,5%).




**Εικόνα 3.2.α.** Δέντρο απόφασης για τα εβδομαδιαία προφίλ συμπεριφοράς των χρηστών.
Πηγή: C.S. Hilas and J.N. Sahalos (2007)

Ο πρώτος κανόνας λέει ότι αν ο χρήστης πραγματοποιεί λιγότερες από 1 κλήσης ανά την ημέρα τότε πρόκειται για έναν νόμιμο χρήστη με εμπιστοσύνη 72%. Σύμφωνα με το δεύτερο κανόνα, αν ο μέσος αριθμός των κλήσεων σε μια εβδομάδα είναι πάνω από 1 (δηλαδή τουλάχιστον 7 δαπανηρές κλήσεις στην εβδομάδα) και η τυπική απόκλιση της διάρκειά τους είναι λιγότερο από 2 λεπτά, τότε ο χρήστης είναι απατεώνας με εμπιστοσύνη 97,5%.

Κατά τον ίδιο τρόπο μελετήθηκε και η ημερήσια αναπαράσταση της συμπεριφοράς του χρήστη. Η αντίστοιχη διαδικασία επιλογής χαρακτηριστικού γνωρίσματος παρουσιάζεται στον Πίνακα 3.2.ε. Με βάση τα αποτελέσματα που μας δίνει ο παρακάτω πίνακας βλέπουμε ότι και οι πέντε μέθοδοι επιλογής χαρακτηριστικών συμφωνούν κατά την επιλογή των πιο σχετικών χαρακτηριστικών.

| Attribute | Feature Selection Method | | | | |
|---|---|---|---|---|---|
| | Best First | Exhaustive Search | Greedy Backward Elimination | Ranker (Information Gain) | Ranker (Gain Ratio) |
| | Number of folds (%) | | | Average Rank (rank) | Average Rank (rank) |
| 1 Calls | 10 (100%) | 10 (100%) | 10 (100%) | 1 ± 0 (1) | 1,1 ± 0,3 (1) |
| 2 Dur | 10 (100%) | 10 (100%) | 10 (100%) | 2,6 ± 0,49 (3) | 2,6 ± 0,49 (3) |

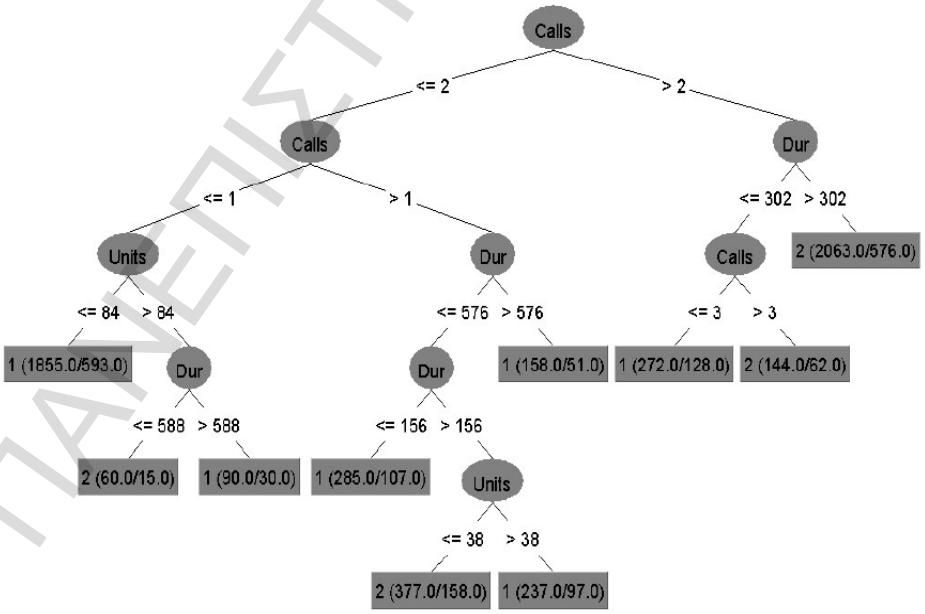
| | | | | | |
|---|---|---|---|---|---|
| 3 Units | 10 (100%) | 10 (100%) | 10 (100%) | 2,4 ± 0 (2) | 2,3 ± 0,64 (2) |
| 4 MaxDur | 0 (0%) | 0 (0%) | 0 (0%) | 5 ± 0 (5) | 4,8 ± 0,4 (5) |
| 5 MaxUnits | 0 (0%) | 0 (0%) | 0 (0%) | 4 ± 0 (4) | 4,2 ± 0,4 (4) |

**Πίνακας 3.2.ε.** Επιλογή χαρακτηριστικού για την καθημερινή παρουσιάση της συμπεριφοράς των χρηστών

Για τα καθημερινά προφίλ συμπεριφοράς των χρηστών χτίστηκαν τρία δέντρα. Για το πρώτο δέντρο (Tree5) δεν εφαρμόστηκε καμία μέθοδος κλαδέματος ή επιλογή χαρακτηριστικού. Η εφαρμογή της επιλογής χαρακτηριστικών έγινε στο δεύτερο δέντρο (Tree6) ενώ για το τρίτο (Tree7) εφαρμόστηκαν τόσο η διαδικασία επιλογής χαρακτηριστικού όσο και το κλάδεμα του δέντρου. Στο τρίτο δέντρο (Tree7) (Εικόνα 3.2.β), δύο είναι οι πιο αποτελεσματικοί κανόνες:

- IF Calls≥ 3 ΚΑΙ Dur> 302sec THEN class = 2 (εμπιστοσύνη: 72%, υποστήριξη: 52%).
- IF Calls = 1 ΚΑΙ Dur <84sec THEN class = 1 (εμπιστοσύνη: 68%, υποστήριξη: 46,7%).

Ο πρώτος κανόνας σημαίνει ότι αν ο χρήστης κάνει πάνω από τρεις χρεώσιμες κλήσεις μέσα σε μια ημέρα με διάρκεια πάνω από 5 λεπτά, τότε αυτός είναι απατεώνας με βεβαιότητα 72%. Ο δεύτερος κανόνας σημαίνει ότι, εάν ο χρήστης κάνει το πολύ μια χρεώσιμη κλήση σε μια μέρα με διάρκεια μικρότερη από 84 δευτερόλεπτα τότε αυτός είναι ο νόμιμος χρήστης με βεβαιότητα 68%.




**Εικόνα 3.2.β.** Δέντρο απόφασης για τα καθημερινά προφίλ συμπεριφοράς των χρηστών.
Πηγή: C.S. Hilas and J.N. Sahalos (2007)

Στη συνέχεια της ενότητας αυτής θα αναφέρουμε την εφαρμογή ενός παραδείγματος ανίχνευσης απάτης στην κινητή τηλεφωνία με τη χρήση ενός νευρωνικού δικτύου. Συνολικά, περισσότερες από 12.000 τηλεφωνικές κλήσεις εξετάστηκαν οι οποίες αποδίδουν περίπου 2500 ημέρες κλήσεων. Η εβδομαδιαία συνάθροιση αυτών των κλήσεων για κάθε χρήστη έδωσε 1100 διανύσματα της εβδομαδιαίας συμπεριφοράς. Το 30-50% των διανυσμάτων στο σύνολο δεδομένων κάθε χρήστη αφορούσαν περιπτώσεις απάτης, ενώ το υπόλοιπο ήταν κανονική χρήση. Στην περίπτωσή μας, επιλέχθηκαν τα δεδομένα από μια βάση δεδομένων που περιέχει την δραστικότητα των κλήσεων περίπου 6.000 χρήστες ενός PBX οργανισμού.

Προκειμένου να χρεώνουν σωστά τους χρήστες, για τις κλήσεις που κάνουν, χρησιμοποιείται ένα σύστημα προσωπικών αριθμών αναγνώρισης (PIN). Κάθε χρήστης έχει ένα μοναδικό PIN. Αν κάποιος (π.χ., ένας απατεώνας) βρει ένα PIN μπορεί να το χρησιμοποιήσει για να κάνει δικές του κλήσεις από οποιαδήποτε τηλεφωνική συσκευή εντός του οργανισμού. Αρκετοί λογαριασμοί χρηστών, που έχουν εξαπατηθεί, έχουν εντοπιστεί. Τρεις από αυτούς παρουσιάζονται στη συνέχεια όπου και οι τρεις περιέχουν παραδείγματα των νόμιμων αλλλά και της δόλιας δραστηριότητας. Οι συγκεκριμένοι λογαριασμοί επελέγησαν επειδή περιέχουν αντιπροσωπευτικά είδη διαφορετικής συμπεριφοράς απατεώνα τα οποία περιγράφεται παρακάτω.

Στο πρώτο παράδειγμα (User1) έχουμε μια περίπτωση όπου ο απατεώνας φανερώνει μια άπληστη συμπεριφορά, που στον κλάδο των τηλεπικοινωνιών αυτό σημαίνει ότι οι απατεώνες έχουν την τάση να μιλούν πολύ ή να κάνουν δαπανηρές κλήσεις. Αφού απέκτησε το PIN του χρήστη, τοποθετεί ένα μεγάλο ποσό υψηλό κόστος κλήσεων προς δορυφορικές υπηρεσίες. Στο δεύτερο παράδειγμα (User2) είναι μια περίπτωση όπου ο απατεώνας δεν πραγματοποίησε σημαντικά δαπανηρές κλήσεις, αλλά χρησιμοποίησε το PIN του χρήστη κατά τη διάρκεια μη εργάσιμων ωρών και ημερών. Τέλος στο τρίτο παράδειγμα (User3) ο απατεώνας φαίνεται να είναι πιο προσεκτικός. Δεν έκανε δαπανηρές κλήσεις και χρησιμοποίησε κυρίως το λογαριασμό κατά τις εργάσιμες ημέρες και ώρες. Μια ενδιαφέρουσα παρατήρηση για την τρίτη περίπτωση είναι ότι το κλεμμένο PIN χρησιμοποιήθηκε από διαφορετικές τηλεφωνικές συσκευές και οι κλήσεις προς συγκεκριμένους προορισμούς δεν συνδέονταν με την τηλεφωνική συσκευή του νόμιμου χρήστη.

Για κάθε χρήστη, κατασκευάστηκαν τρια διαφορετικά τύποι προφίλ. Το πρώτο προφίλ (Profile 1) έχει δημιουργηθεί από τη συσσωρευμένη εβδομαδιαία συμπεριφορά του χρήστη. Το προφίλ αυτό αποτελείται από επτά πεδία, τα οποία είναι η μέση τιμή και η τυπική απόκλιση του αριθμού των κλήσεων ανά εβδομάδα (calls), η μέση τιμή και η τυπική απόκλιση της διάρκειας (dur) των κλήσεων ανά εβδομάδα, ο μέγιστος αριθμός των κλήσεων, η μέγιστη διάρκεια μιας

κλήσης και το μέγιστο κόστος μιας κλήσης (Πίνακας 3.2.ζ) όπου και τα τρία αυτά πεδία υπολογίζονται σε περίοδο μιας εβδομάδας.

| Mean (calls) | Std (calls) | Mean (dur) | Std (dur) | Max (calls) | Max (dur) | Max (cost) |
|---|---|---|---|---|---|---|

**Πίνακας 3.2.ζ.** Profile 1 τηλεφωνικών κλήσεων

Το δεύτερο προφίλ (Profile 2) είναι μία λεπτομερή ημερήσια συμπεριφορά ενός χρήστη το οποίο έχει κατασκευασθεί διαχωρίζοντας τον αριθμό των κλήσεων ανά ημέρα και τις αντίστοιχες διάρκειές τους ανά ημέρα σύμφωνα με τον καλούμενο προορισμό, δηλαδή, εθνικές (nat), διεθνή (int), και κινητής τηλεφωνίας (mob) κλήσεις, καθώς και ο χρόνος της ημέρας, δηλαδή, ώρες εργασίας (w), απογευματινές ώρες (a), και βράδυ (n).

| nat_ calls_w | nat_dur_w | nat_ calls_a | nat_dur_a | nat_ calls_n | nat_dur_n |
|---|---|---|---|---|---|
| mob_calls_w | mob_dur_w | mob_calls_a | mob_dur_a | mob_calls_n | mob_dur_n |
| int_calls_w | int_dur_w | int_calls_a | int_dur_a | int_calls_n | int_dur_n |

**Πίνακας 3.2.η.** Profile 2 τηλεφωνικών κλήσεων

Τέλος, το τρίτο προφίλ (Profile 3) είναι μια συσσωρευμένη ανά ημέρα συμπεριφορά. Αποτελείται από τον αριθμό των κλήσεων και οι αντίστοιχες διάρκειές τους χωρίζονται μόνο σύμφωνα με τον αποκαλούμενο προορισμό, δηλαδή, εθνικών, διεθνών και προς κινητά κλήσεις.

| nat_ calls | nat_dur | mob_calls | mob_dur | int_calls | int_dur |
|---|---|---|---|---|---|

**Πίνακας 3.2.θ.** Profile 3 τηλεφωνικών κλήσεων

Τα τελευταία δύο προφίλ επίσης αθροίζονται ξεχωριστά για κάθε εβδομάδα για να δώσουν ξεχωριστά τα Profile2w και Profile3w. Επομένως για τους τρεις χρήστες που εξετάζονται στο παράδειγμα αυτό έχουμε 5 αναπαραστάσεις προφίλ, και 2 διαφορετικούς τρόπους για να χαρακτηρίσουμε τους λογαριασμούς χρηστών κανονικούς από δόλιους. Έτσι ο κάθε χρήστης έχει 10 διαφορετικά σύνολα δεδομένων και για τα τρία προφίλ συνολικά έχουμε 30 διαφορετικά σύνολα δεδομένων.

Προκειμένου να ελεγχθεί η ικανότητα του κάθε προφίλ να διακρίνει μεταξύ μίας νόμιμης χρήσης από μία δόλια, χρησιμοποιήσαν ένα **νευρωνικό δίκτυο προς τα εμπρός τροφοδότησης** (*feed-forward neural network*). Η συνάρτηση του νευρωνικού δικτύου ορίζεται ως εξής:

$$o_k(x) \equiv z_k = g\left(\sum_{i=1}^{n_H} w_{kj}^{(0)} f\left(\sum_{i=1}^{d} w_{ji}^{(0)} x_i + w_{j0}^{(0)}\right) + w_{k0}^{(1)}\right)$$

όπου το $x_i$ είναι οι $d$ είσοδοι, $w_{kj}$ είναι τα βάρη μεταξύ της εξόδου *y* και το κρυφού στρώματος, $w_{ji}$ είναι τα βάρη από την *i*-οστή είσοδο $x_i$ στον $j$ νευρώνα του κρυφού στρώματος, $n_H$ είναι ο αριθμός των νευρώνων του κρυφού στρώματος, $k$ είναι ο αριθμός των εξωτερικών νευρώνων και $z_k \equiv o_k(x)$ είναι οι $k$ έξοδοι. Η $f(\cdot)$ είναι μία μη γραμμική συνάρτηση ενεργοποίησης[5] καθώς επίσης χρησιμοποιήθηκαν γραμμικές εξόδους, $g(\cdot)$.

Το πρόβλημα αυτό αποτελεί μία μορφή επιβλεπόμενης μάθησης που έχει ως σκοπό να προσαρμόσει τα βάρη έτσι ώστε η απεικόνιση εισόδου-εξόδου να αντιστοιχεί με εκείνα τα ζεύγη εισόδου-εξόδου που έχει παράσχει ο εκπαιδευτής.

Όλα τα δεδομένα μας έχουν κανονικοποιηθεί με τέτοιο τρόπο ώστε για κάθε χαρακτηριστικό ο μέσος να ισούται με το 0 και η τυπική απόκλιση με 1. Επιπλέον χρησιμοποιήθηκε η **κύρια ανάλυση συνιστωσών** (*Principal component analysis*) ώστε να μετατρέψουν τα διανύσματα εισόδου σε ασυσχέτιστες μεταβλητές. Προκειμένου να ελεγχθεί η ικανότητα του κάθε προφίλ να διακρίνει τη νόμιμη χρήση από την απάτη χρησιμοποιήθηκε ένα νευρωνικό δίκτυο προς τα εμπρός τροφοδότησης το οποίο αποτελείται από ένα κρυφό στρώμα και μία γραμμική έξοδο.

Η αξιολόγηση της απόδοσης του νευρωνικού δικτύου γίνεται με τη βοήθεια της αντίστοιχης καμπύλης λειτουργίας χαρακτηριστικού λειτουργίας (*receiver operating characteristic (ROC)*). Μία καμπύλη ROC είναι στην πραγματικότητα μια γραφική αναπαράσταση συναλλαγών μεταξύ των πραγματικών θετικά και των εσφαλμένων θετικά τιμών για κάθε πιθανό σημείο αποκοπής που χωρίζει δύο κατανομές. Στην ανάλυση της καμπύλης ROC, η περιοχή κάτω από αυτή χρησιμοποιείται συχνά ως ένα στατιστικό μέτρο για το πόσο καλά αποδίδει ο ταξινομητής. Επομένως μία τιμή της περιοχής κοντά στο 1 δείχνει μία τέλεια ταξινόμηση ενώ μία τιμή κοντά στο 0,5 υποδεικνύει ότι τα δεδομένα δεν ταξινομήθηκαν με τον πλέον βέλτιστο τρόπο. Οι τιμές ταξινόμησης για όλες τις καμπύλες ROC για τον User1 και για τις πέντε παραστάσεις προφίλ του δίνονται στον ακόλουθο πίνακα (Πίνακα 3.2.ι).

Οι συγκρίσεις έγιναν ανα ζεύγη. Για κάθε τύπο προφίλ και για κάθε χρήστη, η πρώτη περίπτωση περιλαμβάνει το χωρισμό του λογαριασμού του χρήστη σε δύο μέρη, πριν και μετά της πρώτης μέρας εμφάνισης μιας δόλιας δραστηριότητας (στην Εικόνα 3.2.ε. απεικονίζεται με μια σταθερή γραμμή (unchar)). Και η δεύτερη περίπτωση αναφέρεται στον λογαριασμό του χρήστη ο οποίος χωρίστηκε σε κανονική και δόλια δραστηριότητα χρησιμοποιώντας ένα λεπτομερή χαρακτηρισμό της μέρας (στην Εικόνα 3.2.ε. απεικονίζεται με την διακεκομμένη

[5] Περισσότερες πληροφορίες αναφέρονται στο 2$^{ο}$ Κεφάλαιο

γραμμή (char)). Στο σχήμα που ακολουθεί (Εικόνα 3.2.ε.) παρουσιάζονται παραδείγματα καμπύλων ROC για τον ίδιο χρήστη για τα προφίλ Profile1 και Profile3w. Τα σχεδιαγράμματα για τα άλλα προφίλ του ίδιου χρήστη, καθώς και για τους υπόλοιπους χρήστες είναι παρόμοια και παραλείπονται λόγω περιορισμών στο χώρο. Η διαγώνια γραμμή είναι η θεωρητική καμπύλη ROC για τον τυχαίο διαχωρισμό των περιπτώσεων. Η περιοχή κάτω από την καμπύλη χρησιμοποιήθηκε ως το στατιστικό που επιδεικνύει την απόδοση ταξινόμησης κάθε περίπτωσης.

Οι καμπύλες ROC θα πρέπει να κρίνονται με επιφύλαξη. Μια μεγαλύτερη περιοχή κάτω από την καμπύλη δεν συνεπάγεται καλύτερη απόδοση. Στο σχήμα. 3.2.ε οι καμπύλες ROC για το Profile1 συγκρίθηκαν με εκείνες του Profile3w. Το τελευταίο προφίλ υποτίθεται ότι θα δώσει καλύτερα αποτελέσματα, σύμφωνα με τα στοιχεία του Πίνακα 3.2.ι. Ωστόσο, το Profile1 λειτουργεί καλύτερα, καθώς δίνει υψηλό ποσοστό πραγματικών θετικά (80%) για πολύ χαμηλό ποσοστό εσφαλμένων θετικά (2%) (διακεκομμένη γραμμή). Αντίστοιχα, το Profile3w δίνει το ίδιο TP αλλά με ποσοστό εσφαλμένων θετικά 12% (γραμμή με τα αστέρια πάνω).

Το δίκτυο τηλεπικοινωνιών από το οποίο αντλήσαμε τα παραδείγματα μας έχει περισσότερους από 6000 χρήστες. Αυτό σημαίνει ότι ακόμη και ένα ποσοστό μόλις 1% ψευδώς θετικών αποτελεσμάτων μπορεί να δώσει μέχρι και 60 εσφαλμένους συναγερμούς. Σύμφωνα με τις αναλύσεις που έγιναν με βάση τα πέντε προφίλ χρηστών που κατασκευάστηκαν για τον User1, τα αποτελέσματα με τη βοήθεια των νευρωνικών ταξινομητών έδειξαν ότι το Profile1 λειτουργεί καλύτερα από όλα τα υπόλοιπα. Όπως μπορούμε να δούμε και από τον παρακάτω πίνακα το Profile1 παρουσιάζει το υψηλότερο ποσοστό θετικά πραγματικών τιμών με το μικρότερο ποσοστό εσφαλμένων θετικά σε σύγκριση με τα υπόλοιπα προφίλ. Μια άλλη παρατήρηση που μπορούμε να κάνουμε σύμφωνα με τα αποτελέσματα που πήραμε και από τον πίνακα αλλά και από την εικόνα που δίνονται στη συνέχεια είναι ότι ο λεπτομερής χαρακτηρισμός της κάθε τηλεφωνικής κλήσης δίνει καλύτερα αποτελέσματα διάκρισης. Ο διαχωρισμός των δραστηριοτήτων του χρήστη, τόσο με βάση τον προορισμό και με βάση το χρόνο της ημέρας, βοηθά στην αναγνώρισης τέτοιου είδους ενεργειών.

| | *User1* | | | | |
|---|---|---|---|---|---|
| | Profile1 | Profile2 | Profile3 | Profile2w | Profile3w |
| Unchar | 0,7710 | 0,8069 | 0,8238 | 0,8181 | 0,8377 |
| Char | 0,9163 | 0,9134 | 0,8839 | 0,9120 | 0,9262 |

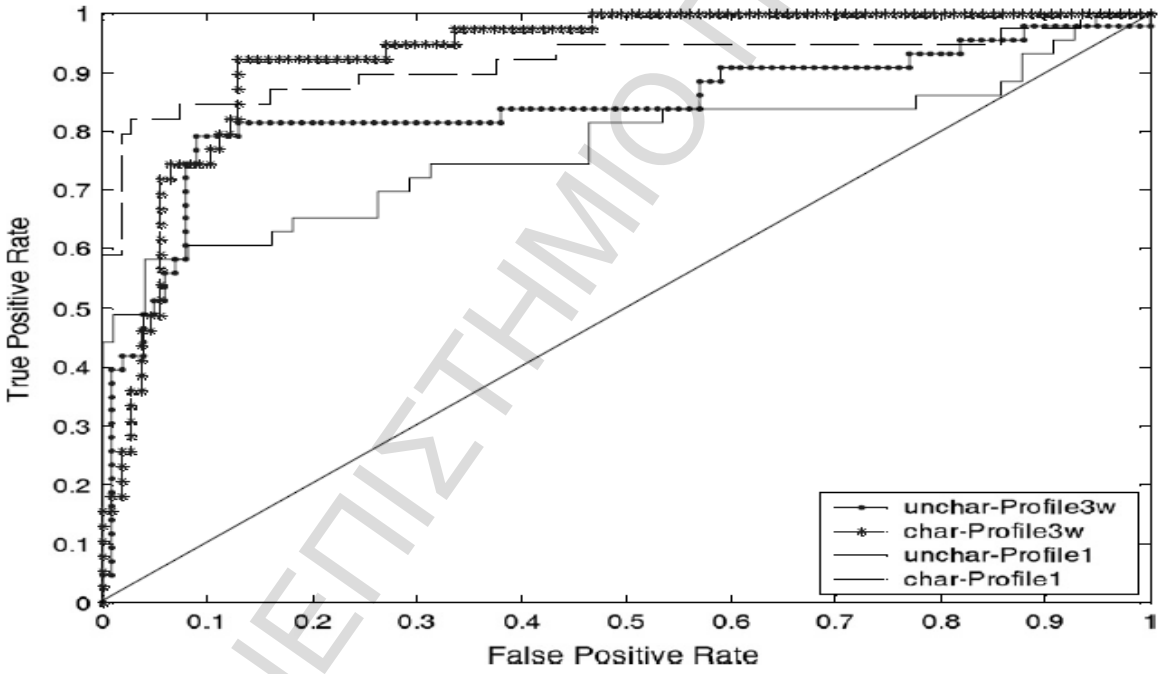
**Πίνακας 3.2.ι.** Περιοχή κάτω από τις καμπύλες ROC για τα 5 προφίλ του User1

| | *User2* | | | | |
|---|---|---|---|---|---|
| | Profile1 | Profile2 | Profile3 | Profile2w | Profile3w |
| Unchar | 0.6765 | 0.7213 | 0.6490 | 0.7241 | 0.6615 |
| Char | 0.8345 | 0.8811 | 0.7790 | 0.8760 | 0.8252 |

**Πίνακας 3.2.κ.** Περιοχή κάτω από τις καμπύλες ROC για τα 5 προφίλ του User2

| | *User3* | | | | |
|---|---|---|---|---|---|
| | Profile1 | Profile2 | Profile3 | Profile2w | Profile3w |
| Unchar | 0.9047 | 0.7971 | 0.7853 | 0.8741 | 0.8536 |
| Char | 0.9297 | 0.7789 | 0.7931 | 0.7298 | 0.8687 |

**Πίνακας 3.2.λ.** Περιοχή κάτω από τις καμπύλες ROC για τα 5 προφίλ του User3




**Εικόνα 3.2.ε.** Καμπύλες ROC, χρησιμοποιώντας τα Profile 1και Profile 3w, δείχνουν την εξισορρόπηση μεταξύ του ποσοστού TP και FP για τον πρώτο χρήστη.

## 3.3 Απάτες Εισβολής Ηλεκτρονικών Υπολογιστών

Η ραγδαία εξέλιξη της τεχνολογίας, η ανάπτυξη της πληροφορικής και η ευρύτατη χρήση του διαδικτύου έχουν επιφέρει επαναστατικές αλλαγές στο σύνολο των καθημερινών δραστηριοτήτων, στην παραγωγική διαδικασία, στις συναλλαγές, στην εκπαίδευση, στη διασκέδαση, ακόμα και στον τρόπο σκέψεως του σύγχρονου ανθρώπου. Μαζί με αυτές τις αλλαγές, οι οποίες κατά κανόνα βελτιώνουν την ποιότητα της ζωής μας, υπεισέρχονται και οι παράμετροι που ευνοούν την ανάπτυξη νέων μορφών εγκληματικότητας. Οι νέες αυτές μορφές εγκληματικότητας χαρακτηρίζονται συνήθως με τον όρο «ηλεκτρονικό έγκλημα».

Ο όρος ηλεκτρονικό έγκλημα ή ηλεκτρονική εγκληματικότητα μπορεί να έχει είτε στενή είτε ευρεία έννοια. Η στενή έννοια της ηλεκτρονικής εγκληματικότητας αναφέρεται στις αξιόποινες πράξεις όπως είναι η ηλεκτρονική απάτη, η χωρίς άδεια απόκτηση δεδομένων, η παραποίηση δεδομένων και η δολιοφθορά δηλαδή εγκλήματα όπου ο ηλεκτρονικός υπολογιστής αποτελεί κύριο μέσο τέλεσης των εγκλημάτων. Αντίθετα η ευρεία έννοια εγκληματικότητας μέσω Η/Υ περιλαμβάνει όλα εκείνα τα αδικήματα για την τέλεση των οποίων ο ηλεκτρονικός υπολογιστής χρησιμοποιείται ως βοηθητικό μέσο.

Είναι γεγονός ότι το ηλεκτρονικό έγκλημα διαπράττεται άμεσα, σε ελάχιστα δευτερόλεπτα. Ο επιτιθέμενος με τη χρήση ενός Η/Υ συνδεδεμένου στο διαδίκτυο, μπορεί να εισβάλει στα υπολογιστικά συστήματα μιας επιχείρησης ή ενός οργανισμού σε οποιοδήποτε σημείο του κόσμου. Δεν απαιτείται η φυσική μετακίνησή του, καθώς οι ενέργειές του μπορούν να ολοκληρωθούν από την οικία του ή άλλο χώρο, με τη χρήση ενός δικτυωμένου προσωπικού υπολογιστή.

Με τον όρο εισβολή ονομάζουμε την ενδεχομένως σκόπιμη μη εξουσιοδοτημένη προσπάθεια για πρόσβαση ή διαχείριση πληροφοριών όπως επίσης και την προσπάθεια αχρήστευσης ή επηρεασμού της αξιοπιστίας ενός συστήματος (Kou *et al*. (2004)). Οι απάτες που γίνονται μέσω εισβολής σε ηλεκτρονικούς υπολογιστές είναι ένας τομέας με αυξημένη δραστηριότητα και η ανίχνευση απάτης από εισβολές Η/Υ αποτελεί μια περιοχή της επιστήμης με ιδιαίτερα έντονη ερευνητική δράση. Οι εισβολείς ηλεκτρονικών υπολογιστών ονομάζονται και χάκερς, από την αγγλική λέξη hack που σημαίνει κομματιάζω.

Ο όρος hacker χαρακτηρίζει ένα κακόβουλο άτομο το οποίο εισβάλει σε υπολογιστικά συστήματα με σκοπό να υποκλέψει ή να καταστρέψει πληροφορίες. Αυτοί έχουν τη δυνατότητα να αλλάζουν πηγαίους κώδικες και να "σπάνε" προγράμματα, μπορούν να ανακαλύπτουν συνθηματικά ασφαλείας ή ακόμη και να υποκλέπτουν ηλεκτρονική αλληλογραφία Τα κίνητρά τους μπορεί να είναι οικονομικά, ιδεολογικά ή και απλώς διασκέδαση και χόμπι. Επομένως αν η

εισβολή σε ένα ηλεκτρονικό υπολογιστή ανακαλυφθεί αρκετά νωρίς μπορεί να αποτραπεί η είσοδος στον hacker. Όμως όπως και στις άλλες περιπτώσεις απάτης έτσι και στην εισβολή ηλεκτρονικών υπολογιστών οι hackers προσαρμόζονται γρήγορα και μόλις ένα ορισμένο είδος εισβολής ανακαλυφθεί και εμποδίστεί πλέον η παραβίαση, τότε οι ίδιοι θα δοκιμάσουν ένα νέο είδος.

Φαινομενικά, η εισβολή σε κάποιο υπολογιστικό σύστημα φαντάζει δύσκολη. Όμως, η άποψη ότι απαιτούνται εξειδικευμένες γνώσεις για την εξαπόλυση τέτοιου είδους επίθεσης, αποτελεί μύθο. Στο διαδίκτυο διατίθενται ελεύθερα εφαρμογές λογισμικού, που επιτρέπουν στους επίδοξους hackers την εισβολή σε δίκτυα και υπολογιστικά συστήματα, τη διασπορά ιών και την πραγματοποίηση πλήθους άλλων ηλεκτρονικών επιθέσεων, καθιστώντας περισσότερο εύκολη την διάπραξη του ηλεκτρονικού εγκλήματος σε σχέση με το συμβατικό. Επιπλέον, το διαδίκτυο προσφέρει μια σειρά από νέες δυνατότητες επικοινωνίας. Το **ηλεκτρονικό ταχυδρομείο** (*e-mail*), τα **εικονικά δωμάτια συζητήσεων** (*chat rooms*) και οι **ομάδες ειδήσεων** (*newsgroups*), επιτρέπουν σε πολλά άτομα ταυτόχρονα να επικοινωνούν γρήγορα, σε πραγματικό χρόνο, χωρίς μετακίνηση, εύκολα και ανέξοδα.

Σύμφωνα με τον Parker (1983) υπάρχουν 4 τύποι εισβολής υπολογιστών:

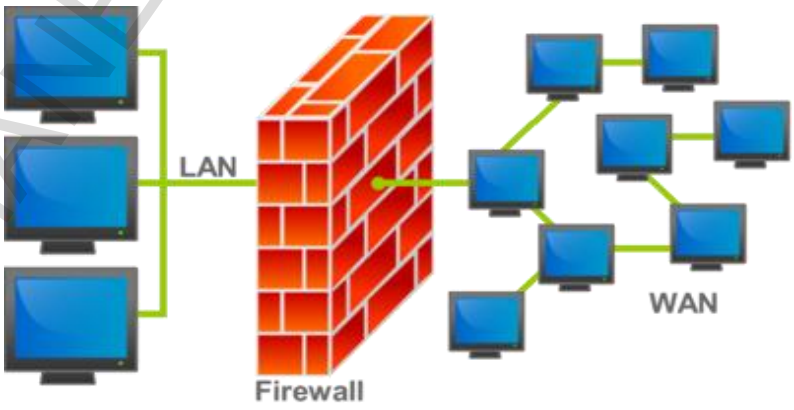
- Ο υπολογιστής μπορεί να αποτελέσει το αντικείμενο της επίθεσης. Είναι δυνατό να καταστραφούν τα πολύτιμα πράγματα και προγράμματα, που φιλοξενεί.
- Να χρησιμοποιηθεί ο υπολογιστής ως εργαλείο για την διάπραξη αδικημάτων (κλοπή, καταπάτηση, παραβίαση).
- Να αξιοποιηθεί συμβολικά ο υπολογιστής ώστε να συμβάλλει αποφασιστικά στην παραπλάνηση, την εξαπάτηση.
- Οι πληροφορίες σε ηλεκτρονική μορφή μπορούν αν αντιγραφούν, τροποποιηθούν, υπονομευθούν ή διαγραφούν, χωρίς οι αναγκαίες ενέργειες να αφήνουν πίσω τους κάποιο φυσικό ίχνος

Όπως μπορούμε να δούμε με βάση όσα αναφέρθηκαν παραπάνω, οι απάτες με τη χρήση ηλεκτρονικών συστημάτων είναι πολλές και γι' αυτό οι μέθοδοι ανίχνευσής τους είναι ιδιαίτερα σημαντικές για την διαφύλαξη των συστημάτων αυτών. Καταρχάς, για την προστασία του υπολογιστή μας και τη διαφύλαξη του δικτύου μας είναι σημαντικό η εγκατάσταση ενός τείχου προστασίας. Στην επιστήμη των υπολογιστών ο όρος «τείχος προστασίας» χρησιμοποιείται για να δηλώσει κάποια συσκευή ή πρόγραμμα που είναι έτσι ρυθμισμένο ώστε να επιτρέπει ή να απορρίπτει πακέτα δεδομένων που περνούν από ένα δίκτυο υπολογιστών σε ένα άλλο.

Η κύρια λειτουργία του είναι η ρύθμιση της κυκλοφορίας δεδομένων ανάμεσα σε δύο δίκτυα υπολογιστών. Συνήθως αυτά τα δύο δίκτυα είναι το διαδίκτυο και το τοπικό/εταιρικό δίκτυο. Ένα τείχος προστασίας παρεμβάλλεται ανάμεσα σε δύο δίκτυα που έχουν διαφορετικό επίπεδο εμπιστοσύνης. Το διαδίκτυο διαθέτει μικρό βαθμό εμπιστοσύνης ενώ το εταιρικό ή το οικιακό δίκτυο έχουν μεγάλο βαθμό εμπιστοσύνης.

Ο σκοπός της τοποθέτησης ενός τείχους προστασίας είναι η πρόληψη επιθέσεων και η αντιμετώπισή τους. Παρόλα αυτά μπορεί να αποδειχθεί άχρηστο εάν δεν ρυθμιστεί σωστά το τείχος προστασίας δρα ως φράχτης ανάμεσα στο διαδίκτυο και στο εσωτερικό δίκτυο ή έναν υπολογιστή και σταματάει διάφορους κινδύνους και επιθέσεις, συμπεριλαμβανομένων και ορισμένων ιών. Τα τείχη προσασίας φιλτράρουν την πληροφορία που εισέρχεται στο δίκτυο ή εξέρχεται από αυτό ούτως ώστε να προστατεύεται το δίκτυο από εισβολείς (π.χ hackers, ορισμένους ιούς κλπ.)

Σύμφωνα με τα παραπάνω, το τείχος προστασίας είναι ένας μηχανισμός που χρησιμοποιείται για να ελέγχει την πρόσβαση από και προς το ιδιόκτητο δίκτυο με απώτερο σκοπό την προστασία του δικτύου. Λειτουργεί σαν μια πύλη από την οποία περνάει όλη η κίνηση δεδομένων από και προς το εξωτερικό δίκτυο (Εικόνα 3.3.α). Στην πύλη εξετάζεται και αποφασίζεται αν θα επιτραπεί ή όχι η διέλευση των δεδομένων, σύμφωνα με την πολιτική ασφάλειας που εφαρμόζει ο οργανισμός του συστήματος. Δεν αποτελεί απλώς ένα σύνολο συνιστωσών λογισμικού ή υλικού, αλλά την τεχνική έκφραση μιας συγκεκριμένης στρατηγικής προστασίας των πόρων ενός οργανισμού. Ωστόσο, αν και η χρήση ενός φράγματος ασφαλείας αποτελεί την πρώτη γραμμή άμυνας ενός υπολογιστή απέναντι στους επίδοξους εισβολείς, δεν αποτελεί πανάκεια για την ασφάλεια του δικτύου. Όπως όλα τα συστήματα ασφάλειας μπορεί εξίσου να παραβιαστεί από κάποιον ικανό εισβολέα. Για την παρεμπόδιση αυτών των εισβολέων εφαρμόζονται διάφορες τεχνικές.




**Εικόνα 3.3.α.** Μία απεικόνιση για την τοποθέτηση ενός τείχου προστασίας στο δίκτυο

Όπως συμβαίνει και στις άλλες μορφές απάτης τις οποίες περιγράψαμε προηγουμένως, μπορούμε να δούμε ότι και στις απάτες εισβολής σε ηλεκρονικά συστήματα εφαρμόστηκαν διάφορες τεχνικές εξόρυξης δεδομένων για την ανίχνευση εισβολής, οι τεχνικές αυτές έχουν το πλεονέκτημα να ανακαλύπτουν τη γνώση εκείνη η οποία είναι χρήσιμη για την περιγραφή της συμπεριφοράς ενός χρήστη ή ενός προγράμματος μέσα από μεγάλα σύνολα δεδομένων ελέγχου. Στατιστικές τεχνικές, τεχνητά νευρωνικά δίκτυα, κοντινότερος γείτονας, τα συστήματα που βασίζονται στην ανίχνευση των ακραίων τιμών και τα κρυμμένα μοντέλα Markov είναι μερικές από τις τεχνικές εξόρυξης δεδομένων που χρησιμοποιούνται ευρέως τόσο για την ανίχνευση εισβολής κακής χρήσης όσο και για την ανίχνευση εισβολής ανωμαλιών.

Η ανίχνευση εισβολής αντιστοιχεί με κατάλληλες τεχνικές οι οποίες χρησιμοποιούνται για να προσδιορίσουν τις επιθέσεις που γίνονται εναντίον των ηλεκτρονικών υπολογιστών καθώς επίσης και των δικτύων. Η ανίχνευση ανωμαλιών αποτελεί το βασικό στοιχείο για την ανίχνευση εισβολής στην οποία διαταρραχές στην φυσιολογική συμπεριφορά των χρηστών υποδηλώνουν την παρουσία εσκεμμένων ή όχι επιθέσεων, ελαττωμάτων, σφαλμάτων, κλπ. Στη συνέχεια της ενότητας αυτής θα περιγράψουμε λεπτομερώς τη συγκριτική μελέτη που διεξήχθει με τη χρήση διάφορων τεχνικών ανίχνευσης ανωμαλιών για την αναγνώριση διαφορετικών εισβολών στο δίκτυο.

Τα δεδομένα που χρησιμοποιήθηκαν για τις αναλύσεις μας βασιζόνταν στο σύνολο δεδομένων συνδέσεων δικτύου του DARPA '98 Intrusion Detection Evaluation Data καθώς επίσης και σε πραγματικά δεδομένα δικτύου του Πανεπιστημίου της Μινεσότας τα οποία χρησιμοποίησαν τις υπάρχουσες τεχνικές αξιολόγησης καθώς και τη χρήση ειδικών συστημάτων μέτρησης οι οποίες είναι κατάλληλες στον εντοπισμό επιθέσεων που αφορούν ένα μεγάλο αριθμό συνδέσεων. Σύμφωνα με τα αποτελέσματα που πήραν από τις αναλύσεις τους διαπίστωσαν ότι ορισμένα από τα συστήματα ανίχνευσης ανωμαλιών τα οποία εφάρμοσαν, κάποια από αυτά ήταν αρκετά ελπιδοφόρα όσον αφορά την ανίχνευση νέων εισβολών τόσο για τα δεδομένα του DARPA '98 όσο και για τα πραγματικά δεδομένα δικτύου. Συγκεκριμένα οι πιο επιτυχημένες τεχνικές που εφαρμόστηκαν για την ανίχνευση ανωμαλιών ήταν σε θέση να επιτύχουν ένα ποσοστό ανίχνευσης κοντά στο 74% για τις πολλαπλές συνδέσεις και ένα ποσοστό γύρω στο 54% για απλές συνδέσεις δικτύων, διατηρώντας παράλληλα το ποσοστό των εσφαλμένων συναγερμών χαμηλά στο 2%. Όταν όμως το ποσοστό αυτό των ψευδών συναγερμών αυξανόταν γύρω στο 4%, τότε το ποσοστό ανίχνευσης έφτανε στο 89% για τις επιθέσεις ξεσπάσματος και το τέλειο ποσοστό του 100% για τις επιθέσεις απλών συνδέσεων. Παρακάτω θα περιγράψουμε περισσότερο αναλυτικά τις τεχνικές αυτές οι οποίες χρησιμοποιήθηκαν για την ανίχνευση εισβολής ανωμαλιών.

*3.3.i Τεχνικές ανίχνευσης ανωμαλιών*

Στο παράδειγμα που χρησιμοποίησαν στην έρευνά τους οι Lazarevic *et al.* (2003) εστιάστηκαν στην εφαρμογή αλγορίθμων διάφορων τεχνικών οι οποίοι βασίζονταν στην ανίχνευση ακραίων τιμών καθώς επίσης στη χρήση αλγορίθμων μη εποπτευόμενων μηχανών στήριξης διανυσμάτων για την ανίχνευση εισβολών δικτύου. Στις επόμενες παραγράφους θα αναφέρουμε συνοπτικά τους αλγόριθμους που χρησιμοποιήθηκαν στην έρευνα την οποία έκαναν και θα παρουσιάσουμε τα αποτελέσματά στα οποία κατέληξαν.

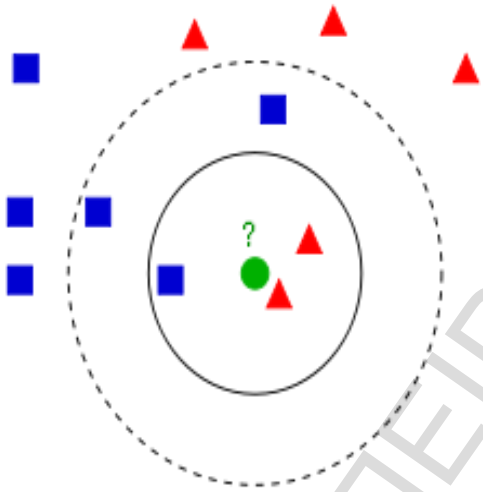
*3.3.ii. Τεχνικές μελέτης ακραίων τιμών για την ανίχνευση ανωμαλιών*

Οι περισσότεροι αλγόριθμοι ανίχνευσης ανωμαλιών απαιτούν ένα σύνολο από όμοια δεδομένα για την εφαρμογή του μοντέλου, και εμμέσως υποθέτουν ότι οι ανωμαλίες μπορούν να αντιμετωπιστούν ως πρότυπα που δεν παρατηρήθηκαν στο παρελθόν. Δεδομένου ότι μια ακραία τιμή μπορεί να οριστεί ως ένα σημείο δεδομένων το οποίο διαφέρει πολύ από τα υπόλοιπα σημεία του συνόλου δεδομένων, χρησιμοποιούνται διάφορα συστήματα ανίχνευσης ακραίων τιμών για να δούμε πόσο αποτελεσματικά τα συστήματα αυτά μπορούν να χειρίστούν το πρόβλημα της ανίχνευσης ανωμαλιών. Ωστόσο, με την αύξηση των διαστάσεων, γίνεται όλο και πιο δύσκολη και ανακριβή η εκτίμηση των πολυδιάστατων κατανομών των σημείων δεδομένων. Εντούτοις οι αλγόριθμοι ανίχνευσης ακραίων τιμών που χρησιμοποιήθηκαν στην έρευνα βασίστηκαν στον υπολογισμό της πλήρους διάστασης των αποστάσεων μεταξύ των σημείων, καθώς και στον υπολογισμό των πυκνοτήτων των τοπικών γειτονιών.

*α) Εύρεση ακραίων τιμών με τη χρήση της απόστασης k-κοντινότερων γειτόνων (Ramaswamy et al. 2000)*

Η προσέγγιση αυτή βασίζεται στον υπολογισμό της Ευκλείδιας απόστασης του *k*-κοντινότερου γείτονα από ένα σημείο Ο. Για ένα δεδομένο *k* και ένα σημείο Ο, το $D^k(O)$ απεικονίζει την απόσταση του σημείου Ο στον *k*- κοντινότερο γείτονα. Επιπλέον η απόσταση αυτή μπορεί να χρησιμοποιηθεί ως μέτρο για τον υπολογισμό ακραίων τιμών του σημείου Ο. Για παράδειγμα, σημεία με μεγαλύτερες τιμές των αποστάσεων $D^k(O)$ και έχοντας επομένως πιο αραιές κλάσεις αντιπροσωπεύουν συνήθως ισχυρότερες ακραίες τιμές από τα σημεία εκείνα που ανήκουν σε πυκνές συστάδες και τα οποία συνήθως έχουν χαμηλότερες $D^k(O)$ τιμές. Δεδομένου ότι μία γενική χρήση της μεθόδου αυτή βασίζεται στην επιλογή των n ακραίων τιμών, η προσέγγιση αυτή προσδιορίζει μια ακραία τιμή, όπως ακολούθως: δεδομένου της τιμής *k* και του

*n*, ένα σημείο Ο μπορεί να οριστεί ως μία ακραία τιμή αν η απόσταση από του από τον *k*-πλησιέστερο γείτονα είναι μικρότερη από (*n*-1) και οχι μεγαλύτερη από τα άλλα σημεία. Με άλλα λόγια οι *n* κορυφές μαζί με τις μέγιστες τιμές $D^k(O)$ θεωρούνται ως ακραίες τιμές.




**Εικόνα 3.3.ii.α.** Σχηματική απεικόνιση *k* -κοντινότερου γείτονα

*β) Προσέγγιση του Κοντινότερου Γείτονα*

Η μέθοδος αυτή είναι παρόμοια με τη μέθοδο του *k*-κοντινότερου γείτονα που αναφέραμε παραπάνω με τη διαφορά ότι εδώ το *k*=1. Καθορίζοντας ένα "κατώφλι ακραίων τιμών" χρησιμεύει στο να καθορίσουν αν το σημείο αποτελεί μια ακραία τιμή ή όχι. Το κατώφλι αυτό βασίζεται μόνο στα δεδομένα εκπαίδευσης και έχει οριστεί 2%. Για να υπολογίστεί το κατώφλι, για όλα τα σημεία των δεδομένων από τα δεδομένα εκπαίδευσης (π.χ "κανονική συμπεριφορά" δεδομένων) υπολογίζονται οι αποστάσεις των κοντινότερων γειτόνων και στη συνέχεια ταξινομόνται. Όλα τα σημεία του συνόλου δεδομένων που οι αποστάσεις των κοντινότερων γείτονων τους είναι μεγαλύτερες από το κατώτατο όριο θεωρούνται ακραίες τιμές.

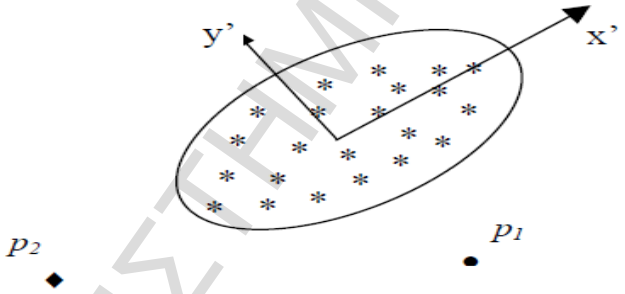
*γ) Ανίχνευση ακραίων τιμών βασισμένη στην απόσταση Mahalanobis*

Δεδομένου ότι τα **δεδομένα εκπαίδευσης** (*training data*) αντιστοιχούν στη "φυσιολογική συμπεριφορά", είναι εύκολο να υπολογιστεί η μέση τιμή και η τυπική απόκλιση των "κανονικών" δεδομένων. Η απόσταση Mahalanobis μεταξύ ενός συγκεκριμένου σημείου $p$ και του μέσου $\mu$ από το κανονικά δεδομένα υπολογίζεται ως εξής:

$$d_M = \sqrt{(p-\mu)^T \cdot \Sigma^{-1} \cdot (p-\mu)}\,,$$

όπου το $\Sigma$ είναι ο πίνακας συνδιακύμανσης των "κανονικών" δεδομένων. Ομοίως με την προηγούμενη προσέγγιση, το όριο υπολογίζεται σύμφωνα με τα πλέον απομακρυσμένα σημεία από τη μέση τιμή των "κανονικών δεδομένων και έχει οριστεί 2% του συνολικού αριθμού των σημείων. Επομένως τα σημεία δεδομένων των οποίων η απόσταση του μέσου των "κανονικών" δεδομένων εκπαίδευσης είναι μεγαλύτερη από το κατώφλι τότε θεωροωρούνται ακραίες τιμές.

Υπολογιζοντας τις αποστάσεις με τη χρήση του μέτρου της Ευκλείδειας απόστασης δεν είναι πάντα επωφελής, ιδίως όταν η κατανομή των δεδομένων είναι παρόμοια με εκείνη που παρουσιάζεται στην Εικόνα 3.3.2.α. Τα σημεία $p_1$ και $p_2$ δεν απέχουν την ίδια απόσταση από το μέσο της κατανομής, όταν υπολογίζονται οι αποστάσεις τους με βάση την Ευκλείδια απόσταση και την απόσταση Mahalanobis. Όταν χρησιμοποιείται το μέτρο της Ευκλείδιας απόστασης, η απόσταση μεταξύ του $p_2$ και του πλησιέστερου γείτονα είναι μεγαλύτερη από την απόσταση του $p_1$ και του κοντινότερου γείτονά του. Αντιθέτως όταν χρησιμοποιείται ως μέτρο η απόσταση Mahalanobis, τότε αυτές οι δύο αποστάσεις είναι ίσες. Είναι προφανές ότι με βάση αυτά τα δύο σενάρια, η προσέγγιση με βάση την απόσταση Mahalanobis είναι επωφελής σε σχέση με τη χρήση της Ευκλείδειας απόστασης.




Εικόνα 3.3.ii.β Πλεονέκτημα της προσέγγισης που βασίζεται στον υπολογισμό της απόστασης Mahalanobis. Πηγή: Lazarevic A. *et al.* (2003)

*δ) Προσέγγιση ανίχνευσης τοπικών ακραίων τιμών βασιζόμενες στην πυκνότητα*

Η βασική ιδέα αυτής της μεθόδου είναι να οριστεί σε κάθε αντικείμενο στο σύνολο δεδομένων ένας βαθμός ο οποίος καλείται **τοπικός παράγοντας έκτροπης** (*local outlier factor (LOF)*) χαρακτηρίζοντάς το καθένα ως ακραία τιμή. Ο συντελεστής αυτός όπως μπορούμε να δούμε και από τον τίτλο που φέρει, βασίζεται στην έννοια της τοπικής πυκνότητας, όπου η τοπικότητα εξαρτάται από το πόσο απομακρυσμένο είναι το αντικείμενο σε σχέση με τη γύρω περιοχή. Συγκρίνοντας την τοπική πυκνότητα ενός αντικειμένου με τις τοπικές πυκνότητες των γειτόνων της, μπορεί κανείς να εντοπίσει περιοχές παρόμοιας πυκνότητας, και σημεία τα οποία έχουν ουσιαστικά χαμηλότερη πυκνότητα από τους γείτονές τους. Αυτά θεωρούνται ακραίες

τιμές. Ο συντελεστής μπορεί να χρησιμοποιηθεί για να βρεθούν ακραίες τιμές οι οποίες φαίνονται να είναι σημαντικές αλλά δεν μπορούσαν να βρεθούν με άλλες προσεγίσεις. Ο αλγόριθμος που χρησιμοποιείται για τον υπολογισμό του *LOF* έχει διάφορα βήματα τα οποία είναι τα ακόλουθα (τα *O, p* που χρησιμοποιούνται παρακάτω αποτελούν αντικείμενα σε κάποιο σύνολο δεδομένων):

a) Αρχικά για ένα αντικείμενο *O* υπολογίζουμε *την k-distance*(*O*) η οποία ας υποθέσουμε ότι είναι η απόσταση *k*-κοντινότερου γείτονα. Δεδομένης της απόστασης αυτής υπολογίζουμε και την απόσταση της γειτονικής περιοχής του αντικειμένου (*k-distance neighborhood*) *O*, η οποία περιέχει κάθε αντικείμενο του οποίου η απόσταση από το *O* δεν είναι μεγαλύτερη από την *k-distance*, και τη συμβολίζουμε με $N_k(O)$

b) Στη συνέχεια υπολογίζουμε την **απόσταση προσσεγγισιμότητας** (*reachability distance*) για κάθε σημείο *O* του παραδείγματος σε σχέση με το σημείο *p*. Πρόκειται για ένα πρόσθετο μέτρο το οποίο χρησιμοποιείται για την παραγωγή πιο σταθερών αποτελεσμάτων μέσα στις συστάδες και υπολογίζεται ως εξής: $reach - dist_k(O,p)$ = max{*k-distance(p), d(O,p)*},όπου το *d(O,p)* είναι η απόσταση ανάμεσα στα δύο σημεία *O* και *p*,

c) Έπειτα υπολογίζουμε την **τοπική πυκνότητα προσσεγγισιμότητας** (*local reachability density*) του αντικειμένου *O* η οποία ορίζεται ως η αντίστροφος της μέσης απόστασης προσσεγγισιμότητας βασιζόμενη στον *MinPts* (η παράμετρος αυτή καθορίζει τον ελάχιστο αριθμό αντικειμένων) κοντινότερο γείτονα των *O* δεδομένων παραδείγματος,

$$lrd_{MinPts}(O) = 1/\left(\frac{\sum_{p \in N_{MinPts}(O)} reach - dist_{MinPts}(O,p)}{|N_{MinPts}(O)|}\right)$$

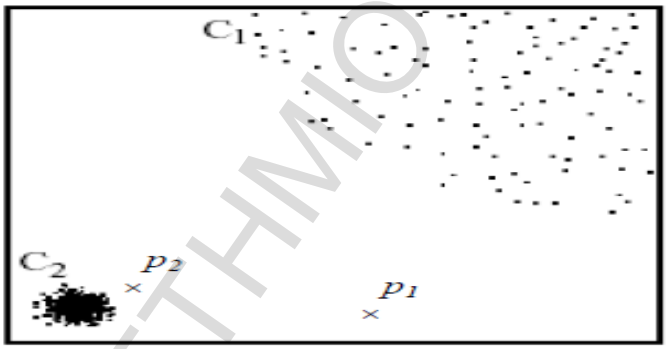
d) Τέλος υπολογίζουμε τον τοπικό παράγοντα έκτροπης (*LOF)* του αντικειμένου *O* ως τον μέσο όρο του λόγου των τοπικών πυκνοτήτων της προσσεγγισιμότητας του αντικειμένου *O* και των *O MinPts-nearest neighbors*. Πιο συγκεκριμένα ο *LOF* δίνεται από τον τύπο:

$$LOF_{MinPts}(O) = \frac{\sum_{p \in N_{MinPts}(O)} \frac{lrd_{MinPts}(p)}{lrd_{MinPts}(O)}}{|N_{MinPts}(O)|}$$

Επομένως, όπως μπορούμε να δούμε, όσο χαμηλότερη είναι η τοπική πυκνότητα προσσεγγισιμότητας του αντικειμένου *O* και όσο μεγαλύτερες είναι οι τοπικές πυκνότητες προσσεγγισιμότητας των *O MinPts-nearest neighbors* τόσο ηψηλότερη είναι η τιμή του *LOF* για το αντικείμενο *O*.

Για την ανάδειξη των πλεονεκτημάτων της προσέγγισης LOF, θεωρούμε ένα απλό σύνολο διδιάστατων δεδομένων το οποίο δίνεται στην Εικόνα 3.3.2.γ. Όπως μπορούμε να δούμε από το σχήμα υπάρχει πολύ μεγαλύτερος αριθμός αντικειμένων στην κλάση $C_1$ από οτι στην κλάση $C_2$,

και η πυκνότητα της κλάσης $C_2$ είναι σημαντικά υψηλότερη από αυτή της κλάσης $C_1$. Συγκεκριμένα το σύνολο των διδιάστατων δεδομένων μας αποτελείται από 502 αντικείμενα όπου τα 400 αντιστοιχούν στην κλάση $C_1$ και τα 100 στην κλάση $C_2$ και έχουμε και άλλα 2 πρόσθετα αντικείμενα τα $p_1$ και $p_2$. Λόγω της χαμηλής πυκνότητας της κλάσης $C_1$ είναι εμφανές ότι για κάθε αντικείμενο $q$ μέσα στη συστάδα $C_1$, η απόσταση μεταξύ του σημείου $q$ και του πλησιέστερου γείτονά του είναι μεγαλύτερη από την απόσταση μεταξύ του $p_2$ και του πλησιέστερου γείτονα από την κλάση $C_2$ και έτσι το σημείο $p_2$ δεν θεωρείται ακραία τιμή. Συνεπώς, η απλή προσέγγιση του κοντινότερου γείτονα βασιζόμενη στον υπολογισμό των αποστάσεων αποτυγχάνει σε δεδομένα αυτής της μορφής. Ωστόσο, το $p_1$ μπορεί να ανιχνευθεί ως ακραία χρησιμοποιώντας μόνο τις αποστάσεις του πλησιέστερου γείτονα. Από την άλλη πλευρά, ο συντελεστής τοπικών ακραίων τιμών είναι σε θέση να συλλάβει και τα δύο αντικείμενα ( $p_1$ και $p_2$) ως ακραίες τιμές που οφείλεται στην πυκνότητα των γειτονικών περιοχών γύρω από τα σημεία.




**Εικόνα 3.3.ii.γ** Πλεονεκτήματα της προσέγγισης LOF.
Πηγή: Breunig M. *et al.* (2000)

*3.3.iii Μη εποπτευόμενες τεχνικές υποστήριξης διανυσμάτων*

Σε αντίθεση με τις επιβλεπόμενες μηχανές στήριξης διανυσμάτων οι οποίες απαιτούν δεδομένα εκπαίδευσης που φέρουν κάποια σήμανση ώστε να δημιουργήσουν τον κανόνα ταξινομησης, ο αλγόριθμος των SVM προσαρμόστηκε σε μη επιβλεπόμενους αλγορίθμους εκμάθησης. Αυτή η μη επιβλεπόμενη μετατροπή δεν απαιτεί τα δεδομένα εκπαίδευσης να φέρουν κάποια σήμανση για να καθορίσουν μία απόφαση. Ενώ ο αλγόριθμος των εποπτευόμενων SVM προσπαθεί να διαχωρίσει κατά το μέγιστο δυνατό τις δύο κατηγορίες δεδομένων σε έναν πολυδιάστατο χώρο χαρακτηριστικών με ένα υπερεπίπεδο, ο αλγόριθμος των μη επιβλεπόμενων μεθόδων προσπαθεί να διαχωρίσει ολόκληρο το σύνολο των δεδομένων εκπαίδευσης από το

αρχικό, δηλαδή να βρεί μια μικρή περιοχή, όπου τα περισσότερα από τα δεδομένα και τα δεδομένα με σήμανση που βρίσκονται σε αυτή την περιοχή ανήκουν σε μια κατηγορία.

Με τη χρήση διαφορετικών τιμών για τις παραμέτρους SVM (παράμετρος διακύμανσης των **ακτινικών συναρτήσεων βάσης** (*radial basis functions* (RBFS)), αναμενόμενος ρυθμός ακραίων τιμών), μπορούν να καταστκευαστούν μοντέλα με διαφορετική πολυπλοκότητα. Για τους πυρήνες RBF με μικρότερη διακύμανση, ο αριθμός των διανυσμάτων υποστήριξης είναι μεγαλύτερος και τα όρια απόφασης είναι πιο πολύπλοκα, οδηγώντας έτσι σε ένα πολύ υψηλό ποσοστό ανίχνευσης αλλά παράλληλα σε ένα υψηλό ποσοστό εσφαλμένων συναγερμών. Από την άλλη πλευρά, λαμβάνοντας υπόψη τους πυρήνες RBF με τη μεγαλύτερη διακύμανση, ο αριθμός των διανυσμάτων υποστήριξης μειώνεται, ενώ τα όρια των περιοχών γίνονται πιο γενικά, με αποτέλεσμα να έχουμε χαμηλότερο ποσοστό ανίχνευσης αλλά και χαμηλότερο ποσοστό εσφαλμένων συναγερμών.

*3.3.iv Πειραματικά αποτελέσματα για το σύνολο δεδομένων του DARPA'98*

Προκειμένου να αξιολογηθεί η απόδοση των αλγορίθμων ανίχνευσης ανωμαλιών σε ένα πραγματικό περιβάλλον, θα παρουσιάσουμε στη συνέχεια της ενότητας αυτής τα αποτελέσματα της αξιολόγησης της εφαρμογής των τεχνικών που αναφέραμε παραπάνω χρησιμοποιώντας το σύνολο δεδομένων DARPA'98 Intrusion Detection Evaluation Data . Το σύνολο δεδομένων που θα ασχοληθούμε περιλαμβάνουν 4 κατηγορίες επιθέσεων οι οποίες είναι οι:

- DoS (Denial of Service) όπως για παράδειγμα ping-of-death, teardrop, smurf, SYN flood, κλπ
- R2L όπου πρόκειται για μη εξουσιοδοτημένη πρόσβαση από ένα απομακρυσμένο μηχάνημα για παράδειγμα, να μαντέψουν τον κωδικό
- U2R όπου πρόκειται μη εξουσιοδοτημένη πρόσβαση σε τοπικά προνόμια υπερχρήστη από έναν τοπικό μη προνομιούχο χρήστη
- PROBING[6] όπου πρόκειται για παρακολούθηση και διερεύνηση όπως για παράδειγμα port-scan, ping-sweep, κλπ.

Επιπλέον τα πρότυπα μέτρησης που χρησιμοποιούνται για την αξιολόγηση των συστημάτων που προτείνονται στο συγκεκριμένο παράδειγμα βασίζονται στο **ποσοστό ανίχνευσης** (*detection rate*) και στο ποσοστό εσφαλμένων συναγερμών και παρουσιάζονται συνοπτικά στον Πίνακα 3.3.iv.α. Το ποσοστό ανίχνευσης υπολογίζεται ως ο λόγος μεταξύ του αριθμού των σωστών εντοπισθέντων επιθέσεων προς τον συνολικό αριθμό των επιθέσεων, ενώ

[6] Περισσότερες πληροφορίες για τα είδη αυτά επιθέσεων αναφέρουν στην έρευνά τους οι Lazarevic *et al.* (2003)

το ποσοστό εσφαλμένων συναγερμών υπολογίζεται ως ο λόγος μεταξύ του αριθμού των κανονικών συνδέσεων που λανθασμένα έχουν ταξινομηθεί ως επιθέσεις προς το συνολικό αριθμό των κανονικών συνδέσεων.

| **Πρότυπα μετρήσεων** (*Standard metrics*) | | **Προβλεπόμενη ετικέτα σύνδεσης** (*Predicted connection label*) | |
|---|---|---|---|
| | | **Κανονική** (*Normal*) | **Εισβολές** (*Intrusions*) |
| **Ετικέτα πραγματικής σύνδεσης** (*Actual connection Label*) | **Κανονική** (*Normal*) | **Αληθώς αρνητικά** (*True negatives*) | **Εσφαλμένος συναγερμός** *(False alarm)* |
| | **Εισβολές** (*Intrusions*) | **Ψευδώς αρνητικά** (*False negatives*) | **Σωστός εντοπισμός επιθέσεων** (*Correctly detected attacks*) |

**Πίνακας 3.3.iv.α** Πρότυπα μετρήσεων για την αξιολόγηση της εισβολής επιθέσεων

Υπάρχουν γενικά δύο τύποι επιθέσεων βάση των οποίων στηριζόμαστε για την ανίχνευση εισβολής σε δίκτυα υπολογιστών: οι επιθέσεις που περιλαμβάνουν απλές συνδέσεις και οι επιθέσεις που περιλαμβάνουν πολλαπλές συνδέσεις (επιθέσεις ξεσπάσματος). Τα πρότυπα-μέτρησεις που αναφέραμε προηγουμένως αντιμετωπίζουν όλους τους τύπους των επιθέσεων παρομοίως αποτυγχάνοντας έτσι να παρέχουν μία επαρκή αξιολόγηση για τις επιθέσεις που περιλαμβάνουν πολλές συνδέσεις δικτύου. Ειδικότερα δεν λαμβάνουν πληροφορίες σχετικά με τον αριθμό των συνδέσεων δικτύων που σχετίζονται με μία επίθεση η οποία έχει ανιχνευθεί σωστά. Συνεπώς, ανάλογα με τον τύπο της επίθεσης, υπάρχουν δύο τύποι ανάλυσης που μπορούν να εφαρμοστούν: η ανάλυση επίθεσης πολλαπλών συνδέσεων για επιθέσεις ξεσπάσματος και η ανάλυση επίθεσης απλών συνδέσεων. Ωστόσο, το πρώτο βήμα για και για τους δύο τύπους ανάλυσης αφορά τον υπολογισμό της τιμής για κάθε σύνδεση δικτύου. Η τιμή αυτή αντιπροσωπεύει την πιθανότητα ότι η συγκεκριμένη σύνδεση δίκτυου που εξετάζεται σχετίζεται με μια εισβολή.

Δεδομένου ότι το σύνολο των διαθέσιμων δεδομένων είναι τεράστιο, δοκιμάστηκαν ακολουθίες κανονικών εγγραφών συνδέσεων προκειμένου να δημιουργηθεί το κανονικό σύνολο δεδομένων το οποίο είχε την ίδια κατανομή με τα αρχικά δεδομένα των κανονικών συνδέσεων. Στη συνέχεια χρησιμοποιήθηκε το κανονικό σύνολο δεδομένων για την εφαρμογή των συστημάτων ανίχνευσης ανωμαλιών και έπειτα εξετάστηκε πόσο καλά μπορούν να ανιχνευθούν οι επιθέσεις χρησιμοποιώντας τα προτεινόμενα συστήματα.

Συγκεκριμένα στις αναλύσεις που έκαναν χρησιμοποίησαν τις TCP συνδέσεις από τις 5 εβδομάδες των δεδομένων εκπαίδευσης από τις οποίες εξετάστηκε ένα δείγμα 5000 εγγραφών δεδομένων οι οποίες αποτελούσαν κανονικές συνδέσεις. Για λόγους ελέγχου χρησιμοποιήθηκαν συνδέσεις που σχετίζονταν με επιθέσεις από τα δεδομένα των πρώτων 5 εβδομάδων με σκοπό να προσδιοριστεί το ποσοστό ανίχνευσης. Επιπλέον έλαβαν υπόψην ένα τυχαίο δείγμα 1000 εγγραφών συνδέσεων οι οποίες βασίζονταν σε κανονικά δεδομένα με σκοπό να ανιχνεύσουν το ποσοστό των αρνητικά εσφαλμένων. Είναι σημαντικό να τονίσουμε ότι το δείγμα αυτό που χρησιμοποιείται για λόγους ελέγχου έχει την ίδια κατανομή με το αρχικό σύνολο κανονικών συνδέσεων.

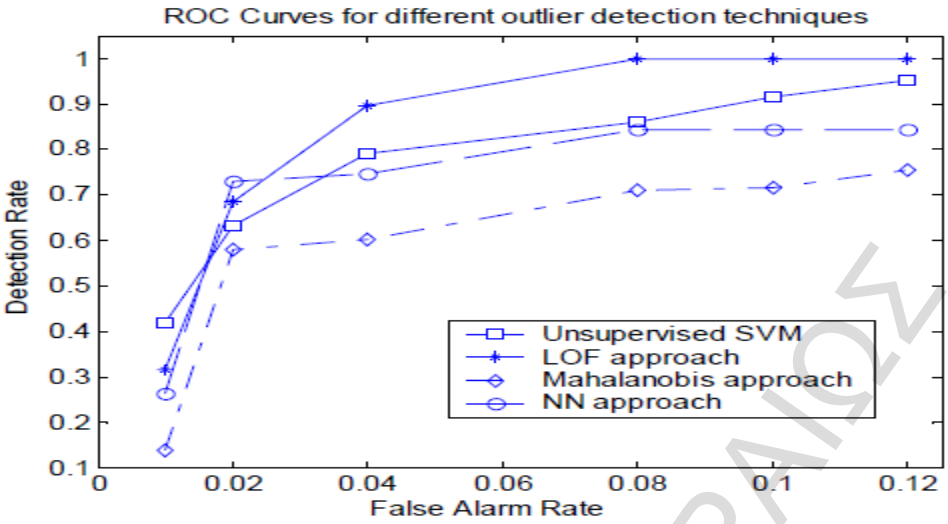
Η λειτουργία της ανίχνευσης ανωμαλιών για τα συστήματα που αναφέραμε παραπάνω εξετάστηκε ξεχωριστά για μεικτές επιθέσεις ξεσπάσματος καθώς επίσης και για επιθέσεις ξεσπάσματος ή μη. Σε όλα τα πειράματα που έγιναν το ποσοστό των ακραίων τιμών στα δεδομένα εκπαίδευσης κυμαινόταν από 1% έως 12%. Είναι ενδιαφέρον να τονίσουμε ότι το μεγαλύτερο ποσοστό εσφαλμένων συναγερμών δημιουργήθηκε κατά την ανίχνευση κανονικών συνδέσεων από το σύνολο δεδομένων.

Τα πειράματα αρχικά διεξήχθησαν στα ξεσπάσματα επιθέσεων και τα αποκτηθέντα ποσοστά ανίχνευσης επιθέσεων για τα τέσσερα συστήματα ανίχνευσης ανωμαλιών παρουσιάζονται στον ακόλουθο πίνακα (Πίνακας 3.3.iv.β.). Θεωρούμε ότι ένα ξέσπασμα ανιχνεύεται, αν το αντίστοιχο ποσοστό ανίχνευσης ξεσπάσματος είναι πάνω από 50%. Όπως μπορούμε να δούμε στον πίνακα (όπου εξετάζονται συνολικά 19 επιθέσεις ξεσπάσματος) τα πιο δημοφιλή συστήματα ανίχνευσης ακραίων τιμών ήταν αυτά των κοντινότερων γειτόνων και των προσεγγίσεων LOF. Η προσέγγιση των κοντινότερων γειτόνων ήταν ικανή να ανιχνεύσει 14 ξεσπάσματα επιθέσεων ενώ η δεύτερη προσέγγιση ανίχνευσε μέχρι 13 επιθέσεις. Όσον αφορά τώρα τις άλλες δύο προσεγγίσεις που εφαρμόστηκαν βλέπουμε ότι οι μη εποπτευόμενες μέθοδοι υποστήριξης διανυσμάτων ήταν λίγο χειρότερες από τις δύο προηγούμενες προσεγγίσεις δείχνοντας μία σημαντική ικανότητα στην ανίχνευση εισβολής δικτύων, ενώ η προσέγγιση της απόστασης *Mahalanobis* ήταν ικανή να ανιχνεύσει μόνο 11 πολλαπλές συνδέσεις επιθέσεων. Η χαμηλή απόδοση της προσέγγισης αυτής οφειλόταν στο γεγονός ότι η κανονική συμπεριφορά μπορεί να διαθέτει περισσότερους τύπους και δεν μπορούσαν να χαρακτηριστούν με μία ενιαία κατανομή την οποία χρησιμοποιήσαμε στα πειράματά μας. Επομένως, προκειμένου να εξαλείψουμε αυτό το πρόβλημα υπάρχει η ανάγκη να γίνει διαμέριση της κανονικής συμπεριφοράς σε περισσότερες από μία παρόμοιες κατανομές και να ανιχνεύσουμε τις ανωμαλίες σύμφωνα με τις αποστάσεις *Mahalanobis* για κάθε μία από τις κατανομές.

| Burst position | burst length (# of connections) | Attack type and category | *LOF approach* | *NN approach* | *Mahalanobis-based approach* | *Unsupervised SVM approach* |
|---|---|---|---|---|---|---|
| Week1, burst1 | 15 | neptune (DOS) | 15 (100%) | 15 (100%) | 4 (26.7%) | 15 (100%) |
| Week2, burst1 | 50 | guest (U2R) | 49 (98%) | 49 (98%) | 49 (98%) | 48 (96%) |
| Week2, burst2 | 102 | portsweep (probe) | 31 (30.3%) | 63 (61.7%) | 25 (24.5%) | 48 (47.1%) |
| Week2, burst3 | 898 | ipsweep (probe) | 158 (17.6%) | 428 (47.7%) | 369 (41.1%) | 375 (41.8%) |
| Week2, burst4 | 1000 | back (DOS) | 752 (75.2%) | 62 (6.2%) | 44 (4.4%) | 825 (82.5%) |
| Week3, burst1 | 15 | satan (probe) | 0 (0%) | 0 (0%) | 0 (0%) | 1 (6.7%) |
| Week3, burst2 | 137 | portsweep (probe) | 15 (10.9%) | 118 (86.1%) | 84 (61.3%) | 115 (83.9%) |
| Week3, burst3 | 105 | nmap (probe) | 61 (58.1%) | 105 (100%) | 105 (100%) | 97 (92.4%) |
| Week3, burst4 | 1874 | nmap (probe) | 1060 (57%) | 1071 (57.1%) | 993 (53%) | 850 (45.4%) |
| Week3, burst5 | 5 | imap (r2l) | 4 (80%) | 5 (100%) | 4 (80%) | 5 (100%) |
| Week3, burst6 | 17 | warezmaster (u2r) | 16 (94.1%) | 15 (88.2%) | 15 (88.2%) | 16 (94.1%) |
| Week4, burst1 | 86 | warezclient (u2r) | 33 (38.4%) | 38 (44.2%) | 38 (44.2%) | 42 (48.8%) |
| Week4, burst2 | 6104 | satan (probe) | 5426 (89%) | 5558 (91.1%) | 5388 (88.3%) | 5645 (92.5%) |
| Week4, burst3 | 1322 | pod (DOS) | 957 (72.4%) | 969 (73.3%) | 680 (51.4%) | 645 (48.8%) |
| Week4, burst4 | 297 | portsweep (probe) | 221 (74.4%) | 259 (87.2%) | 230 (77.4%) | 271 (91.2%) |
| Week4, burst5 | 2304 | portsweep (probe) | 1764 (76.6%) | 1809 (79%) | 1095 (47.5%) | 1969 (85.5%) |
| Week5, burst1 | 3067 | satan (probe) | 2986 (97.4%) | 3022 (99%) | 2983 (97%) | 2981 (97.2%) |
| Week5, burst2 | 5 | ffb (r2l) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) |
| Week5, burst3 | 1021 | portsweep (probe) | 937 (92%) | 978 (98%) | 938 (92%) | 942 (92.3%) |
| Total | 18424 | - | 13/19 | 14/19 | 11/19 | 12/19 |
| Detection rate | | | 68.4% | 73.7% | 57.9% | 63.2% |

**Πίνακας 3.3.iv.β.** Ποσοστά ανίχνευσης για τα είδη επιθέσεων που εξετάστηκαν για κάθε μία από τις 4 προσεγγίσεις που εφαρμόστηκαν καθώς επίσης ο αριθμός των συνδέσεων από τις επιθέσεις ξεσπάσματος που επιτυχώς διαπιστώθηκαν ως επιθέσεις. Πηγή: Lazarevic A. *et al.* (2003

Για να γίνει η σύγκριση όλων των συστημάτων ανίχνευσης ακραίων τιμών, υπολογίστηκαν οι καμπύλες ROC για όλους τους προτεινόμενους αλγορίθμους. Αυτές απεικονίζονται στην Εικόνα 3.3.iv.γ., όπου μπορούμε να παρατηρήσουμε το ρυθμό ανίχνευσης καθώς επίσης και το ποσοστό εσφαλμένων συναγερμών τα οποία, όπως μπορούμε να δούμε, διαφέρουν όταν χρησιμοποιούνται διαφορετικά κατώτατα όρια. Σύμφωνα με τα αποτελέσματα του παρακάτω σχήματος είναι εμφανές ότι η προσέγγιση LOF όσον αφορά την ανίχνευση ανωμαλιών φέρει καλύτερα αποτελέσματα σε σχέση με τις υπόλοιπες προσεγγίσεις για μεγαλύτερα ποσοστά ποσοστά εσφαλμένων συναγερμών (μεγαλύτερο από 2%) ενώ η προσέγγιση αυτή είναι ελαφρώς χειρότερη από την προσέγγιση του κοντινότερου γείτονα για χαμηλότερα ποσοστά εσφαλμένων συναγερμών (1% και 2%).




**Εικόνα 3.3.iv.γ.** Καμπύλες ROC οι οποίες παρουσιάζουν την απόδοση των αλγορίθμων ανίχνευσης ανωμαλιών για τις επιθέσεις ξεσπάσματος. Πηγή: Lazarevic A. *et al.* (2003)

Στον Πίνακα 3.3.iv.δ. αναφέρονται πρόσθετοι μετρητές για την αξιολόγηση των ξεσπασμάτων επιθέσεων, πιο συγκεκριμένα το εμβαδόν επιφάνειας και ο χρόνος απόκρισης. Όσο μικρότερη είναι η επιφάνεια μεταξύ των δύο αυτών καμπυλών επίθεσης, τόσο καλύτερος είναι ο αλγόριθμος ανίχνευσης εισβολής. Ωστόσο, η ίδια η επιφάνεια δεν είναι επαρκής να συλλάβει πολλές σχετικές πτυχές των αλγορίθμων ανίχνευσης εισβολής (π.χ. πόσες συνδέσεις σχετίζονται με την επίθεση, πόσο γρήγορος είναι ο αλγόριθμος ανίχνευσης εισβολής, κλπ.). Επομένως, θα πρέπει να χρησιμοποιηθούν πρόσθετες μετρήσεις, προκειμένου να υποστηριχθεί η βασική μέτρηση της επιφάνειας κάτω από την καμπύλη επίθεσης. Υποθέτουμε ότι ο συνολικός αριθμός των συνδέσεων στο δίκτυο σύμφωνα με το σύνολο δεδομένων είναι $N$, ο οποίος ισούται με το άθροισμα του συνολικού αριθμού των κανονικών συνδέσεων δικτύου ($N_n$) και του συνολικού αριθμού συνδέσεων του δικτύου που σχετίζονται με τις εισβολές ($N_i$). Ο αριθμός ($n_{fa}$) αντιστοιχεί στον αριθμό των κανονικών συνδέσεων δικτύου οι οποίοι έχουν υψηλότερη τιμή από το προκαθορισμένο όριο (διακεκομμένη γραμμή στην Εικόνα 3.3.iv.δ.) και επομένως ταξινομήθηκαν εσφαλμένα ως δίκτυα εισβολής. Τώρα, οι επιπλέον μετρήσεις μπορεί να ορισθούν ως εξής

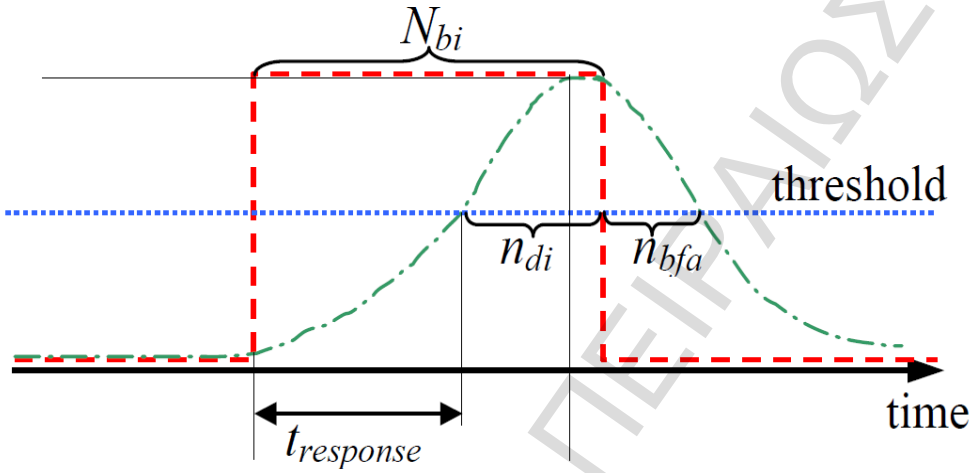
- **Ρυθμός ανίχνευσης ξεσπάσματος** (*Burst detection rate* (bdr)) ορίζεται για κάθε ξέσπασμα και αντιπροσωπεύει την αναλογία μεταξύ του συνολικού αριθμού των συνδέσεων δικτύου εισβολής $n_{di}$ που έχουν τιμή μεγαλύτερη από αυτή του κατωφλιού (διακεκομμένη γραμμή στην Εικόνα 3.3.iv.δ) προς τον συνολικό αριθμό των συνδέσεων δικτύου εισβολής μέσα στα χρονικά διαστήματα επίθεσης ($N_{bi}$). Πιο συγκεκριμένα:

$$bdr = n_{di} / N_{bi.},$$

όπου

$$\Sigma_{all\_bursts}\ N_{bi} = N_i$$

- **Ο χρόνος απόκρισης** (*response time*) αντιπροσωπεύει το χρονικό διάστημα που μεσολάβησε από την στιγμή που άρχισε η επίθεση μέχρι τη στιγμή όπου η πρώτη σύνδεση με το δίκτυο έχει τιμή υψηλότερη από ό, τι η προκαθορισμένη τιμή κατωφλίου.




| **Μετρητές** | **Ορισμός** |
|---|---|
| bdr | Ρυθμός ανίχνευσης ξεσπάσματος = $n_{di}$ / $N_{bi.}$ |
| $n_{di}$ | αριθμός των συνδέσεων δικτύου εισβολής $n_{di}$ που έχουν τιμή μεγαλύτερη από αυτή του κατωφλιού |
| $n_{bfa}$ | αριθμός των κανονικών συνδέσεων που ακολουθούν επίθεση και οι οποίες έχουν ταξινομηθεί εσφαλμένα ως παρεμβατικές |
| $t_{response}$ | Χρόνος απόκρισης - χρόνος για να φτάσει στο προκαθορισμένο κατώφλι |

**Εικόνα 3.3.iv.δ.** Πρόσθετοι μετρητές για την για την αξιολόγηση ανίχνευσης των επιθέσεων ξεσπάσματος. Πηγή: Lazarevic A. *et al.* (2003)

Είναι σημαντικό να σημειωθεί ότι το εμβαδόν επιφανείας του Πίνακα 3.3.iv.γ.[7] κανονικοποιήθηκε, έτσι ώστε το συνολικό εμβαδόν επιφανείας να διαιρείται με το συνολικό αριθμό των συνδέσεων από το αντίστοιχο ξεσπάσμα επίθεσης. Δεδομένου ότι, σε διαφορετικά χρονικά διαστήματα, εμπλέκονται διαφορετικές επιθέσεις έκρηξης αποφασίστηκε να μετρηθεί ο χρόνος απόκρισης καθώς και ο αριθμός των συνδέσεων. Λαμβάνοντας υπόψην τους επιπλέον μετρητές αξιολόγησης, υπολογίστηκε ο ρυθμός ανίχνευσης. Στον Πίνακα 3.3.iv.γ. θεωρούμε ότι έχει ανιχνευθεί μια επίθεση ξεσπάσματος εάν το κανονικοποιημένο εμβαδόν επιφάνειας είναι μικρότερο από 0.5. Σύμφωνα με τα αποτελέσματα που έχουμε βλέπουμε οι 2 πιο επιτυχημένοι

[7] Όπως φαίνεται στον ακόλουθο πίνακα (Πίνακα 3.3.iv.γ) υπάρχει ένα τυπογραφικό λάθος όσον αφορά το ποσοστό ανίχνευσης για την προσέγγιση των SVM από 84.2% θα έπρεπε να ήταν 63.2%

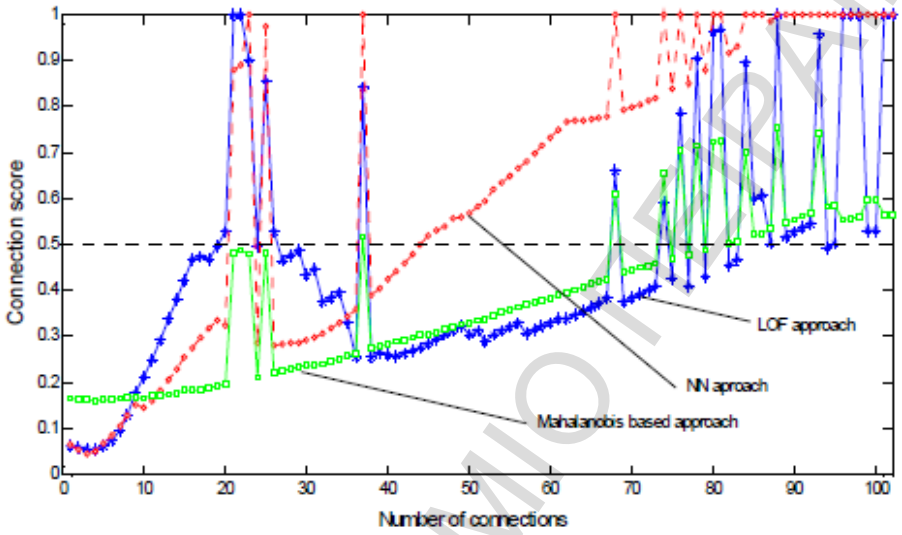
αλγόριθμοι ανίχνευσης εισβολής ήταν πάλι οι προσεγγίσεις των NN και LOF, σύμφωνα με τις οποίες ανιχνεύθηκαν 15 και 14 επιθέσεις αντιστοίχως. Κατά τη χρήση των προτεινόμενων πρόσθετων μετρήσεων, η προσέγγιση Mahalobis ήταν και πάλι κατώτερη από την NN προσέγγιση, ενώ από την άλλη πλευρά, η μη εποπτευόμενη προσέγγιση των SVM κατάφερε να επιτεύξει και πάλι ελαφρώς χειρότερο ποσοστό ανίχνευσης από αυτές των LOF και NN.

| Burst position (burst length) | Attack type and category | *LOF approach* | | *NN approach* | | *Mahalanobis-based* approach | | *Unsupervised SVM* | |
|---|---|---|---|---|---|---|---|---|---|
| | | SA | $t_{response}$ | SA | $t_{response}$ | SA | $t_{response}$ | SA | $t_{response}$ |
| Week1, burst1 | neptune (DOS) | 0.03 | 1 | 0.22 | 1 | 0.25 | 1 | 0.02 | 1 |
| Week2, burst1 | guest (u2r) | 0.22 | 1 | 0.01 | 1 | 0.03 | 1 | 0.04 | 1 |
| Week2, burst2 | portsweep (probe) | 0.5 | 20 | 0.38 | 21 | 0.54 | 37 | 0.23 | 15 |
| Week2, burst3 | ipsweep (probe) | 0.61 | 2 | 0.5 | 1 | 0.55 | 2 | 0.41 | 1 |
| Week2, burst4 | back (DOS) | 0.3 | 3 | 0.74 | 3 | 0.82 | 5 | 0.37 | 2 |
| Week3, burst1 | satan (probe) | 0.89 | - | 0.94 | - | 0.95 | - | 0.69 | 9 |
| Week3, burst2 | portsweep (probe) | 0.8 | 30 | 0.2 | 1 | 0.32 | 4 | 0.28 | 2 |
| Week3, burst3 | nmap (probe) | 0.3 | 2 | 0 | 1 | 0.1 | 3 | 0.09 | 2 |
| Week3, burst4 | nmap (probe) | 0.33 | 13 | 0.34 | 1 | 0.52 | 5 | 0.27 | 3 |
| Week3, burst5 | imap (r2l) | 0.14 | 2 | 0.0004 | 1 | 0.2 | 2 | 0.03 | 1 |
| Week3, burst6 | warezmaster (u2r) | 0.08 | 1 | 0.12 | 1 | 0.15 | 1 | 0.07 | 1 |
| Week4, burst1 | warezclient (u2r) | 0.56 | 1 | 0.58 | 1 | 0.69 | 2 | 0.52 | 1 |
| Week4, burst2 | satan (probe) | 0.12 | 10 | 0.08 | 13 | 0.11 | 19 | 0.06 | 7 |
| Week4, burst3 | pod (DOS) | 0.34 | 1 | 0.34 | 1 | 0.59 | 28 | 0.32 | 1 |
| Week4, burst4 | portsweep (probe) | 0.48 | 17 | 0.13 | 21 | 0.39 | 37 | 0.12 | 16 |
| Week4, burst5 | portsweep (probe) | 0.2 | 1 | 0.41 | 1 | 0.54 | 4 | 0.19 | 1 |
| Week5, burst1 | satan (probe) | 0.06 | 21 | 0.02 | 38 | 0.08 | 47 | 0.03 | 14 |
| Week5, burst2 | ffb (r2l) | 0.86 | - | 0.89 | - | 0.93 | - | 0.73 | - |
| Week5, burst3 | portsweep (probe) | 0.49 | 8 | 0.04 | 8 | 0.06 | 12 | 0.05 | 9 |
| Total: 18424 | Detection rate | 14/19 (73.7%) | | 15/19 (78.9%) | | 10/19 (52.63%) | | 12/19 (84.2%) | |

**Πίνακας 3.3.iv.γ.** Σύγκριση των συστημάτων ανίχνευσης ανωμαλιών όταν εφαρμόζεται σε όλες τις εκρήξεις επιθέσεων με βάση τα πρότυπα μέτρησης του χρόνου απόκρισης και του εμβαδού επιφάνειας. Πηγή: Lazarevic A. *et al.* (2003)

Ωστόσο, υπάρχουν σενάρια όπου τα δύο συστήματα που αναφέραμε παραπάνω έχουν διαφορετική συμπεριφορά ανίχνευσης. Για παράδειγμα, το ξέσπασμα που είναι σκιαγραφημένο με γκρι πλαίσιο στον Πίνακα 3.3.iv.γ. (Burst2, Week 2) χρησιμοποιώντας το ρυθμό ανίχνευσης ξεσπάσματος για την αξιολόγηση των προσεγγίσεων, ανιχνεύθηκε βάση της προσέγγισης των NN και οχι της LOF. Στην Εικόνα 3.3.η. απεικονίζεται η ανίχνευση της έκρηξης αυτής χρησιμοποιώντας τις προσεγγίσεις των NN, LOF και Mahalanobis. Όπως φαίνεται σε αυτό είναι φανερό ότι η προσέγγιση LOF διαθέτει μικρότερο ποσοστό ανίχνευσης ξεσπάσματος καθώς έχει ένα μικρότερο αριθμό συνδέσεων που βρίσκονται πάνω από το όριο σε σχέση με την προσέγγιση των NN αλλά επίσης διαθέτει μία ελαφρώς καλύτερη απόδοση από αυτή της προσσέγγισης των NN. Επιπλέον σύμφωνα με την Εικόνα 3.3.iv.η. προκύπτει ότι οι δύο αυτές προσεγγίσεις για

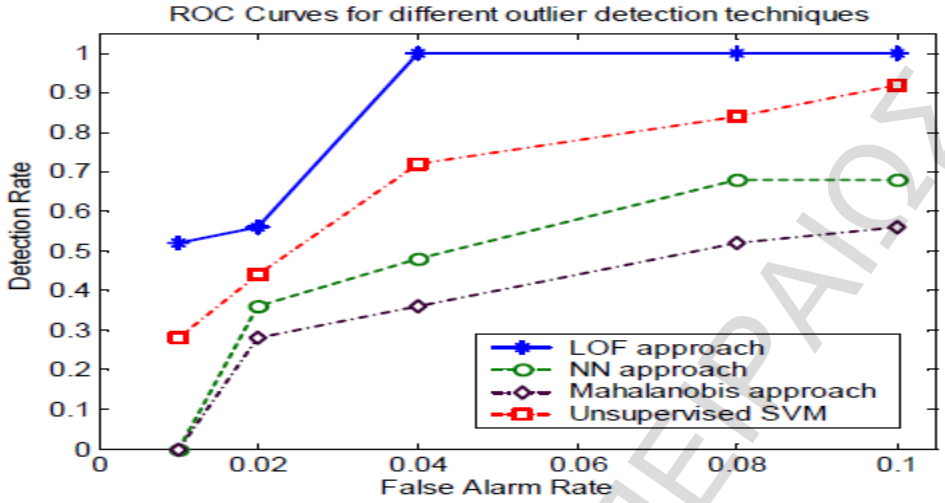
καθορισμένο όριο τα δύο συστήματα έχουν παρόμοιο χρόνο απόκρισης. Ωστόσο, κατά την ανίχνευση αυτής της εκρηκτικής επίθεσης, η προσέγγιση των κοντινότερων γειτόνων ήταν ανώτερη σε σχέση με τις άλλες δύο προσεγγίσεις. Η κυριαρχία της προσέγγισης NN σε σχέση με αυτή της LOF πιθανότατα έγκειται στο γεγονός ότι οι συνδέσεις αυτού του τύπου της επίθεσης βρίσκονται σε περιοχές τις οποίες η προσέγγιση LOF δεν είναι σε θέση να τις ανιχνεύσει λόγω της χαμηλής πυκνότητας, ενώ οι αποστάσεις από τους κοντινότερους γείτονές τους ήταν αρκετά υψηλές και συνεπώς η προσέγγιση NN ήταν σε θέση να τις προσδιορίσει ως ακραίες τιμές.




**Εικόνα 3.3.η.** Ανίχνευση της έκρηξης (burst 2, week 2) χρησιμοποιώντας τις δύο προσεγγίσεις των NN και LOF. Πηγή: Lazarevic A. *et al.* (2003)

Κατά την ανίχνευση των επιθέσεων ξεσπάσματος, πολύ συχνά υπάρχουν σενάρια στα οποία οι κανονικές συνδέσεις αναμιγνύονται με τις συνδέσεις από τις επιθέσεις ξεσπάσματος και έτσι καθιστούν το έργο του εντοπισμού επιθέσεων πιο περίπλοκο. Σε αυτές τις περιπτώσεις η προσέγγιση LOF είναι καταλληλότερη από ότι η προσέγγιση των NN και αυτό οφείλεται στο γεγονός ότι οι συνδέσεις που σχετίζονται με επιθέσεις είναι πολύ πιο κοντά σε πυκνές περιοχές από ότι αυτές που βασίζονται σε κανονική συμπεριφορά και συνεπώς η προσέγγιση NN δεν είναι σε θέση να τις ανιχνεύει μόνο ανάλογα με την απόσταση. Για παράδειγμα, το ξέσπασμα 4 την εβδομάδα 2 περιλαμβάνει 1000 συνδέσεις, αλλά μέσα στο χρονικό διάστημα επίθεσης υπάρχουν επίσης 171 κανονικές συνδέσεις (Εικόνα 3.3.iv.θ). Ο Πίνακας 3.3.iv.β δείχνει ότι η προσέγγιση LOF ήταν σε θέση να ανιχνεύσει 752 συνδέσεις που σχετίζονταν με αυτή την επίθεση, ενώ η προσέγγιση NN εντόπισε μόνο 62 από αυτές. Σε τέτοιες περιπτώσεις, η παρουσία των κανονικών συνδέσεων συνήθως προκαλεί τις χαμηλές τιμές κορυφών σε σχέση με τις συνδέσεις επιθέσεων, μειώνοντας έτσι το ποσοστό ανίχνευσης ξεσπάσματος και την αύξηση του εμβαδού επιφανείας

(Εικόνα 3.3.iv.θ). Επιπλέον, ένας μεγάλος αριθμός φυσιολογικών συνδέσεων ταξινομήθηκαν εσφαλμένα ως συνδέσεις που σχετίζονται με τις επιθέσεις, αυξάνοντας έτσι το ποσοστό ψεύτικων συναγερμών.




**Εικόνα 3.3.iv.θ** Καμπύλες ROC που δείχνουν την απόδοση των αλγορίθμων ανίχνευσης ανωμαλιών για ξεχωριστές επιθέσεις συνδέσεων. Πηγή: Lazarevic A. *et al.* (2003)

Η απόδοση των συστημάτων ανίχνευσης για την ανίχνευση ανωμαλιών επιθέσεων ενιαίας σύνδεσης στηρίχτηκε με βάση τη χρήση των καμπύλων ROC για όλους τους προτεινόμενους αλγορίθμους. Στην Εικόνα 3.3.iv.θ απεικονίζεται πως μεταβάλλεται το ποσοστό ανίχνευσης όταν ο ρυθμός εσφαλμένων συναγερμών κυμαίνεται από 1% έως 12%. Σύμφωνα με τα αποτελέσματα αυτά βλέπουμε ότι και πάλι η προσέγγιση LOF ήταν ανώτερη από όλες τις άλλες τεχνικές για όλες τις τιμές του ποσοστού ψευδών συναγερμών. Επομένως μπορούμε να συμπεράνουμε ότι το σύστημα LOF μπορεί να είναι καταλληλότερο σε σχέση με τα άλλα συστήματα για την ανίχνευση ανωμαλιών ξεχωριστών συνδέσεων επίθεσης ειδικά για εισβολές R2L. Όπως φαίνεται στον πίνακα που ακολουθεί (Πίνακας 3.3.iv.δ) παρουσιάζεται το **ποσοστό ανίχνευσης** (*Detection Rate*) για όλους τους επιμέρους τύπους ξεχωριστών επιθέσεων όταν το ποσοστό ψευδών συναγερμών είχε καθοριστεί 2%.

Είναι προφανές ότι οι προσεγγίσεις LOF, NN και SVM έχουν καλύτερες επιδόσεις σε σχέση με αυτές των Mahalanobis για όλους τους τύπους επίθεσεων. Η προσέγγιση LOF είναι σαφώς καλύτερη από τις μη επιβλεπόμενες μεθόδους των SVM και αυτές των προσεγισεων NN ειδικά για τις επιθέσεις R2L, όπου η προσέγγιση LOF ήταν σε θέση να ανιχνεύσει 7 από τις 11 επιθέσεις, ενώ η προσέγγιση των SVM και NN ήταν σε θέση να ανιχνεύσουν μόνο τρεις και μία επιθέσεις αντίστοιχα. Αυτή η υψηλή απόδοση που προσφέρει η προσέγγιση των LOF σε σύγκριση με αυτές των προσεγγίσεων NN ίσως μπορεί να εξηγηθεί από το γεγονός ότι η πλειοψηφία των ξεχωριστών επιθέσεων σύνδεσης βρίσκονται κοντά σε πυκνές περιοχές των

κανονικών δεδομένων και συνεπώς δεν είναι ορατές ως ακραίες τιμές από την προσέγγιση των NN.

| Number of attacks | Attack type and category | *LOF approach* | *NN approach* | *Mahalanobis based approach* | *Unsupervised SVM approach* |
|---|---|---|---|---|---|
| 13 | U2R | 6 (46.2%) | 7 (53.8%) | 5 (38.5%) | 7 (76.9%) |
| 11 | R2L | 7 (63.7%) | 1 (9.1%) | 1 (9.1%) | 3 (63.7 %) |
| 1 | DOS | 1 (100%) | 1 (100%) | 1 (100%) | 1 (100 %) |
| Detection rate | | 14 / 25 (56.0%) | 9 / 25 (36.0%) | 8/25 (28 %) | 11 /25 (44 %) |

**Πίνακας 3.3.iv.δ** Αριθμός των ανιχνευθέντων επιθέσεων και ποσοστά ανίχνευσης για την ανίχνευση ξεχωριστών επιθέσεων σύνδεσης. Πηγή: Lazarevic A. *et al.* (2003)

Συνοψίζοντας τα παραπάνω αποτελέσματα που αναλύσαμε στην Ενότητα 3.3, μπορούμε να συμπεράνουμε ότι οι τεχνικές ανίχνευσης ανωμαλιών με τη χρήση ακραίων τιμών φέρουν ικανοποιητικά αποτελέσμα στον κλάδο ανίχνευσης εισβολής σε ηλεκτρονικά συστήματα υπολογιστών. Συγκεκριμένα οι καμπύλες ROC που χρησιμοποιήθηκαν έδειξαν ότι πιο πολύ ικανοποιητικές προσεγγίσεις για την ανίχνευση εισβολών (με βάση το σύνολο δεδομένων του DAPRA'98) ήταν αυτή των LOF. Όσον αφορά τις τεχνικές των SVM τα αποτελέσματα που έλαβαν από τις αναλύσεις που έκαναν στην έρευνα που πραγματοποίησαν οι Lazarevic *et al.* (2003)οι διαπίστωσαν ότι ήταν πολύ ελπιδοφόρες για την ανίχνευση νέων εισβολών καθώς είχαν πολύ υψηλό ποσοστό ανίχνευσης, παράλληλα όμως παρουσίαζαν πολύ υψηλό ποσοστό ψευδών συναγερμών. Συνεπώς, θα πρέπει να διεξαχθούν μελλοντικά έρευνες για τις τεχνικές αυτές ώστε να διατηρηθεί ένα υψηλό ποσοστό ανίχνευσης μειώνοντας παράλληλα το ποσοστό ψεύτικων συναγερμών. Όσον αφορά την προσέγγιση Mahalanobis διερευνάται η ιδέα προσδιορισμού διάφορων τύπων κανονικής συμπεριφοράς μετρώντας τις αποστάσεις για κάθε τύπο συμπεριφοράς με σκοπό να εντοπιστούν οι ακραίες τιμές. Γενικότερα ο μακροπρόθεσμος στόχος των Lazarevic *et al.* (2003) είναι να συνεχίσουν τις έρευνές τους σχετικά με την ανάπτυξη ενός συνολικού πλαισίου ενάντια στις επιθέσεις και τις απειλές που αφορούν συστήματα ηλεκτρονικών υπολογιστών.

# ΚΕΦΑΛΑΙΟ 4

# Ανίχνευση Ιατρικής και Επιστημονικής Απάτης



## 4.1 Εισαγωγή

Όπως έχουμε αναφέρει και στην ιστορική αναδρομή του 1ου Κεφαλαίου, η απάτη είναι τόσο παλιά όσο και η ίδια η ανθρωπότητα. Η φύση της αλλάζει με το χρόνο και η έκτασή της είναι άγνωστη εκ των προτέρων καθώς μπορεί να απλώνεται σε διάφορους τομείς της κοινωνικής ζωής. Η ανακάλυψη της απάτης παραλληλίζεται με την εύρεση μιας βελόνας στα άχυρα υπό την έννοια της μετακίνησης μέσα σε μεγάλες μάζες δεδομένων προς εύρεση κάτι σπάνιου. Τα φαινόμενα απάτης τείνουν να κυριαρχήσουν τις τελευταίες δεκαετίες σε κάθε τομέα. Η ραγδαία εξέλιξη της τεχνολογίας σε συνδυασμό με την αύξηση της καταναλωτικής ισχύος και τις προηγμένες μεθόδους επικοινωνίας, έχει οδηγήσει σε πολλαπλά είδη απάτης όπως αυτά που προαναφέραμε σε προηγούμενα κεφάλαια. Παρόλα αυτά όμως απάτες μπορούμε να συναντήσουμε και σε χώρους πολύ διαφορετικούς από αυτούς που μέχρι τώρα εξετάσαμε. Παραδείγματος χάρη στο χώρο της υγείας έχουμε διάφορες περιπτώσεις απάτης (*medical fraud*) και σε διάφορα επίπεδα. Για παράδειγμα, έχουν αναφερθεί πολλές απάτες που αφορούν σε κλινικές δοκιμές, τα αποτελέσματα των οποίων αλλοιώνονται για να εξυπηρετήσουν τα συμφέροντα των επιστημόνων ή των φαρμακοβιομηχανιών.

Στον τομέα της υγείας πολυάριθμα είναι τα παραδείγματα εξαπάτησης των ασφαλιστικών φορέων από ιδιώτες και μη. Το 1999 επιβεβαιώθηκε, κατόπιν ερευνών, ότι το οργανωμένο έγκλημα εμπλέκεται άμεσα πλέον στις απάτες στο χώρο της υγείας. Συνήθεις πρακτικές είναι:

- η πλαστή υπερχρέωση δημόσιων και ιδιωτικών ασφαλιστικών φορέων από ιατρούς για υποτιθέμενες παρεχόμενες πολύπλοκες και δαπανηρές ιατρικές παροχές ή προϊόντα,
- κατευθυνόμενες συνταγογραφήσεις προς τα προϊόντα μιας συγκεκριμένης εταιρίας,
- χρέωση για υπηρεσίες ή προϊόντα που ποτέ δεν παρασχέθηκαν,
- πλαστή εικόνα της κατάστασης του ασθενούς,

- παραποίηση της ταυτότητας του ασθενούς για ασφαλιστική κάλυψη άλλου ασθενούς που δεν διαθέτει ασφαλιστική κάλυψη,
- σκόπιμη απόκρυψη ή παραποίηση πληροφοριών που αφορούν την κλινική εικόνα ή το ιστορικό του ασθενούς,
- ψευδή δήλωση στοιχείων του ασθενούς,
- παραποίηση του χρόνου παροχής ιατρικών υπηρεσιών ή προϊόντων,
- παραπομπή για μη αναγκαίες υπηρεσίες ή προϊόντα,
- χωριστή χρέωση για υπηρεσίες ή προϊόντα που φυσιολογικά καλύπτονται από μία χρέωση,
- διπλή χρέωση για την ίδια παρεχόμενη υπηρεσία ή προϊόν,
- χρήση κωδικού που δεν αντιστοιχεί στην ιατρική πράξη ή στο προϊόν που παρασχέθηκε, πλαστά έγγραφα για υποτιθέμενη ανάγκη διακομιδής του ασφαλισμένου από και προς το νοσοκομείο με ασθενοφόρο, κ.ά.

Επιπλέον οι ιατρικές απάτες σχετίζονται συχνά και με τις **ασφαλιστικές απάτες** (*insurance frauds*) όπως είναι οι ψευδείς δηλώσεις τραυματισμών κατά τη διάρκεια τροχαίων ατυχημάτων (Brockett *et al*., (1998)).

Συνήθης στόχος απάτης είναι ακόμα και άτομα με ανίατες ή σε τελικό στάδιο ασθένειες, τα οποία αναζητώντας ένα θαύμα, πέφτουν θύματα ιατρών που με μη συμβατικές μεθόδους θεραπείας αφενός εκθέτουν τη ζωή τους σε άμεσο κίνδυνο και αφετέρου υπερχρεώνουν τους ασφαλιστικούς τους φορείς για υπηρεσίες και προϊόντα ανακόλουθα της κατάστασης των ασφαλισμένων.

Ο ανταγωνισμός ανάμεσα στα νοσοκομεία και τους ασφαλιστικούς φορείς έχει αποδειχθεί αρκετά δαπανηρός. Οι οργανισμοί υγείας πρέπει να έχουν μηχανισμούς συμμόρφωσης που να μπορούν να εφαρμοσθούν στην ανίχνευση της απάτης. Εξάλλου η τεχνολογία έχει διευκολύνει την αυξανόμενη εμπλοκή των καταναλωτών στη βιομηχανία υγείας. Η πρόοδος στην τεχνολογία και στο λογισμικό μέσω προγραμμάτων που κινούνται μέσα στα εκατομμύρια αιτημάτων αποζημίωσης, αποδεικνύονται πολύ χρήσιμα. Αυτά τα προγράμματα εξόρυξης δεδομένων μπορούν να αποκαλύπτουν επαναλήψεις, ανωμαλίες ή και να συσχετίζουν άτομα με παράνομες δραστηριότητες. Πολύτιμο εργαλείο στα χέρια των ερευνητών είναι μια περιεκτική βάση δεδομένων (η αποθήκη δεδομένων-*data warehouse*) αιτημάτων που με προγράμματα εξόρυξης δεδομένων, συλλέγουν και αναλύουν πληροφορίες για τους ασφαλιστικούς φορείς και το κράτος.

Η ανίχνευση της απάτης προϋποθέτει τεράστιο αριθμό ομάδων δεδομένων, επομένως η επεξεργασία τους απαιτεί όχι απλά καινοτόμες στατιστικές μεθόδους αλλά και πολύ γρήγορους

αλγόριθμους. Τα στατιστικά εργαλεία είναι ποικίλα, καθώς η προέλευση των δεδομένων είναι ποικίλη και διαφοροποιημένη σε μέγεθος και είδος ανάλογα με την εφαρμογή. Τέτοια εργαλεία βασίζονται στην σύγκριση των τιμών των παρατηρούμενων δεδομένων με τις αναμενόμενες τιμές. Μπορεί να είναι μοναδικές τιμές που αντιστοιχούν σε κάποιο προφίλ συμπεριφοράς ή και τιμές πολλών μεταβλητών. Τέτοια προφίλ συμπεριφοράς μπορεί να βασίζονται σε προγενέστερη μελέτη της συμπεριφοράς του υπό μελέτη συστήματος, ή σε αντιπαραβολή του με άλλα παρόμοια συστήματα. Μία από τις δυσκολίες των συστημάτων ανίχνευσης απάτης είναι η ύπαρξη πολλών νόμιμων εγγραφών για κάθε μία εγγραφή απάτης. Ένα σύστημα που ανιχνεύει 99% των νόμιμων εγγραφών ως νόμιμες και το 99% των απατηλών ως απατηλές, θεωρείται πολύ ικανοποιητικό. Όμως, αν μόνον μία στις χίλιες ήταν απατηλή τότε σε κάθε 100 που θα προτείνει το σύστημα ως απατηλές μόνον οι 9 θα είναι πραγματικά έτσι. Αυτό θα οδηγεί σε έναν διεξοδικό έλεγχο των 9%, με αποτέλεσμα την αύξηση του κόστους. Ουσιαστικά η απάτη μπορεί να μειωθεί όσο περισσότερο επιθυμεί κάποιος, λαμβάνοντας όμως υπόψη του και την αντιστοιχία κόστους ανίχνευσης και του προκύπτοντος **συνολικού οφέλους** (*cost effectiveness*). Στην πράξη γίνεται συνήθως ένα είδος «εμπορικού συμβιβασμού» ανάμεσα στο επιθυμητό συνολικό κέρδος από την ανίχνευση και στο κόστος εντοπισμού της απάτης.

Στο προηγούμενο κεφάλαιο (Κεφάλαιο 3) αναφερθήκαμε για την **Εξόρυξη Γνώσης από Δεδομένα** (*Data Mining*) η οποία εφαρμόζεται σε διάφορους τομείς. Ένας από αυτούς είναι και ο τομέας της υγείας καθώς η συλλογή δεδομένων για την πραγματοποίηση των ερευνών γίνεται από βάσεις δεδομένων οι οποίες έχουν τεράστιο μέγεθος. Αν όλος αυτός ο όγκος των δεδομένων δεν επεξεργαστεί με κατάλληλο τρόπο τότε θα είναι άχρηστα. Ο στόχος αυτών των τεχνικών είναι η ανακάλυψη νέων προτύπων μεταξύ των δεδομένων και η εξαγωγή χρήσιμων πληροφοριών. Όσο η τεχνολογία θα εξελίσσεται, όλο και περισσότεροι οργανισμοί, εταιρίες και επιστήμες, θα έχουν την ανάγκη να χρησιμοποιήσουν την εξόρυξη γνώσης καθώς ο όγκος των δεδομένων που θα διαχειρίζονται για την εξαγωγή συμπερασμάτων ολοένα και θα αυξάνεται.

Σημαντικό ρόλο στην επεξεργασία των ιατρικών δεδομένων έπαιξε η πληροφορική, η οποία έχει εδραιωθεί πλέον στα περισσότερα συστήματα υγείας–ιατρικής περίθαλψης ανά τον κόσμο. Η χρησιμοποίηση ηλεκτρονικών υπολογιστών στα περισσότερα νοσοκομεία, αλλά και σε υπόλοιπους οργανισμούς που έχουν σχέση με την ιατρική περίθαλψη των ανθρώπων έδωσε τη δυνατότητα να αποθηκευτεί μεγάλος όγκος ιατρικών δεδομένων και να υπάρχει εύκολη πρόσβαση σε αυτά (*medical databases*). Τα δεδομένα αυτά, που αποθηκεύονται πλέον σε ψηφιακή μορφή, αφορούν εγγραφές ασθενών στο αρχείο, τις ασθένειες του κάθε ανθρώπου, τα φάρμακα που του χορηγούνται, τη θεραπεία που έχει ακολουθηθεί, δημογραφικά στοιχεία κτλ.

Καθώς λοιπόν ο όγκος των ιατρικών δεδομένων είναι τεράστιος παρουσιάζει δυσκολίες η μελέτη τους έτσι ώστε να εξαχθεί κάποια χρήσιμη πληροφορία για να πάρουμε μία απόφαση και για να μπορέσουν να επεξεργαστούν τα δεδομένα απαιτούνται καινοτόμες μεθόδους. Με τις υπάρχουσες μεθόδους ανάλυσης είναι εξαιρετικά δύσκολο να υπάρξει κάποια γνώση, γι' αυτό και είναι αναγκαία η υλοποίηση μεθόδων μέσα από υπολογιστικά συστήματα για να πραγματοποιηθεί μία σωστή ανάλυση των δεδομένων (Lavrac, 1999). Αυτό, έχει σαν αποτέλεσμα να είναι προτιμότερη η χρήση των τεχνικών εξόρυξης γνώσης για τα ιατρικά δεδομένα. Η εξόρυξη από ιατρικά δεδομένα είναι ένα από τα πιο ενδιαφέροντα και δύσκολα πεδία της εξόρυξης γνώσης. Τα δεδομένα που χρειάζονται για την εφαρμογή της εξόρυξης είναι τόσα πολλά, περίπλοκα και ετερογενή που καθιστούν την εξόρυξη μία πρόκληση για κάθε αναλυτή (Delen, 2009).

Η εξόρυξη γνώσης από δεδομένα δεν είναι κάτι το καινούριο και στον τομέα της υγείας γίνεται ολοένα και πιο δημοφιλής αν όχι όλο και περισσότερο απαραίτητη καθώς δίνει τη δυνατότητα να επεξεργάζεται και να εξάγει συμπεράσματα βάση ενός μεγάλου όγκου δεδομένων. Διάφοροι παράγοντες έχουν κίνητρο τη χρήση των εφαρμογών εξόρυξης δεδομένων στον τομέα της υγείας. Η ύπαρξη της ιατρικής ασφαλιστικής απάτης και κατάχρησης, για παράδειγμα, έχει οδηγήσει πολλούς ασφαλιστικούς φορείς υγειονομικής περίθαλψης να προσπαθούν να μειώσουν τις ζημίες τους με τη χρήση εργαλείων εξόρυξης δεδομένων για να βοηθήσει να βρουν και να παρακολουθήσουν τους δράστες. Ωστόσο πρόσφατα, υπήρξαν αναφορές σχετικά με επιτυχείς εφαρμογές εξόρυξης δεδομένων σε απάτη της υγειονομικής περίθαλψης και ανίχνευση κακοποίησης. Ένας άλλος παράγοντας είναι ότι οι τεράστιες ποσότητες δεδομένων που παράγονται από συναλλαγές της υγειονομικής περίθαλψης είναι πολύ πολύπλοκες και ογκώδες να υποβληθούν σε επεξεργασία γι' αυτό το λόγο η εξόρυξη δεδομένων μπορεί να βελτιώσει τη λήψη αποφάσεων από την ανακάλυψη προτύπων και τάσεων σε μεγάλες ποσότητες πολύπλοκων δεδομένων. Τέτοια ανάλυση έχει γίνει όλο και απαραίτητη καθώς οικονομικές πιέσεις ενέτειναν την ανάγκη στους οργανισμούς υγείας να λαμβάνουν αποφάσεις με βάση την ανάλυση των κλινικών και οικονομικών δεδομένων. Επομένως οι διορατικές ιδέες που αποκτήθηκαν από την εξόρυξη δεδομένων μπορεί να επηρεάσουν το κόστος, τα έσοδα και τη λειτουργική αποτελεσματικότητα, διατηρώντας παράλληλα ένα υψηλό επίπεδο φροντίδας.

Όσον αφορά τώρα την απάτη και την κατάχρηση στον τομέα της υγείας, χρησιμοποιούνται τεχνικές εξόρυξης δεδομένων σύμφωνα με τις οποίες δημιουργούνται πρότυπα ώστε να ανιχνευθεί η απάτη και η κατάχρηση μέσα από ασυνήθιστα ή μη φυσιολογικά πρότυπα των απαιτήσεων από γιατρούς, εργαστήρια, κλινικές, ή από διάφορους άλλους. Μεταξύ άλλων αυτές

οι εφαρμογές έχουν την ικανότητα να μπορούν να επισημάνουν ακατάλληλες συνταγές, δόλιες ασφαλιστικές ή ακόμη και δόλιους ιατρικούς ισχυρισμούς. Τέτοια παραδείγματα χρησιμοποιήσης των τεχνικών εξόρυξης δεδομένων έχουμε τα εξής: το **Γραφείο Ιατρικής Απάτης της Utah** (*Utah Bureau of Medicaid Fraud* ) έχει εξορύξει εκατομμύρια ιατρικές συνταγές, χειρουργικές επεμβάσεις, και προγράμματα θεραπείας για τον εντοπισμό ασυνήθιστων μοτίβων και την αποκάλυψη απάτης. Επίσης η **Επιτροπή Αυστραλιανής Υγειονομικής Ασφάλισης** (*Australian Health Insurance Commission*) έσωσε δεκάδες εκατομμύρια δολάρια από δόλιες δηλώσεις χρησιμοποιώντας την εφαρμογή των τεχνικών αυτών. Ένα ακόμη επιτυχημένο παράδειγμα χρήσης της εξόρυξης δεδομένων είναι **το σύστημα ανίχνευσης της Ιατρικής Απάτης και Κατάχρησης του Τέξας** (*Texas Medicaid Fraud and Abuse Detection System*), η οποία σε λιγότερο από ένα χρόνο έχει ανακάμψει 2,2 εκατομμύρια δολάρια.

Στη συνέχεια του κεφαλαίου αυτού θα παρουσιάσουμε μία έρευνα που έγινε σχετικά με τη χρήση τεχνικών εξόρυξης δεδομένων για την ανίχνευση απάτης στον τομέα της υγειονομικής περίθαλψης χρησιμοποιώντας τα τιμολόγιά τους σχετικά με την παροχή διαβητικών υπηρεσιών εξωτερικών ιατρείων. Αυτή η έρευνα στηρίχτηκε στο σύστημα της Εθνικής Ασφάλισης Υγείας που υπάρχει στο Ταϊβάν. Τα μοντέλα που χρησιμοποιήθηκαν στην έρευνα αυτή ήταν η λογιστική παλινδρόμηση, τα νευρωνικά δίκτυα και τα δέντρα ταξινόμησης.

Όσον αφορά την ακρίβεια του κάθε μοντέλου, σύμφωνα με τα αποτελέσματα που λήφθησαν τα δέντρα ταξινόμησης ήταν καλύτερα σε σχέση με τις άλλες δύο μεθόδους. Τα υψηλά ποσοστά της σωστής αναγνώρισης υποδεικνύουν ότι οι μεταβλητές που επιλέχτηκαν μπορούν να προσδιορίσουν τα νοσοκομεία υποβολής παράτυπων ιατρικών απαιτήσεων.

Κατά την τελευταία δεκαετία στο Ταϊβάν, από τότε που η κυβέρνηση υιοθέτησε ένα εθνικό σύστημα φροντίδας υγείας, οι ιατρικές δαπάνες και τα ποσοστά χρησιμοποίησής του έχουν εκτοξευθεί στα ύψη. Το νέο καθεστώς υποκίνησε επίσης έναν πολλαπλασιασμό σε αναφορές των χρόνιων ασθενειών. Μεταξύ των χρόνιων παθήσεων, ο **σακχαρώδης διαβήτης** (*DM*) κυριαρχεί στην επιβάρυνση των εθνικών ιατρικών δαπανών. Στις Ηνωμένες Πολιτείες και μόνο, η θεραπεία του διαβήτη τύπου 2 κοστίζει πάνω από 100 δισεκατομμύρια δολάρια ετησίως. Ανάμεσα στους ηλικιωμένους Αμερικανούς, το ποσοστό της νόσου ανέρχεται στο 28% των εθνικών δαπανών υγείας. Σε παγκόσμιο επίπεδο, οι ιατρικές δαπάνες έχουν αυξηθεί από 170 εκατ. ευρώ σε 4.440 εκατομμύρια δολάρια από το 1967 ως το 1999.

Στην Ταϊβάν ο διαβήτης αποτελεί την κυρίαρχη χρόνια πάθηση μεταξύ των ηλικιωμένων ατόμων εξαιτίας των αλλαγών στις διατροφικές τους συνήθειες και στον τρόπο ζωής. Η ασθένεια έχει καταλάβει την τέταρτη θέση μεταξύ των κύριων αιτιών θανάτου στο Ταϊβάν από το 1987.

Σύμφωνα με το Υπουργείο Υγείας στο Ταϊβάν ο αριθμός των ασθενών του διαβήτη που εντάχθηκαν στο εθνικό σύστημα υγειονομικής περίθαλψης έχει αυξηθεί σε 360 χιλιάδες, 1.8% του συνόλου των ασφαλισμένων. Από το 2001 έως το 2003, ωστόσο, οι ιατρικοί ισχυρισμοί σχετικά με την πάθηση του διαβήτη αυξήθηκαν από 23 έως 28 εκατομμύρια NT $ ($ 1 = 32 Νέα Ταϊβάν δολάρια). Το ποσό αυτό αντιπροσωπεύει το 8.1% των εξόδων της υγειονομικής περίθαλψης του έθνους, ένα κόστος που απέχει αρκετά σε σχέση με το ποσοστό του πληθυσμού που πάσχει από την ασθένεια αυτή. Η συνολική ετήσια αίτηση ανά διαβητικό ασθενή κατά μέσο όρο είναι 10 εκατομμύρια NT $ (περίπου 0,3 εκατομμύρια δολάρια). Η ανίχνευση της απάτης και κατάχρησης εξακολουθεί να αποτελεί σημαντικό έργο στην εξοικονόμηση κόστους.

Η μελέτη αυτή χαρακτηρίζει ως "διαβητικούς ασθενείς" εκείνους για τους οποίους ο διαβήτης έχει διαγνωστεί ως κύρια ή δευτερεύουσα νόσο. Δημιουργήθηκε ένα τυχαίο δείγμα από το **Ινστιτούτο Έρευνας Εθνικής Υγείας** (*National Health Research Institute,* (*NHI*)) χρησιμοποιώντας τη βάση δεδομένων του NHI, η οποία περιέχει 1.050.979 διαβητικούς ασθενείς και 17.668 παρόχους υγειονομικής περίθαλψης. Το δείγμα των παρόχων υγειονομικής περίθαλψης που εμπλέκονταν με δόλιους ισχυρισμούς χαρακτηρίζεται από εκείνους των οποίων οι συμβάσεις έχουν λήξει. Τέσσερα νοσοκομεία και κλινικές στο δείγμα τιμωρήθηκαν με αυτό τον τρόπο, τρία εκ των οποίων (που σχετίζονται με ένα σύνολο 189 δόλιες απαιτήσεις) βρίσκονται σε περιοχές που διέπονται από το BNHI κεντρικό υποκατάστημα. Στη μελέτη χρησιμοποιούνται μόνο οι πάροχοι υγειονομικής περίθαλψης από το κεντρικό υποκατάστημα για την κατασκευή του μοντέλου ανίχνευσης της απάτης. Μεταξύ άλλων παρόχων, η κατηγορία αυτή περιέχει 1275 συμβεβλημένα νοσοκομεία σε καλή κατάσταση. Η βάση δεδομένων που δημιουργήθηκε περιλαμβάνει πληροφορίες σχετικά με τους ασθενείς που πάσχουν από διαβήτη, τις διαγνώσεις τους, τις απαιτήσεις τους, καθώς και τους παρόχους υγειονομικής περίθαλψης που υπέβαλαν τις απαιτήσεις.

Σύμφωνα με τον **Οργανισμό Ασφάλισης Υγείας της Αμερικής** (*Health Insurance Association of America*) οι περισσότερες ιατρικές απάτες σχετίζονται με τη διάγνωση (43%) και τις υπηρεσίες χρέωσης (34%). Για την μελέτη που εξέτασαν οι Fen-May Liou *et al*. (2008) με βάση το σύνολο δεδομένων που επιλέχτηκε από τη βάση δεδομένων NHI, επιλέχτηκαν εννέα μεταβλητές που σχετίζονται με έξοδα για χρήση στα μοντέλα ανίχνευσης. Όλες έχουν προηγουμένως βρεθεί χρήσιμες στην ανίχνευση περιπτώσεων απάτης. Ο μέσος και η τυπική απόκλιση για κάθε μεταβλητή παρουσιάζονται στον Πίνακα 4.1, χωριστά για τους συνήθεις παρόχους και τους τρεις παρόχους που εμπλέκονται σε παράνομες δραστηριότητες.

Στην έρευνα αυτή που θα περιγράψουμε χρησιμοποιήθηκε το SPSS Clementine 7 για την εφαρμογή τριών τεχνικών εξόρυξης δεδομένων χωρίζοντας το σύνολο σε training και testing. Ωστόσο τα αποτελέσματα που θα αναφέρουμε είναι με βάση το testing σύνολο. Το μοντέλο που επιλέχτηκε, αρχικά εκπαιδεύεται στα training data και έπειτα η εγκυρότητά του εξετάζεται με χρήση των testing data. Τα αποτελέσματα από το testing σύνολο συγκρίνονται με εκείνα από το training, όπου θα πρέπει να είναι αρκετά κοντά έτσι ώστε οι αλγόριθμοι να μπορούν να προβλέπουν σωστά την περίπτωση απάτης στα δεδομένα μας.

Οι αλγόριθμοι που χρησιμοποιήθηκαν στην έρευνα αυτή είναι οι ακόλουθοι: η λογιστική παλινδρόμηση, το νευρωνικό δίκτυο, και ένα δέντρο ταξινόμησης. Και οι τρεις αυτοί αλγόριθμοι αποδίδουν καλύτερα όταν αναφερόμαστε σε ισόρροπα δεδομένα, δηλαδή όταν οι δόλιες και μη περιπτώσεις είναι ίσες. Αυτό μπορεί να επιτευχθεί είτε με την επανάληψη των δεδομένων απάτης ή μειώνοντας το δείγμα μη δόλιων περιπτώσεων μέχρι οι δύο πληθυσμοί να επιτύχουν τη ζητούμενη αναλογία. Το Clementine 7 παρέχει τη δυνατότητα μια επιλογής στον **κόμβο κατανομής** (*Distribution Node*) για τη δημιουργία ιστογραμμάτων συχνότητας. Η συχνότητα επιλογής ανάμεσα στα δείγματα των δόλιων και μη περιπτώσεων ήταν σταθμισμένη ώστε να εξασφαλιστεί μια ισορροπημένη κατανομή μεταξύ αυτών δύο ομάδων. Στη συνέχεια θα περιγράψουμε συνοπτικά τις μεθόδους που χρησιμοποιήθηκαν και θα αναφέρουμε και τα αποτελέσματα που έλαβαν.

| **Μεταβλητή** | **Νόμιμα Νοσοκομεία** | | **Δόλια Νοσοκομεία** | |
|---|---|---|---|---|
| | **Μέσος** | **Τυπική Απόκλιση** | **Μέσος** | **Τυπική Απόκλιση** |
| μέσος όρος ημερών του φαρμάκου να λήξει | 7,72 | 5,60 | 7,39 | 1,50 |
| μέσο κόστος φαρμάκου ανά ασθενή | 221,63 | 274,06 | 208,25 | 88,13 |
| μέσος όρος διαβούλευσης και επεξεργασίας τελών | 358,71 | 176,88 | 259,58 | 113,69 |
| μέσος όρος ιατρικών δαπανών | 265,42 | 42,93 | 265,00 | 43,46 |
| ο μέσος όρος τελών παροχής υπηρεσίας | 24,48 | 8,13 | 30,01 | 11,21 |
| το μέσο όρο ιατρικών δαπανών | 548,04 | 408,33 | 584,75 | 145,92 |
| το μέσο ποσό που ζητήθηκε | 487,99 | 394,69 | 511,81 | 131,97 |
| μέσο κόστος φαρμάκου ανά ασθενή ανά ημέρα | 28,81 | 27,79 | 33,82 | 10,07 |
| μέσος όρος των ιατρικών δαπανών ανά ημέρα | 134,37 | 92,29 | 173,13 | 73,56 |

**Πίνακας 4.1.** Πίνακας περιγραφικών στατιστικών για κανονικά και δόλια νοσοκομεία

*α. Λογιστική παλινδρόμηση*

Όπως αναφέραμε και στο δεύτερο κεφάλαιο (Ενότητα 2.2) θα αναφέρουμε και εδώ μερικά πράγματα σχετικά με τη λογιστική παλινδρόμηση, η οποία πρόκειται για μια μη γραμμική μέθοδος που αφορ την μοντελοποίηση δυαδικών εξαρτημένων μεταβλητών. Η λογιστική παλινδρόμηση χρησιμοποιείται σε πολλές διαφορετικές περιπτώσεις για την εξαγωγή συμπερασμάτων, ωστόσο σημαντική απήχηση έχει στον τομέα της υγείας. Ένα παράδειγμα αποτελούν οι κλινικές δοκιμές όπου πρέπει να συγκρίνουμε τα αποτελέσματα διαφορετικών θεραπειών τα οποία έχουν δυαδική μορφή. Για τη βελτίωση του μοντέλου εξετάζεται η σημασία κάθε μεταβλητής. Η εξαρτημένη μεταβλητή παίρνει δύο μόνο τιμές, οι οποίες αντιστοιχούν σε δύο ενδεχόμενα. Για παράδειγμα, θα μπορούσαν να ορισθούν ως αληθής ή ψευδής ένας ισχυρισμός ή ως επιτυχία ή αποτυχία όταν αναφερόμαστε σχετικά με κάποια θεραπεία. Οι τιμές των μεταβλητών αποτελούν μία αυθαίρετη κωδικοποίηση των δύο ενδεχομένων, συνήθως 0 και 1. Εάν ορίσουμε την τιμή $y$=1 σαν «επιτυχία» και την τιμή $y$=0 σαν «αποτυχία», τότε η $y$ είναι τ.μ της κατανομής Bernoulli, δηλαδή $y$~B(p), με μέση τιμή E($y$)=$p$ και διασπορά V($y$)=$p$(1-$p$).

Σε πολλές περιπτώσεις η τ.μ $y$ ενδέχεται να εξαρτάται από κάποιες επεξηγηματικές μεταβλητές. Η εξάρτηση της $y$ από τις επεξηγηματικές μεταβλητές $x$ (ανεξάρτητες μεταβλητές ή συμμεταβλητές) εισάγεται μέσω της εξάρτησης της πυκνότητας επιτυχίας $p$ από τις $x$. Η αντιστοιχία μεταξύ της πιθανότητας ($p$) και ενός διανύσματος παραγόντων επιρροής ($x$) μπορεί να διατυπωθεί ως εξής:

$$P = \frac{e^{f(x)}}{1+e^{f(x)}} \ ,$$

Η πιθανότητα αποτυχίας δίνεται αντίστοιχα ως εξής:

$$1 - P = \frac{1}{1+e^{f(x)}} \ ,$$

Αν διαιρέσουμε τώρα αυτές τις δύο σχέσεις που γράψαμε παραπάνω θα έχουμε τα odds ratio:

$$odds = \frac{P}{1-P} = e^{f(x)}$$

Η γραμμική συνάρτηση των odds ratio γράφεται ως εξής:

$$logit(P) = \log(\frac{P}{1-P}) = \beta_0 + \beta_1 X_1 + \cdots + \beta_\kappa X_\kappa$$

Η συνάρτηση της λογιστικής παλινδρόμησης έχει το πλεονέκτημα ότι ερμηνεύεται πιο εύκολα. Η μέθοδο μεγίστης πιθανότητας χρησιμοποιείται συνήθως για να βρεθεί το καλύτερο μοντέλο που διακρίνει τις δυο ομάδες. **Βηματική λογιστική παλινδρόμηση** (*Stepwise logistic regression*) πραγματοποιήθηκε σε κάθε μεταβλητή ξεχωριστά για τον εντοπισμό των πλέον αποτελεσματικών παραγόντων τα αποτελέσματα δίνονται στον πίνακα που ακολουθεί (Πίνακας 4.2).

| **Μεταβλητή** | **Πιθανότητα Μέγιστης Πιθανοφάνειας για το μειωμένο μοντέλο** | **p-value σε επίπεδο σημαντικότητας 1%** |
|---|---|---|
| Σταθερά | 2,562.454 | 0.000*** |
| μέσος όρος ημερών του φαρμάκου να λήξει | 1,449.564 | 0.000*** |
| μέσο κόστος φαρμάκου ανά ασθενή | 1,448.026 | 0.000*** |
| μέσος όρος διαβούλευσης και επεξεργασίας τελών | 1,439.871 | 0.012*** |
| μέσος όρος ιατρικών δαπανών | 2,053.640 | 0.000*** |
| ο μέσος όρος τελών παροχής υπηρεσίας | 2,041.844 | 0.000*** |
| το μέσο όρο ιατρικών δαπανών | 1,765.307 | 0.000*** |
| το μέσο ποσό που ζητήθηκε | 1,794.374 | 0.000*** |
| μέσο κόστος φαρμάκου ανά ασθενή ανά ημέρα | 1,508.997 | 0.000*** |
| μέσος όρος των ιατρικών δαπανών | 1,433.659 | 0.748 |

**Πίνακας 4.2.** Αποτελέσματα της Stepwise λογιστικής παλινδρόμησης

Όπως φαίνεται και από τον παραπάνω πίνακα (Πίνακα 4.2) οι οκτώ από τις εννέα μεταβλητές βρέθηκαν να έχουν σημαντική προβλεπτική ικανότητα. Η μεταβλητή "**μέσος όρος των ιατρικών δαπανών**" (*average medical expenditure*) έχει *p*-value>0.01 επομένως δεν είναι στατιστικά σημαντική για το μοντέλο μας. Στη συνέχεια αυτές οι οχτώ μεταβλητές που θεωρήθηκαν στατιστικά σημαντικές χρησιμοποιήθηκαν για να δημιουργήσουν ένα πλήρες μοντέλο λογιστικής παλινδρόμησης. Το τελικό μας μοντέλο που χρησιμοποιήθηκε για την εκτίμηση του λογαρίθμου της σχετικής πιθανότητας της επιτυχίας της δίτιμης μεταβλητής *Y* θα είναι της μορφής:

$$logit(P) = \log(\frac{P}{1-P}) = 2{,}562.454 + 1{,}449.564X_1 + \cdots + 1{,}508.997X_8,$$

Όπου $X_1$ είναι ο μέσος όρος ημερών του φαρμάκου να λήξει, $X_2$ είναι το μέσο κόστος φαρμάκου ανά ασθενή κ.ο.κ. Το $\beta_0$ είναι η σταθερά του μοντέλου που δημιουργήθηκε και ισούται με $2{,}562.454$, το $\beta_1$ είναι ο συντελεστής παλινδρόμησης για τη μεταβλητή $X_1$ και ισούται με $1{,}449.564$ κ.ο.κ. Καθένας από τους συντελεστές που αναφέρονται στον Πίνακα 4.2 ερμηνεύει τη μεταβολή του λογαρίθμου της σχετικής πιθανότητας αν μεταβάλουμε κατά μία μονάδα την ανεξάρτητη μεταβλητή $X_i$, εφόσον οι υπόλοιπες μεταβλητές παραμείνουν σταθερές. Επομένως αν αυξήσουμε τη μεταβλητή $X_1$ κατά μία μονάδα κρατώντας σταθερές τις υπόλοιπες μεταβλητές τότε θα αυξηθεί κατά 1,449.564 ο λογάριθμος της πιθανότητας ότι ο ισχυρισμός πως κάποιος ασθενής πάσχει από διαβήτη είναι αληθής. Τα ίδια ισχύουν και για τους υπόλοιπους συντελεστές.

Το ποσοστό ανίχνευσης των δόλιων νοσοκομείων είναι 100% (τρία στα τρία έχουν ανιχνευθεί σωστά), ενώ το σωστό ποσοστό ταυτοποίησης για κανονικά νοσοκομεία είναι 84.6% τα αποτελέσματα αυτά παρουσιάζονται στον Πίνακα 4.3. Το σωστό ποσοστό αναγνώρισης για το σύνολο του δείγματος είναι 92.2%.

*β. Νευρωνικά Δίκτυα*

Η τεχνολογία νευρωνικών δικτύων αποτελείται από ένα συνδυασμό μεθόδων, σχεδιασμένων να παρέχουν μέγιστη αξιολόγηση κινδύνου και ικανότητα διαχείρισης και έχουν εξελισιμότητα ώστε να «σηκώνουν» αύξηση στον όγκο των δεδομένων. Τα νευρωνικά δίκτυα είναι ιδανικά για την διαχείριση μεγάλων ποσοτήτων δεδομένων, όπως αυτά στον τομέα της υγείας, για αναγνώριση και συνδυασμό περίπλοκων, ασαφών ή ημιτελών πληροφοριών. Ένα νευρωνικό δίκτυο αποθηκεύει πληροφορίες κατά τον ίδιο τρόπο με τον ανθρώπινο εγκέφαλο, δηλαδή κάθε τμήμα πληροφορίας συσχετίζεται με άλλα τμήματα, έχουν επομένως ευλυγισία και διαμορφωσιμότητα, καθότι οι λύσεις που προτείνει ένα τέτοιο σύστημα, ρυθμίζονται ώστε να «στεγάσουν» τις υπάρχουσες ροές και τροποποιούνται ώστε να ενσωματώνουν νέες. Αυτό διευκολύνει την αναγνώριση σημαντικών και προβλεπτικών τάσεων και προτύπων εντός των δεδομένων των ασφαλιστικών φορέων και εντοπίζει ανωμαλίες ενδεικτικές κινδύνου για απάτη.

Τα νευρωνικά δίκτυα δεν προσδιορίζουν τη σημαντικότητα των μεμονωμένων μεταβλητών, και αυτό μπορεί να θεωρηθεί μειονέκτημα. Στο Clementine τα νευρωνικά δίκτυα έχουν **ανάλυση ευαισθησίας** (*sensitivity analysis*). Η επιλογή *Sensitivity Analysis* παρέχει ένα μέτρο σχετικής βαρύτητας για κάθε εκτιμητή του δικτύου και είναι βοηθητική στην αξιολόγηση αυτών των μεταβλητών. Η επιλογή αυτή δείχνει τις μεταβλητές εκείνες οι οποίες είναι πιο σημαντικές για την ταξινόμηση.

Τα αποτελέσματα που προέκυψαν από το συκεκριμένο παράδειγμα δείχνει την κατάταξη της σχετικής σπουδαιότητας κάθε μεταβλητής στην ταξινόμηση των στοιχείων η οποία είναι: ο μέσος όρος τελών παροχής υπηρεσίας, το μέσο κόστος διάγνωσης, ο μέσος όρος των ιατρικών δαπανών ανά ημέρα, ο μέσος όρος ημερών του φαρμάκου να λήξει, το μέσο κόστος φαρμάκου ανά ασθενή, το μέσο κόστος φαρμάκου ανά ασθενή ανά ημέρα, το μέσο διαβούλευσης και επεξεργασίας τελών, το μέσο όρο ιατρικών δαπανών, και το μέσο ποσό που ζητήθηκε. Για τον αλγόριθμο αυτό το σωστό ποσοστό αναγνώρισης των τιμών για το σύνολο του δείγματος δοκιμής καθώς επίσης και για τα κανονικά νοσοκομεία ειδικότερα είναι 95.73 και 91.47%, αντίστοιχα.

*γ. Δέντρα ταξινόμησης*

Τα δέντρα ταξινόμησης είναι μία από τις πιο γνωστές τεχνικές κατηγοριοποίησης η οποία χρησιμοποιείται ευρέως στην εξόρυξη γνώσης. Κατά την εφαρμογή ενός δέντρου ταξινόμησης κατασκευάζεται ένα δέντρο του οποίου τα φύλλα αναπαριστούν την κατηγοριοποίηση, όπου το καθένα αποτελεί και μία κλάση, και οι διακλαδώσεις αναπαριστούν τους διαχωρισμούς που πραγματοποιούνται κάθε φορά για να γίνει η κατηγοριοποίηση. Κάθε κόμβος (όπου ο πρώτος αποτελεί και τη ρίζα του δέντρου) χωρίζει τα δεδομένα σε δύο ή περισσότερες υπο-ομάδες. Η συγκεκριμένη μέθοδος χρησιμοποιεί κάποια δεδομένα εκπαίδευσης από περιπτώσεις με τις οποίες το δέντρο ταξινόμησης δημιουργήθηκε, αρχικά για γενίκευση και αξιολόγηση της αξιοπιστίας των κανόνων που εξάγονται από το δέντρο ταξινόμησης και κατά δεύτερο λόγο για να βελτιώσει τη συλλογή των κανόνων σε έναν ολοκληρωμένο κανόνα ο οποίος θα δώσει μια λύση. Οι λύσεις και οι κανόνες που θα προκύψουν από ένα δέντρο ταξινόμησης είναι εύκολα κατανοητές ακόμη και από έναν που δεν είναι ειδικός στον τομέα αυτό, έτσι η τεχνική των δέντρων ταξινόμησης είναι αρκετά διαδεδομένη στον τομέα της εξόρυξης γνώσης. Υπάρχουν πολλοί αλγόριθμοι για τα δέντρα ταξινόμησης με τους πιο διαδεδομένους να είναι ο C4.5 και ο ID3.

Η διαδικασία ταξινόμησης σταματά όταν στους "κόμβους φύλλα" έχει επιτευχθεί η κατηγοριοποίηση. Μετά τη δημιουργία ενός δέντρου ταξινόμησης με βάση τα δεδομένα εκπαίδευσης, οι κανόνες ή τα μοτίβα στα δεδομένα είναι προφανή και μπορούν να εφαρμόζονται με απλό αλγόριθμο ανίχνευσης. Κάθε δυνατό μονοπάτι από τον κόμβο ρίζα προς τον κόμβο φύλλο αντιπροσωπεύει μια σειρά από κανόνες ταξινόμησης. Δύο από τους κανόνες που οδηγούν σε τερματικό κόμβο προς την "απάτη" απεικονίζονται παρακάτω:

(1) Κανόνας 1:

IF "average days of drug dispense" <=6.0070958,

Then "contract status" = regular → σταμάτα σε αυτόν τον κόμβο

IF "average days of drug dispense" >6.0070958,

Then "contract status" = irregular → συνέχισε στον επόμενο κόμβο

(2) Κανόνας 2:

IF "average medical expenditure" <=416.23779,

Then "contract status" = regular → σταμάτα σε αυτόν τον κόμβο

IF "average medical expenditure" >416.23779,

Then "contract status" = regular → συνέχισε στον επόμενο κόμβο

.

.

.

Η αλληλουχία συνεχίζεται μέχρις ότου αποκτηθεί μια βέλτιστη πρόβλεψη. Όσον αφορά τώρα η ερμηνεία που μπορούμε να δώσουμε από τους κανόνες που αναφέραμε παραπάνω είναι η ακόλουθη: για παράδειγμα, ο κανόνας 1 έχει την εξής λογική: αν ο μέσος όρος ημερών του φαρμάκου να λήξει ("average days of drug dispense" ) είναι μικρότερος από 6 τότε το συμβόλαιο χαρακτηρίζεται ως κανονικό. Αν πάλι ο μέσος όρος ημερών του φαρμάκου να λήξει ("average days of drug dispense" ) είναι μεγαλύτερος από 6 τότε το συμβόλαιο χαρακτηρίζεται ως δόλιο και συνεχίζεται η ανάλυση μας προς το επόμενο κόμβο. Ο κανόνας 2 μας λέει ότι αν ο μέσος όρος ιατρικών δαπανών ("average medical expenditure") είναι μικρότερος από 416 τότε το συμβόλαιο χαρακτηρίζεται ως κανονικό. Ωστόσο αν ο μέσος όρος ιατρικών δαπανών ("average medical expenditure") είναι μεγαλύτερος από 416 τότε το συμβόλαιο χαρακτηρίζεται ως δόλιο και συνεχίζεται η ανάλυση μας προς το επόμενο κόμβο. Αυτή η μορφή των κανόνων μας επιτρέπει να εστιάσουμε περισσότερο σε ένα συγκεκριμένο συμπέρασμα από το να χρειάζεται να αναλύσουμε ολόκληρο το δένδρο.

Στην ανάλυση που έγινε με βάση τον αλγόριθμο του δέντρου ταξινόμησης, βρέθηκε ότι οι πιο σημαντικές μεταβλητές στις οποίες θα πρέπει να στηριχτεί κάποιος για την εμφάνιση δόλιου ιισχυρισμού σχετικά με την εμφάνιση του διαβήτη, από την πιο σημαντική στη λιγότερο, είναι: ο μέσος όρος ημερών του φαρμάκου να λήξει, ο μέσος όρος των συνολικών ιατρικών δαπανών, και ο μέσος όρος διαβούλευσης και επεξεργασίας τελών. Όσον αφορά τώρα για τον αλγόριθμο δέντρου ταξινόμησης, το ποσοστό ανίχνευσης για δόλια νοσοκομεία είναι 100%. Το ποσοστό σωστής αναγνώρισης για το σύνολο δεδομένων και το σύνολο για τα κανονικά νοσοκομεία είναι 99.37 και 98.73%, αντίστοιχα (Πίνακας 4.3).

| **Δείγματα** | **Κανονικά Νοσοκομεία** | **Δόλια Νοσοκομεία** | **Σύνολο** | **Ακρίβεια** | **Μέσος όρος ακρίβειας** |
|---|---|---|---|---|---|
| **Λογιστική Παλινδρόμηση** | | | | | =(1.069+1.267)/2.534 |
| *Κανονικά Νοσοκομεία* | 1069 | 198 | 1267 | 84.6 | = 92.18% |
| *Δόλια Νοσοκομεία* | 0 | 1.267 | 1267 | 100.0 | |
| **Νευρωνικά Δίκτυα** | | | | | =(1.159+1.267)/2.534 |
| *Κανονικά Νοσοκομεία* | 1159 | 108 | 1267 | 91.47 | = 95.73% |
| *Δόλια Νοσοκομεία* | 0 | 1.267 | 1267 | 100.0 | |
| **Δέντρα Ταξινόμησης** | | | | | =(1.251+1.267)/2.534 |
| *Κανονικά Νοσοκομεία* | 1251 | 16 | 1267 | 98.73 | = 99.37% |
| *Δόλια Νοσοκομεία* | 0 | 1.267 | 1267 | 100.0 | |

**Πίνακας 4.3.** Αποτελέσματα πρόβλεψης για τις 3 μεθόδους εξόρυξης δεδομένων

Και οι τρεις αλγόριθμοι κατόρθωσαν να επιτύχουν ένα ποσοστό ορθής αναγνώρισης 100% για τα δόλια ιδρύματα. Ωστόσο, τα ποσοστά ακριβείας τους για τους κανονικούς παρόχους είναι διαφορετική. Το μοντέλο του δέντρου ταξινόμησης έχει το μικρότερο ποσοστό σφάλματος (1%) κατά την ταξινόμηση κανονικών παρόχων, στη συνέχεια ακολουθεί το μοντέλο νευρωνικού δικτύου (9%) και το μοντέλο λογιστικής παλινδρόμησης (15%).

Στη συνέχεια θα αναφέρουμε ένα άλλο παράδειγμα σχετικά με την εφαρμογή εξόρυξης δεδομένων στον τομέα της υγείας. Ας υποθέσουμε ότι στο πλαίσιο του προγράμματος διαχείρισης της υγειονομικής περίθαλψης του, ο HealthOrg (πρόκειται για μια πλασματική οργάνωση της υγειονομικής περίθαλψης) ενδιαφέρεται να ανακαλύψει πώς ορισμένες μεταβλητές συνδέονται με την εμφάνιση του διαβήτη. Ο σκοπός της εφαρμογής εξόρυξης δεδομένων είναι να εντοπιστούν τα άτομα υψηλού κινδύνου ώστε να τους σταλούν μηνύματα τους σχετικά με την εμφάνιση του διαβήτη.

Το σύνολο δεδομένων που εξετάστηκε από την αποθήκη δεδομένων της HealthOrg στην έρευνα αυτή, περιέχει τις ακόλουθες επτά μεταβλητές οι οποίες παρουσιάζουν ιδιαίτερο ενδιαφέρον: το φύλο, την ηλικία, το δείκτη μάζας σώματος (BMI), αναλογία περιφέρειας (WHR), το κάπνισμα, ο αριθμός των φορών που ένας ασθενής ασκείται ανά εβδομάδα, και η εμφάνιση του διαβήτη, το οποίο είναι η μεταβλητή-στόχος και πρόκειται για κατηγορική μεταβλητή σύμφωνα με την οποία δείχνει αν ένα άτομο εμφανίζει ή όχι διαβήτη. Το σύνολο δεδομένων περιλαμβάνει 262, ή 12.78 % των ατόμων που εξετάστηκαν και ήταν θετικοί στην εμφάνιση του διαβήτη και 1778 άτομα, ή 87.22 %, τα οποία ήταν μη-διαβητικοί.

Εξετάζοντας το έργο των Breault *et al* (2002) για τα δεδομένα εξόρυξης σε ένα σύνολο δεδομένων για διαβητικούς, ο HealthOrg αποφάσισε ότι το δέντρο απόφασης είναι η κατάλληλη τεχνική εξόρυξης δεδομένων που θα πρέπει να χρησιμοποιήσει για να μάθει πώς ορισμένες μεταβλητές συνδέονται με την έναρξη του διαβήτη. Τα Δέντρα απόφασης έχουν το πλεονέκτημα της ευκολίας της ερμηνείας και της απεικόνισης. Τα αποτελέσματα που έλαβαν χρησιμοποιώντας το λογισμικό SPSS Clementine συνοψίζονται στον ακόλουθο πίνακα (Πίνακα 4.4) όπου παρουσιάζονται οι σημαντικές μεταβλητές για την εμφάνιση του διαβήτη καθώς το *p*-value< 0.01.

| **Μεταβλητές** | **$X^2$ (p-value)** |
|---|---|
| 1/ Ηλικία | 463,84 (<0.0001) |
| 2/ Δείκτης μάζας σώματος<br>Δείκτης μάζας σώματος | 184,99 (<0.0001)<br>98,66 (<0.0001) |
| 3/ Ηλικία<br>Αναλογία περιφέρειας<br>Αναλογία περιφέρειας | 26.72 (<0.0001)<br>32.33 (<0.0001)<br>40.34 (<0.0001) |
| 4/ Δείκτης μάζας σώματος<br>Αριθμός φορών ένας ασθενής ασκείται ανά εβδομάδα<br>Ηλικία | 17.63 (<0.0001)<br>10.19 ( 0.0014)<br>10.49 ( 0.0013) |

**Πίνακας 4.4.** Πίνακας σημαντικών μεταβλητών

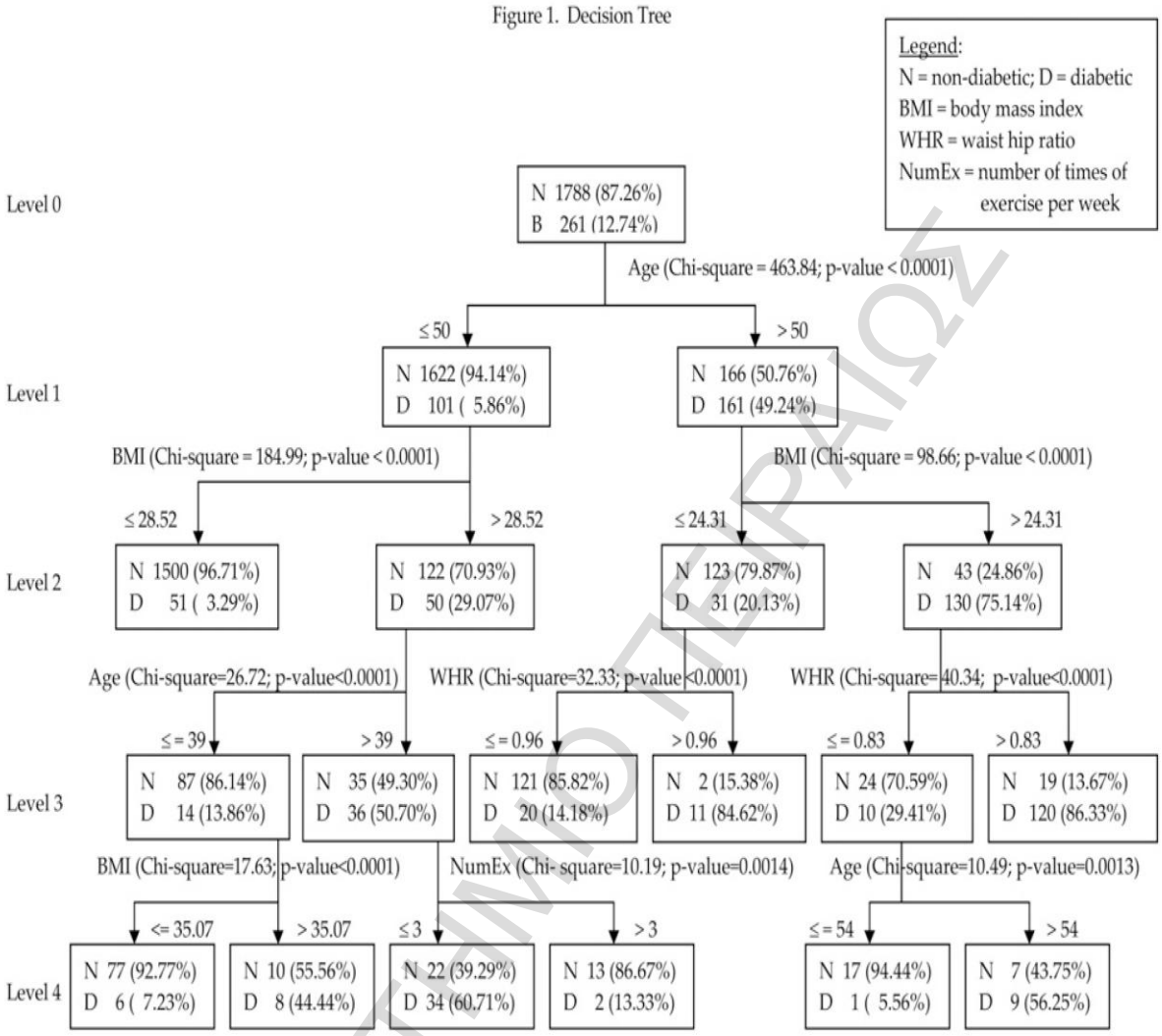
Στην συνέχεια με βάση τις σημαντικότερες μεταβλητές στις οποίες κατέληξαν ότι είναι απαραίτητες για την εμφάνιση του διαβήτη σχηματίστηκε το δέντρο απόφασης. Στο σχήμα που ακολουθεί (Σχήμα 4.1) απεικονίζεται το τελικό δέντρο απόφασης στο οποίο κατέληξε η ανάλυσή μας. Ωστόσο πριν προχωρήσουμε στην ερμηνεία των αποτελεσμάτων είναι σημαντικό να αξιολογήσουμε την απόδοση που προσφέρει το δέντρο ταξινόμησης, με άλλα λόγια τα ποσοστά ακρίβειας του δέντρου αυτού. Όπως φαίνεται και στον Πίνακα 4.5, τα ποσοστά ακρίβειας για τις περιπτώσεις των μη-διαβητικών ατόμων είναι πολύ υψηλό. Ειδικότερα όπως μπορούμε να δούμε από τα 1788 μη-διαβητικά άτομα το δέντρο απόφασης ταξινομεί σωστά τους 1728, ένα ποσοστό ακρίβειας του 96.64 %. Περαιτέρω, για τα 1808 άτομα που έχουν ταξινομηθεί από το δέντρο απόφασης ως μη διαβητικές περιπτώσεις, τα 1728 άτομα (95.58 %), είναι πραγματικά μη διαβητικοί. Οι τιμές ακρίβειας για την πρόβλεψη των διαβητικών περιπτώσεων είναι χαμηλότερες, ωστόσο ακόμη και αυτό το ποσοστό ακρίβειας μπορεί να θεωρηθεί επαρκές για το στόχο της εφαρμογής εξόρυξης δεδομένων. Ειδικότερα, για τα 262 διαβητικά άτομα στο σύνολο των δεδομένων, το δέντρο απόφασης κατατάσσει σωστά τους 182, δηλαδή ένα ποσοστό ακρίβειας 69.47 %. Περαιτέρω, για τα 242 άτομα ταξινομημένα από το δέντρο απόφασης ως διαβητικές περιπτώσεις, οι 182 (75.21 %), είναι πραγματικά διαβητικοί.

| **Ταξινόμηση** | | | |
|---|---|---|---|
| **Πραγματική Κατάσταση** | **Μη- διαβητικοί** | **Διαβητικοί** | **Σύνολο** |
| **Μη- διαβητικοί** | 1728 | 60 | 1788 |
| **Διαβητικοί** | 80 | 182 | 262 |
| **Σύνολο** | 1808 | 242 | 2050 |

**Πίνακας 4.5.** Πίνακας ταξινόμησης

Κοιτάζοντας τα αποτελέσματα του Πίνακα 4.5 μπορούμε να υπολογίσουμε οριμένα ποσοστά ακρίβειας. Αυτά είναι τα ακόλουθα:

1. Ποσοστό σωστής ταξινόμησης πραγματικών περιπτώσεων ατόμων που πάσχουν από διαβήτη: Sensitivity = 182/262 = 69.47%
2. Ποσοστό σωστής ταξινόμησης πραγματικών περιπτώσεων ατόμων μη διαβητικών: Specificity = 1728/1788 = 96.64%
3. Ποσοστό ατόμων που πάσχουν από διαβήτη που σωστά προβλέφθηκαν ως περιπτώσεις διαβητικών: True Positive = 182/242 = 75.21%
4. Ποσοστό ατόμων μη διαβητικών που σωστά προβλέφθηκαν ως περιπτώσεις μη διαβητικών: True Negative = 1728/1808 = 95.58%




**Σχήμα 4.2.** Δέντρο απόφασης.
Πηγή: Hian C. K. and Gerald T. (2005)

Η αρχή του δένδρου μας δείχνει το συνολικό αριθμό και τα ποσοστά και των 2 κατηγοριών της μεταβλητής Diabetic Σύμφωνα με αυτά που βλέπουμε από το δέντρο απόφασης που σχηματίστηκε, παρατηρούμε ότι η μεταβλητή Age αποτελεί τη πρώτη διακλάδωση στο δένδρο. Είναι ο πιο σημαντικός παράγοντας που σχετίζεται με την εμφάνιση των διαβητικών, με τα άτομα ηλικίας άνω των 50 ετών να δείχνουν σημαντικά υψηλότερο κίνδυνο διαβήτη σε σύγκριση με τα άτομα της ίδιας ηλικίας. Συγκεκριμένα βλέπουμε ότι αν Age<=50 τότε έχουμε 1622 άτομα μη διαβητικούς και 101 διαβητικούς, ενώ αν Age>50 τότε έχουμε 161 και 166 άτομα διαβητικούς και μη αντίστοιχα. Η μεταβλητή BMI αποτελεί την επόμενη διακλάδωση επομένως αποτελεί τον επόμενο πιο σημαντικό παράγοντα ο οποίος συνδέεται με την έναρξη του διαβήτη. Ειδικότερα, όπως φαίνεται και από το παραπάνω σχήμα, τα άτομα που έχουν ηλικία μικρότερη από 50 με BMI μικρότερο από 28.52 έχουν ένα πολύ χαμηλό κίνδυνο εμφάνισης του διαβήτη (ένα ποσοστό

μόνο 3.29 % για όσους ανήκουν σε αυτή την ομάδα). Περαιτέρω, η αύξηση των επιπέδων του BMI σχετίζεται με αυξανόμενο κίνδυνο διαβήτη. Για τα άτομα ηλικίας άνω των 50 με BMI μεγαλύτερο από 24.31, ο κίνδυνος εμφάνισης είναι 75.14 %. Όπως φαίνεται στο Επίπεδο 3 του Σχήματος 4.1, η μεταβλητή WHR είναι ο επόμενος πιο σημαντικός παράγοντας, η αύξηση του WHR συνδέεται με την αύξηση του κινδύνου εμφάνισης του διαβήτη. Για παράδειγμα, τα άτομα υψηλού κινδύνου στη βάση δεδομένων είναι εκείνα άνω των 50 ετών με δείκτη μάζας σώματος πάνω από 24.31 και WHR μεγαλύτερο από 0.83-με ποσοστό κινδύνου όσοι ανήκουν σε αυτή την ομάδα 86.33 %. Με τον ίδιο τρόπο μπορούν να ερμηνευθούν και οι υπόλοιποι κόμβοι στο δέντρο απόφασης. Αυτό το δέντρο απόφασης μπορεί να βοηθήσει την HealthOrg στον εντοπισμό των ατόμων εκείνων υψηλού κινδύνου εμφάνισης του διαβήτη ώστε να μπορούν να κοινοποιούνται σε αυτούς κατάλληλα μηνύματα. Για παράδειγμα, η HealthOrg θα μπορούσε να ξεκινήσει μια διαφημιστική εκστρατεία με βάση την υγεία σχετικά με το να ενημερώσει τους ανθρώπους ότι μεγάλο BMI και WHR αποτελούν παράγοντες κινδύνου οι οποίοι σχετίζονται με την εμφάνιση του διαβήτη. Επιπλέον, θα μπορούσε να ανιχνεύσει μέσω από τις διάφορες ασθενείς ομάδες στον εντοπισμό ατόμων στους οποίους θα έπρεπε να δώσουν περαιτέρω συμβουλές ή ακόμη να χρειαζόντουσαν επιπλέον ιατρικές εξετάσεις.

Το δέντρο ταξινόμησης θα μπορούσε να χρησιμοποιηθεί για την ανίχνευση απάτης. Σύμφωνα με αυτό διαπυστώσαμε ποιες μεταβλητές βρέθηκαν σημαντικές για την εμφάνιση του διαβήτη. Όπως αναφέραμε και παραπάνω κύρια μεταβλητή θεωρείται η ηλικία. Τα άτομα που είναι μεγαλύτερα των 50 ετών είναι πιο πιθανό να είναι διαβητικοί από οτι εκείνους που είναι μικρότερης ηλικίας. Επομένως θα πρέπει να είμαστε πιο προσεχτικοί στα άτομα αυτά για την δήλωση ψευδών ισχυρισμών. Το ίδιο συμβαίνει και με τις υπόλοιπες μεταβλητές που κρίθηκαν σημαντικές.

# ΞΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ


1. Agrawal R., Srikant R. (1994), *Fast Algorithms for Mining Association Rules Proceedings of the 20th Very large Data Bases Conference*, pp 487-499.
2. Bhattacharyya S., Jha S., Kurian K., Tharakunnel J., Westland C. (2011), *Data mining for credit card fraud: A comparative study*, pp 602-613.
3. Bolton R. and Hand D. (2002). *Statistical fraud detection: A review*. Statistical Science, vol. 17, no. 3, pp 235–255.
4. Brause, R., Langsdorf, T., Hepp, M. (1999), *Neural Data Mining for Credit Card Fraud detection*, Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence, pp 477-484.
5. Breunig M. M., Kriegel H.-P., Ng R. T., Sander J. (2000). *LOF: Identifying Density-Based Local Outliers*, Proceedings of the ACM SIGMOD Conference, pp 93-104
6. Chen R.C., Chen T.S., Chien Y.E., Yang Y.R. (2005), *Novel Questionnaire-Responded Transaction Approach with SVM for Credit Card Fraud Detection*, Lecture Notes in computer Science (LNCS), pp 916-921.
7. Delen D. (2009), *Analysis of cancer data: a data mining approach*, Expert Systems,Vol. 26, pp 100-112
8. Fawsett T., Haimowitz I., Provost F., and Stolfo S. (1998) *Approaches to Fraud Detection and Risk Management*. AAAI Press, Menlo Park, CA, vol. 19, no.3, pp 83-90.
9. Ferreira, P., Alves, R., Belo, O., and Ribeiro, J. (2007). *Detecting telecommunications fraud based on signatures clustering analysis*. In: Proceedings of Business Intelligence Workshop of 13th Portuguese Conference on Artificial Intelligence, pp 286-299.
10. Friedman, J.H. (1999), *Random Forests*. Stanford: Statistics Department, Stanford University.
11. Gerstner W., Germond A., Hasler M., and Nicoud J. (1997) *Conference on Artificial Neural Networks*. Springer, Berlin, vol.1327, pp 1274.
12. Geyer-Schulz, A., Hahsler, M. (2002), *Evaluation of recommender algorithms for an internet information broker based on simple association rules and on the repeat-buying theory.* In Proceedings of the WebKDD Workshop: Web Mining for Usage Patterns & User Profiles, pp 100-114.
13. Haykin, S. (1999), *Neural Networks, A Comprehensive Foundation*. Prentice Hall International, Inc, New Jersey.

14. Hian Chye Koh and Gerald Tan (2005), *Data Mining Applications in Healthcare,* Journal of healthcare information management, Vol. 19, pp 64-72.
15. Hilas C. S., and Sahalos J. N. (2007). *An application of decision trees for rule extraction towards telecommunications fraud detection*. In B. Apolloni et al. (Eds.): KES 2007/ WIRN 2007, Lecture Notes in Artificial Intelligence, vol. 4693, Part II, Berlin-Heidelberg: Springer – Verlag, 2007, pp 1112–1121.
16. Hilas C. S., and Sahalos J. N., *Testing the fraud detection ability of different user profiles by means of FF-NN classifiers,* in: S. Kollias et al. (Eds.), vol. 4132, Part II, Berlin Heidelberg Springer –Verlag, 2006, pp 872–883.
17. Hilas C. S., and Mastorocostas. P A. (2008). *An Application of Supervised and Unsupervised Learning Approaches to Telecommunications Fraud Detection*, Knowledge-Based Systems, vol. 21, 2008, pp 721 - 726.
18. Jing L., Kuei-Ying H., Jionghua J., Jianjun S. (2008), *A survey on statistical methods for health care fraud detection,* Health Care Management Science, Vol. 11, pp 275-287.
19. King-Fung Pun, J. (2011), *Improving Credit Card Fraud Detection using a Meta-Learning Strategy*, M.A.Sc. Chemical Engineering and Applied Chemistry University of Toronto.
20. Lavrac N. (1999), *Selected techniques for data mining in medicine*, Artificial Intelligence in Medicine, pp 3–23.
21. Lazarevic A., Ertoz L., Kumar V., Ozgur A., Srivastava J. (2003). *A Comparative Study of Anomaly Detection Schemes in Network Intrusion Detection,* In Proc. of SIAM International Conference on Data Mining (SDM), pp 19-78.
22. Liou F., Tang Y., and Chen J. (2008), *Detecting hospital fraud and claim abuse through diabetic outpatient services,* Health Care Management Science 11, pp 353-358
23. Parker D. (1983). *Fighting Computer Crime,* Charles Scribner's Sons, United States, pp 416.
24. Schonlau M. and Theus M. (2000). Detecting Masquerades in Intrusion Detection Based on Unpopular Commands, Information Processing Letters, vol. 76, pp 33-38.
25. Tan P., Steinbach M., and Kumar V. (2005), *Introduction to Data Mining*, Addison-Wesley. (ISBN:0-321-32136-7).
26. Usman J., Alan E. P. Cameron (2006). *Deceit and fraud in medical research*. International Journal of Surgery , vol. 4, no. 2, pp 122-126.

27. Wu X. and Carlsson M.(2010). *Detecting data fabrication in clinical trials from cluster nalysis perspective*. *Pharmaceutical Statistics*, vol. 10, no. 3, pp.257-264.
28. Yuan Y. and Shaw, M.J. (1995), *Induction of fuzzy decision trees*. Fuzzy Sets and Systems 69 pp 125–139.

## ΕΛΛΗΝΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Πολίτης Κ. (2010), *Γενικευμένα Γραμμικά Μοντέλα*, Πανεπιστημιακές Σημειώσεις του ΠΜΣ στην «Εφαρμοσμένη Στατιστική», Πανεπιστήμιο Πειραιά.
2. Dunham M. (2004), *Data Mining, Εισαγωγικά και Προηγμένα Θέματα Εξόρυξης Γνώσης από Δεδομένα,* επιμέλεια ελλ. έκδοσης Βερύκιος Β. και Θεοδωρίδης Ι, Εκδόσεις Νέων Τεχνολογιών.

## ΙΣΤΟΣΕΛΙΔΕΣ

http://en.wikipedia.org

http://www.dtreg.com/svm.htm

http://mines.humanoriented.com/classes/2010/fall/csci568/portfolio_exports/lguo/decisionTree.html

http://www.thehistoryof.net/history-of-credit-cards.html

http://news.bbc.co.uk/2/hi/business/7289856.stm

http://www.fraudaid.com/money-laundering.htm

http://www.wisegeek.com/what-are-the-different-money-laundering-methods.htm

http://www.ukpayments.org.uk/