Τεχνητή νοημοσύνη και ασφάλεια συστημάτων πληροφορικής.

Θ. Σπύρου, Ι. Δαρζέντας

Πανεπιστήμιο Αιγαίου
Τμήμα Μαθηματικών
Ερευνητικό Εργαστήρι Σάμου
83200 Καρλόβασι, Σάμος

e-mail: tsp@aegean.gr
idarz@aegean.gr

Περίληψη

Στην εργασία αυτή παρουσιάζεται μια γενική εικόνα των τρόπων με τους οποίους διάφορες μεθοδολογίες και τεχνικές από το χώρο της Τεχνητής Νοημοσύνης συντελούν στην αντιμετώπιση των διάφορων κατηγοριών απειλών προσβολής και παρεισφρήσεων (intrusions) σε συστήματα πληροφορικής. Περιγράφονται εκείνα τα συστήματα που ήδη υπάρχουν και χρησιμοποιούν μεθοδολογίες και τεχνικές Τεχνητής Νοημοσύνης και γίνεται σύγκριση της αποτελεσματικότητάς τους. Ιδιαίτερη προσοχή δίνεται στη συζήτηση δυο προσεγγίσεων που πρόσφατα έχουν εισαχθεί στο χώρο των συστημάτων ασφάλειας (των νευρωνικών δικτύων και των μοντέλων προθέσεων των χρηστών). Τέλος συζητείται κριτικά το μέλλον και οι προοπτικές της χρήσης της Τεχνητής Νοημοσύνης στο συγκεκριμένο χώρο των συστημάτων ασφάλειας πληροφοριακών συστημάτων.

Λέξεις Κλειδιά: Συστήματα ασφάλειας, Τεχνητή Νοημοσύνη, Συστήματα Ανίχνευσης παρεισφρήσεων.

T. Spyrou, J. Darzentas, Artificial Intelligence in Information Systems Security, In *Information Security, E. Kiountouzis, N. Alexandris (Eds.)*, GCS, 1995.

1. Εισαγωγή

Η ανάπτυξη πληροφοριακών συστημάτων στην εποχή μας είναι ραγδαία, τόσο για συστήματα υποστήριξης μεμονωμένων ατόμων όσο και για μεγαλύτερες ομάδες χρηστών που βρίσκονται στους ίδιους χώρους ή και σε διαφορετικές χώρες ή και ηπείρους.

Τα συστήματα αυτά εξυπηρετούν διάφορες κατηγορίες σκοπών, από τους πιο απλούς, όπως παροχή πληροφορίας γενικής φύσης, μέχρι τους πλέον σύνθετους όπως εξυπηρέτηση θεμελιακών αναγκών μεγάλων οργανισμών. Αυτή η ραγδαία ανάπτυξη των συστημάτων πληροφορικής γέννησε την ανάγκη μελέτης του βαθμού ασφάλειας των συστημάτων αυτών σε συνάρτηση με διάφορους παράγοντες όπως η σημασία του σκοπού που εξυπηρετείται από το σύστημα, τα επίπεδα ανοχής σε θέματα ασφάλειας, οι επιβαρύνσεις από τις πρόσθετες διαδικασίες ασφάλειας (οικονομικές αλλά και σε υπολογιστική ισχύ επιβαρύνσεις) κ.λ.π.

Διάφορες θεωρίες, μεθοδολογίες και τεχνικές ασφάλειας συστημάτων πληροφορικής έχουν προταθεί και μελετηθεί διεξοδικά [5,6,15,19]. Με βάση αυτές τις θεωρίες, μεθοδολογίες και τεχνικές ασφάλειας έχουν παραχθεί ή βρίσκονται ακόμα σε εξέλιξη διάφορα συστήματα προστασίας συστημάτων πληροφορικής [8,11,12,16,17,20,29,31]. Στην παρούσα εργασία γίνεται μια περιγραφή και ανάλυση εκείνων των προσπαθειών που χρησιμοποιούν μεθόδους και τεχνικές από το χώρο της Τεχνητής Νοημοσύνης.

Αρχικά παρουσιάζεται μια γενική εικόνα των τρόπων αντιμετώπισης των διάφορων κατηγοριών απειλών, κατόπιν δίνεται μια γενική εικόνα των υπαρχόντων συστημάτων που χρησιμοποιούν τεχνικές Τεχνητής Νοημοσύνης και ακολουθεί σύγκριση της αποτελεσματικότητας των μεθοδολογιών αυτών. Στη συνέχεια συζητούνται δυο προσεγγίσεις που πρόσφατα έχουν εισαχθεί στο χώρο των συστημάτων ασφάλειας (των νευρωνικών δικτύων και των μοντέλων πρόθεσης χρηστών) και τέλος γίνονται κρίσεις για το μέλλον και τις προοπτικές της χρήσης της Τεχνητής Νοημοσύνης στο χώρο των συστημάτων ασφάλειας πληροφοριακών συστημάτων.

2. Παρεισφρήσεις και γενικοί τρόποι αντιμετώπισης.

Στην εργασία αυτή υιοθετούνται οι όροι οι σχετικοί με την ασφάλεια των συστημάτων που χρησιμοποιήθηκαν στα πλαίσια της εκτέλεσης του Ευρωπαϊκού ερευνητικού προγράμματος SECURENET [29]. Ετσι με τον όρο *κακόβουλη επίθεση* (*malicious attack*), εννοούμε την κακόβουλη προσπάθεια για προσβολή ενός συστήματος και της λειτουργικότητας του μέσω μιας σειράς μη επιτρεπτών, ύπουλων και κακής πρόθεσης δραστηριοτήτων. Οι δραστηριότητες αυτές κατατάσσονται σε δύο πλατιές κατηγορίες. Είναι η κατηγορία των *ιών* (*viruses*) και η κατηγορία των *παρεισφρήσεων* (*intrusions*).

Η αντιμετώπιση των κακόβουλων επιθέσεων γίνεται με δυο βασικούς τρόπους. Είτε με διαδικασίες ανίχνευσής τους και άμεσης λήψης μέτρων αντιμετώπισής τους, ή με διαδικασίες πρόβλεψης και προστασίας των συστημάτων μέσω προληπτικής επέμβασης. Και στις δύο περιπτώσεις πρέπει είτε να ανιχνευθεί η παρουσία των αιτιών της εν δυνάμει ή ήδη εκδηλωμένης κακόβουλης επίθεσης ή να αναγνωριστεί η επικίνδυνη δράση που αναπτύσσεται στα πλαίσια αυτής της κακόβουλης επίθεσης.

Ετσι έχουμε δύο μεγάλες κατηγοριοποιήσεις ανίχνευσης, την *ανίχνευση δια της παρουσίας* (*detection by appearence*) και την *ανίχνευση δια της συμπεριφοράς* (*detection by behaviour*).

Αρκετοί μηχανισμοί έχουν προταθεί και υλοποιηθεί υιοθετώντας την προσέγγιση της ανίχνευσης δια της παρουσίας. Τέτοιοι είναι οι σαρωτές αρχείων (file scanners) και ελεγκτές ακεραιότητας (integrity checkers) [3,4], αλλά και μηχανισμοί προστασίας μέσω λογισμικού για αυτοάμυνα (self-defence) [34], αλλαγής ελέγχου (software change control) [10] ανοχής σφάλματος (software fault-

tolerance) [2] κλπ. Το χαρακτηριστικό αυτών των παραδοσιακών μηχανισμών είναι ότι ανιχνεύουν την επίθεση ανιχνεύοντας την παρουσία του επιτιθέμενου ή την παρουσία κάποιου αποτελέσματος μιας επίθεσης. Εχουν αποδειχτεί αρκετά ικανοποιητικοί για αντιμετώπιση ορισμένου είδους επιθέσεων και εξελιγμένες μορφές τους υπόσχονται ακόμα καλύτερα αποτελέσματα.

Τεχνικές όμως τεχνητής νοημοσύνης χρησιμοποιούνται σχεδόν αποκλειστικά για την *ανίχνευση δια της συμπεριφοράς* (detection by behaviour). Τρεις είναι οι βασικές εφαρμόσιμες μορφές μέσω των οποίων επιχειρείται η ανίχνευση δια της συμπεριφοράς: Τα *έμπειρα συστήματα - συστήματα κανόνων*, τα *συστήματα νευρωνικών δικτύων* και τα *συστήματα που βασίζονται σε μοντέλα χρήσης.* Στη συνέχεια κάθε μια από τις τρεις αυτές κατηγορίες θα εξεταστούν ξεχωριστά.

2.1. Τα έμπειρα συστήματα.

Τα *έμπειρα συστήματα*, τις περισσότερες φορές με τη μορφή συστημάτων κανόνων, είναι και τα πιο συνηθισμένα συστήματα αυτής της κατηγορίας, που υλοποιούν δηλαδή ανίχνευση δια της συμπεριφοράς. Η βασική φιλοσοφία αυτών των συστημάτων έγκειται στο πρόβλημα ελέγχου επιτρεπτότητας με βάση τρεις θεμελιώδεις οντότητες καθορισμού επιτρεπτότητας πρόσβασης, Τα υποκείμενα (subjects), τα αντικείμενα (objects) και τα δικαιώματα πρόσβασης (rights). Ο έλεγχος επιτρεπτότητας μετασχηματίζεται σε πρόβλημα ισχύος προτάσεων της μορφής: “Επιτρέπεται στο υποκείμενο υ δικαίωμα δ πάνω στο αντικείμενο α;”. Τριάδες της μορφής (υ,δ,α) τροφοδοτούν μια συνάρτηση φ(χ,ψ,ω) που δίνει σαν τιμή το μέτρο της επιτρεπτότητας ή όχι της άσκησης δικαιώματος δ από το υποκείμενο υ στο αντικείμενο α. Η συνάρτηση μπορεί να υλοποιηθεί απλά σαν ένας πίνακας επιτρεπτότητας ή και με πολύ πιο σύνθετες μορφές από σύνολα λογικών κανόνων (μοντέλα επιτρεπτότητας). Τις τελευταίες δύο δεκαετίες έχουν διαμορφωθεί και αναπτυχθεί αρκετά τέτοια μοντέλα επιτρεπτότητας. Ουσιαστικά, με τα σύνολα κανόνων που υλοποιούν τέτοια μοντέλα, ελέγχονται πράξεις σε ένα σύστημα για το αν αποτελούν μέρος συμπεριφοράς μη επιτρεπτής, ύποπτης, ή μη συνηθισμένης.

Η ανάπτυξη τέτοιων έμπειρων συστημάτων ξεκίνησε ιστορικά σαν φυσικό αποτέλεσμα της ανάγκης για έξυπνη επεξεργασία των *δεδομένων ελέγχου* (*audit data*) που προέρχονται από την παρακολούθηση και καταγραφή των δραστηριοτήτων που πραγματοποιούνται σε ένα σύστημα και των αποτελεσμάτων τους. Η επεξεργασία των τεράστιων όγκων δεδομένων που συγκεντρώνονται από την παρακολούθηση αυτών των δραστηριοτήτων (*auditing*) πραγματοποιείτο αρχικά με αλγοριθμικούς τρόπους, όπου αλγοριθμικής μορφής φίλτρα περιόριζαν τον προς εξέταση όγκο. Γρήγορα όμως αναπτύχθηκαν εξελιγμένοι αλγόριθμοι (όπως στατιστικοί) που μεταφράζουν τα απλά δεδομένα σε χρήσιμη, επεξεργάσιμη πληροφορία. Πέρα από αυτούς τους αλγόριθμους όμως, η έξυπνη επεξεργασία των δεδομένων για εξαγωγή συμπερασμάτων σχετικά με την ασφάλεια του συστήματος έφερε στην επιφάνεια τα έμπειρα συστήματα. Γνωστά σενάρια παρείσφρησης μπορούν εύκολα να εντοπιστούν με έμπειρα συστήματα και τα περισσότερα σύγχρονα συστήματα ασφάλειας έχουν ένα αντίστοιχο τμήμα που υλοποιεί τέτοιας μορφής έμπειρα συστήματα.

Μέχρι αυτή τη στιγμή μπορούμε να ισχυριστούμε ότι τα έμπειρα συστήματα παρουσιάζουν ικανοποιητικά αποτελέσματα στις περιπτώσεις ανίχνευσης γνωστών σεναρίων παρείσφρησης, στις περιπτώσεις κάλυψης του εύτρωτου εκείνων των τμημάτων των συστημάτων που είναι ευπρόσβλητα, στην παροχή προστασίας από δραστηριότητες που έχουν σκοπό να υπερκεράσουν τις διαδικασίες και την πολιτική ασφάλειας ενός συστήματος και γενικά για κάθε περίπτωση που μπορεί να χαρακτηριστεί εκ των προτέρων σαν ύποπτη δραστηριότητα.

2.2 Τα συστήματα νευρωνικών δικτύων

Η βάση της χρήσης των συστημάτων νευρωνικών δικτύων έγκειται στα εξής τρία χαρακτηριστικά τους: Προσαρμοστικότητα, δηλαδή ικανότητα να αναδιοργανώνονται δυναμικά, γενικευσιμότητα, δηλαδή ικανότητα να μπορούν να αποδώσουν ικανοποιητικά σε όχι καλά ορισμένες

περιπτώσεις ή σε καταστάσεις θορύβου και τέλος ικανοποιητική αντιμετώπιση χρονικά εξελισσόμενων καταστάσεων. Η κύρια χρήση τους συναντάται στην ανίχνευση και κατηγοριοποίηση των παρεισφρήσεων. Αυτό το πετυχαίνουν προσπαθώντας να εντοπίσουν αντικανονικές και παράτυπες δραστηριότητες μετά από σύγκριση με υπάρχουσες μορφές που έχουν αναγνωριστεί σαν κανονικές διαδικασίες ή επίσης προσπαθώντας να ταξινομήσουν και να κατηγοριοποιήσουν τα είδη και την εξάπλωση ιών.

Συνοπτικά μπορούμε να θεωρήσουμε ότι τα συστήματα νευρωνικών δικτύων χρησιμοποιούνται σαν αποτελεσματικοί κατηγοριοποιητές/ ταξινομητές προτύπων μορφών (patterns) ακόμα και σε περιπτώσεις που η αντίληψη της αντίστοιχης έννοιας αυτών των μορφών δεν είναι ξεκάθαρη.

2.3 Τα συστήματα μοντέλων

Τα συστήματα που χρησιμοποιούν μοντέλα για την ανίχνευση δια της συμπεριφοράς βασίζονται στο ότι δεν αντιμετωπίζουν την κακόβουλη επίθεση σαν αποτέλεσμα μιας δραστηριότητας ή μιας σειράς δραστηριοτήτων αλλά σαν αποτέλεσμα μιας γενικότερης αλληλεπίδρασης με ένα σύστημα. Το κύριο χαρακτηριστικό τους είναι η συνολικότερη γνώση για τη χρήση του συστήματος. Αυτή η συνολικότερη γνώση είναι συνήθως οργανωμένη σε μεγάλες οντότητες, τα καθήκοντα, (tasks) και οι συγκρίσεις που πραγματοποιούν δεν είναι μόνο στο επίπεδο των θεμελιωδών πράξεων οι οποίες σχηματίζουν ένα καθήκον/εργασία, αλλά συνολικά σε ολόκληρο το καθήκον/εργασία. Με αυτό τον τρόπο δίνεται η δυνατότητα αντιμετώπισης των πράξεων σαν μέρος ενός συνόλου όπου συνδυασμοί επιτρεπτών πράξεων μπορούν να μας δώσουν μια ύποπτη ή και μη επιτρεπτή εργασία.

Αντιπροσωπευτικά παραδείγματα αυτής της φιλοσοφίας είναι το Task-Based Authorisation [30], το model-based intrusion detection [9] και το Intention modelling based detection [25].

Στη συνέχεια δίνεται μια συνοπτική περιγραφή των συστημάτων που αυτή τη στιγμή είναι υλοποιημένα ή που η υλοποίησή τους βρίσκεται υπό εξέλιξη. Δεν αποτελεί σκοπό αυτής της εργασίας η αναδρομή στα υπάρχοντα συστήματα με μια λεπτομερή περιγραφή τους. Ο αναγνώστης μπορεί να μελετήσει αναλυτικά στη σχετική βιβλιογραφία που παρατίθεται τις ιδιαιτερότητες του κάθε συστήματος ξεχωριστά. Εδώ η περιγραφή εστιάζεται στα κομμάτια των συστημάτων που ενσωματώνουν κάποιες από τις τεχνικές Τεχνητής Νοημοσύνης που αναφέρθηκαν πιο πάνω έτσι ώστε να προετοιμαστεί ο χώρος για μια κριτική αντιπαράθεση των τεχνικών αυτών ως προς τα σημεία που αλληλοεπικαλύπτονται ή/και αλληλοσυμπληρώνονται.

3. Συστήματα που χρησιμοποιούν προσεγγίσεις Τεχνητής Νοημοσύνης

Το IDES [16,17,19,20,21,22,23] είναι ένα πραγματικού χρόνου έμπειρο σύστημα ανίχνευσης παρεισφρήσεων. Παρατηρεί τη συμπεριφορά των χρηστών ενός συστήματος και μαθαίνει τι είναι συνηθισμένο για ανεξάρτητους χρήστες, για ομάδες χρηστών και για το ίδιο το σύστημα. Είναι το πιο ολοκληρωμένο πακέτο αυτή τη στιγμή που συνθέτει διάφορες τεχνολογίες για την εξυπηρέτηση του σκοπού του και βρίσκεται σε διαρκή εξέλιξη για τη βελτιστοποίησή του και ενσωμάτωση νέων τεχνολογιών.

Το IDES βασίζεται σε δύο αρχές: i. οι παρεισφρήσεις, επιτυχείς ή επιχειρούμενες μπορούν να ανιχνευθούν όταν εξεταστεί λεπτομερώς η συμπεριφορά χρηστών που κάποια στιγμή φάνηκε να διαφοροποιείται από τη συνηθισμένη (δηλαδή από νόρμες συμπεριφοράς όπως αυτές έχουν σχηματιστεί από ιστορικά δεδομένα). ii. γνωστοί τρόποι διείσδυσης και κακής χρήσης ενός συστήματος μπορούν να ανιχνευθούν από ένα σύστημα κανόνων με τη μορφή ενός έμπειρου συστήματος. Το IDES καλύπτει και τις δύο αυτές αρχές βασικά χρησιμοποιώντας σύγχρονους εξελιγμένους στατιστικούς αλγορίθμους για την ανίχνευση ανωμαλιών [18,14] αλλά και

συμπληρωματικά έχοντας υλοποιήσει ένα έμπειρο σύστημα που ανιχνεύει γνωστά σενάρια παρεισφρήσεων [22,23]. Ο έλεγχος του υπό παρακολούθηση συστήματος γίνεται μέσω των *εγγραφών ελέγχου* (*audit records*). Το υπό παρακολούθηση σύστημα παρέχει στο IDES αυτές τις εγγραφές ελέγχου ενώ το λογισμικό του ίδιου του IDES είναι εγκατεστημένο και εκτελείται σε διαφορετικό υπολογιστή. Εκεί σχηματίζονται και διατηρούνται τα προφίλ των χρηστών και των ομάδων χρηστών του συστήματος. Με τον όρο προφίλ εννοείται η περιγραφή με κάποιο συγκεκριμένο τρόπο της κανονικής (αναμενόμενης) συμπεριφοράς των παρακολουθουμένων. Τα προφίλ αυτά ενημερώνονται καθημερινά (μια φορά την ημέρα) και έτσι το IDES μαθαίνει και προσαρμόζεται στα διάφορα χαρακτηριστικά των χρηστών. Τα προφίλ αυτά είναι στατιστικά προφίλ δηλαδή η περιγραφή του χαρακτήρα γίνεται με τη βοήθεια στατιστικής πάνω σε διάφορες χαρακτηριστικές μετρήσεις.

Το τμήμα του έμπειρου συστήματος χρησιμοποιείται σαν συμπληρωματική μέθοδος για την ανίχνευση των intrusions. Υπάρχουν περιπτώσεις όπου συγκεκριμένες ενέργειες μπορεί να θεωρούνται επιτρεπτές και όχι ύποπτες για κάποιον χρήστη επειδή ο συγκεκριμένος χρήστης παραπλανεί το σύστημα κάνοντας τις ενέργειές του να παρουσιάζονται σαν ενέργειες που πραγματοποιούνται στα πλαίσια κανονικής μη ύποπτης συμπεριφοράς (στατιστικά). Στο IDES μια σειρά ενεργειών ενεργοποιεί το σύστημα κανόνων ανεξάρτητα από το αν είναι τμήμα κανονικής ή όχι συμπεριφοράς. Χρησιμοποιώντας κλασικούς μηχανισμούς χρήσης κανόνων το σύστημα αποφαίνεται για την ομοιότητα των τρεχουσών πράξεων με κάποιο από τα γνωστά σενάρια παρείσφρησης.

Στο σύστημα IDES, σαν επέκταση του έμπειρου συστήματος αναπτύσσεται παράλληλα και ένα τμήμα ανίχνευσης παρεισφρήσεων με τη βοήθεια μοντέλων συμπεριφοράς (model-based intrusion detection) [9]. Το βασικό χαρακτηριστικό αυτού του τμήματος σε σχέση με αυτό του έμπειρου συστήματος είναι ότι ενώ το έμπειρο σύστημα προδιαθέτει για ανάπτυξη ανεξάρτητων, όχι καλά συσχετισμένων (ad hoc) κανόνων, το νέο τμήμα ενσωματώνει συγκεκριμένα μοντέλα μη επιτρεπτών δραστηριοτήτων. Αυτά τα μοντέλα περιγράφουν σενάρια παρεισφρήσεων και συνοδεύονται από ένα μηχανισμό καθορισμού και ανίχνευσης παρεισφρήσεων ανάμεσα από υποθετικά σενάρια που βασίζεται σε συμπερασματολογία σχετικά με την πιθανότητα να ισχύει ή να μην ισχύει το υποθετικό σενάριο συνολικά.

Η χρήση της τεχνολογίας των νευρωνικών δικτύων έχει αρχίσει να σχεδιάζεται και να υλοποιείται σαν συμπληρωματική μέθοδος ανίχνευσης παρεισφρήσεων για το IDES [17]. Εδώ γίνεται προσπάθεια εκμετάλλευσης των χαρακτηριστικών των συστημάτων νευρωνικών δικτύων για την αντιμετώπιση των προβλημάτων που ήδη αντιμετωπίζονται στο IDES όπως η ανάγκη για ακριβείς στατιστικές κατανομές, η δυσκολία αξιολόγησης των μετρήσεων που θεωρούνται αποτελεσματικές για την ανίχνευση, το υψηλό κόστος που απαιτείται για την ανάπτυξη και διαχείριση κλασσικών στατιστικών αλγορίθμων και οι δυσκολίες που παρουσιάζονται σχετικά με τις ομαδοποιήσεις χρηστών σε μεγάλης κλίμακας συστήματα με εκατοντάδες χρήστες.

Ενα δεύτερο κλασσικό σύστημα λογισμικού είναι το Wisdom and Sense [32,33] που προσπαθεί επίσης να αναγνωρίσει διαδικασίες που αποκλίνουν από ήδη γνωστές πρότυπες μορφές. Είναι από τα πρώτα συστήματα που προχώρησαν ένα επίπεδο πιο πέρα από τον απλό έλεγχο πρόσβασης και υλοποιούν συνθετότερους ελέγχους κατηγοριοποίησης συνόλου πράξεων των χρηστών ενός συστήματος (transactions) σε σχέση με τη διαφοροποίηση αυτών των συνόλων πράξεων από παρόμοια σύνολα σχηματισμένα από επεξεργασμένα ιστορικά δεδομένα. Για τη λειτουργία του χρησιμοποιεί και αυτό τα αρχεία καταγραφής δραστηριοτήτων του συστήματος (*αρχεία ελέγχου* / *audit trails*). Το σύστημα βασίζεται στο γεγονός ότι οποιαδήποτε ανωμαλία χρήσης θα παράγει εγγραφές στα αρχεία καταγραφής δραστηριοτήτων μέσω των οποίων (εγγραφών) μπορεί να αναγνωριστεί. Φυσικά θεωρείται ότι τα αρχεία καταγραφής δραστηριοτήτων (αρχεία ελέγχου) θα πρέπει να εμπλουτιστούν ως προς το είδος των στοιχείων που συλλέγονται. Και αυτό το σύστημα αντιμετωπίζει το πρόβλημα του τεράστιου και μη επεξεργάσιμου (ως προς την αποθηκευσιμότητα και αμεσότητα προσπέλασης) όγκου ιστορικής φύσης πληροφορίας για τις πράξεις των χρηστών του συστήματος. Η πρόκληση που τίθεται είναι η αφαιρετική (abstraction) και περιληπτική (summarise) αναπαράσταση της πληροφορίας αυτής ώστε με ελάχιστες παραδοχές να είναι εφικτή η σύγκριση με

τρέχουσες πράξεις του χρήστη.

Το σύστημα έχει σχεδιαστεί αρχικά για εφαρμογές σε προβλήματα διαχείρισης πυρηνικών υλικών και οι θεωρητικές ρίζες του συστήματος βρίσκονται σε μια εμπειρική προσέγγιση λύσης προβλημάτων.

Η καινοτομία που εισάγει είναι ο τρόπος με τον οποίο αναπαρίστανται τα ιστορικά δεδομένα που καταγράφουν τις πράξεις των χρηστών δηλαδή τα προφίλ του χαρακτήρα των χρηστών. Εδώ το σύστημα αντιμετωπίζει τα δεδομένα σαν *κατηγορικά, μη μετρικά* (*categorical, non-metric*) δεδομένα και όχι σαν στατιστικές μετρήσεις μεγεθών που αντιστοιχούν σε πραγματικά φαινόμενα. Ετσι πετυχαίνει τη διατήρηση του νοήματος των συγκεκριμένων πράξεων του χρήστη μέχρι την τελική επεξεργασία ελέγχου ύποπτης συμπεριφοράς. Οι εγγραφές ελέγχου δραστηριοτήτων (audit records) περνούν μια επεξεργασία με αποτέλεσμα την αυτόματη παραγωγή κανόνων που περιγράφουν τη συμπεριφορά των χρηστών. Το "δάσος κανόνων" όπως ονομάζεται το σύστημα αυτό των κανόνων δεν καθορίζεται από πρότυπες μορφές κανόνων (rule templates) και το σύστημα δε χρειάζεται τέτοιους σαν αρχική κατάσταση εκκίνησης παραγωγής του δάσους. Ούτε έχει ανάγκη από κανόνες που πρέπει να εισάγει ο έμπειρος διαχειριστής ασφάλειας του συστήματος. Αντίθετα το σύστημα κανόνων που παράγεται αυτόματα έχει μορφή που είναι εύκολα επεξεργάσιμη από τον άνθρωπο διαχειριστή έτσι ώστε αυτός να μπορεί να διαγράψει ή να διαμορφώσει κάποιον κανόνα που φαίνεται να μην απαιτείται ή ακόμα και να προσθέσει κανόνες που απαιτούνται και που το σύστημα δεν έχει μόνο του παράγει. Υπάρχει αυτόματος έλεγχος της συντακτικής ορθότητας και λογικής συμβατότητας κάθε νέου κανόνα που εισάγεται με αυτούς που ήδη υπάρχουν στη βάση κανόνων. Το σύστημα κανόνων περιέχει αρκετό πλεονασμό αλλά αυτός ο πλεονασμός θεωρείται απαραίτητος γιατί είναι αδύνατη η πρόβλεψη εκ των προτέρων του επιπέδου λεπτομέρειας που απαιτείται για τον έλεγχο κάποιας σειράς διαδικασιών και πράξεων του χρήστη. Αλλος λόγος της ύπαρξης αυτού του πλεονασμού είναι η ανάγκη για πραγματοποίηση του ελέγχου ακόμα και με έλλειψη κάποιων δεδομένων.

Αλλο σημαντικό χαρακτηριστικό του συστήματος είναι η δυνατότητα καθορισμού της εκτίμησης για τη σημαντικότητα του κάθε κανόνα σε σχέση με τους άλλους κανόνες μέσα στη βάση κανόνων. Κάθε κανόνας δηλαδή χαρακτηρίζεται για το πόσο επιζήμια μπορεί να είναι η παραβίασή του από κάποιον χρήστη.

Το κύριο πρόβλημα που αντιμετωπίζεται από το σύστημα είναι το μεγάλο πλήθος κανόνων που αναγκάζουν στη δημιουργία περίπλοκων δομών για τη διαχείρισή τους. Το σύστημα στη σημερινή μορφή του έχει να επιδείξει σημαντικές επιδόσεις τόσο στον τρόπο και στην ταχύτητα διαχείρισης του δάσους κανόνων για την ανίχνευση intrusions όσο και στα θέματα παραγωγής αυτών των κανόνων από τα αρχεία καταγραφής των δραστηριοτήτων των χρηστών (audit trails).

Το Time-based Inductive learning [31] είναι ένα άλλο σύστημα το οποίο αποκτά γνώση μέσω της εμπειρίας και με βάση αυτή τη γνώση δημιουργεί κανόνες για το πώς εξελίσσεται χρονικά η συμπεριφορά του χρήστη. Χρησιμοποιώντας αυτούς τους κανόνες έχει τη δυνατότητα να ανιχνεύει ανωμαλίες χρήσης.

Και σε αυτό το σύστημα χρησιμοποιούνται σύνολα κανόνων για την ανίχνευση κακής χρήσης ενός συστήματος. Η σημαντική διαφορά εδώ είναι ότι τα βασικά χαρακτηριστικά των πράξεων των χρηστών που εκμεταλλεύεται το σύστημα είναι οι ιδιαιτερότητες της χρονικής εξέλιξής τους. Μία ενότητα (session) σε ένα αρχείο αντιμετωπίζεται εδώ σαν ένα επεισόδιο (episode) με τη λογική ότι για να προβλεφτεί ένα συγκεκριμένο είδος γεγονότος, σκοπός είναι να ανακαλυφθεί από παρατηρήσεις ένα σύνολο χρονικά συσχετιζόμενων συνθηκών από τις οποίες προκύπτει (με κάποια βεβαιότητα) ότι θα επακολουθήσει το προς πρόβλεψη γεγονός. Ετσι οι εγγραφές στο αρχείο θεωρούνται σαν χρονικά συσχετιζόμενες και αναζητείται αυτή η χρονική συσχέτιση. Ακολουθιακοί συσχετισμοί όπως επόμενο γεγονός, προηγούμενο γεγονός κ.λ.π. είναι καλά ορισμένοι και με βάση τέτοιους ορισμούς αναλαμβάνει το τμήμα δημιουργίας προφίλ συμπεριφοράς χρηστών να παράγει τα αντίστοιχα μοντέλα χρήσης.

Και σε αυτό το σύστημα οι κανόνες δημιουργούνται αυτόματα σε περιόδους ενημέρωσης των βάσεων κανόνων από τα αρχεία καταγραφής δραστηριοτήτων των χρηστών (audit trails) αλλά υπάρχει η δυνατότητα να επέμβει ο άνθρωπος διαχειριστής ασφάλειας του συστήματος για να διαμορφώσει υπάρχοντες κανόνες ή να προσθέσει νέους.

Και εδώ η ανίχνευση ύποπτων δραστηριοτήτων γίνεται με την ανακάλυψη αποκλίσεων των τρεχόντων πράξεων από τα ήδη δημιουργημένα και αναπαραστημένα μέσω των κανόνων προφίλ.

Το σύστημα αυτό έχει παρουσιάσει επίσης ικανοποιητικά αποτελέσματα ειδικά στις περιπτώσεις χρηστών που παρουσιάζουν σημαντικές ποικιλίες μεταβολών κατά τη διάρκεια χρήσης του συστήματος.

Η επαγωγικής φύσης προσέγγιση που ακολουθεί βασισμένη σε αναπαράσταση μέσω εκφράσεων λογικής, το καθιστούν να θεωρείται σαν αναγκαίο συμπληρωματικό μοντέλο στα άλλα συστήματα μοντέλων αναλογιών, όπως τα στατιστικά.

Το σύστημα που βασίζεται στο Pattern-oriented Intrusion Detection [27] αναλύει μορφές ροής δεδομένων και δικαιώματα αντικειμένων σε σχέση με διαδικασίες πρόσβασης. Ετσι έχει τη δυνατότητα να ανακαλύπτει λειτουργικά προβλήματα ασφάλειας και να εξακριβώνει συγκεκριμένες μορφές παρεισφρήσεων. Και αυτό το μοντέλο παίζει συμπληρωματικό ρόλο στην ασφάλεια των συστημάτων υπολογιστών και δικτύων παράλληλα με τα συστήματα που είναι βασισμένα σε μοντέλα στατιστικής. Η βασική διαφορά του από τα προηγούμενα είναι ότι ενδιαφέρεται για την αναγνώριση παρεισφρήσεων που έχουν άμεση σχέση με το γενικό πλαίσιο πράξεων του χρήστη (context-dependent intrusions). Με το μοντέλο αυτό παρέχεται ένας ακριβής ορισμός των intrusions μέσω του φορμαλιστικού ορισμού των καταστάσεων του συστήματος (system states), των μεταβολών αυτών των καταστάσεων (state transitions), και των άμεσων σχέσεών τους (direct relations) και επίσης με τη βοήθεια κανόνων που συνιστούν έμμεσες σχέσεις μεταξύ των υποκειμένων και των αντικειμένων που αναγνωρίζονται σε ένα σύστημα. Τα υποκείμενα και αντικείμενα αυτά ορίζονται αυστηρά σαν οι δύο τύποι κόμβων ενός γράφου (protection graph). Τα υποκείμενα είναι οι ενεργοί κόμβοι (πχ. διαδικασίες ή χρήστες) που ενεργοποιούν λειτουργίες, μεταφέρουν πληροφορία και μπορούν να μεταβάλλουν δικαιώματα και δικαιοδοσίες. Τα αντικείμενα είναι οι παθητικοί κόμβοι (πχ. αρχεία, directories) και δεν έχουν τη δυνατότητα να είναι άμεσες αιτίες ενεργοποίησης λειτουργιών. Το σύστημα έχει να επιδείξει ικανοποιητική συμπεριφορά λειτουργίας στην ανίχνευση χρήσης προγραμμάτων χωρίς πρόθεση από το χρήστη και στην διασπορά και εξάπλωση ιών.

Το MIDAS [26] είναι ένας φλοιός - έμπειρο σύστημα με δυνατότητα επεξεργασίας ενός ικανοποιητικού αριθμού συμπερασματολογιών (150 inferences per second) και βασίζεται και αυτό στις αρχές του μοντέλου Denning [6] δηλαδή στην αναπαράσταση της συμπεριφοράς των δρώντων σε ένα σύστημα μέσω στατιστικών μετρήσεων και κανόνων για αξιοποίηση αυτού του είδους μετρήσεων. Επιπροσθέτως το MIDAS κατηγοριοποιεί τους ευρεστικούς κανόνες σε τέσσερις κατηγορίες: Αμεσους (immediate) που δρούν χωρίς μεγάλη ανάγκη την πληροφορία που είναι ιστορικά συγκεντρωμένη στα στατιστικά προφίλ, εύρεσης ανωμαλιών (anomaly) που χρησιμοποιούν άμεσα την πληροφορία των στατιστικών προφίλ, καθολικών (system wide) με τους οποίους καθορίζεται καθολικά η ομαλή συμπεριφορά και λειτουργία του συστήματος, και ευαισθησίας (sensitive path) όπου εξετάζεται μια ακολουθία πράξεων του χρήστη. Η τελευταία κατηγορία βρίσκεται ακόμα υπό ανάπτυξη στο σύστημα.

Το σύστημα που βασίζεται στο task-based authorisation [30] είναι μια άλλη προσπάθεια που κάνει χρήση αποτελεσμάτων από το χώρο της θεωρητικής μελέτης των προβλημάτων των βάσεων δεδομένων. Οι περιορισμοί που μελετήθηκαν από τους ερευνητές των βάσεων δεδομένων σχετικά με την αποτελεσματικότητα των κλασσικών μοντέλων διεργασιών (transaction models) μεταφέρθηκαν στο χώρο των συστημάτων ασφάλειας σε χώρους κατανεμημένων εργασιών μεγάλης διάρκειας (long-lived, distributed activities). Ετσι το σύστημα αυτό θεωρεί τις εργασίες (tasks) με τα εξής χαρακτηριστικά: Είναι μακράς διάρκειας, περιλαμβάνουν πολλαπλές υποεργασίες (subtasks), υπάρχουν κατευθυντήριες γραμμές για την εκτέλεσή τους και είναι κατανεμημένα στο χώρο και στο

χρόνο. Με το σύστημα αυτό γίνεται προσπάθεια να αντιμετωπιστούν προβλήματα όπως ποιες είναι οι κατάλληλες αφαιρετικές διαδικασίες για τη διαχείριση του ελέγχου επιτρεπτότητας των εργασιών, με ποιες ομαδοποιήσεις οι σχετικές δραστηριότητες μπορούν συνολικά να εξεταστούν και με ποιο τρόπο μπορούν να ελεγθούν οι κατανομές αρμοδιοτήτων.

Αλλα συστήματα που στη βασική τους θεωρητική προσέγγιση καλύπτονται από αυτά που έχουν ήδη αναφερθεί είναι το Intrusion Detection Agent [1] που είναι ένα σύστημα λογισμικού που βασίζεται στη λειτουργία ενός συγκριτή μορφών βάσει κανόνων (rule-based pattern matcher) και αξιοποιεί και αυτό στατιστικά προφίλ, το σύστημα Network Security Monitor [24] που είναι ένα έξυπνο σύστημα παρακολούθησης δικτύου που επίσης χρησιμοποιεί αρχεία ελέγχου (audit trails) και στατιστικά προφίλ, το HAYSTACK [28], το DIDS [7] και το NADIR [13] που είναι επίσης συστήματα για ανίχνευση ανωμαλιών χρήσης σε ένα σύστημα χρησιμοποιώντας και αυτά αρχεία ελέγχου (audit trails) και έμπειρο σύστημα κανόνων (rule-based).

Τέλος τα τελευταία δύο χρόνια βρίσκεται υπό εξέλιξη το SECURENET [25,29] που είναι ένα έξυπνο σύστημα ανίχνευσης και προστασίας συστημάτων δικτύων από παρεισφρήσεις. Το SECURENET έχει σχεδιαστεί για τη προστασία δικτυακού περιβάλλοντος και το βασικό χαρακτηριστικό του είναι ότι συνδυάζει την πλειοψηφία των μεθοδολογιών που προαναφέρθηκαν προσπαθώντας επιπλέον να υλοποιήσει τρεις βασικές προσεγγίσεις στα θέματα ασφάλειας που περιγράφονται περιληπτικά στη συνέχεια.

Ο βασικός σχεδιασμός του συστήματος είναι τέτοιος ώστε αυτό να λειτουργεί σε δικτυακό περιβάλλον ολοκληρωμένων, ευρείας ζώνης επικοινωνιών (Integrated Broadband Communications IBC). Το SECURENET με βασικό σκοπό την ολοκληρωμένη προστασία ενός δικτύου έχει τρεις θεμελιώδεις στόχους: την ανίχνευση μιας επίθεσης (detection of attack) την κατηγοριοποίηση της επίθεσης σε πραγματικό χρόνο (classification of attack) και την επιλογή και εφαρμογή των κατάλληλων μέτρων κατά της επίθεσης που εκδηλώνεται (countermeasurs). Ετσι στο σύστημα περιλαμβάνονται όλα τα τμήματα από την παρακολούθηση του δικτύου (monitoring), μέχρι την τελική συμβουλή στον διαχειριστή ασφάλειας του δικτύου για την αντιμετώπιση πιθανού κινδύνου.

Διάφορες τεχνικές χρησιμοποιούνται στα διάφορα κομμάτια, αλλά εδώ θα εξεταστούν τα τμήματα που υιοθετούν τεχνικές από την περιοχή της τεχνητής νοημοσύνης.

Για την ανίχνευση δια της συμπεριφοράς και εδώ χρησιμοποιούνται τα αρχεία καταγραφής δραστηριοτήτων των χρηστών (audit trails). Η τεχνολογία των συστημάτων νευρωνικών δικτύων χρησιμοποιείται σαν πρώτο στάδιο στις διαδικασίες ανίχνευσης και κατηγοριοποίησης των διαφόρων πιθανών παρεισφρήσεων. Η τεχνολογία των έμπειρων συστημάτων αποτελεί το κεντρικό πυρήνα ανίχνευσης των επιθέσεων ενστερνιζόμενη τις διάφορες θεωρητικές αρχές που προέρχονται από τις υπάρχουσες μοντελοποιήσεις αλλά και τη νέα μοντελοποίηση προθέσεων των χρηστών που εισάγεται από το συγκεκριμένο σύστημα.

Το SECURENET χρησιμοποιεί την τεχνολογία των νευρωνικών δικτύων γιατί το σύστημα αναγνώρισης παρεισφρήσεων πρέπει να χαρακτηρίζεται i. από προσαρμοστικότητα ώστε δυναμικά να αυτοτροποποιείται και να καλύπτει και συμπεριλαμβάνει περιπτώσεις δραστηριοτήτων χρηστών που δεν είχαν αρχικά υπολογιστεί, ii. από δυνατότητα γενικευσιμότητας ώστε να μπορεί να τα βγάλει πέρα σε περιπτώσεις υψηλού θορύβου ή ασάφειας και iii. από τη δυνατότητα επεξεργασίας χρονικά εξαρτώμενων μορφών δραστηριότητας.

Ετσι στο σύστημα SECURENET τα νευρωνικά δίκτυα χρησιμοποιούνται για να φιλτράρουν τις ιεραρχικά δομημένες εγγραφές δραστηριοτήτων των χρηστών (audit records). Κάθε μια από αυτές τις εγγραφές περιέχει πεδία που μπορούν να θεωρηθούν σαν τυχαίες μεταβλητές που η κατανομή τους περιγράφεται από ένα νόμο συναρτήσεων κατανομής (distribution function law). Τέτοιας μορφής νόμοι που αντιστοιχούν στην επίδραση της αντίστοιχης δραστηριότητας στην τρέχουσα κατάσταση του δικτύου, είναι δύσκολο να υλοποιηθούν βασιζόμενοι μόνο στην δειγματοληψία τιμών σχετιζόμενων με τις τυχαίες μεταβλητές. Τα νευρωνικά δίκτυα, έχοντας τα χαρακτηριστικά που προαναφέρθηκαν, μπορούν αποτελεσματικά να προσεγγίσουν τέτοιας μορφής νόμους

κανονικότητας της δραστηριότητας του δικτύου (regular activity), αφού φυσικά θεωρείται ότι οι πράξεις που πραγματοποιούνται στο δίκτυο κατευθύνονται από κάποια λογική. Η χρήση των recurent νευρωνικών δικτύων με τη δυνατότητα που έχουν να χειρίζονται χρονικές ακολουθίες κρατώντας στην εσωτερική τους μνήμη παρελθοντικά γεγονότα, και τη δυνατότητά τους να κρατούν τους νόμους κανονικότητας της συμπεριφοράς του δικτύου στην long term μνήμη τους, βοηθά στο να ξεπεραστούν προβλήματα που έχουν σχέση με την είσοδο δεδομένων σε συνάρτηση με το χρόνο (temporal input).

Μια νέα μορφή μοντελοποίησης χρησιμοποιείται επίσης στο SECURENET, η μοντελοποίηση των προθέσεων των χρηστών του συστήματος. Εδώ χρησιμοποιούνται θεωρητικά ευρήματα από χώρους όπως αυτός της εφαρμοσμένης ψυχολογίας (Applied Psychology) και της επιστήμης της νόησης (Cognitive Science). Τέτοια ευρήματα είναι οι φορμαλισμοί μοντελοποίησης νοητικών εργασιών (Cognitive Task Modelling CTM) και των γνωσιακών δομών εργασίας (Task Knowledge Structures TKS). Τα CTM αποτελούν μια προσέγγιση στην μοντελοποίηση των χρηστών συστημάτων ηλεκρονικών υπολογιστών και ασχολούνται με το κτίσιμο προσεγγιστικών περιγραφών της νοητικής δραστηριότητας που συντελείται στο νου του χρήστη κατά τη διάρκεια εκτέλεσης συγκεκριμένης εργασίας. Τα δε TKS περιγράφουν πώς η γνώση που χρειάζεται για την εκτέλεση μιας εργασίας είναι δομημένη, παρέχοντας έτσι μια πλούσια αναπαράσταση της γνώσης που σχετίζεται με τις διάφορες παρατηρούμενες συμπεριφορές όταν επιτελούνται οι σχετικές εργασίες (tasks).

Παράλληλα από το σύστημα αυτό εισάγεται η χρήση πιο σύγχρονων θεωριών του χώρου της Τεχνητής Νοημοσύνης όπως αυτές των ορθολογικών πράξεων (rational action) της αναγνώρισης πλάνου (plan recognition) και των προθέσεων των χρηστών (intention theories) βάσει των οποίων συνθέτονται τα μοντέλα προθέσεων (intention models).

Η απαιτούμενη γνώση αναπαρίσταται με κλασσικές μορφές αναπαράστασης (frames) καθώς και με σύνολα κανόνων. Αυτά καθορίζουν τον τρόπο λειτουργίας του συνθέτη (synthesiser) και του συγκριτή (comparator) τα οποία είναι τα δύο βασικά τμήματα της αρχιτεκτονικής του συστήματος. Ο συνθέτης συνθέτει οργανωμένες και καλά δομημένες μορφές σύνθετων πράξεων από το σύνολο των παρατηρούμενων ενεργειών των χρηστών ενός συστήματος και ο συγκριτής (comparator) πραγματοποιεί τις συγκρίσεις των τρεχουσών πράξεων του χρήστη (αποτελέσματα του συνθέτη) με τα ήδη διαμορφωμένα μοντέλα των εργασιών που μπορούν να εκτελεστούν. Αποτέλεσμα αυτής της σύγκρισης είναι η δημιουργία εκτίμησης για το πόσο ύποπτη είναι η πρόθεση του χρήστη που παρακολουθείται. Η ανάπτυξη του συστήματος βρίσκεται σε εξέλιξη.

4. Συμπεράσματα.

Από την επισκόπηση των συστημάτων που εξετάστηκαν πιο πάνω προκύπτει ότι υπάρχει γενική τάση να σχεδιάζονται και να υλοποιούνται συστήματα ασφάλειας που προστατεύουν πληροφοριακά συστήματα, λαμβάνοντας υπόψη το μέγιστο δυνατόν των απειλών. Η διαδικασία εξέλιξης των συστημάτων ασφάλειας εκφράζεται κυρίως από την τάση του ενστερνισμού νέων τεχνικών και μεθοδολογιών που παράλληλα με τις ήδη υπάρχουσες δίνουν τη δυνατότητα κάλυψης κάποιων απειλών επιπλέον αυτών που ήδη υπάρχουν.

Απειλές για την ασφάλεια ενός συστήματος μπορεί να οφείλονται σε οποιονδήποτε ή οτιδήποτε χρησιμοποιεί κάποιους από τους πόρους (resources) του συστήματος. Σαν αιτίες πιθανών απειλών μπορούν λοιπόν να θεωρηθούν οι κανονικοί και νόμιμοι χρήστες, οι χρήστες που αποκτούν δικαίωμα χρήσης του συστήματος παράνομα και τέλος τα προγράμματα που μπορούν με κάποιο τρόπο να εκτελεστούν στο σύστημα. Πιθανές απειλές μπορούν να προέλθουν και από παράγοντες έξω από το σύστημα προκαλώντας εξωτερικής φύσης επιθέσεις αλλά η ασφάλεια σε αυτή την περίπτωση έχει να κάνει με τη φυσική προφύλαξη του συστήματος

Οπως έχει τονιστεί αυτή τη στιγμή θεωρείται ανέφικτη η πλήρης προστασία των σημερινών συστημάτων πληροφορικής με αυτοματοποιημένα συστήματα ασφάλειας. Ενα σύστημα ασφάλειας

για να είναι αποτελεσματικό πρέπει να έχει καθορίσει τους σκοπούς και στόχους του ως προς τι ακριβώς προστατεύει το αντίστοιχο πληροφοριακό σύστημα και πρέπει να ενσωματώνει όλες εκείνες τις τεχνικές που αφορούν στην ικανοποίηση των στόχων του.

Υιοθετώντας αυτή την αντίληψη όλες οι τεχνικές που προαναφέρθηκαν είναι απαραίτητες για την προστασία ενός πληροφοριακού συστήματος. Από τα βασικά στάδια προστασίας συστημάτων (απλοί έλεγχοι πρόσβασης - access control) έως τις τελευταίες γραμμές άμυνας ενός συστήματος (διαρκής έλεγχος και παρακολούθησή του / auditing) οι τεχνικές της Τεχνητής Νοημοσύνης έχουν να προσφέρουν σημαντική βοήθεια βελτιώνοντας τις διαδικασίες ασφάλειας. Αυτό προϋποθέτει γενίκευση της χρήσης των μεθόδων της τεχνητής νοημοσύνης και χρήση του συνόλου των εργαλείων που αυτή παρέχει.

Η χρήση των έμπειρων συστημάτων με τη μορφή κανόνων είναι αυτή τη στιγμή το πιο διαδεδομένο και χρησιμοποιημένο εργαλείο από τα περισσότερα συστήματα ασφάλειας. Στην εξέλιξή τους όμως αυτά τα συστήματα προβλέπεται να χρησιμοποιούν όλο και περισσότερο τα χαρακτηριστικά των νευρωνικών δικτύων και αρχές από τις θεωρίες των ορθολογικών πράξεων (rational actions), της αναγνώρισης πλάνου (plan recognition) και των προθέσεων των χρηστών (intentions). Για την εφαρμογή των αρχών αυτών των θεωριών με ικανοποιητικό αποτέλεσμα απαιτείται άμεσα κατάλληλη διεπιστημονική συνεργασία με ερευνητές από τους αντίστοιχους χώρους της τεχνητής νοημοσύνης.

Το ερευνητικό πρόγραμμα SECURENET που αναφέρθηκε προηγουμένως αποτελεί αντιπροσωπευτικό παράδειγμα αυτής της τάσης, όπου ήδη χρησιμοποιούμενες αρχές ανίχνευσης παρεισφρήσεων συμπληρώνονται με στοιχεία από τις νέες στο χώρο των συστημάτων ασφάλειας θεωρίες των προθέσεων των χρηστών και σε συνδυασμό με τα πλεονεκτήματα των νευρωνικών δικτύων υπόσχεται θετικότερα αποτελέσματα για τη βελτίωση της ασφάλειας των συστημάτων πληροφορικής.

Σύσ	Ονομα	Ανάπτυξη	Χώρα	Κατάσταση	Επιδόσεις	Τεχνολογίες ΤΝ	ΒΔ	Δ	ΝΔ	Πόροι
IDES	Intrusion Detection Expert System	SRI	Αμερική	Ενεργό	Πραγματικού χρόνου	*	*	*		US Navy, SPAWAR
W&S	Wisdom and Sence	LosAlamos Nat Lab	Αμερική	Ενεργό		*	*			US Govmn
TBIL	Time-based Inductive Learning	DEC. U. of Illinois	Αμερική	Μερική Ανάπτυξη	Πραγματικού χρόνου	*	*			US Govmn
POIDM	Rattern-Oriented Intrusion Detection Model	University of Maryland	Αμερική	Υπό Ανάπτυξη		*				US Govmn
IDA	Intrusion Detection Agent	University of Hamburg	Γερμανία	Υπό ανάπτυξη		*				German Govmn
NSM	Network Security Monitor	Lawrence Livermore	Αμερική	Μερική Ανάπτυξη		*		*		US Govmn
	MIDAS	Natl. Comp. Sec. Centre	Αμερική	Ενεργό		*				US Govmn
	HAYSTACK	Tractor Inc.	Αμερική	Ενεργό		*				Tracor Inc.
	NADIR	Los Alamos Natl Lab.	Αμερική	Υπό Ανάπτυξη		*		*		US Govmnt
DIDS	Distributed Intrusion Detection System	Lawrence Livermore & Tractor	Αμερική	Υπό Ανάπτυξη		*		*		US Govmn &Trac
SECURENET	An Intelligent System for Preventing and Detecting Attacks in Open Networks	Expertnet, Dassault, Cnet, CCC, Faw-Ulm, Un.Aegean U. of Oulu	ΕΟΚ	Υπό Ανάπτυξη	Πραγματικού χρόνου	*	*	*	*	ΕΟΚ

Πίνακας 1: *Συγκριτική παρουσίαση συστημάτων ασφάλειας.*
ΤΝ: τεχνητή νοημοσύνη, Δ: δίκτυα, ΝΔ: νευρωνικά δίκτυα, ΒΔ: βάσεις δεδομένων

Αναφορές

[1] Brunnstein K., Fischer-Hubner S. and Swimmer M., Concepts of an Expert System for Virus Detection, *In Proc. of the Sec'91 IFIP*, 1991

[2] Chen L., Avizienis A., "N-version programming: a fault tolerance approach to reliability of software operations", *in Proc. of FTCS-8*, pp. 3-9, 1978.

[3] Cohen F., "A cryptographic checksum for integrity protection", *Computers and Security*, Vol. 6, no 5, pp. 505-510, 1987.

[4] Davida G.L., Desmedt Y. G., Matt B.J., "Defending systems against viruses through cruptographic authentication", *in the Proc. of the 1989 IEEE Symposioum on Computer Security and Provacy*, pp. 312-324, 1989.

[5] Debar H., Becker M. and Siboni D., A Neural Network Component for an Intrusion Detection System, *in Proc. of the 1992 IEEE Symposium on Research in Computer Security and Privacy*, pp. 240-250, 1992.

[6] Denning D.E., An Intrusion-Detection Model, *IEEE Transactions on Softwre Engineering*, 13(2), 1987.

[7] Dias G., Leviyy K. and Mukherjee B. "Modelling Attacks on Computer Systems: Evaluating Vulnerabilities and Forming a Basis for Attack Detection", in Proceedings, SRI Intrusion Detection Workshop 5, pp. 296-304, 1990.

[8] Fox K., Henning R. and Reed J., A Neural network approach towards intrusion detection, *in Proc. of the 1990 Symposium on Research in Security and Privacy*, pp. 125-134, 1990.

[9] Garvey T.D. and Lunt T.F. "Model-Based Intrusion Detection", *In proceedings of the 14th National Computer Security Conference*, Washington DC., 1991.

[10] Gritzalis D., Information Systems Security, *Greek Computer Society Monograph Series*, 1989.

[11] Guinier D., Biological versus Computer Viruses: A Better Understanding for a Better Defence, *in ACM SIGSAC Review*, Vol.7, no 2, pp. 1-15, 1989.

[12] Guinier D., Computer Virus Identification by neural networks, *in ACM SIGSAC review*, Vol. 9, No 4, pp. 49-59, 1991

[13] Jackson K.A., Dubois D.H. and Stallings C.A., An Expert System Application for Network Intrusion Detection*, in Proc 14th National Computer Security Conference*, pp. 215-225, 1991

[14] Javitz H.J. and Valdes A. "The SRI IDES Statistical Anomaly Detector" In Proceedings 1991 IEEE Symposioum on Security and Privacy, 1991.

[15] Kerr S., Using AI to Improve Security, *Datamation*, Feb 1990.

[16] Lunt T.F., Tamatu A., Gilham F., Jagannathan R., Jalali C., Javitz H., Valdes A. and Neumann P., A Real-time Intrusion Detection Expert System, *Technical Report SRI-CSL-90-05*, Stanford Research Institute, 1990

[17] Lunt T.F., Tamatu A., Gilham F., Jagannathan R., Jalali C. and Neumann P., A Real-time Intrusion Detection Expert System (IDES), *Final Technical Report*, Stanford Research Institute, 1992.

[18] Lunt T.F. "Using Ststistics to Track Intruders", *In proceedings of the Joint Statistical Meetings of the American Statistical Association*, 1990

[19] Lunt T.F., Automated audit trail analysis and intrusion detection: A survey, *In Proceedings of the 11th National Computer Security Conference*, 1988

[20] Lunt T.F., A Survey of Intrusion Detection Techniques, *Computers and Security*, 12, pp. 405-418, 1993

[21] Lunt T.F., IDES: An Intelligent System for Detecting Intruders, *In Proceedings of the Symposioum: Computer Security, Threat and Countermeasures*, 1990.

[22] Lunt T.F., Jagannathan R., A Prototype Real-Time Intrusion-Detection Expert System, *In Proceedings of the 1988 IEEE Symposium on Security and Privacy*, 1988

[23] Lunt T.F., Jagannathan R., Lee R., A. Whitehurst, Knowledge-Based Intrusion Detection*, In Proceedings of the 1989 AI Systems in Government Conference*, 1989

[24] Mansur D., Network Security Monitor, *in Presentation Notes for IDES Workshop 3*, 1988

[25] Spyrou T. and Darzentas J. "User Intention Modelling", *Technical Report*, *R2113* SECURENET II Project, Deliverable 4, RACE, 1994

[26] Sebring M., Shellhouse E., Hanna M. and Whitehurst A., Expert Systems in Intrusion Detection:

A Case Study, *in Proc. 11th National Computer Security Conference*, 1988
[27] Shieh S.W. and Gligor V., A Pattern-Oriented Intrusion Detection Model and its Applications, *in Proc. 1991 IEEE Symposioum on Research in Computer Security and Privacy*, IEEE Computer Society Press, pp. 327-342, 1991
[28] Smaha S.E., Haystack: An Intrusion Detection System, *in Proc. 12th National Computer Security Conference*, pp. 37-44, 1988.
[29] Spyrakis P., Katsikas S., Gritzalis D., Allegre F., Darzentas J., Gigante C., Karagiannis D., Putkonen H. and Spyrou T., "SECURENET: A Network-Oriented Intelligent Intrusion, Prevention and Detection System, *10th IFIP International Information Security Conference*, 1994
[30] Thomas R. K., R. S. Sandhu, Task-based Authorisation: A Paradigm for Flexible and Adaptable Access Control in Distributed Applications, *16th National Computer Security Conference*, pp. 409-415, 1993
[31] Teng H.S., Chen K. and Lu S.C.-Y., Adaptive Real-time Anomaly Detection Using Inductively Generated Sequential Patterns, *in Proc. 1990 IEEE Symposioum on Research in Computer Security and Privacy*, IEEE Computer Society Press, pp. 278-284, 1990
[32] Vaccaro H.S. and Liepins G.E., Anomaly Detection: Purpose and Framework, *in Proc. 12th National Computer Security Conference*, 280-289, 1989
[33] Vaccaro H.S. and Liepins G.E., Detection of Anomalous Computer Session Activity, *In Proc. of the 1989 IEEE Symposioum on Security and Privacy*, 1989
[34] Yau S., Cheung R., "Design of self checking software", *in Proc. of the IEEE Conference on Reliable Software*, pp. 450-457, 1975.