ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΠΙΣΤΗΜΗ ΤΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Διπλωματική Εργασία
Μεταπτυχιακού Διπλώματος Ειδίκευσης

«*Εφαρμογή Πλαισίου Διαχείρισης Εμπιστοσύνης σε δίκτυα Pub/Sub*»

Γεώργιος Τσιώλης

Επιβλέπων: Ιωάννης Μαριάς
2ος Αξιολογητής: Γ. Ξυλωμένος

ΑΘΗΝΑ, ΙΟΥΛΙΟΣ 2010

ΠΕΡΙΕΧΟΜΕΝΑ

Περιεχόμενα Πινάκων

Περίληψη

Στα δίκτυα τύπου Publish/subscribe κάθε οντότητα (subscriber) εγγράφεται ώστε να ανακτήσει περιεχόμενο και υπηρεσίες από σμήνος ομότιμων (publishers). Σε τέτοιους είδους πληροφορια-κεντρικές εφαρμογές η εμπιστοσύνη αποκτά νέα ερμηνεία, εφόσον αποτιμάται όχι μόνο ο παραγωγός αλλά και ο καταναλωτής του περιεχομένου. Στόχος της διπλωματικής είναι να εξετάσει την εφαρμογή ενός μοντέλου διαχείρισης εμπιστοσύνης που αποτιμά κακόβουλη συμπεριφορά καταναλωτών περιεχομένου σε δίκτυα Pub/Sub.

Δίνεται μια γενική περιγραφή συστημάτων publish/subscribe και ειδικότερα του PSIRP (Publish Subscribe Internet Routing Paradigm) ενός συστήματος που ερευνά την αρχιτεκτονική του Internet από την αρχή και βασίζεται σε publish/subscribe αρχιτεκτονική. Επίσης περιγράφονται βασικά θέματα ασφάλειας σε συστήματα publish/subscribe ενώ δίνεται και μια γνωστή έκθεση ταξινόμησης επιθέσεων DoS (Denial of Service). Στη συνέχεια αναφέρονται οι επιλεγμένες επιθέσεις που απασχολούν το συγκεκριμένο μοντέλο.

Η αποτίμηση στηρίζεται στη θεωρία Dempster-Shafer με την οποία μπορεί να προσδιοριστεί και η αβεβαιότητα στη λήψη απόφασης. Περιγράφεται αναλυτικά η θεωρία και το μαθηματικό μοντέλο που προκύπτει ενώ υλοποιείται σενάριο εφαρμογής και αποτίμησης οριοθετώντας μετρικές απόδοσης της εφαρμογής

Τα αποτελέσματα αναλύονται στο τέλος τόσο για κεντρικοποιημένα συστήματα όσο και για κατανεμημένα συστήματα και εξάγονται τα συμπεράσματα τα οποία επιβεβαιώνουν την εκτίμηση μας για το μοντέλο.

Abstract

In Publish/subscribe networks each entity subscribed to consume content and services from swarm of publishers. In such information – centrened applications trust has a new meaning because both publishers and subscribers are being evaluated. The goal of this work is to examine a model of trust management that evaluates the behaviour of the subscribers in publish/subscribe networks

A short description of publish/subscribe paradigm will be given and a more elaborated analysis of the PSIRP (Publish Subscribe Internet Routing Paradigm) system, a research project for a clean state Internet architecture based on publish/subscribe paradigm. There is a also a short analysis of security issues and a summary of a known taxonomy report for DoS (Denial of Service) attacks in publish/subscribe systems. The attacks that our model is most interested are also described in a bit more detail.

The evaluation will be based on Dempster-Shafer theory where it is possible to incorporate the uncertainty of information into decision. There is a detailed analysis of the theory itself and the mathematical model of our approach and there is also scenarios implemented for the purpose of evaluation

Result will be analysed for both centralised and distributed systems and there are also conclusions that support our estimations on the model.

Γενικά για Pubsub

Ο σχεδιασμός του Internet και των πρωτοκόλλων του προήλθε από την ανάγκη διαμοιρασμού πόρων οι οποίοι ήταν ιδιαίτερα ακριβοί. Το μοντέλο επικοινωνίας βασίστηκε στην διασύνδεση δύο οντοτήτων με σκοπό την ανταλλαγή δεδομένων για χρήση κάποιας υπηρεσίας ή αποθήκευσης δεδομένων. Η αλήθεια είναι ότι έχουν περάσει πάνω από 40 χρόνια περίπου από τότε και η συγκεκριμένη αρχιτεκτονική αποδείχθηκε ιδιαίτερα επιτυχημένη ενώ ο σχεδιασμός ήταν απλός. Τα πακέτα IP ουσιαστικά τις διευθύνσεις του αποστολέα και του παραλήπτη καθώς επίσης και τα δεδομένα προς ανταλλαγή. Γενικά η αρχή του IP είναι απλή και βασίζεται στο ότι δεν έχει απαιτήσεις από το επόμενο επίπεδο OSI και δεν παρέχει αξιόπιστη μετάδοση, ούτε παράδοση με τη σειρά αποστολής αλλά βασίζεται στην καλύτερη δυνατή προσπάθεια μετάδοσης χωρίς περιορισμούς.

Σήμερα όμως οι βασικές ανάγκες έχουν αλλάξει. Οι υπολογιστικές μηχανές αποτελούν ένα φθηνό αγαθό στο οποίο έχουν πρόσβαση όλοι και οι υπηρεσίες που κάποτε παρεχόταν μόνο από μεγάλες εταιρείες έχουν σχεδόν μηδενίσει το κόστος τους συμπεριλαμβανομένου του αποθηκευτικού χώρου και της επεξεργαστικής ισχύος. Σε συνδυασμό με την άνθηση του Internet αυτό οδήγησε πολύ κόσμο να ασχοληθεί με τον χώρο της πληροφορικής, νέες ανάγκες δημιουργήθηκαν και νέες λύσεις προέκυψαν πάνω στην παλιά αρχιτεκτονική. Πολλές φορές οι λύσεις αυτές δούλεψαν (για πολλά χρόνια) ικανοποιητικά ενώ άλλες φορές αντιμετώπισαν ένα πρόβλημα δημιουργώντας ένα άλλο. Πολλοί υποστηρίζουν ότι ο βασικός λόγος είναι ότι ο σχεδιασμός του Internet έγινε χωρίς να μπορεί να υπάρχει πρόβλεψη για τις σημερινές ανάγκες και λύσεις. Επιπλέον είναι γενικά αποδεκτό ότι η αξία του διαδικτύου βρίσκεται στο περιεχόμενο που μπορεί να βρει κάποιος και όχι στο που ακριβώς θα το βρει. Άρα θα μπορούσαμε να πούμε ότι δεν μας ενδιαφέρει το «που» αλλά το «τι». Η διαφοροποίηση των αναγκών όπως και η συγκεκριμένη αλλαγή αντίληψης οδηγεί σε μερικά θέματα που επηρεάζουν τους χρήστες όπως η διαθεσιμότητα, η ασφάλεια και η εξάρτηση από το που βρίσκεται το περιεχόμενο. Ο Jacobson (παπερ11) υποστηρίζει ότι η ιδέα του «επώνυμου» περιεχομένου είναι καλύτερη για τις σημερινές ανάγκες σε σχέση με την ιδέα των «επώνυμων» δικτύων. Η πρακτική υλοποίηση που υποστηρίζει συνοψίζεται στο σύστημα CCN το οποίο ακολουθεί αρκετά το publish/subscribe παράδειγμα. Μια άλλη πρόταση που υποστηρίζει τη σημασία του περιεχομένου σε σχέση με τα σημεία σύνδεσης και τοποθεσίας αυτού είναι η πρόταση του PSIRP (Publish Subscribe Internet Routing Paradigm) η οποία αναλύεται σε επόμενη ενότητα.

Το publish/subscribe είναι ένα παράδειγμα αρχιτεκτονικής που βασίζεται στην ιδέα της επικοινωνίας που είναι προσανατολισμένη στο περιεχόμενο αντίθετα με την ιδέα που επικρατεί σήμερα στο Internet που προσανατολίζεται στις δικτυακές συνδέσεις από άκρο σε άκρο.

Οι οντότητες που παρέχουν περιεχόμενο καθορίζουν και τον τρόπο αναζήτησης και πρόσβασης σε αυτό. Οι οντότητες που τους ενδιαφέρει να αποκτήσουν περιεχόμενο εκδηλώνουν ενδιαφέρον με βάση κάποια κριτήρια. Το σύστημα είναι υπεύθυνο να ταιριάξει το περιεχόμενο που υπάρχει και έχει καταχωρηθεί με αυτό που εκδηλώθηκε ενδιαφέρον και στη συνέχεια να φέρει σε επαφή τις ενδιαφερόμενες οντότητες με διάφορους τρόπους (απευθείας, με χαλαρή σύνδεση, σύγχρονα/ασύγχρονα κλπ.). Αυτό επίσης επιτυγχάνεται με την ιδέα των ειδοποιήσεων όπου γίνεται ασύγχρονα το ταίριασμα περιεχομένου/κριτηρίων και παρέχει μια χαλαρή σύνδεση μεταξύ των οντοτήτων που παρέχουν περιεχόμενο με τις οντότητες που το καταναλώνουν.

Γενικά ένα στοιχειώδες σύστημα publish/subscribe αποτελείται από τους δημιουργούς/παροχείς περιεχομένου που ονομάζονται publishers, τους καταναλωτές/χρήστες περιεχομένου που ονομάζονται subscribers και τις ενδιάμεσες υπηρεσίες που αποτελούν την υποδομή του συστήματος που ονομάζονται brokers. Το σύστημα παρέχει καταρχήν τρεις βασικές υπηρεσίες για την εκδήλωση ενδιαφέροντος περιεχομένου (subscribe), την ανακοίνωση του περιεχομένου που υπάρχει (advertise) και την παροχή περιεχομένου στο σύστημα (publish). Οι δύο τελευταίες συνδυάζονται σε μία σε πολλά publish/subscribe συστήματα.

Τα publish/subscribe χωρίζονται σε δύο κατηγορίες, αυτά που βασίζονται στο θέμα (topic bases) και αυτά που βασίζονται στο περιεχόμενο (content based). Στην πρώτη κατηγορία οι publishers καθορίζουν σε ποιο θέμα ανήκει το περιεχόμενο που βάζουν στο σύστημα από μια κατηγορία ή ομάδα θεμάτων που καθορίζει το ίδιο το σύστημα. Οι subscribers στη συνέχεια πρέπει να δηλώσουν ενδιαφέρον για τις ομάδες που θέλουν και αρχίζουν να λαμβάνουν όλες τις ειδοποιήσεις που δημιουργούνται όταν παρέχεται νέο περιεχόμενο στα προκαθορισμένα θέματα. Στα συστήματα που βασίζονται στο περιεχόμενο, οι subscribers εξειδικεύουν το ενδιαφέρον τους με κριτήρια ή φίλτρα και λαμβάνουν ειδοποιήσεις που ταιριάζουν με τις τιμές που έχουν δώσει σε αυτά τα κριτήρια. Επίσης δεν υπάρχουν προκαθορισμένα θέματα αλλά προκύπτουν από το τι δηλώνουν σαν πεδία χαρακτηρισμού (metadata, tags) οι παροχείς

περιεχομένου.

Η υλοποίηση ενός συστήματος publish/subscribe στηρίζεται στην ιδέα του broker ο οποίος είναι εκείνη η οντότητα στο δίκτυο η οποία έχει την ευθύνη να δρομολογήσει τα συμβάντα και το περιεχόμενο στο δίκτυο μεταξύ των publishers και των subscribers. Η υλοποίηση αυτή μπορεί να γίνει με κεντρικοποιημένο τρόπο ή με κατανεμημένο τρόπο. Στον κεντρικοποιημένο τρόπο υπάρχει ένας μόνο broker ο οποίος λαμβάνει ειδοποιήσεις για νέο περιεχόμενο και τις ταιριάζει με όλα τα αιτήματα ενδιαφέροντος όλων των καταναλωτών περιεχομένου στο σύστημα. Στον κατανεμημένο τρόπο υπάρχουν περισσότεροι brokers οι οποίοι είναι υπεύθυνοι για ένα μέρος μόνο των αιτημάτων ενδιαφέροντος και κατ' επέκταση ένα μέρος των subscribers. Στον κατανεμημένο τρόπο είναι δυνατό να υπάρξει και προώθηση μηνυμάτων μεταξύ των brokers όταν αυτό κρίνεται απαραίτητο και υποστηρίζεται από το σύστημα. Δηλαδή μπορεί να γίνει αρχική δημιουργία ενός μηνύματος σε έναν broker και αυτός με τη σειρά του να προωθήσει το μήνυμα είτε σε όλους τους υπόλοιπους είτε μόνο στους γείτονες του ανάλογα με τον σχεδιασμό και την πολιτική που ακολουθείται. Επιπλέον μπορούν να υπάρχουν και κριτήρια ή φίλτρα ή κάποιος αλγόριθμος ταιριάσματος για το ποιοι γείτονες πρέπει να λάβουν την προώθηση μειώνοντας έτσι το φόρτο του δικτύου.

Γενικά εξαιτίας των ομοιοτήτων που έχουν τα P2P συστήματα με τα συστήματα publish/subscribe υπάρχουν υλοποιήσεις πάνω από P2P overlays με τις πιο γνωστές να είναι πάνω από Pastry και CAN. Στο Pastry ο κάθε κόμβος έχει μια μοναδική ταυτότητα και η δρομολόγηση γίνεται με βάση έναν άπληστο αλγόριθμο. Το κάθε μήνυμα δρομολογείται στον κόμβο που βρίσκεται πιο κοντά αριθμητικά με βάση το κλειδί. Το σύστημα SCRIBE επιπλέον δημιουργεί ένα δίκτυο multicast πάνω από το Pastry και βασίζεται σε αυτό για την βελτιστοποίηση της δρομολόγησης των πακέτων. Το σύστημα CAN δημιουργεί ένα εικονικό χώρο Χ διαστάσεων που προκύπτει ως το καρτεσιανό γινόμενο συντεταγμένων που χωρίζεται δυναμικά σε όλους τους κόμβους στο σύστημα. Ένα μειονέκτημα και των δύο συστημάτων είναι ότι έχουν υψηλό κόστος συντήρησης (σε πόρους) το οποίο αυξάνει με τον αριθμό των ομάδων κόμβων που δημιουργούνται και επίσης αυξάνει με τον αριθμό των δυναμικών αλλαγών στο σύστημα όπως η προσθήκη και η αφαίρεση κόμβων. Επιπλέον επειδή βασίζονται στο συγκεκριμένο τρόπο δρομολόγησης είναι κατάλληλα μόνο για publish/subscribe που βασίζονται στο θέμα και όχι στο περιεχόμενο. Μόνο όταν το P2P είναι ένα super peer τότε μπορεί σαν ελάχιστο να υλοποιηθεί μια αρχιτεκτονική content-based publish/subscribe σαν αναλογία ενός P2P δικτύου.

PSIRP

Όπως περιγράψαμε πιο πριν, η παρούσα αρχιτεκτονική που έχει επικρατήσει και που βασίζεται στο IP λειτουργεί αρκετά ικανοποιητικά για τις περισσότερες σημερινές ανάγκες αλλά έχει και κάποιους περιορισμούς. Η πιο αξιοσημείωτη είναι ότι αυτός που κατέχει την πληροφορία συνήθως βρίσκεται σε πλεονεκτικότερη θέση αφού πρέπει να υπάρχει εμπιστοσύνη για να υπάρχει και πρόσβαση στη πληροφορία. Το δίκτυο με τη σειρά του αποδέχεται ότι του στέλνει ο αποστολέας και κάνει την καλύτερη δυνατή προσπάθεια να στείλει την πληροφορία στον παραλήπτη. Αυτό οδηγεί σε προβλήματα με αυτόκλητη κίνηση στο δίκτυο (spam) κατανεμημένες επιθέσεις άρνησης υπηρεσίας (DDoS)

Ένα από τα επιχειρήματα του PSIRP είναι ότι τα προβλήματα που ανακύπτουν λόγω της υπάρχουσας αρχιτεκτονικής έχουν βαθύτερα αίτια και στηρίζονται στο γεγονός ότι η επικοινωνία προσανατολίζεται στα άκρα σύνδεσης στο δίκτυο δηλαδή στις οντότητες που επικοινωνούν, τον αποστολέα και τον παραλήπτη. Αυτή η αρχιτεκτονική, υποστηρίζουν ότι, θέτει επιπλέον περιορισμούς στην τοπολογία των δικτύων χωρίς μάλιστα να υπάρχει ομοιογένεια. Με την αύξησης μιας πληθώρας υπηρεσιών που έχουν επίκεντρο το ίδιο το περιεχόμενο η σημερινή αρχιτεκτονική δεν παρουσιάζει καμία ευελιξία. Επίσης υποστηρίζουν ότι εξαιτίας αυτού το Internet δεν έχει εκμεταλλευτεί πλήρως τις δυνατότητες που έχει.

Επιπλέον οι ειδικοί στον κόσμο (όπως υποστηρίζει το PSIRP) έχουν αρχίσει να συμφωνούν ότι χρειάζεται μια αναδιάρθρωση του Internet και των βασικών τεχνολογιών του για να αντεπεξέλθει στις σημερινές προκλήσεις. Ο στόχος του PSIRP (όπως και παρόμοιων ερευνητικών έργων πχ. CCN) είναι να αντιμετωπίσει τα προβλήματα αυτά ερευνώντας ένα νέο εναλλακτικό μοντέλο επικοινωνίας που κέντρο του είναι η ίδια η πληροφορία. Έτσι η ιδέα είναι να προσεγγίσουμε κάθε σενάριο επικοινωνίας με τις έννοιες της παραγωγής πληροφορίας, της ανάκτησης πληροφορίας και τελικά την κατανάλωση πληροφορίας μεταξύ οντοτήτων που επικοινωνούν μεταξύ τους. Η έννοια της πρόθεσης είναι σημαντική για την δημιουργία μιας επικοινωνίας όπου τις οντότητες τις ενδιαφέρει η συμφωνία των κοινών προθέσεων (δηλαδή τι περιεχόμενο θέλουν) και όχι τα άκρα του δικτύου.

Υπάρχουν αρκετά παραδείγματα εφαρμογών που η ιδέα τους είναι βασικά publish/subscribe όπως η προώθηση ενημερώσεων λογισμικού, ενημερώσεις ειδήσεων, γενική μετάδοση πολυμεσικού περιεχομένου όπως βίντεο και ήχος και άλλα. Έτσι φαίνεται θετική η

διερεύνηση για μία νέα αρχιτεκτονική στο Internet η οποία να βασίζεται στο παράδειγμα publish/subscribe όπως εγγενή οι εφαρμογές αυτές.

Δικτυακή αρχιτεκτονική

Η εννοιολογική αρχιτεκτονική του PISRP αποτελείται από τις ακόλουθες οντότητες

- **αναγνωριστικά/ταυτότητες**: όπως έχει ήδη συζητηθεί, υπάρχουν τρεις ή τέσσερις κλάσεις ταυτοτήτων, ταυτότητες προώθησης, ταυτότητες εύρους, ταυτότητες ραντεβού και ταυτότητες εφαρμογής. ΟΙ ταυτότητες εφαρμογής αναλύονται στις ταυτότητες ραντεβού και τελικά στις ταυτότητες προώθησης και ονομάζονται πολλές φορές και ετικέτες. Αυτή η ανάλυση γίνεται από τα πρωτόκολλα και οι εφαρμογές ασχολούνται κυρίως με τις ταυτότητες εφαρμογών αλλά και τις ταυτότητες ραντεβού όταν θέλουν σύνδεση με το δίκτυο.
- **δεδομένα και μέτα-δεδομένα**: Μία ταυτότητα ραντεβού, σχετίζεται με ένα σύνολο δεδομένων ρητά ορισμένο και αποτελείται από μια ή περισσότερες εκδόσεις. Τα σύνολα αυτά μπορεί να έχουν συσχετισμένα μέτα-δεδομένα τα οποία μπορούν να περιλαμβάνουν πληροφορίες όπως μετα-δεδομένα εύρους ή άλλες χρήσιμα δεδομένα για τους παραλήπτες ή για το δίκτυο. Τα ίδια τα μετα-δεδομένα θεωρούνται δεδομένα και είναι ενσωματωμένα μέσα στην έκδοση/περιεχόμενο που σχετίζονται ή εναλλακτικά παρέχονται σαν αυτόνομη έκδοση περιεχόμενο στην ταυτότητα ραντεβού.
- **Πληροφορίες εύρους**: οι πληροφορίες εύρους είναι συσχετισμένες με μία έκδοση περιεχομένου μέσω της ταυτότητας εύρους. Η έννοια του εύρους αποφασίζει τα στοιχεία του συστήματος ραντεβού τα οποία μπορούν να έχουν πρόσβαση στα δημοσιευμένα δεδομένα και εμμέσως καθορίζουν το δίκτυο στο οποία τα δεδομένα αυτά ανήκουν. Μια έκδοση περιεχομένου μπορεί να ανήκει (ή να έχει) σε ένα ή περισσότερες ταυτότητες εύρους.
- **publishers και subscribers**: είναι ουσιαστικά οι οντότητες οι οποίες δημοσιεύουν περιεχόμενο ή καταναλώνουν περιεχόμενο αντίστοιχα.
- **περιοχές κυριότητας (domain)**: είναι οι δικτυακές περιοχές που συνιστούν μια δικτυακή ομάδα αυτόνομη με τις πολιτικές της και μπορούν να συνδεθούν μεταξύ τους με της διαδικτυακή αρχιτεκτονική.

Διεργασίες

Είναι απαραίτητο να αναφέρουμε τις παρακάτω διεργασίες του μοντέλου που μόλις περιγράψαμε:

- **Ραντεβού**: ουσιαστικά αναλύει τις ταυτότητες ραντεβού σε ταυτότητες προώθησης μέσα στο καθορισμένο και δεδομένο εύρος. Το εύρος καθορίζει ουσιαστικά το μέρος του δικτύου που είναι διαθέσιμο και χρησιμοποιείται από το δίκτυο. Η ταυτότητα ραντεβού και εύρους καθορίζουν μοναδικά την συγκεκριμένη έκδοση περιεχομένου και τα μετα-δεδομένα της όπως και τις συσχετισμένες πολιτικές. Υπάρχουν τρεις βασικές τοπολογίες
    - Τοπική σύνδεση
    - Εσωτερική (intra-domain) περιοχή κυριότητας
    - Εξωτερική ή διαδικτυακή (inter-domain) περιοχή κυριότητας
- **Εσωτερική δρομολόγηση και προώθηση**: αναφέρεται στη παράδοση των δεδομένων σε μία διαχειριστική περιοχή κυριότητας και βασίζεται σε τοπικές πολιτικές κυρίως.
- **Εξωτερική ή διαδικτυακή δρομολόγηση και προώθηση**: αναφέρεται στην παράδοση δεδομένων στα παγκόσμια δίκτυα και γενικά διατρέχει περισσότερες περιοχές ή δίκτυα. Διαχειρίζεται από την διαδικασία ραντεβού και εφαρμόζονται οι διαδικτυακές πολιτικές.
- **Προώθηση και μεταφορά**: αναφέρεται στην μεταφορά δεδομένων μεταξύ subscribers και publishers.
- **Προσωρινή αποθήκευση**: είναι μια δικτυακή διεργασία που παρέχεται από το τοπικό σύστημα ή από οποιοδήποτε σύστημα από τα διασυνδεδεμένα στο σχετικό δίκτυο
- **Δικτυακή προσάρτηση**: είναι η διεργασία που είναι υπεύθυνη για την ανακάλυψη των σημείων σύνδεσης ή προσάρτησης στο δίκτυο και την ρύθμιση των επιμέρους μερών για την επιτυχημένη πρόσβαση στο δίκτυο.

Δικτυακές υπηρεσίες

Στο PSIRP παρέχει τρία είδη υπηρεσιών για την πρόσβαση των εφαρμογών στο δίκτυο ή την πρόσβαση των εργαλείων διαχείρισης του δικτύου με το ίδιο το δίκτυο.

- **publisher/sender**: η διεπαφή υποστηρίζει ουσιαστικά την δημοσίευση περιεχομένου και καθορίζει πως τα δεδομένα αποστέλλονται πάνω από το δίκτυο και τι στοιχεία προσφέρονται από το δίκτυο γι αυτό το σκοπό. Κάθε δημοσίευση περιεχομένου

σχετίζεται με μία απλή ετικέτα και ονομάζεται ταυτότητα ραντεβού και περιέχει προαιρετικά και άλλα δεδομένα (μετα-δεδομένα). Ο ρόλος των μετα-δεδομένων θα μπορούσε να καθορίζει για παράδειγμα μία πολιτική χρήσης.

- **Subscriber/receiver**: η διεπαφή ορίζει μία επικοινωνία που βασίζεται στον παραλήπτη (ή καταναλωτή της πληροφορίας) μέσω μίας ταυτότητας ραντεβού και αρχίζει ουσιαστικά με την δήλωση της συνδρομής (εκδήλωση ενδιαφέροντος για περιεχόμενο). Όπως και προηγουμένως ο παραλήπτης/καταναλωτής μπορεί να δηλώσει προαιρετικά μετα-δεδομένα μαζί με την δήλωση της συνδρομής.
- **Υπηρεσίες δικτύου**: Παρέχει υπηρεσίες διαχείρισης και ελέγχου καθώς επίσης και εργαλεία μέτρησης. Βασική είναι η υποστήριξη υπηρεσιών καταλόγου απαραίτητες για την εκκίνηση επικοινωνιών και υλοποιούνται από το υποσύστημα ραντεβού. Από την άποψη της διαχείρισης έχουμε δύο φάσεις, την φάση ραντεβού και την φάση προώθησης. Στην πρώτη γίνεται εγκαθίδρυση των δεδομένων προώθησης και στην δεύτερη η μεταφορά δεδομένων με τα στοιχεία της πρώτης. Η φάση ραντεβού είναι η πιο αργή γιατί διαχειρίζεται πολιτικές και επεξεργασία μετα-δεδομένων.

Περιγραφή προβλημάτων ασφαλείας

Ένα πολλά υποσχόμενο παράδειγμα για εφαρμογές μελλοντικού Internet ή ακόμα και νέα αρχιτεκτονική για το μελλοντικό Internet είναι το παράδειγμα του publish/subscribe. Γενικά ένα publish/subscribe σύστημα αποτελείται από τους publishers (εκδότες περιεχομένου), τους subscribers (τους καταναλωτές περιεχομένου) και τους brokers ή κόμβους δρομολογητές. Επιπλέον όπως περιγράψαμε πιο πριν στο PSIRP υπάρχει και η έννοια του scope που ομαδοποιεί οντότητες, περιεχόμενο και πολιτικές ασφαλείας. Το περιεχόμενο που δημοσιεύεται από τους εκδότες είναι περιορισμένο ανάλογα με το scope το οποίο είναι με τη σειρά του μια οντότητα από μόνο του.

Επίσης για λόγους ταυτοποίησης και εξουσιοδότησης όπως για λόγους πρόσβασης κατά την διάρκεια σύνδεσης μιας οντότητας στο δίκτυο, η κάθε οντότητα έχει ένα μοναδικό χαρακτηριστικό το οποίο είναι γενικό και δεν εξαρτάται από το που βρίσκεται στο δίκτυο.

Οι publishers είναι ουσιαστικά παροχείς περιεχομένου και «διαφημίζουν» πληροφορίες, υπηρεσίες και περιεχόμενο. Οι subscribers είναι καταναλωτές τέτοιων πληροφοριών και εκφράζουν το ενδιαφέρον τους για περιεχόμενο που έχει δημοσιευτεί. Οι brokers ταιριάζουν το ενδιαφέρον των subscribers με το περιεχόμενο των publishers και δρομολογούν την σύνδεση ώστε το περιεχόμενο να προωθηθεί στους subscribers. Επίσης ο κόμβος που γίνεται τελικά το ταίριασμα αυτό και είναι υπεύθυνος για την ενεργοποίηση της σύνδεσης και το μονοπάτι που θα ακολουθήσει το περιεχόμενο προς τον subscriber λέγεται Rendezvous point ή σημείο ραντεβού.

Το κεντρικό σημείο σε ένα publish/subscribe σύστημα είναι το περιεχόμενο ή όπως ονομάζεται ατομικά, μία έκδοση. Το περιεχόμενο όπως έχει ήδη αναφερθεί, περιγράφεται από το όνομα του και επιπλέον πεδία (metadata) έτσι ώστε να είναι δυνατή η αναζήτησή τους από τους subscribers. Τα στοιχεία και τα επιπλέον πεδία περιγραφής του περιεχομένου (ή της έκδοσης) καθορίζουν ουσιαστικά και την δρομολόγηση σύμφωνα με τους κανόνες σχεδιασμού.

Η μοντελοποίηση εμπιστοσύνης σε ένα σύστημα συναντάται με διάφορες μορφές στις προτεινόμενες λύσεις. Κάποιοι χρησιμοποιούν την εμπιστοσύνη μέσω πινάκων όπου σε ένα πίνακα καταχωρούν αυτούς που γενικά θεωρούνται αξιόπιστοι και σε άλλο πίνακα αυτούς

που δεν θεωρούνται αξιόπιστοι. Αυτό γίνεται και σε συνδυασμό ή με γενικές πολιτικές όπως δεν επιτρέπεται πρόσβαση σε αυτούς που δεν είναι στον πρώτο πίνακα (pessimistic) ή επιτρέπεται πρόσβαση σε αυτούς που δεν είναι στον δεύτερο πίνακα (optimistic). Αποτίμηση εμπιστοσύνης μπορεί να γίνει και με την μορφή κάποιου κοινού κωδικού ή εμπιστευτικής πληροφορίας έτσι ώστε να αποφεύγονται επιθέσεις άρνησης υπηρεσίας αφού αποκλείονται οι μη έμπιστες οντότητες. Για την αποφυγή επιθέσεων που η κακόβουλη οντότητα προσπαθεί να εμφανιστεί με περισσότερες ταυτότητες στο σύστημα υπάρχουν προτάσεις υλοποίησης με έμπιστα δένδρα δρομολόγησης. Γενικά οι λύσεις που προτείνονται έχουν πλεονεκτήματα και μειονεκτήματα αλλά και θέματα κλιμάκωσης σε μεγάλα και ανώνυμα δίκτυα.

Για να σχεδιαστεί μια υποδομή διαχείρισης εμπιστοσύνης είναι απαραίτητο να μοντελοποιηθεί και να αποτιμηθεί η συμπεριφορά των οντοτήτων που συμμετέχουν στο δίκτυο. Στη βιβλιογραφία μπορεί κάποιος να βρει αναφορές για αποτίμηση εμπιστοσύνης στους εκδότες περιεχομένου καθώς και σχήματα πολιτικής επιβολής αυτών των μοντέλων. Το ίδιο ισχύει και για τις οντότητες που απαρτίζουν το δίκτυο υποδομής δηλαδή τους brokers όπου υπάρχουν διάφορα συστήματα που καθορίζουν τον βαθμό ασφάλειας και εμπιστοσύνης μαζί με τις υποδομές δημοσίου κλειδιού. Σχετικά με τους subscribers υπάρχει περιγραφή προβλημάτων που μπορούν να προκαλέσουν αλλά δεν υπάρχει κάποια αποτίμηση συμπεριφοράς που μπορεί να αποτελέσει βάση δημιουργίας μια πολιτικής εμπιστοσύνης ή ασφάλειας.

Στη παρούσα εργασία γίνεται μια προσπάθεια να αποτιμηθεί η συμπεριφορά καταναλωτών πληροφορίας/περιεχομένου με βάση ένα μαθηματικό μοντέλο ευλογοφάνειας. Το μοντέλο αυτό δίνει τη δυνατότητα να υπάρξει μια πρώτη προσέγγιση συμπεριφοράς των subscribers δίνοντας ταυτόχρονα την δυνατότητα περαιτέρω διερεύνησης τόσο για την ίδια τη συμπεριφορά όσο και τις πολιτικές που μπορούν να αξιοποιηθούν με βάση τα αποτελέσματα. Υπάρχουν παραδείγματα για συμπεριφορές που μπορούν να οδηγήσουν σε επιθέσεις άρνησης υπηρεσιών (DoS attacks), επιθέσεις που προέρχονται είτε από publishers είτε από subscribers.

Επίσης αναφέρονται και κάποιες καλές συμπεριφορές όπως η ταυτοποίηση των κόμβων μπορεί να επιτευχθεί με υποδομή δημοσίου κλειδιού, η ακεραιότητα των μηνυμάτων (και του περιεχομένου) μπορεί να επαληθευτεί με ψηφιακές υπογραφές και έλεγχο πρόσβασης ενώ η λογοδοσία αναφέρεται σε εφαρμογές καταγραφής συμβάντων και δοσοληψιών.

Δεν θα γίνει αναφορά καθόλου σε επιθέσεις που σχετίζονται πρωτίστως με publishers όπως εισαγωγής στο σύστημα περιεχομένου που δεν ανταποκρίνεται στην περιγραφή του, είτε ημιτελούς περιεχομένου ή μη παράδοση περιεχομένου που έχει γίνει αναφορά. Αυτά οδηγούν σε υπερφόρτωση του δίκτυο και την μη ικανοποίηση των subscribers. Επίσης δεν θα γίνει αναφορά σε περιπτώσεις όπου γίνεται «διαφήμιση» περιεχομένου που δεν ενδιαφέρει κανένα subscriber ή που ενδιαφέρει πάρα πολλούς (spam). Ακόμα δεν γίνεται αναφορά στο επίπεδο εμπιστοσύνης στην υποδομή του συστήματος, δηλαδή στους brokers. Το σύστημα θεωρείται ασφαλές και αξιόπιστο για την μελέτη μας και αυτό είναι κοντά στην πραγματικότητα όσο αφορά στο σύστημα PSIRP με τις προδιαγραφές αρχιτεκτονικής και υλοποίησης.

Security Requirements

Υπάρχουν δύο διαφορετικές απαιτήσεις ασφάλειας σε ένα publish/subscribe σύστημα. Αυτές είναι:

- η εφαρμογή (application), που συνίσταται από τους publishers και τους subscribers και μπορεί να μην εμπιστεύονται ο ένας τον άλλο αλλά και το ίδιο το δίκτυο και οι απαιτήσεις υποδομής όπου δεν υπάρχει εμπιστοσύνη απέναντι στην υποδομή και
- η υποδομή (infrastructure), που αποτελείται από το publish/subscribe δίκτυο που παρέχει υπηρεσίες στην εφαρμογή. Η υποδομή μπορεί να μην εμπιστεύεται τους publishers και τους subscribers και όλα τα συστατικά μέρη της υποδομής να μην εμπιστεύονται το ένα το άλλο.

Generic Issues

Στα συστήματα publish/subscribe υπάρχουν κάποια θέματα ασφαλείας διαφορετικά από αυτά που εμφανίζονται στα άλλα κατανεμημένα συστήματα όπου υπάρχουν domains διαχείρισης. Αυτά είναι:

- Authentication.
  Η ταυτοποίηση εγκαθιστά την ταυτότητα του δημιουργού μιας ενέργειας.
- Information integrity
  Η ακεραιότητα της πληροφορίας πετυχαίνεται μέσω των ηλεκτρονικών υπογραφών.
- Subscription integrity
  Σε ένα publish/subscribe σύστημα οι συνδρομές των χρηστών διατηρούνται στο δίκτυο και για αυτό θα πρέπει να προστατεύονται από μη εξουσιοδοτημένες αλλαγές.
- Service integrity

Η ακεραιότητα των υπηρεσιών μπορεί να βρεθεί σε κίνδυνο αν βρεθεί κάποιο ελάττωμα ή αδυναμία στο επίπεδο υποδομής.

– User Anonymity

  Η ανωνυμία των χρηστών μπορεί να επιτευχθεί με ποικίλες τεχνικές ανωνυμίας που αναπτύσσονται σε κατανεμημένα συστήματα.

Confidentiality

Σε ένα publish/subscribe εισάγονται τρία επίπεδα εμπιστευτικότητας:

– Information confidentiality

  Όταν δημοσιεύεται μια πληροφορία εμπεριέχει ευαίσθητα περιεχόμενα, εκδότες, και συνδρομητές που επιθυμούν να παραμείνει η πληροφορία εμπιστευτική από το publish.subscribe σύστημα.

– Subscription confidentiality

  Οι συνδρομές των χρηστών μπορούν να αποκαλύπτουν ευαίσθητες πληροφορίες σχετικά με τον χρήστη, όπου και σε αυτήν την περίπτωση ο συνδρομητής επιθυμεί να παραμείνουν αυτές οι πληροφορίες εμπιστευτικές.

– Publication confidentiality

  Στα συστήματα publish/subscribe υπάρχουν εφαρμογές στις οποίες οι εκδότες δεν ενδιαφέρονται να γνωρίζουν την ταυτότητα των συνδρομητών. Υπάρχουν όμως και άλλες εφαρμογές στις οποίες είναι πολύ σημαντικό να μείνουν εμπιστευτικές οι δημοσιεύσεις από συνδρομητές που δεν είναι νόμιμοι.

Accountability

Η λογοδοσία των συνδρομητών στα συστήματα publish/subscribe μπορεί να γίνει με τους παρακάτω τρόπους:

– One-of-band solutions

  Ο πιο εμφανής τρόπος λογοδοσίας είναι να χρεώνουν οι εκδότες τους συνδρομητές πουλώντας ένα κλειδί το οποίο αποκρυπτογραφεί τα επιλεγμένα δεδομένα.

– Infrastructure-based accountability

  Αν τόσο οι εκδότες όσο και οι συνδρομητές εμπιστεύονται την υποδομή publish/subscribe, τότε η υποδομή μπορεί να χρεώσει τους συνδρομητές ανάλογα με το ποσό της πληροφορίας που λαμβάνουν και να πληρώνει τους εκδότες αναλόγως με την

πληροφορία που παρέχουν.

– Auditability

  Αν χρησιμοποιείται η υποδομή για την χρέωση και την πληρωμή των συνδρομητών και των εκδοτών αντίστοιχα τότε θα πρέπει να κρατείται αρχείο παρακολούθησης / καταγραφής.

Availability

Όπως σε όλα τα συστήματα επικοινωνίας υπάρχει περίπτωση λόγο κάποιων επιθέσεων να μην είναι διαθέσιμο το σύστημα. Για το λόγο αυτό θα πρέπει να ληφθούν κάποια προληπτικά μέτρα.

– Limited publication

  Η περιορισμένη έκδοση είναι ένα μέτρο που απλά μειώνει το μέγεθος κάθε έκδοσης.

– CPU-cycle-based payment scheme

  Σενάριο πληρωμής με βάση τη χρήση της CPU

– Customized publication control

  Επιτρέπει στους συνδρομητές να καθορίσουν ποιοι εκδότες μπορούν να εκδώσουν πληροφορίες.

Περιγραφή taxonomy

'Το παρακάτω είναι περίληψη του «A Taxonomy for Denial of Service Attacks in Content-based Publish/Subscribe Systems» και έχει συμπεριληφθεί για λόγους πληρότητας'

Η φύση των CPS συστημάτων να διατηρούν κάποια δεδομένα για την κατάσταση των οντοτήτων τους (stateful) έχει αθροιστικά αποτελέσματα σε επιθέσεις DoS.

Σε αντίθεση με τη σταθερή μορφή των πακέτων που χρησιμοποιείται στα χαμηλότερα επίπεδα του δικτύου, τα CPS μηνύματα είναι πολύ ευέλικτα ως προς το περιεχόμενο. Παρόλα αυτά, το όφελος της ευελιξίας και της σημασιολογίας των μηνυμάτων οδηγεί σε πιθανότητα για DoS επιθέσεις. Αυτό συμβαίνει γιατί ο φόρτος μιας επίθεσης μπορεί να είναι πολύ διαφορετικός από κάποιας άλλης με αποτέλεσμα η επεξεργασία αιτημάτων να διαφοροποιείται πολύ.

Οι χρόνοι απόκρισης δεν επανέρχονται μόλις σταματάει η DoS επίθεση επειδή τα κακόβουλα μηνύματα συσσωρεύονται στην ουρά αναμονής και συνεχίζουν να γεμίζουν με φόρτο τον broker. Στην πραγματικότητα, οι χρόνοι απόκρισης δεν επανέρχονται σε κανονικούς ρυθμούς παρά μέχρι να αρχίσει να επανέρχεται και η ουρά αναμονής στους κανονικούς ρυθμούς.

Taxonomy Summary

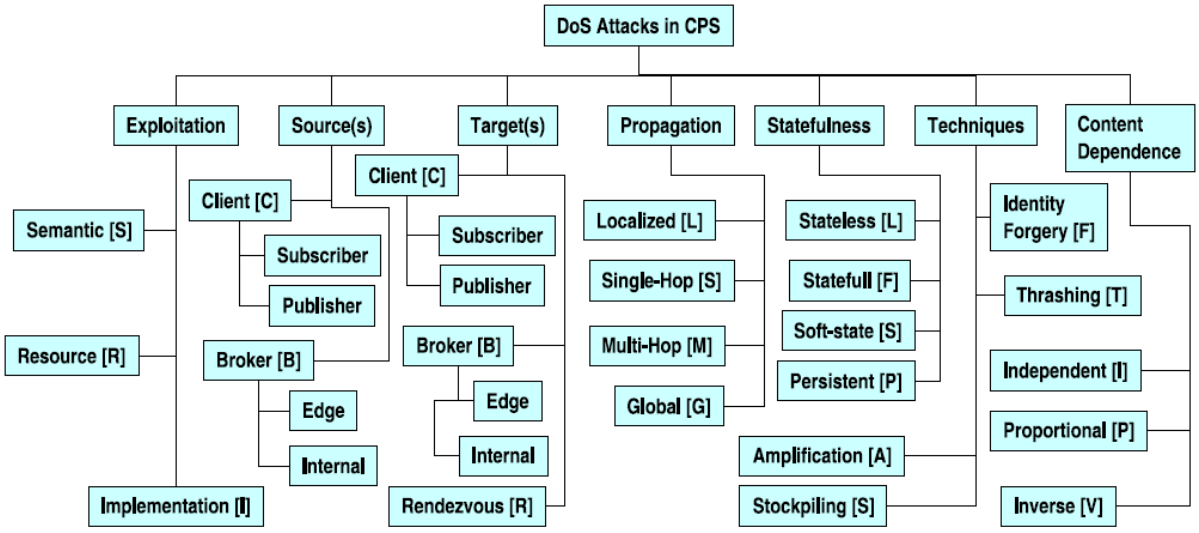


Figure 5: Taxonomy tree for Denial of Service attacks in CPS systems.

Κάθε κατηγορία είναι σημαντική γιατί κάθε κατηγορία επίθεσης απαιτεί διαφορετικό μηχανισμό άμυνας. Οι κατηγορίες που χρησιμοποιούνται στην ταξινόμησή είναι επίσης ανεξάρτητες η μία από την άλλη έτσι ώστε μια επίθεση να μπορεί να χαρακτηριστεί από μία ή

περισσότερες κατηγορίες.

Exploitation Type

Παρατηρούμε ότι μια επίθεση μπορεί να εκμεταλλευτεί διαφορετικές αδυναμίες σε ένα σύστημα έτσι ώστε να επιτύχει DoS. Σε μια επίθεση πάντα εμποδίζεται η νόμιμη προσπέλαση στις υπηρεσίες του συστήματος. Παρόλα αυτά, αν υπάρχουν διαφορετικοί τρόποι εκμετάλλευσης, πράγμα που σημαίνει ότι χρησιμοποιήθηκαν εντελώς διαφορετικές στρατηγικές επίθεσης από τον εχθρό, χρειάζεται να έχει το σύστημα διαφορετικούς μηχανισμούς άμυνας.

Αυτή η κατηγορία είναι παρόμοια με την ‘load-based versus logic-based’ κατηγοριοποίηση η οποία υπάρχει και έχει ήδη μελετηθεί για DoS.

Resource Limitations (R)

Η στρατηγική που βλέπει στον περιορισμό των πόρων είναι πολύ γνωστή από τις υπάρχουσες DoS επιθέσεις. Όμως, υπάρχουν μοναδικοί τρόποι που μπορεί ένας εχθρός να έχει σαν στόχο τον περιορισμό των πόρων σε ένα CPS σύστημα. Αρχικά, η ανάγκη να εκτελεστεί ένα ταίριασμα βασιζόμενο στο περιεχόμενο και να διατηρηθεί η κατάσταση σε ένα CPS σύστημα οδηγεί σε νέες στρατηγικές επίθεσης DoS για να προκληθεί υπερφόρτωση στο δίκτυο, στην επεξεργαστική ισχύ και στη μνήμη. Μικρότερο πλήθος από προσεχτικά δημιουργημένα μηνύματα μπορεί να είναι αποδοτικό για ένα CPS σύστημα.

Γενικά, πολύ δύσκολο για το σύστημα να αμυνθεί στις επιθέσεις που σχετίζονται με τους πόρους και για το λόγο αυτό συνήθως χρησιμοποιούνται ευριστικές αμυντικές προσεγγίσεις.

Semantic Weaknesses (S)

Μια DoS επίθεση μπορεί να χρησιμοποιήσει τη σημασιολογική αδυναμία σε ένα σύστημα λόγο ελαττωμάτων εννοιολογικού και θεωρητικού σχεδιασμού. Αυτά τα ελαττώματα μπορούν να γίνουν εργαλεία εκμετάλλευσης για να επιτευχθεί DoS χωρίς όμως να δημιουργείται ασυνήθιστη χρήση πόρων ή να παραβιάζεται η σωστή συμπεριφορά του συστήματος που πρέπει να υπάρχει από τον σχεδιασμό του.

Implementation Flaws (I)

Μια επίθεση DoS μπορεί να έχει σαν στόχο αδυναμίες εφαρμογής οι οποίες είναι

συγκεκριμένες για μια ορισμένη υλοποίηση του συστήματος. Οι σημασιολογικές αδυναμίες και οι αδυναμίες εφαρμογής πολλές φορές καλούνται με ένα όνομα ‘επιθέσεις DoS που βασίζονται στη λογική’ (logic-based). Εείναι σημαντικό να διαχωριστούν οι DoS επιθέσεις που έχουν σαν στόχο τις αδυναμίες που αφορούν τη σημασιολογία (έννοιες που είναι εφαρμόσιμες σε διαφορετικά συστήματα που όμως χρησιμοποιούν την ίδια σημασιολογία) με τις επιθέσεις εφαρμογής (που αφορούν μια συγκεκριμένη υλοποίηση του συστήματος).

Attack Source(s)

Οι επιθέσεις είναι μία ή περισσότερες οντότητες που ενεργούν υπό τον έλεγχο κάποιου εχθρού. Ο εχθρός μπορεί να στέλνει διαρκώς μηνύματα από αυτές τις οντότητες ή μπορεί να τις χρησιμοποιεί για να καταστρέφουν υπάρχοντα μηνύματα. Στα CPS συστήματα, οι επιθέσεις μπορούν να προέρχονται από πελάτες, brokers ή αν χρειάζονται πολλές οντότητες από έναν συνδυασμό των δύο. Είναι σημαντικό να κατατάξουμε τις οντότητες από τις οποίες προέρχεται μια επίθεση γιατί οι επιθέσεις που προέρχονται από διαφορετικές πηγές χρειάζονται διαφορετικούς μηχανισμούς άμυνας.

Οι περισσότερες DoS επιθέσεις προέρχονται από ένα δίκτυο κατανεμημένων πελατών σε μια μορφή κατανεμημένων DoS επιθέσεων (DDoS – Distributed DoS). Σπανίως, οι επιθέσεις προέρχονται από το εσωτερικό της δικτυακής υποδομής. Παρόλα αυτά, στα CPS συστήματα δεν μπορούμε να αγνοήσουμε την πιθανότητα ανοιχτών ή συνεργαζόμενων υποδομών, στις οποίες έμπιστοι brokers θα ενεργούν μαζί με μη έμπιστους brokers. Για το λόγο αυτό, τα CPS συστήματα αναπτύσσονται σε overlay δίκτυα που τρέχουν στην άκρη της υποδομής του διαδικτύου και για το λόγο αυτό είναι πιο ευάλωτα από τους συμβατικούς δρομολογητές.

Client-Sourced (C)

Στις περισσότερες περιπτώσεις, οι επιθέσεις που προέρχονται από τους πελάτες είναι πιο εύκολο να εξαρτηθούν από τις επιθέσεις που προέρχονται από τους brokers, αφού οι πελάτες είναι ουσιαστικά οι τελικοί χρήστες του συστήματος και πιθανώς να είναι λιγότερο αυστηρό το σύστημα γι’ αυτούς. Δηλαδή, ένας εχθρός είναι πιο πιθανό να αποκτήσει νόμιμη άδεια πελάτη σε ένα CPS σύστημα από το να αποκτήσει άδεια broker. Επίσης, εύλογα αναμένουμε ότι θα υπάρξουν πολλοί περισσότεροι πελάτες από ότι brokers και ότι στα συστήματα που ανήκουν μπορούν να είναι λιγότερο ασφαλή. Αυτό σημαίνει ότι οι κατανεμημένες επιθέσεις που στηρίζονται στην κλίμακα (π.χ. DDoS) είναι πιο πιθανό να στηρίζονται στους πελάτες και γι’ αυτό είναι ακόμα πιο πιθανό να εκτεθούν αυτοί οι πελάτες. Παρόλα αυτά, οι επιθέσεις που στηρίζονται

στους πελάτες έχουν το μειονέκτημα ότι μέσω των πελατών μπορούν να επιτεθούν στο σύστημα απλά και μόνο στέλνοντάς του μηνύματα. Δηλαδή είναι πιο πιθανό να γίνει επίθεση στέλνοντας μεγάλο όγκο μηνυμάτων.

Broker-Sourced (B)

Γενικά, είναι πιο δύσκολο για έναν εχθρό να επιτεθεί στο σύστημα μέσω ενός broker, έτσι οι σχεδιαστές των CPS συστημάτων θα πρέπει να ανησυχούν περισσότερο για τις επιθέσεις μέσω πελατών. Από την άλλη πλευρά όμως, αν κάποιος καταφέρει να επιτεθεί στο σύστημα μέσω ενός broker, τότε η επίθεση θα είναι πιο σοβαρή αφού οι brokers έχουν μεγαλύτερο έλεγχο στο σύστημα από ότι οι πελάτες.

Attack Terget(s)

Στηριζόμενοι στην ιδιαίτερη φύση των επιθέσεων DoS, ένας εχθρός μπορεί να έχει σαν στόχο brokers, πελάτες ή και τους δύο. Αν σε μια επίθεση εμπλακεί μια οντότητα που λαμβάνει μηνύματα όπου κανονικά τα μηνύματα θα διανέμονταν κατευθείαν σε αυτή, τότε λέμε ότι η επίθεση είχε σαν στόχο την συγκεκριμένη οντότητα. Θα πρέπει να σημειωθεί ότι υπάρχει διαφορά μεταξύ των οντοτήτων που γίνονται στόχος και αυτών που επηρεάζονται (γίνονται θύματα) από μια επίθεση. Τυπικά, οι συνδρομητές πάντα θα επηρεάζονται από επιθέσεις DoS, αλλά οι συνδρομητές δεν είναι πάντα ο άμεσος στόχος των επιθέσεων.

Client Targeted(C)

Από τον ορισμό μας για τους στόχους μιας επίθεσης, υπάρχει μόνο ένας τρόπος να είναι στόχος οι συνδρομητές : παρεμποδίζοντας την παράδοση μηνυμάτων ανακοίνωσης (N) από τον τελικό broker στον πελάτη. Αναλόγως με την επίθεση, αυτό μπορεί να επιτευχθεί εμποδίζοντας την ανακοίνωση, καταστρέφοντας τον ίδιο τον συνδρομητή ή ακόμα αλλοιώνοντας την κατάσταση της συνδρομής του πελάτη που φιλοξενείται από έναν εξωτερικό broker. Έχοντας σαν στόχο έναν πελάτη δε σημαίνει απαραιτήτως ότι η επίθεση θα γίνει στον ίδιο τον πελάτη. Η επίθεση σε έναν πελάτη που φιλοξενείται από κάποιον εξωτερικό broker έχει σαν στόχο και επηρεάσει όλους τους πελάτες που φιλοξενούνται από τον συγκεκριμένο broker.

Θα πρέπει επίσης να επισημανθεί ότι στη βασική γραμμή ενός CPS συστήματος, τα μηνύματα δεν διανέμονται στον εκδότη (publisher) στο CPS επίπεδο, έτσι δεν είναι πιθανό για μια επίθεση να έχει σαν στόχο publishers στην έννοια ότι έχουμε ορίσει τον στόχο μιας επίθεσης. Η καταστροφή ενός publisher έχει σαν αποτέλεσμα να αποτρέπεται η λήψη μηνυμάτων από έναν εξωτερικό broker, πράγμα που σημαίνει ότι

μια τέτοια επίθεση έχει σαν στόχο τον publisher που φιλοξενείται από έναν εξωτερικό broker.

Στις επιθέσεις που έχουν σαν στόχο πελάτες είναι πιο εύκολο να αμυνθούμε αφού το μόνο που χρειάζεται να εξασφαλίσουμε είναι το τελικό βήμα (από τον ορισμό μας) θεωρώντας ότι η δρομολόγηση μέσα στο δίκτυό μας είναι ασφαλής.

Broker-Targeted (B)

Οι επιθέσεις που έχουν σαν στόχο τους brokers είναι πολύ λιγότερο περιοριστικές από αυτές που έχουν σαν στόχο τους πελάτες. Από τη στιγμή που ένας broker είναι μέρος της υποδομής της δρομολόγησης, είναι πιθανό μια τέτοια επίθεση να επηρεάσει πολύ μεγαλύτερο αριθμό οντοτήτων. Όμως, ο εχθρός έχει μικρότερο έλεγχο πάνω στις οντότητες στις οποίες έχει επιτεθεί.

Γενικά, το να γίνει στόχος έναν broker μπορεί να επιτευχθεί ανακόπτοντας την πορεία μηνυμάτων, δρομολογώντας σε λάθος διευθύνσεις μηνύματα ή αλλοιώνοντας την κατάσταση. Το να γίνει στόχος ένας εσωτερικός broker σημαίνει ότι μπορεί να επηρεαστεί η ικανότητά του να εκτελεί ταιριάσματα και δρομολογήσεις ενώ η επίθεση σε εξωτερικούς brokers μπορεί επιπλέον να επηρεάσει την ικανότητά τους να φιλοξενούν πελάτες ή να εξυπηρετούν κατάλληλα τους φιλοξενούμενους πελάτες.

Στις επιθέσεις που έχουν σαν στόχο τους brokers, είναι πιο δύσκολο να αμυνθούμε αφού η λύση πρέπει να είναι γερή κάτω από πολλές παραλλαγές της απώλειας μηνυμάτων κατά τη διάρκεια της δρομολόγησης.

Randezvous-Targeted (R)

Ειδικά για τα CPS συστήματα που είναι χτισμένα σε δομημένα επίπεδα, μπορούμε να ξεχωρίσουμε σημεία αναφοράς (randezvous points). Παρόλο που τα σημεία αναφοράς μπορεί να είναι ή να μην είναι brokers τα ίδια, θα χρησιμοποιούμε τους όρους σημείο αναφοράς (renzdevous point) και broker αναφοράς (randevous broker) σαν συνώνυμους από εδώ και πέρα. Στις προσεγγίσεις δρομολόγησης με randevous, η δρομολόγηση μηνυμάτων προς τους randevous brokers βασίζεται σε αποτελέσματα που προέρχονται από hash συναρτήσεις των μηνυμάτων. Αν και αυτή η προσέγγιση αποφεύγει την ανάγκη για τις πλημμύρες (flooding) ανακοινώσεων, οι randevous brokers έχουν γίνει τα γενικά σημεία αναφοράς, όχι πολύ διαφορετικά από τους DNS servers και για το λόγο αυτό αποτελούν στόχους για επιθέσεις DoS. Σαν στόχους, οι randevous brokers είναι ειδική περίπτωση brokers που εκτελούν τη δρομολόγηση κάποιων θεμάτων.

Attack Propagation

Οι υπάρχουσες DoS επιθέσεις τυπικά δημιουργούνται από κακόβουλους πελάτες προς servers στόχους και είναι point-to-point. Στις περισσότερες επιθέσεις τα πακέτα απεικονίζονται από έναν τρίτο server ή δρομολογητή. Όμως, η ανάγκη να επαναληφθεί η συνδρομή στα συστήματα υπονοεί ότι τα κακόβουλα μηνύματα μπορούν να διανεμηθούν η να δρομολογηθούν μαζί με σωστά και νόμιμα μηνύματα. Από τη στιγμή που το παράδειγμα διανομής μηνυμάτων είναι ένα σημαντικό ιδιαίτερο χαρακτηριστικό γνώρισμα των CPS συστημάτων, η διάδοση των μηνυμάτων σε όλο ένα σύστημα συμπεριφέρεται πολύ διαφορετικά ανάλογα με το αν χρησιμοποιείται όπως αυτό της ελεγχόμενης υπερφόρτωσης (πλημμύρας) , randevous, ή το πολλαπλής διανομής.

Localized (L)

Μια επίθεση χαρακτηρίζεται ως τοπική (localized) αν κακόβουλα μηνύματα δεν διαδίδονται πέρα από τον αρχικό broker που βρίσκεται πρώτος. Για παράδειγμα, μια βασική επίθεση υπερφόρτωσης που αποτελείται από τυχαίες εκδόσεις που δεν ταιριάζουν με κάποια από τις συνδρομές, δεν θα διαδοθεί πέρα από τον πρώτο εξωτερικό broker που λαμβάνει αυτές τις εκδόσεις. Όμως, θα πρέπει να σημειωθεί ότι οι επιπτώσεις από μια localized επίθεση μπορεί να μην περιοριστεί με τον ίδιο τρόπο. Αναλόγως με τη σημασιολογία διάδοσης των αναιρέσεων subscriptions, τα κακόβουλα unsubscription μηνύματα μπορεί να μη διαδοθούν ποτέ σε κάποιον άλλο broker, πράγμα που σημαίνει ότι η επίθεση είναι localized. Όμως θα σταματήσουν οι αναφορές στον κάτοχο της συνδρομής που έχει απομακρυνθεί, πράγμα που σημαίνει ότι οι επιπτώσεις δεν είναι τοπικές (localized). Θα πρέπει επίσης να σημειωθεί ότι οι επιθέσεις DoS σε κεντρικοποιημένα (centralized) συστήματα publish/subscribe είναι απαραιτήτως τοπικές (localized).

Single-Hop (S)

Σε μερικά συστήματα, η διάδοση επίθεσης μπορεί να είναι περιορισμένη σε ένα βήμα, όταν η κατανομή των μηνυμάτων συμβαίνει μόνο σε δύο brokers (τον αρχικό πρώτο broker και τον επόμενο). Το εύρος της επίθεσης εξαρτάται από το πόσο απλωτά εκπέμπεται το μήνυμα και είναι εξίσου ανησυχητικό για τις δύο πηγές και για τις σημασιολογικές επιθέσεις.

Η διάδοση επίθεσης μπορεί να είναι περιορισμένη σε ένα βήμα από τον κατανεμημένο μηχανισμό που χρησιμοποιείται. Όπως για παράδειγμα, οι brokers σε ένα σύστημα.

Multi-Hop (M)

Γενικά, τα κακόβουλα μηνύματα μπορούν να διανέμονται μέσω ενός αυθαίρετου αριθμού brokers πέρα από την αρχική εισαγωγή brokers σε επιθέσεις πολλών βημάτων. Αυτές οι επιθέσεις απαιτούν περισσότερη προσοχή από ότι οι localized επιθέσεις, αφού το φιλτράρισμα είναι κάτι που θα πρέπει να αποφευχθεί από έναν εχθρό σε ένα CPS σύστημα.

Γενικά, τα subscriptions διαδίδονται όταν ταιριάζουν με διαφημίσεις και τα publications διαδίδονται όταν ταιριάζουν με τα subscriptions. Έτσι, οι επιθέσεις που χρησιμοποιούν subscription μηνύματα μπορεί λανθασμένα να επιστρέψουν πίσω στον Publishing broker, επηρεάζοντας έτσι όλους τους brokers που βρίσκονται κατά μήκος του συγκεκριμένου μονοπατιού. Σε αυτή την περίπτωση, μπορεί ένας εχθρός να χρησιμοποιήσει τη φυσική διάδοση των μηνυμάτων του συστήματος για να στοχεύσει έναν publisher broker μόνο από το περιεχόμενο. Σε αδόμητα CPS συστήματα, αν δεν υπάρξει ταίριασμα τα μηνύματα είτε καταστρέφονται είτε παραμένουν στους εξωτερικούς brokers. Όμως, σε δομημένα CPS συστήματα η σημασιολογία της διάδοσης είναι διαφορετική λόγο της χρήσης των randevous brokers. Τα μηνύματα αντί να καταστραφούν ή να παραμείνουν στους εξωτερικούς brokers, διαδίδονται ανεπιφύλακτα προς έναν ή περισσότερους randevous brokers. Συνεπώς, η κακόβουλη κίνηση σε ένα δομημένο CPS σύστημα δεν μπορεί να είναι τοπική (localized) παρά μόνο αν η κακόβουλη αυτή κίνηση εισαχθεί κατευθείαν σε έναν randevous broker.

Εκτός του εύρους, είναι πιθανό μια multi-hop επίθεση να χαρακτηρίζεται από το βάθος.

Global (G)

Τελικά, οι γενικευμένες (Global) επιθέσεις συμβαίνουν όταν η κακή κίνηση μεταδίδεται σε κάθε broker στο σύστημα. Τα συστήματα τα οποία υποστηρίζουν την global πλημμύρα μηνυμάτων είναι προφανώς ευπαθείς σε αυτές τις επιθέσεις. Αλλά είναι επίσης δυνατό να κάνουν επιπλέον επιθέσεις αν τα κακόβουλα μηνύματα φτιαχτούν έτσι ώστε να ταιριάξουν με τα global ενδιαφέροντα. Για παράδειγμα αν υπάρχει τουλάχιστον ένας subscriber σε κάθε broker που να κάνει subscribe σε κάποιο κοινό χώρο publication τότε μια έκδοση (publication) η οποία μπαίνει σε αυτό το χώρο θα μπορούσε να μεταδοθεί σε κάθε broker στο σύστημα. Τα CPS προσπαθούν να αποφύγουν τη μετάδοση global μηνυμάτων σε κανονικές καταστάσεις.

Content Dependance

Ένα σημαντικό χαρακτηριστικό των CPS συστημάτων είναι η δρομολόγηση μηνυμάτων ανάλογα με το περιεχόμενο τους. Εφόσον το περιεχόμενο ουσιαστικά αποφασίζει πως θα χειριστούν τα μηνύματα και πως θα παραδοθούν ένας εχθρός μπορεί δραστικά να αλλάξει πως μια επίθεση συμπεριφέρεται απλά αλλάζοντας το περιεχόμενο των μηνυμάτων αυτής της κακόβουλης επίθεσης. Η μετάδοση και οντότητες που είναι οι στόχοι είναι παραδείγματα χαρακτηριστικών επίθεσης η οποία ελέγχεται ουσιαστικά από διευθύνσεις host σε πολλές υπάρχουσες επιθέσεις DoS οι οποίες ελέγχονται από περιεχόμενο μηνυμάτων σε CPS συστήματα. Το αν μια επίθεση ενισχύεται από το περιεχόμενο ή όχι η ακριβής φύση ενός κακόβουλου μηνύματος περιεχομένου το οποίο χρησιμοποιείται θα αποφασίσει ουσιαστικά αν οι υπάρχοντες μηχανισμοί άμυνας είναι εφαρμόσιμοι.

Independent (I)

Οι επιθέσεις δεν εξαρτώνται στο περιεχόμενο των μηνυμάτων τα οποία μπαίνουν από έναν εχθρό. Επειδή δεν γίνεται χρήση περιεχομένου σε αυτή την περίπτωση, είναι πιο κατάλληλες οι παραδοσιακές τεχνικές ασφαλείας που βασίζονται στην ταυτοποίηση και εξουσιοδότηση.

Proportional (P)

Αν η αποτελεσματικότητα μιας επίθεσης εξαρτάται από τα μηνύματα περιεχομένου που θα φτιάξει ο εχθρός, δηλαδή να τροποποιήσει τα μηνύματα και το περιεχόμενό τους και έτσι να επηρεάσει τη δρομολόγηση, τότε η επίθεση μπορεί να είναι είτε ανάλογη είτε αντιστρόφως ανάλογη σε κάποιο μέτρο από την πολυπλοκότητα που έχει αυτό το μήνυμα περιεχομένου. Στα σημερινά CPS συστήματα ο αριθμός των γραμμών σε ένα μήνυμα είναι ενδεικτική της πολυπλοκότητας ενός μηνύματος.

Για να αμυνθεί το σύστημα σε τέτοιες επιθέσεις περιεχομένου χρειάζεται να γνωρίζει το περιεχόμενο του μηνύματος. Όμως η ανίχνευση εισβολής που βασίζεται στο περιεχόμενο είναι ένα πολύ δύσκολο πρόβλημα. Μια ακριβής μέτρηση της πολυπλοκότητας του περιεχομένου ίσως να μπορεί να βοηθήσει να προβλεφθούν οι επιπτώσεις και να υπάρξει μια άμυνα απέναντι σε αυτές τις επιθέσεις.

Inversely Proportional (V)

Αν η σοβαρότητα μιας επίθεσης αυξάνει όταν ο χρήστης χρησιμοποιεί μεγαλύτερο όγκο κακόβουλων μηνυμάτων λιγότερης όμως πολυπλοκότητας τότε η επίθεση

χαρακτηρίζεται ως επίθεση αντιστρόφως ανάλογη της πολυπλοκότητα του περιεχομένου. Αρχικά μπορεί να φαίνεται ότι όλοι οι πόροι της επίθεσης εξαρτώνται αναλογικά από την πολυπλοκότητα του περιεχομένου αλλά αυτό εξαρτάται από το πως χτίζονται αυτές οι επιθέσεις αφού ο φόρτος μπορεί να προκληθεί βάζοντας πολύ μεγάλο ποσοστό γενικών subscriptions. Για παράδειγμα, στα σημερινά CPS συστήματα τα γενικά subscriptions συνήθως τείνουν να προσελκύσουν πολλά publications έχοντας χαμηλή πολυπλοκότητα. Συνεπώς οι ανακοινώσεις μπορούν να προσελκύσουν περισσότερα subscriptions έχοντας υψηλότερα σχήματα πολυπλοκότητας και να καλύψουν μεγαλύτερο εύρος subscriptions.

Οπότε, αν τα συστήματα γνωρίζουν το περιεχόμενο ή να έχουν μία ακριβής μέτρηση της πολυπλοκότητας του περιεχομένου θα τους βοηθήσει στην άμυνα απέναντι σε αυτές τις επιθέσεις.

Statefulness of effects

Αφού ο εχθρός σταματήσει όλη την άμεση δραστηριότητα μιας επίθεση DoS τότε είναι πιθανόν τα θύματα να συνεχίσουν να υποφέρουν από τις επιπτώσεις της επίθεσης για ένα χρονικό διάστημα. Οι επιπτώσεις της επίθεσης μπορεί τελικά να εξαφανιστούν ή το σύστημα μπορεί να εκτελέσει κάποιο μηχανισμό αποκατάστασης για να ξαναέρθει στη σωστή του κατάσταση. Ο βαθμός με τον οποίο οι επιπτώσεις μιας επίθεσης DoS παραμένουν στο σύστημα δίνει ένα είδος μέτρου της προσπάθειας που χρειάζεται ένας εχθρός να δημιουργήσει μια επίθεση.

Εκτός από της επιθέσεις που είναι βασισμένες στη λογική και δημιουργούν μεγάλα προβλήματα στα συστήματα, οι υπάρχουσες επιθέσεις DoS είναι κυρίως stateless, δηλαδή δεν υπάρχει διατήρηση της κατάστασης στο σύστημα. Όμως εξαιτίας της διαδεδομένης πρακτικής για διατήρηση της κατάστασης στα συστήματα CPS, οι επιθέσεις με διατήρηση της κατάστασης μπορεί να είναι τελικά προσβάσιμες στους εχθρούς.

Stateless (L)

Οι επιπτώσεις από επιθέσεις χωρίς διατήρηση της κατάστασης υπάρχουν όσο το σύστημα ενεργά επεξεργάζεται κακόβουλα μηνύματα τα οποία έχουν παραχθεί από έναν εχθρό. Οι επιπτώσεις από τέτοιες επιθέσεις τελικά εξαφανίζονται κατά τη διάρκεια της κανονικής διαδικασίας χωρίς να χρειάζεται το σύστημα να πάρει επιπλέον μέτρα για αποκατάσταση του συστήματος. Οι υπάρχουσες DoS επιθέσεις σε μη CPS συστήματα έχουν υπάρξει κυρίως σαν επιθέσεις χωρίς διατήρηση της κατάστασης στο σύστημα. Αυτό συμβαίνει γιατί ο κάθε εχθρός θα πρέπει ενεργά να

διατηρεί της συνθήκες DoS συνήθως με ένα υψηλό ρυθμό ή με υψηλό όγκο κακόβουλης κίνησης. Δεν χρειάζεται κανένας ειδικός μηχανισμός επαναφοράς του συστήματος. Το σύστημα μπορεί εύκολα να αμυνθεί σε επιθέσεις stateless απλά κόβοντας όλη την κακόβουλη κίνηση.

Stateful (F)

Σε αντίθεση, οι επιπτώσεις των stateful επιθέσεων παραμένουν στο σύστημα ακόμα και αφού έχει τελειώσει η επεξεργασία των κακόβουλων μηνυμάτων που έχουν παραχθεί από κάποιον εχθρό. Όμως για την άμυνα απέναντι σε τέτοιες επιθέσεις χρειάζονται ειδικοί μηχανισμοί αποκατάστασης που είναι επιπλέον της κανονικής επεξεργασίας των μηνυμάτων. Υπάρχουν γενικά δύο ήδη καταστάσεων τα οποία χρησιμοποιούνται στα CPS συστήματα. Αυτά είναι: η κατάσταση δρομολόγησης (routing state) και η κατάσταση ανίχνευσης κάποιων γεγονότων (complex event detection state).

Η κατάσταση δρομολόγησης (routing state) διατηρείται από τους brokers για την επαλήθευση, το ταίριασμα και τη δρομολόγηση των μηνυμάτων. Στα CPS συστήματα οι πελάτες μπορούν να τροποποιούν την κατάσταση δρομολόγησης ενός brokers στέλνοντας μηνύματα. Τα ποιο κοινά μηνύματα τα οποία επιτρέπουν να αλλάζει αυτή η κατάσταση δρομολόγησης τα subscriptions, οι αναφορές, η αναίρεση subscription, η αναίρεση αναφορών και πιθανών ορισμός τύπων και εξειδικευμένα μηνύματα ενημέρωσης. Αυτά τα μηνύματα είναι υποψήφια στο να ενεργοποιήσουν stateful επιθέσεις. Η κακόβουλη απομάκρυνση τέτοιων μηνυμάτων (για παράδειγμα κάνοντας αναίρεση ενός subscription) είναι δυνατόν να διακόψει τις αναγγελίες έπ' αόριστον. Επίσης η κακόβουλη εισαγωγή τέτοιων μηνυμάτων μπορεί να είναι stateful επίθεση αν η απόδοση του συστήματος μειώνεται όσο αυξάνεται ο όγκος της κατάστασης που διατηρείται. Τέτοιες αλλοιώσεις της κατάστασης δρομολόγησης τυπικά μπορούν να διορθωθούν ρητά από ένα σύστημα αφού έχει ανιχνευτεί η επίθεση.

Η κατάσταση ανίχνευσης κάποιων γεγονότων (complex event detection state) διατηρείται από τους brokers για σχετικά publications που σχετίζονται με subscription που διατηρούν κάποια κατάσταση. Άρα τα ίδια τα subscriptions μπορεί να διατηρούν κάποια κατάσταση. Στα συστήματα που υποστηρίζουν την κατάσταση ανίχνευσης κάποιων γεγονότων τα publications είναι επίσης συνυποψήφια για να δημιουργηθούν τέτοιες επιθέσεις με διατήρηση της κατάστασης. Πρέπει να σημειωθεί ότι ανάλογα με το σχεδιασμό ενός τέτοιου συστήματος μπορεί είναι ή μπορεί να μην είναι απαραίτητο να διατηρούνται τέτοια publications στο σύστημα.

Είναι αρκετά δύσκολο για το σύστημα να αμυνθεί σε τέτοιες επιθέσεις αφού πρέπει να χρησιμοποιηθούν κάποιες συγκεκριμένες τεχνικές αποκατάστασης.

Soft-state (S)

Ανάμεσα στις stateful και stateless επιθέσεις οι επιπτώσεις από μια soft-state επίθεση επίσης παραμένουν στο σύστημα αφού τελειώσει η επεξεργασία των κακόβουλων μηνυμάτων. Όμως τέτοιες επιπτώσεις απομακρύνονται αυτόματα από μηχανισμούς που λήγουν, που έχουν δηλαδή κάποια χρονική ετικέτα και όταν περάσει ο χρόνος τους λήγουν. Σχετικά με την κατάσταση δρομολόγηση αυτοί οι μηχανισμοί που λήξης τυπικά χρησιμοποιούνται για σκοπούς ανοχής σε λάθη.

Όταν η λήξη των δεδομένων που χρησιμοποιούνται δεν είναι πολύ συχνό φαινόμενο τότε οι soft-state επιθέσεις μοιάζουν με τις stateful επιθέσεις.

Persistent (P)

Ένα CPS σύστημα υποστηρίζει subscriptions σε ιστορικά δεδομένα ή σε κατάσταση που έχει να κάνει με αποκατάσταση από κάποια μόνιμη μονάδα αποθήκευσης για επίτευξη ανοχής σφαλμάτων. Σε αυτά τα συστήματα υπάρχει μια ευκαιρία για τους κακούς για να αλλοιώσουν ή να εισάγουν κακόβουλα μηνύματα στη μόνιμη μονάδα αποθήκευσης. Όταν αυτά ανακτηθούν ξανά τότε τα κακόβουλα μηνύματα προφανώς μπορούν να προκαλέσουν επιπτώσεις DoS. Οι επιθέσεις αυτές θεωρούνται μόνιμες και τα αποτελέσματά τους μπορούν να ενεργοποιηθούν απλά και μόνο ανακτώντας τα κακόβουλα δεδομένα από τη μονάδα αποθήκευσης.

Component Techniques

Αν ένας εχθρός ανακαλύψει αποτελεσματικές τεχνικές για να δημιουργήσει DoS επιθέσεις, τότε αυτές θα χρησιμοποιηθούν σε μια ποικιλία διαφορετικών επιθέσεων. Buffer overflow, packet reflection και IP spoofing είναι παραδείγματα τεχνικών που υπάρχουν αυτή τη στιγμή για να δημιουργούνται σε DoS επιθέσεις σε μη CPS συστήματα. Αυτά δεν είναι από μόνα τους επιθέσεις αλλά χρησιμοποιούνται σαν τεχνικές για να δημιουργηθούν σε επιθέσεις. Το να αμυνθείς σε τέτοιες επιθέσεις είναι πολύ δύσκολο γιατί είναι πολύ γενικές Για τα CPS συστήματα αναγνωρίζουμε τεχνικές όπως identity forgery (πλαστογράφηση ταυτότητας), thrashing, amplification και stockpiling.

Οι τεχνικές που αναγνωρίζουμε σε αυτή την κλάση δεν είναι αμοιβαίες αποκλειόμενες αφού σε μια DoS επίθεση μπορούν να χρησιμοποιηθούν πολλές από αυτές τις τεχνικές.

Identity Forgery (F)

Στα κατανεμημένα CPS συστήματα, υπάρχουν γενικά δύο ταυτότητες που σχετίζονται με το κάθε μήνυμα: η ταυτότητα προέλευσης του γνήσιου μηνύματος και η ταυτότητα της οντότητας που διανέμει το μήνυμα. Ένας εχθρός μπορεί να προσπαθήσει να δημιουργήσει μια επίθεση παραποιώντας είτε μία από τις δύο ή και τις δύο αυτές ταυτότητες.

Thrashing (T)

Το thrashing καταναλώνει πόρους στο σύστημα προκαλώντας έτσι γρήγορη και συχνή αλλαγή κατάστασης (ανάλογο με το thrashing στη διαχείριση μνήμης). Το thrashing είναι μια ειδική περίπτωση για πιο γενικές επιθέσεις σε πόρους, έτσι η αποδοτικότητά του εξαρτάται από το πόσος φόρτος συστήματος σε όρους επεξεργασίας και κίνησης στο δίκτυο.

Amplification (A)

Το Amplification δίνει σε έναν εχθρό την δυνατότητα να δημιουργήσει παράδοση πολλαπλών μηνυμάτων σε μία μόνο οντότητα απλά εισάγοντας ένα κακόβουλο μήνυμα. Υπόψη η ενίσχυση δεν χρησιμοποιεί multicasting το οποίο απλά παραδίδει ένα μήνυμα σε πολλές οντότητες. Πολλά χαρακτηριστικά στο παράδειγμα του publish/subscribe είναι ευαίσθητα για αυτόν τον τύπο επίθεσης. Για παράδειγμα μία νέα ανακοίνωση μπορεί να προσελκύσει πολλά αδρανή subscriptions ή μια αναίρεση subscription μπορεί να προκαλέσει προώθηση πολλαπλών subscriptions

Stockpiling(S)

Για να είναι επιτυχής κάποιες επιθέσεις μπορεί να εξαρτώνται από κάποια συγκεκριμένη προεργασία ώστε αν δημιουργηθούν οι κατάλληλες συνθήκες. Έτσι ο εχθρός μπορεί πριν ξεκινήσει μια τέτοια επίθεση να πρέπει να φτιάξει την κατάσταση που χρειάζεται. Τέτοια κακόβουλη κίνηση μπορεί να χρειάζεται να είναι stateful αλλά δεν είναι τεχνικό μέρος της επίθεσης γιατί δεν προκαλεί κάποια αποτελέσματα DoS επίθεσης από μόνη της. Αν ανακαλυφθεί και αντιμετωπιστεί η κακόβουλη κατάσταση που έχει φτιαχτεί τότε το σύστημα μπορεί να αποφύγει τις επιθέσεις DoS συνολικά. Χρησιμοποιούμε τον όρο stockpiling για να περιγράψουμε τις επιθέσεις που δημιουργούν την κακόβουλη κατάσταση σαν προεργασία ή για να ενεργοποιήσουν τις πραγματικές DoS επιθέσεις. Για παράδειγμα ο εχθρός μπορεί πρώτα να εισάγει έναν

αριθμό από κακόβουλες ανακοινώσεις στο σύστημα οι οποίες αυξάνουν σημαντικά τα αποτελέσματα της επίθεσης με πλημμύρα publications.

Η αδυναμία σε stockpiling επιθέσεις εισάγεται επιτρέποντας τους χρήστες να βάζουν μηνύματα τα οποία παραμένουν στο σύστημα.

Επιθετικές ενέργειες σε δίκτυα publish/subscribe

Με βάση τα παραπάνω είναι σημαντικό να αναγνωρίσουμε τις ενέργειες που μπορεί να αποτελέσουν πρόβλημα ασφαλείας από συνδρομητές.

Γενικά έχουν δοθεί παραδείγματα για συμπεριφορές που μπορούν να οδηγήσουν σε επιθέσεις άρνησης υπηρεσιών (DoS attacks), επιθέσεις που προέρχονται είτε από publishers είτε από subscribers. Στην προηγούμενη ενότητα αναλύονται σε βάθος επιθέσεις σε publish/subscribe συστήματα και με βάση τη βιβλιογραφία επιλέγουμε κάποιες επιθέσεις που μπορούν να γίνουν από τους συνδρομητές και αποτελούν αντικείμενο περαιτέρω διερεύνησης ως προς την συμπεριφορά των συνδρομητών/χρηστών.

Ένας τύπος επίθεσης είναι εφαρμόζοντας περίπλοκες αναζητήσεις εκδήλωσης ενδιαφέροντος για περιεχόμενο. Τέτοιου είδους αναζητήσεις είναι δυνατό να υπερφορτώσουν τις ενδιάμεσες οντότητες, δηλαδή τους brokers και να συμβάλουν στη σταδιακή αποδόμηση βασικών λειτουργιών του ίδιου του συστήματος. Περίπλοκες αναζητήσεις θα αναγκάσουν τις ενδιάμεσες οντότητες να χρησιμοποιήσουν τόσο υπολογιστικούς πόρους σε μεγάλο ρυθμό όσο ενδεχομένως και πόρους αποθηκευτικής ισχύος ή και δικτυακούς πόρους.

Ένας άλλος τύπος επίθεσης που εξετάζεται και είναι ιδιαίτερος σε συστήματα και αρχιτεκτονικές publish/subscribe είναι η δημιουργία ψευδών συνδρομών/αναζητήσεων ή αναζητήσεων που δεν πρόκειται να μπουν ποτέ στο σύστημα. Για παράδειγμα μία συνδρομή θα μπορούσε να είναι «δώσε μου το περιεχόμενο της εφημερίδας New York Times» το οποίο ακόμα και αν δεν είναι διαθέσιμο εκείνη τη περίοδο κάποια χρονική στιγμή στο μέλλον ενδεχομένως να είναι διαθέσιμο. Όμως μία συνδρομή του τύπου «δώσε μου το περιεχόμενο της εφημερίδας New York Times1» προφανώς οφείλεται είτε σε λάθος εκ παραδρομής είτε σε κακόβουλη ενέργεια που σε συνδυασμό με άλλες παρόμοιες μπορούν να παρουσιάσουν κίνδυνο. Το πρόβλημα οφείλεται στο γεγονός ότι το σύστημα πρέπει να κρατήσει στοιχεία για την αναζήτηση έτσι ώστε να ενημερώσει τον subscriber όταν η συγκεκριμένη έκδοση περιεχομένου γίνει διαθέσιμη (στο παράδειγμά μας δεν θα γίνει ποτέ). Αυτό οδηγεί αναπόφευκτα στη δέσμευση πόρων αλλά και στη καθυστέρηση νέων αναζητήσεων που θα ακολουθήσουν έχοντας ένα μεγαλύτερο αριθμό συνδρομών που πρέπει να ελεγχθούν.

Είναι ενδιαφέρον να παρουσιαστεί και ένα φαινόμενο όπου είναι σύνηθες σε P2P συστήματα όπου μία οντότητα εμφανίζεται στο δίκτυο με πολλές ταυτότητες, πλαστές στις περισσότερες

περιπτώσεις, δημιουργώντας στο σύστημα την αίσθηση ότι λειτουργούν ταυτόχρονα πολλοί χρήστες. Με αυτό τον τρόπο γίνεται προσπάθεια χειραγώγησης του συστήματος, της πολιτικής ασφάλειας/εμπιστοσύνης και της ίδιας της δυνατότητας του συστήματος να ρυθμίζει την χρήση και των διαμοιρασμό πόρων ισότιμα και δίκαια. Αυτά τελικά μπορούν να οδηγήσουν το σύστημα σε μερική ή πλήρη απορύθμιση και την μείωση της ικανοποίησης που έχουν οι χρήστες από το σύστημα. Η συμπεριφορά αυτή μπορεί να αποφευχθεί αν υπάρχει κεντρικός διαχειριστής κάτι που είναι δύσκολο σε πολύ μεγάλα συστήματα και επίσης με τη βοήθεια συστήματος υποδομής δημοσίου κλειδιού που έχει και αυτό αντικειμενικές δυσκολίες όσο αφορά στην κλιμάκωση.

Ένας subscriber έχει τη δυνατότητα να εκδηλώσει ενδιαφέρον για μία έκδοση που εν δυνάμει περιλαμβάνει όλο τον λογικό χώρο περιεχομένου, πχ μια ακραία υποθετική έκφραση θα ήταν «φέρε μου όλο το Internet». Αυτό θα έχει σαν αποτέλεσμα σημαντική υπερφόρτωση τόσο των κόμβων προώθηση όσο και των κόμβων rendezvous points και επίσης την αύξηση της κίνησης του δικτύου όταν δεν υπάρχει πολιτική διαμοιρασμού δικτυακών πόρων.

Είναι δυνατό ένας subscriber να προκαλέσει πρόβλημα στο σύστημα δημιουργώντας με μεγάλη συχνότητα subscriptions. Αυτό προκαλεί υπερφόρτωση στο σύστημα καθώς επίσης και κατανάλωση πόρων όπως μνήμη και αποθηκευτικός χώρος.

Ένα άλλο πρόβλημα που μπορεί να δημιουργήσει ο subscriber είναι να κάνει dettach από το subscription αφού ξεκινήσει η διαδικασία αποστολής. Σε πολλά publish/subscribe συστήματα η επικοινωνία γίνεται ασύγχρονα και η παράδοση περιεχομένου γίνεται από άλλο κόμβο από αυτόν που παρέχει το περιεχόμενο. Ο κόμβος αυτό λέγεται rendezvous point και σε αυτόν τον κόμβο μεταφέρεται το περιεχόμενο που μπορεί να θέλει κάποιος subscriber. Αν τελικά ο subscriber διακόψει την λήψη αφού οριστικοποιηθεί η μεταφορά και μεταφερθεί το περιεχόμενο στο κόμβο ραντεβού τότε θα έχει καταναλώσει πόρους για την εναρξη της επικοινωνία καθώς επίσης και κίνηση στο δίκτυο αφού η μεταφορά στο σημείο ραντεβού έγινε χωρίς λόγο. Στα συστήματα αυτά που επιτρέπουν τη λειτουργία αυτή προτείνεται να μην επιτρέπουν σε έναν subscriber να κάνει dettach πάνω από κάποιες φορές (συνήθως 2).

Όταν το σύστημα δεσμεύει πόρους και επεξεργαστική ισχύ για να καταχωρήσει subscriptions τότε ένας κακός μπορεί να δημιουργεί συνεχόμενα (ακόμα και τα ίδια) subscriptions και στη συνέχεια να τα αναιρεί αμέσως. Αυτό, ανάλογα με την υλοποίηση στο σύστημα δημιουργεί σοβαρά προβλήματα φόρτου.

Συνοπτικά οι ενέργειες που χαρακτηρίζονται ως ύποπτες σε ένα publish/subscribe σύστημα και μπορούν υπό προϋποθέσεις να δημιουργήσουν πρόβλημα είναι:

- y1 = complex queries → Χρήση πολλών πόρων (large) ή metadata.
- y2 = queries frequent→ Μεγάλη συχνότητα queries.
- y3 = bad queries → Χρήση τέτοιων ερωτήσεων προκαλούν εξαντλητική χρήση.
- y4 = μεγάλος αριθμός subscription που δε χρειάζεται στην πραγματικότητα
- y5 = Sybil attacks → αριθμός επιθέσεων με διαφορετικό όνομα /ταυτότητα
- y6 = Dettach → διακοπή σύνδεσης/ εγκατάλειψη σύνδεσης στο τελευταίο βήμα.
- y7 = subscribe/unsubscribe → διαδοχικά subscribe/unsubscribe που προκαλούν φόρτο στο σύστημα

Evidence Theory – Dempster-Shafer

Στην ενότητα αυτή αναπτύσσεται η θεωρία Dempster-Shafer και είναι μια μαθηματική θεωρία συναρτήσεων πεποίθησης (ή πίστης). Η θεωρία αυτή προσφέρει έναν εναλλακτικό τρόπο της μαθηματικής αναπαράστασης της αβεβαιότητας.

Η θεωρία του Dempster-Shafer είναι γνωστή και ως θεωρία συναρτήσεων πεποίθησης. Η συγκεκριμένη θεωρία βασίζεται σε λογισμό με αριθμητικές τιμές *πεποίθησης* (*belief*), δηλαδή πίστης για την ισχύ κάποιου υποθετικού συμπεράσματος για το οποίο υπάρχουν κάποιες ενδείξεις (γεγονότα). Επίσης, δεν απαιτείται η συλλογή όλων των απλών και των υπό συνθήκη πιθανοτήτων. Η θεωρία του Dempster-Shafer έχει σαν βασική ιδέα την εξαγωγή βαθμών πεποίθησης για μια ερώτηση από τις υποκειμενικές πιθανότητες μιας άλλης συναφούς ερώτησης. Με τον τρόπο αυτό, οι συναρτήσεις πεποίθησης μας δίνουν τη δυνατότητα να εξάγουμε βαθμούς πεποίθησης για μια ερώτηση. Όσο πιο συναφείς είναι οι δύο ερωτήσεις τόσο πιο κοντά στην πραγματικότητα θα είναι οι πιθανότητες. Παρακάτω περιγράφουμε μερικά πλεονεκτήματα της θεωρίας σε σχέση με τη θεωρία του Bayes

- είναι ένας μαθηματικός τρόπος να συνδυασθούν τα στοιχεία από πολλούς παρατηρητές, χωρίς την ανάγκη γνώσης των εκ των προτέρων ή υποθετικών πιθανοτήτων, όπως συμβαίνει στην Bayesian προσέγγιση.
- δεν απαιτείται a priori γνώση μιας υπόθεσης, γεγονός που την καθιστά κατάλληλη για την ανίχνευση ανώμαλων συμπεριφορών που εμφανίζονται για πρώτη φορά, σε ένα δίκτυο κόμβων. Αντιθέτως, για να χρησιμοποιηθεί ο κανόνας του Bayes, απαιτείται a priori γνώση.
- στον κανόνα του Bayes, η απουσία άλλων ενδείξεων για τις δυνατές εκδοχές, τις καθιστά όλες τις εκδοχές ισοπίθανες, αντιθέτως στην θεωρία Dempster – Shafer η απουσία κάποιων ενδείξεων θέτει την πιθανότητα κάθε εκδοχής κάπου στο διάστημα [0,1].
- αποδίδεται τιμή και στην άγνοια μιας κατάστασης. Αντιθέτως, με τον κανόνα του Bayes, δεν επιτρέπεται η απόδοση πιθανότητας στην άγνοια άρα δεν ορίζεται η αβεβαιότητα

Βασικά Στοιχεία

– Πλαίσιο διάκρισης (frame of discernment): είναι το σύνολο Θ των διακριτών και

αμοιβαία αποκλειόμενων υποθέσεων που αφορούν το πρόβλημα

- 2^{Θ}: Το σύνολο των υποσυνόλων του Θ
  - Κάθε στοιχείο του 2^{Θ} αντιστοιχεί σε διαζευγμένες προτάσεις.
  - Στοιχεία του Θ που δεν ανήκουν σε ένα στοιχείο του 2^{Θ}, κάνουν σαφή την άρνηση του αντίστοιχου υποθετικού συμπεράσματος.
  - Το κενό υποσύνολο {} αντιστοιχεί στην περίπτωση που όλα τα υποθετικά συμπεράσματα είναι ψευδή (*null hypothesis*).
- Η βασική κατανομή πιθανότητας (basic probability assignment - bpa) είναι μία απεικόνιση: m: $2^{\Theta} \rightarrow [0,1]$ δηλαδή το μέτρο της πεποίθησης που υπάρχει για το κατά πόσο ισχύει το υποθετικό συμπέρασμα που εκφράζεται με το συγκεκριμένο στοιχείο του Θ.
  - Υποκειμενική ποσότητα που δε μοιράζεται στα επιμέρους στοιχεία κάθε στοιχείου του 2^{Θ}.
  - Π.χ. αν m({A, B})=0.3, τότε αυτή η πεποίθηση δε μοιράζεται στα {A} και {B} αλλά αφορά το {A, B}.
  - Για το στοιχείο {} ισχύει m({})=0
  - Δεδομένου ότι το αληθές υποθετικό συμπέρασμα βρίσκεται κάπου μέσα στα στοιχεία του 2^{Θ}, ισχύει: $\sum_{x \in 2^{\Theta}} m(x) = 1$
  - Η ποσότητα m(X) εκφράζει το πόσο ισχυρή είναι η πεποίθηση για το ότι ένα συγκεκριμένο στοιχείο του Θ ανήκει στο X αλλά όχι σε κάποιο από τα τυχόν υποσύνολα του X.
- Η συνολική πεποίθηση (*belief*) ότι ένα στοιχείο του Θ ανήκει στο X καθώς και στα τυχόν υποσύνολα του X συμβολίζεται με $bel(x) = \sum_{Y \subseteq X} m(Y)$. Ονομάζεται συνάρτηση πεποίθησης και αναπαριστά την βαρύτητα των στοιχείων που υποστηρίζουν την ισχύ της υπόθεσης X.
- Αν m1 και m2 δύο ανεξάρτητες εκτιμήσεις (βασικές κατανομές πιθανότητας) που αποδίδουν κάποιο βαθμό πεποίθησης στα στοιχεία του 2^{Θ}, τότε αυτές συνδυάζονται σε μία τρίτη εκτίμηση $m_3 = m_1 \oplus m_2$ με τρόπο που ορίζεται με τον κανόνα του Dempster:

$$m_3(B) = m_1(B) \oplus m_2(B) = \frac{\sum_{i,j:A_i \cap A_j = B} m_1(A_i) m_2(A_j)}{1 - \sum_{i,j:A_i \cap A_j = \varnothing} m_1(A_i) m_2(A_j)}$$

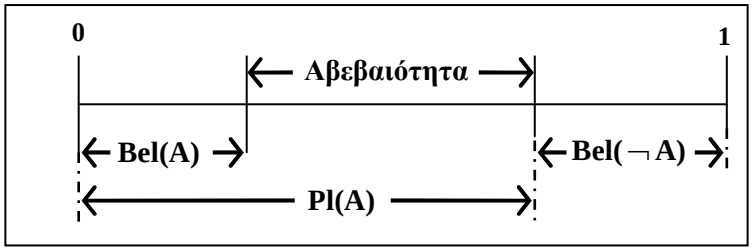
Μπορούμε να συνδυάσουμε περισσότερες από δύο συναρτήσεις πεποίθησης, ανά

ζεύγη και με οποιαδήποτε σειρά.

– Η συνάρτηση ευλογοφάνειας λαμβάνει υπόψη της όλα τα στοιχεία που σχετίζονται με τον συλλογισμό Α (δηλαδή, και αυτά που υποστηρίζουν ότι ισχύει, αλλά και αυτά που δηλώνουν αβεβαιότητα για την ισχύ του). Επομένως:

$$Pl(A) = \sum_{j:B_j \cap A \neq \varnothing} m(B_j) \text{ , ή } \quad Pl(A) = 1 - Bel(\neg A)$$

Για το υποσύνολο Α, τα Bel(A) και Pl(A) αντιπροσωπεύουν το ανώτερο και το κατώτερο όριο πεποίθησης αντίστοιχα, ενώ το διάστημα [Bel(A), Pl(A)] αντιπροσωπεύει την έκταση της πεποίθησης για την υπόθεση Α. Οι σχέσεις μεταξύ των τιμών του Bel(A), του Pl(A) και της αβεβαιότητας εμφανίζονται στο σχήμα που ακολουθεί.



Παράδειγμα:

- Αν Θ={A,B,C} το σύνολο των πιθανών ασθενειών που μπορούν να διαγνωστούν.
- Πιθανές διαγνώσεις 2^{Θ}={{},{A},{B},{C},{A,B},{A,C},{B,C},{A,B,C}}
- m({{},{A},{B},{C},{A,B},{A,C},{B,C},{A,B,C}}) = 1
  - υποδηλώνει τη βεβαιότητα ότι η διάγνωση βρίσκεται κάπου στα στοιχεία του 2^{Θ} αλλά ελλείψει άλλων ενδείξεων δεν είναι δυνατό να δοθεί ιδιαίτερη βαρύτητα σε κάποιο
  - Bays: θα έπρεπε κάθε στοιχείο του 2^{Θ} να θεωρηθεί ισοπίθανο
- Έστω ότι γίνεται διαθέσιμη επιπλέον πληροφορία (π.χ. πραγματοποιούνται ιατρικές εξετάσεις) και προκύπτει ότι η ασθένεια είναι μία από τις Α ή Β με βαθμό πίστης 0.7.
  - m_1({A<B}) = 0.7
  - m_1({{},{A},{B},{C},{A,B},{A,C},{B,C},{A,B,C}}) = 0.3

- Δηλαδή, η έλλειψη πίστης σε ένα από τα υποθετικά συμπεράσματα του 2^{Θ} ισοδυναμεί αυτόματα με ισόποσο βαθμό πίστης στα υπόλοιπα στοιχεία του 2^{Θ} , χωρίς όμως να δίνεται ιδιαίτερη προτίμηση σε κάποιο από αυτά.
- Bays: απαιτείται ο υπολογισμός μεγάλου αριθμού υπό συνθήκη πιθανοτήτων, κάτι που είναι υπολογιστικά ακριβό και πολλές φορές αδύνατο.

- Πώς μπορεί να συνδυαστούν δύο ανεξάρτητες εκτιμήσεις (π.χ. δύο ιατρών) σε μία;

Συνδυασμός Διαγνώσεων

- Έστω ότι δύο ιατροί εξετάζουν ανεξάρτητα τον ασθενή και δίνουν την εκτίμησή τους m1 και m_2 αντίστοιχα, για την αρρώστια από την οποία αυτός πάσχει.

	Γιατρός 1		**Γιατρός 2**	
Δυνατές περιπτώσεις διάγνωσης	**m_1**	**bel_1**	**m_2**	**bel_2**
{A}	0.05	0.05	0.15	0.15
{B }	0	0	0	0
{C}	0.05	0.05	0.05	0.05
{A,B}	0.15	0.2	0.05	0.2
{A,C}	0.1	0.2	0.2	0.4
{B,C}	0.05	0.1	0.05	0.1
{A,B,C}	0.6	1	0.5	1

Πίνακας 1 – Παράδειγμα εκτιμήσεων D-S

$$bel(x) = \sum_{Y \subseteq X} m(Y)$$

- Π.χ. $Bel_1(\{A,B\}) = m_1(\{A,B\}) + m_1(\{A\}) + m_1(\{B\}) = 0.015 + 0.05 + 0 = 0.2$
- Οι δύο ανεξάρτητες εκτιμήσεις m1 και m2 μπορούν να συνδυαστούν σε μία τρίτη m_3 χρησιμοποιώντας τον κανόνα των Dempster-Shafer.

$$m_3(B) = m_1(B) \oplus m_2(B) = \frac{\sum_{i,j:A_i \cap A_j = B} m_1(A_i) m_2(A_j)}{1 - \sum_{i,j:A_i \cap A_j = \varnothing} m_1(A_i) m_2(A_j)}$$

Παράδειγμα Συνδυασμός Εκτιμήσεων

$m_3=m_1 \oplus m_2$		m_1						
		{A}	{B}	{C}	{A,B}	{A,C}	{B,C}	{A,B,C}
m_2		0.05	0	0.05	0.15	0.1	0.05	0.6
{A}	0.15	{A} .0075	{ } 0	{ } .0075	{A} .0225	{A} .015	{ } .0075	{A} .09
{B}	0	{ } .0	{B} 0	{ } .0	{B} .0	{ } .0	{B} .0	{B} .0
{C}	0.05	{ } .0025	{ } 0	{C} .0025	{ } .0075	{C} .005	{C} .0025	{C} .03
{A,B}	0.05	{A} .0025	{B} 0	{ } .0025	{A,B} .0075	{A} .005	{B} .0025	{A,B} .03
{A,C}	0.2	{A} .01	{ } 0	{C} .01	{A} .03	{A,C} .02	{C} .01	{A,C} .012
{B,C}	0.05	{ } .0025	{B} 0	{C} .0025	{B} .0075	{C} .005	{B,C} .0025	{B,C} .03
{A,B,C}	0.5	{A} .025	{B} 0	{C} .025	{A,B} .075	{A,C} .05	{B,C} .025	{A,B,C} .3

Δυνατές περιπτώσεις διάγνωσης	**m_3**	**Bel_3**
{A}	0.21	0.21
{B }	0.01	1
{C}	0.09	0.09
{A,B}	0.12	0.34
{A,C}	0.20	0.50
{B,C}	0.06	0.16
{A,B,C}	0.31	1

Πίνακας 2 – Παράδειγμα συνδυασμού εκτιμήσεων

- Η αρχική εκτίμηση ότι η ασθένεια είναι μία από τις Α ή Β αποδυναμώθηκε
  - Η διάγνωση βρίσκεται μάλλον στο σύνολο {A,C}
  - Επειδή αρχίζει να διαφαίνεται ότι η τελική διάγνωση είναι η Α.
- Η παραπάνω συνδυασμένη εκτίμηση μπορεί να συνδυαστεί εκ νέου με μία άλλη εκτίμηση (π.χ. 3ου γιατρού)

Μαθηματικό Μοντέλο Dempster-Shafer

Ο στόχος μας είναι να δημιουργήσουμε ένα μοντέλο αποτίμησης της συμπεριφοράς των συνδρομητών ενός συστήματος publish/subscribe αξιοποιώντας την θεωρία Dempster-Shafer. Ο λόγος η θεωρία αυτή λαμβάνει υπόψη και την έννοια της αβεβαιότητας αντίθετα με τις περισσότερες προσεγγίσεις.

Σύμφωνα με τα βασικά θα πρέπει καταρχήν να ορίσουμε το πλαίσιο διάκρισης, στη συνέχεια να καθορίσουμε τη συνάρτηση καταχώρησης της βασικής πιθανότητας και με βάση αυτή να υπολογίσουμε τον βαθμό πίστης και ευλογοφάνειας. Σαν μέτρο προσέγγισης θα μπορούσαμε να επιλέξουμε τον βαθμό πίστης ή την ευλογοφάνεια. Μελετώντας τα δεδομένα προτιμούμε την ευλογοφάνεια που ουσιαστικά «επιδοτεί» τον καλό (ή τον κακό) με την αβεβαιότητα. Θα αναλύσουμε παρακάτω

Έτσι στο πρώτο βήμα ορίζουμε το πλαίσιο διάκρισης Θ το οποίο περιέχει δύο στοιχειώδεις υποθέσεις Θ = {Good, Bad} και είναι ουσιαστικά το ζητούμενο του μοντέλου μας. Ανάλογα με το αποτέλεσμα από την διαδικασία που θα ορίσουμε θέλουμε επίσης να αποτιμήσουμε αυτό το αποτέλεσμα δηλαδή το κατά πόσο χαρακτηρίσαμε κάποιον good/bad και κάναμε ή όχι λάθος. Ιδιαίτερα μας ενδιαφέρουν οι περιπτώσεις που έχουμε αποκλείουμε έναν καλό συνδρομητή (false positive) ή επιτρέπουμε έναν κακόβουλο συνδρομητή (false negative)

Έτσι όπως ορίσαμε το πλαίσιο διάκρισης Θ το PowerSet ορίζεται ως εξής
2^{Θ} = { 0, {Good}, {Bad}, {Good,bad} } και περιέχει όπως ήδη εξηγήσαμε $2^{|\Theta|}$ στοιχεία. Το σύνολο ισχύος (Powerset) 2^{Θ} περιλαμβάνει όλες τις δυνατές υποθέσεις,

Γενικά η μόνη γνώση που έχουμε για τους subscribers είναι οι ενέργειες που κάνουν στο σύστημα. Έτσι είναι λογικό να αξιοποιήσουμε τις ενέργειες ως evidence/στοιχεία για το αν ένας subscriber είναι τελικά κακόβουλος ή καλός. Έτσι το πρόβλημά μας είναι τώρα πως θα μετατρέψουμε τις ενέργειες (evidence) του χρήστη (το σύνολό τους πιθανόν) σε μία τιμή trust και πιο συγκεκριμένα σε μία πιθανότητα ή μαζική τιμή. Δηλαδή θέλουμε τρόπο να δώσουμε τιμές στα m(Good) και m(Bad) όπως ζητά η θεωρία και τιμές επίσης στο m(Good,Bad) ώστε να χαρακτηρίσουμε και την αβεβαιότητα. Αν δεν υπάρχει αβεβαιότητα προφανώς καταλήγουμε σε απλό πιθανοτικό μοντέλο.

Το σημαντικότερο βήμα στο μοντέλο μας είναι ο τρόπος που θα κάνουμε τις εκτιμήσεις που ζητά η θεωρία, δηλαδή το πώς θα μετατρέψουμε τα evidence/στοιχεία που έχουμε σε πιθανότητα. Επίσης θα πρέπει η διαδικάσία/συνάρτηση που θα κάνουμε την μετατροπή να έχει ορισμένες ιδιότητες.

- Όταν αυξάνονται οι καλές ενέργειες θα πρέπει να αυξάνεται η πιθανότητα του συνδρομητή να είναι καλός
- Όταν αυξάνονται οι ύποπτες ενέργειες θα πρέπει να αυξάνεται η πιθανότητα του συνδρομητή να είναι κακός
- Όταν γίνεται καταμερισμός των ύποπτων ενεργειών σε περισσότερες από μία θα πρέπει να μειώνεται η πιθανότητα να είναι κακός.

Καταρχήν πρέπει να προσδιορίσουμε τις βασικές υποθέσεις που κάνουμε για το σύστημα.

- Η πιο σημαντική από αυτές είναι ότι το σύστημα μπορεί να αναγνωρίσει αν μια ενέργεια είναι ύποπτη και ποια ακριβώς ενέργεια είναι αυτή (i = 1,2,…,7) από τις 7 ενέργειες που έχουμε ξεχωρίσει ότι μας ενδιαφέρουν.
- Επίσης ένας κακός χρήστης εκτελεί μόνο ένα τύπο από τις «ύποπτες» ενέργειες και δεν αξιοποιεί υβριδικές επιθέσεις αλλά όμως εκτελεί και κάποιες καλές ενέργειες.
- Οι καλοί χρήστες εκτελούν και αυτοί «ύποπτες» ενέργειες όταν το χρειάζονται

Για την ανάγκη του υπολογισμού των εκτιμήσεων ορίζουμε τα παρακάτω.

- Το σύνολο των ύποπτων ενεργειών που κάνει ένας συνδρομητής ως:

$$TY = \sum Y_i \text{ όπου } i = 1, 2, .., 7$$

- Τις καλές ενέργειες που αναλογούν στις Y_i ύποπτες ενέργειες του συνόλου x των καλών ενεργειών.

$$KY_i = x * Y_i / TY$$

  Οι ύποπτες ενέργειες υπολογίζονται αναλογικά από το σύνολο των υπόπτων ενεργειών. Αν για παράδειγμα ένας κακός χρήστης έχει κάνει 10% ύποπτες ενέργειες

Y_1 από τις συνολικές ύποπτες ενέργειες τότε του αναλογούν 10% καλές ενέργειες από το σύνολο των καλών ενεργειών *x*

- Την πιθανότητα να είναι καλός ο χρήστης με βάση τις καλές ενέργειες που αντιστοιχούν στην ενέργεια Y_i

$$GY_i = KY_i / (KY_i + TY)$$

- Την πιθανότητα να είναι κακός ο χρήστης με βάση το προηγούμενο

$$P(Y_i) = 1 - GY_i$$

Έτσι για παράδειγμα έστω χρήστης που έχει κάνει x=10 καλές ενέργειες και TY=10 ύποπτες ενέργειες και Y_1=6 τότε με βάση τα παραπάνω έχουμε:

KY_1 = 6

GY_1 = 6 / (6 + 10) = 0.375

$P(Y_1)$ = 0.625

Τώρα είμαστε έτοιμοι να ορίσουμε τις συναρτήσεις βασικής πιθανότητας. Ο υπολογισμός της πιθανότητας να είναι κακός ο συνδρομητής σύμφωνα με όσα ορίσαμε είναι το γινόμενο των πιθανοτήτων να είναι κακός για κάθε μια από τις ύποπτες ενέργειες:

$$m(bad) = \prod P(Y_i)$$

Η πιθανότητα να είναι καλός είναι ίση με το λόγο των καλών ενεργειών που έχει κάνει στο σύστημα προς τις συνολικές ενέργειες που έχει κάνει στο σύστημα (*total = x + TY*):

$$m(good) = x / total$$

Η πιθανότητα που απομένει να είναι είτε καλός είτε κακός δηλαδή η πιθανότητα της αβεβαιότητας (ή της σχετικής άγνοιας) και είναι ίση με:

$$m(bad, good) = 1 - m(bad) - m(good)$$

Με βάση τις παραστάσεις που ακολουθούν υπολογίζονται τα Belief και Plausibility

Στην προηγούμενη ενότητα ορίσαμε το Belief ως:

$$bel(x) = \sum_{Y \subseteq X} m(Y)$$

έτσι έχουμε:

$$bel(bad) = \sum_{Y \subseteq \{bad\}} m(Y) = m(bad)$$

και

$$bel(good) = \sum_{Y \subseteq \{good\}} m(Y) = m(good)$$

και

$$bel(bad, good) = \sum_{Y \subseteq \{bad, good\}} m(Y) = m(bad) + m(good) + m(bad, good)$$

Επίσης σύμφωνα με την προηγούμενη ενότητα έχουμε:

$$Pl(A) = \sum_{j:B_j \cap A \neq \varnothing} m(B_j)$$

και άρα:

$$Pl(bad) = \sum_{j:B_j \cap \{bad\} \neq \varnothing} m(B_j) = m(bad) + m(bad, good)$$

και

$$Pl(good) = \sum_{j:B_j \cap \{good\} \neq \varnothing} m(B_j) = m(good) + m(bad, good)$$

και

$$Pl(bad, good) = \sum_{j:B_j \cap \{bad, good\} \neq \varnothing} m(B_j) = m(bad, good)$$

Πρέπει να σημειωθεί ότι με τον τρόπο που περιγράψαμε οι ύποπτες ενέργειες θεωρούνται ισοδύναμες ως προς το πρόβλημα που πιθανόν να δημιουργήσουν στο σύστημα.

Πειραματικό μοντέλο

Το μοντέλο που αναπτύσσουμε βασίζεται στη θεωρία Dempster-Shafer και στους συγκεκριμένους ορισμούς που περιγράψαμε στη προηγούμενη ενότητα. Για τις ανάγκες του πειραματικού μοντέλου δημιουργήθηκε ένα πρόγραμμα σε Java το οποίο προσπαθεί να εξομοιώσει τις βασικές λειτουργίες που μας ενδιαφέρουν. Δημιουργεί τυχαία brokers και subscribers και παράγει με τυχαίο τρόπο ενέργειες που βασίζονται στους subscribers. Οι brokers καταγράφουν αυτές τις ενέργειες σε αρχείο το οποίο επεξεργαζόμαστε σε Excel.

Το πρόγραμμα τρέχει με παράμετρο τον αριθμό των κινήσεων που γίνεται από όλους τους subscribers σε κάθε broker. Επιπλέον σε αρχείο μπορούμε να «πειράξουμε» τις υπόλοιπες παραμέτρους τους συστήματος όπως:

- Αριθμός συνδρομητών
- Αριθμός brokers
- Ποσοστό καλών συνδρομητών στο σύστημα
- Ποσοστό καλών ενεργειών των καλών συνδρομητών
- Ποσοστό κακών ενεργειών των κακών συνδρομητών
- Αριθμός κινήσεων ανά μονάδα χρόνου

Καταρχήν το πρόγραμμα εξομοίωσης ακολουθεί τις υποθέσεις που περιγράψαμε στην προηγούμενη ενότητα και αρχικά δημιουργεί τους brokers και στη συνέχεια τους subscribers. Κάθε φορά που δημιουργείται ένας subscriber έχει μία πιθανότητα να είναι κακός με βάση την παράμετρο που έχουμε δώσει. Στη συνέχεια παράγονται οι ενέργειες των συνδρομητών με τυχαίο τρόπο σύμφωνα με τις παραμέτρους που έχουμε δώσει παραπάνω.

Πιο συγκεκριμένα το πρόγραμμα παράγει δεδομένα της μορφής :

TIME	BROKER	SUB	ISGOOD	ACTION
1	Bid000	SUB5	GOOD	Y1
1	Bid000	SUB2	GOOD	X
1	Bid000	SUB10	GOOD	Y1
1	Bid000	SUB21	GOOD	X
1	Bid000	SUB18	GOOD	Y5
1	Bid000	SUB15	GOOD	Y5
1	Bid000	SUBBAD3	BAD	Y6

1	Bid000	SUB1	GOOD	X
1	Bid000	SUB25	GOOD	Y1
…	…	…	…	…

Πίνακας 3 – Παράδειγμα καταγραφής δεδομένων broker

Όπου x είναι καλή ενέργεια και Yi είναι ύποπτη ενέργεια.

Η στήλη SUB έχει το όνομα του subscriber και η στήλη TIME έχει την χρονική στιγμή που δημιουργείται η συγκεκριμένη κίνηση. Η στήλη ISGOOD περιέχει την ένδειξη BAD ή GOOD για κάθε συνδρομητή και η στήλη ACTION την ενέργεια που κάνει. Τέλος η στήλη BROKER περιέχει το όνομα/ταυτότητα του broker. Η ένδειξη ISGOOD (BAD ή GOOD) δεν χρησιμοποιείται από τον broker πουθενά παρά μόνο όταν γίνεται η αποτίμηση του μοντέλου όπου συγκρίνουμε αυτό που εκτιμήσαμε με αυτό που ήταν στην πραγματικότητα. Επίσης παρόλο που το όνομα των κακών περιέχει την λέξη BAD, ο broker έχει άγνοια γι αυτό.

Όταν θέτουμε σαν παράμετρο Αριθμό Brokers = 1 τότε τα δεδομένα που παράγονται εξομοιώνουν ουσιαστικά ένα κεντρικοποιημένο σύστημα όπου όλες οι ενέργειες ενημερώνουν μια κεντρική οντότητα, τον μοναδικό broker που υπάρχει στο σύστημα. Ένα τέτοιο κεντρικοποιημένο σύστημα έχει πολλές ομοιότητες με ένα τοπικό δίκτυο μιας επιχείρησης όπου υπάρχει μόνο ένα scope και μόνο ένας broker. Όταν ο αριθμός των brokers είναι μεγαλύτερος τότε το σύστημα είναι κατανεμημένο και οι εκτιμήσεις των brokers συνδυάζονται με τον κανόνα του Dempster.

Ο broker μπορεί να κρατάει στοιχεία για κάθε subscriber και να τα ανανεώνει με κάθε ενέργεια που κάνει στο σύστημα ο συγκεκριμένος subscriber. Αυτό δεν είναι μακριά από αυτό που συμβαίνει στην πραγματικότητα όπου τα publish/subscribe συστήματα κρατάνε πληροφορίες για τους subscribers (συνήθως για τα subscribtions που έχουν). Οι τύποι που περιγράφηκαν στην προηγούμενη ενότητα είναι σχετικά απλοί και δεν απαιτούν μεγάλη πολυπλοκότητα.

Μια σημαντική παρατήρηση για το κατανεμημένο μοντέλο είναι ότι ο κακός subscriber έχει παρόμοια συμπεριφορά σε όλους τους brokers που είναι συνδεδεμένος. Πχ. αν οι ενέργειες που κάνει είναι 80% ύποπτες σε ένα broker τότε θα είναι το ίδιο και στους υπόλοιπους.

Επίσης το κατανεμημένο μοντέλο μπορεί να λειτουργήσει ως κεντρικοποιημένο αν τα στοιχεία που ανταλλάσσουν οι brokers δεν είναι οι εκτιμήσεις τους αλλά τα δεδομένα τους για τους subscribers. Γι αυτό το λόγο στα αποτελέσματα των πειραμάτων σε κατανεμημένο σύστημα έχουμε συμπεριλάβει και την εκτίμηση πιθανότητας του κεντρικοποιημένου ώστε

να μπορεί να γίνει σύγκριση.

Τέλος το μοντέλο μας θα λειτουργούσε ακόμα και αν οι subscribers μπορούσαν να εισάγουν ταυτόχρονα διαφορετικούς τύπους κινήσεων (υβριδικές επιθέσεις) όπως για παράδειγμα Y1,Y2.

Π1. Έστω ένα παράδειγμα στοιχείων που συγκέντρωσε ο broker Bid000 για τον subscriber SUB1:

SUB	BROKER	X	Y1	Y2	Y3	Y4	Y5	Y6	Y7	total
SUB1	Bid000	250	19	13	13	12	8	17	16	348

Σύμφωνα με όσα έχουμε πει θα πρέπει να ισχύει:

SUB	***m(bad)***	***m(good)***	***m(bad,good)***	***bel(bad)***	***bel(good***	***bel(bad,good)***	***Pl(bad)***	***Pl(good)***	***Pl(bad,good)***
SUB1	0,12	0,72	0,17	0,12	0,72	1	0,28	0,88	0,17

Παρατηρούμε ότι οι ύποπτες ενέργειες είναι καταμερισμένες και ότι το σύστημα υπολογίζει ότι υπάρχει αβεβαιότητα (για το αν ο subscriber είναι κακός ή καλός) ίση με 0.17. Αυτό είναι λογικό γιατί το σύστημα βλέπει ότι υπάρχουν αρκετές ύποπτες κινήσεις (28%) αλλά δεν είναι σίγουρο αφού αυτές είναι καταμερισμένες.

Π2. Έστω ένα παράδειγμα κακού subscriber:

SUB	BROKER	X	Y1	Y2	Y3	Y4	Y5	Y6	Y7	total
SUBBAD4	Bid000	45					281			326

Βλέπουμε ότι όπως το ορίσαμε στις υποθέσεις μας στην προηγούμενη ενότητα ο κακός subscriber κάνει μόνο ένα τύπο ύποπτων ενεργειών. Το αποτέλεσμα είναι:

SUB	***m(bad)***	***m(good)***	***m(bad,good)***	***bel(bad)***	***bel(good***	***bel(bad,good)***	***Pl(bad)***	***Pl(good)***	***Pl(bad,good)***
SUBBAD4	0,86	0,14	0,00	0,86	0,14	1	0,86	0,14	0,00

Παρατηρούμε ότι δεν υπάρχει αβεβαιότητα αφού ο subscriber κάνει κατ' επανάληψη μία ενέργεια και έτσι χαρακτηρίζεται ως κακός.

Π3. Έστω στο ίδιο παράδειγμα «πειράζουμε» τις ενέργειες του subscriber έτσι ώστε το 1/3 των καλών ενεργειών μετατρέπονται σε ύποπτες και κατανέμονται ομοιόμορφα στις υπόλοιπες ύποπτες ενέργειες (εκτός από την κύρια που επαναλαμβάνει)

SUB	BROKER	X	Y1	Y2	Y3	Y4	Y5	Y6	Y7	total
SUBBAD4	Bid000	30	3	3	3	2	281	2	2	326

Το αποτέλεσμα είναι:

SUB	***m(bad)***	***m(good)***	***m(bad,good)***	***bel(bad)***	***bel(good***	***bel(bad,good)***	***Pl(bad)***	***Pl(good)***	***Pl(bad,good)***
SUBBAD4	0,91	0,09	0,00	0,91	0,09	1	0,91	0,09	0,00

Βλέπουμε ότι το ποσοστό ότι είναι κακός αυξήθηκε και η αβεβαιότητα μειώθηκε παρόλο που έγινε επιμερισμός των ύποπτων ενεργειών. Αυτό σημαίνει ότι είναι πιο σημαντική η αύξηση των συνολικών ύποπτων ενεργειών από τον επιμερισμό τους.

Π4. Έστω στο ίδιο παράδειγμα «πειράζουμε» τις ενέργειες του subscriber έτσι ώστε το 1/3 περίπου των κακών ενεργειών κατανέμονται ομοιόμορφα στις υπόλοιπες ύποπτες ενέργειες (εκτός από την κύρια που επαναλαμβάνει). Στο παράδειγμα αφαιρέσαμε 90 ενέργειες από την κύρια ύποπτη και τις κατανείμαμε ομοιόμορφα (ανά 15)

SUB	BROKER	X	Y1	Y2	Y3	Y4	Y5	Y6	Y7	total
SUBBAD4	Bid000	45	15	15	15	15	191	15	15	326

Το αποτέλεσμα είναι:

SUB	***m(bad)***	***m(good)***	***m(bad,good)***	***bel(bad)***	***bel(good***	***bel(bad,good)***	***Pl(bad)***	***Pl(good)***	***Pl(bad,good)***
SUBBAD4	0,85	0,14	0,01	0,86	0,14	1	0,86	0,14	0,01

Βλέπουμε ότι το ποσοστό ότι είναι κακός μειώθηκε ελάχιστα και η αβεβαιότητα αυξήθηκε ελάχιστα. Αυτό σημαίνει ότι δεν αρκεί ένας απλός καταμερισμός στις ύποπτες ενέργειες για να ξεγελάσει το σύστημα όταν οι συνολικές ύποπτες ενέργειες είναι υψηλές.

Αποτελέσματα

Χρησιμοποιήσαμε πολλά διαφορετικά σενάρια για να δούμε πως δουλεύει το μοντέλο μας. Εκτός από τα λογικά και αναμενόμενα, χρησιμοποιήσαμε και σενάρια που είναι σχεδόν αδύνατο να υπάρξουν απλά και μόνο για να ελέγξουμε πως λειτουργεί το μοντέλο μας ακόμα και σε ακραίες καταστάσεις.

Στους παρακάτω πίνακες περιγράφουμε συνοπτικά τα σενάρια που χρησιμοποιήσαμε και που αναλύονται στις σελίδες που ακολουθούν για το κεντρικοποιημένο μοντέλο και για το κατανεμημένο μοντέλο. Με (*) εμφανίζονται τα ακραία σενάρια σύμφωνα με τις υποθέσεις μας.

	Σενάριο 1	**Σενάριο 2**	**Σενάριο 3**	**Σενάριο 4 ***
Ποσοστό καλών χρηστών	90%	80%	80%	80%
Συμπεριφορά καλών (καλή)	70%	50%	80%	35%
(κακή)	*30%*	*50%*	*20%*	*65%*
Συμπεριφορά κακών (κακή)	85%	70%	50%	85%
(καλή)	*15%*	*30%*	*50%*	*15%*
Αριθμός Brokers	1	1	1	1
Συνολικές ενέργειες/Broker	10000	10000	10000	10000
Ενέργειες ανά μονάδα χρόνου	100	100	100	100
Αριθμός συνδρομητών	30	30	30	30
Απόδοση συστήματος	*Καλή*	*Καλή*	*Καλή*	*Με λάθη/Οριακή*
Λάθη (M.O)	*Κανένα*	*Κανένα*	*Κανένα*	*4 False Positive*
	Κεντρικοποιημένο	*Κεντρικοποιημένο*	*Κεντρικοποιημένο*	*Κεντρικοποιημένο*

Πίνακας 4 - Σενάρια κεντρικοποιημένου μοντέλου

	Σενάριο 5	**Σενάριο 6**	**Σενάριο 7 ***
Ποσοστό καλών χρηστών	80%	80%	80%
Συμπεριφορά καλών (καλή)	70%	50%	80%
(κακή)	*30%*	*50%*	*20%*
Συμπεριφορά κακών (κακή)	85%	70%	35%
(καλή)	*15%*	*30%*	*65%*
Αριθμός Brokers	3	3	3
Συνολικές ενέργειες/Broker	10000	10000	10000
Ενέργειες ανά μονάδα χρόνου	100	100	100
Αριθμός συνδρομητών	30	30	30
Απόδοση συστήματος	*Καλή*	*Καλή/Οριακή*	*Κακή*
Λάθη (M.O)	*Κανένα*	*Κανένα*	*100% False Negative*
	Κατανεμημένο	*Κατανεμημένο*	*Κατανεμημένο*

Πίνακας 5 - Σενάρια κατανεμημένου μοντέλου

Ακολουθούν μερικά ενδεικτικά σενάρια τα οποία εξετάστηκαν μαζί με τους αναλυτικούς πίνακες δεδομένων και αποτελεσμάτων.

Σενάριο 1

Στο συγκεκριμένο σενάριο περιγράφουμε ένα ρεαλιστικό παράδειγμα με βάση τις υποθέσεις μας. Οι καλοί χρήστες στο σύστημα αντιστοιχούν στο 90% των συνολικών χρηστών. Η συμπεριφορά των καλών χρηστών είναι 70% καλές ενέργειες και 30% ύποπτες ενέργειες. Ενώ η συμπεριφορά των κακών χρηστών είναι 85% ύποπτες ενέργειες και 15% καλές ενέργειες. Εδώ πρέπει να επισημανθεί η υπόθεση που έχουμε κάνει για το σύστημα ότι δηλαδή ένας κακός χρήστης επαναλαμβάνει την ίδια ύποπτη ενέργεια και μόνο.

Τα δεδομένα που κατέγραψε ο broker συγκεντρωτικά είναι:

		ΚΑΛΕΣ	***ΥΠΟΠΤΕΣ***							
SUBSCRIBER	***BROKER***	***X***	***Y1***	***Y2***	***Y3***	***Y4***	***Y5***	***Y6***	***Y7***	***Total***
SUB1	Bid000	250	19	13	13	12	8	17	16	348
SUB10	Bid000	247	8	12	18	9	23	14	15	346
SUB11	Bid000	219	17	14	10	12	15	16	17	320
SUB12	Bid000	219	15	7	19	13	14	18	17	322
SUB13	Bid000	248	13	18	9	19	18	15	19	359
SUB14	Bid000	248	9	14	18	12	14	13	19	347
SUB15	Bid000	205	14	12	13	14	21	17	13	309
SUB16	Bid000	242	13	10	16	11	18	17	7	334
SUB17	Bid000	231	12	15	12	12	19	14	19	334
SUB18	Bid000	267	14	12	16	17	9	16	16	367
SUB19	Bid000	250	11	16	17	21	12	4	15	346
SUB2	Bid000	203	21	14	9	16	10	11	8	292
SUB20	Bid000	232	13	15	15	17	11	10	15	328
SUB21	Bid000	222	14	14	14	16	13	23	16	332
SUB22	Bid000	224	10	12	10	22	9	15	6	308
SUB23	Bid000	234	15	10	12	18	12	14	16	331
SUB24	Bid000	234	14	19	14	7	17	10	20	335
SUB25	Bid000	229	10	19	12	19	15	12	18	334
SUB3	Bid000	232	9	15	17	12	10	11	19	325
SUB4	Bid000	233	15	19	14	22	16	11	15	345
SUB5	Bid000	238	17	11	20	11	14	20	13	344
SUB6	Bid000	256	19	18	21	14	16	10	21	375
SUB7	Bid000	245	14	9	18	7	11	13	11	328
SUB8	Bid000	226	21	13	6	15	14	10	16	321
SUB9	Bid000	241	15	19	13	9	10	7	16	330
SUBBAD1	Bid000	58							289	347
SUBBAD2	Bid000	63	260							323
SUBBAD3	Bid000	43				291				334
SUBBAD4	Bid000	45					281			326
SUBBAD5	Bid000	41				269				310
	Grand Total:	**6125**	**612**	**350**	**356**	**917**	**630**	**338**	**672**	**10000**

Πίνακας 6 – Δεδομένα Σεναρίου 1

Με βάση τα παραπάνω δεδομένα που δημιούργησε με τυχαίο τρόπο η εξομοίωση υπολογίστηκαν από το σύστημα οι τιμές των βασικών πιθανοτήτων, οι τιμές πίστης και οι τιμές ευλογοφάνειας. Τα στοιχεία αυτά εμφανίζονται αναλυτικά παρακάτω:

SUB	m(bad)	m(good)	m(bad,good)	Bel(bad)	Bel(good	Bel(bad,good)	Pl(bad)	Pl(good)	Pl(bad,good)
SUB1	0,12	0,72	0,17	0,12	0,72	1	0,28	0,88	0,17
SUB10	0,12	0,71	0,16	0,12	0,71	1	0,29	0,88	0,16
SUB11	0,15	0,68	0,16	0,15	0,68	1	0,32	0,85	0,16
SUB12	0,16	0,68	0,16	0,16	0,68	1	0,32	0,84	0,16
SUB13	0,15	0,69	0,16	0,15	0,69	1	0,31	0,85	0,16
SUB14	0,12	0,71	0,17	0,12	0,71	1	0,29	0,88	0,17
SUB15	0,18	0,66	0,16	0,18	0,66	1	0,34	0,82	0,16
SUB16	0,11	0,72	0,17	0,11	0,72	1	0,28	0,89	0,17
SUB17	0,14	0,69	0,16	0,14	0,69	1	0,31	0,86	0,16
SUB18	0,11	0,73	0,17	0,11	0,73	1	0,27	0,89	0,17
SUB19	0,11	0,72	0,16	0,11	0,72	1	0,28	0,89	0,16
SUB2	0,14	0,70	0,16	0,14	0,70	1	0,30	0,86	0,16
SUB20	0,13	0,71	0,17	0,13	0,71	1	0,29	0,87	0,17
SUB21	0,17	0,67	0,16	0,17	0,67	1	0,33	0,83	0,16
SUB22	0,11	0,73	0,16	0,11	0,73	1	0,27	0,89	0,16
SUB23	0,13	0,71	0,17	0,13	0,71	1	0,29	0,87	0,17
SUB24	0,14	0,70	0,16	0,14	0,70	1	0,30	0,86	0,16
SUB25	0,15	0,69	0,16	0,15	0,69	1	0,31	0,85	0,16
SUB3	0,12	0,71	0,17	0,12	0,71	1	0,29	0,88	0,17
SUB4	0,16	0,68	0,16	0,16	0,68	1	0,32	0,84	0,16
SUB5	0,14	0,69	0,16	0,14	0,69	1	0,31	0,86	0,16
SUB6	0,15	0,68	0,16	0,15	0,68	1	0,32	0,85	0,16
SUB7	0,09	0,75	0,17	0,09	0,75	1	0,25	0,91	0,17
SUB8	0,13	0,70	0,16	0,13	0,70	1	0,30	0,87	0,16
SUB9	0,10	0,73	0,17	0,10	0,73	1	0,27	0,90	0,17
SUBBAD1	0,83	0,17	0,00	0,83	0,17	1	0,83	0,17	0,00
SUBBAD2	0,80	0,20	0,00	0,80	0,20	1	0,80	0,20	0,00
SUBBAD3	0,87	0,13	0,00	0,87	0,13	1	0,87	0,13	0,00
SUBBAD4	0,86	0,14	0,00	0,86	0,14	1	0,86	0,14	0,00
SUBBAD5	0,87	0,13	0,00	0,87	0,13	1	0,87	0,13	0,00

Πίνακας 7 –Αποτελέσματα Σεναρίου 1

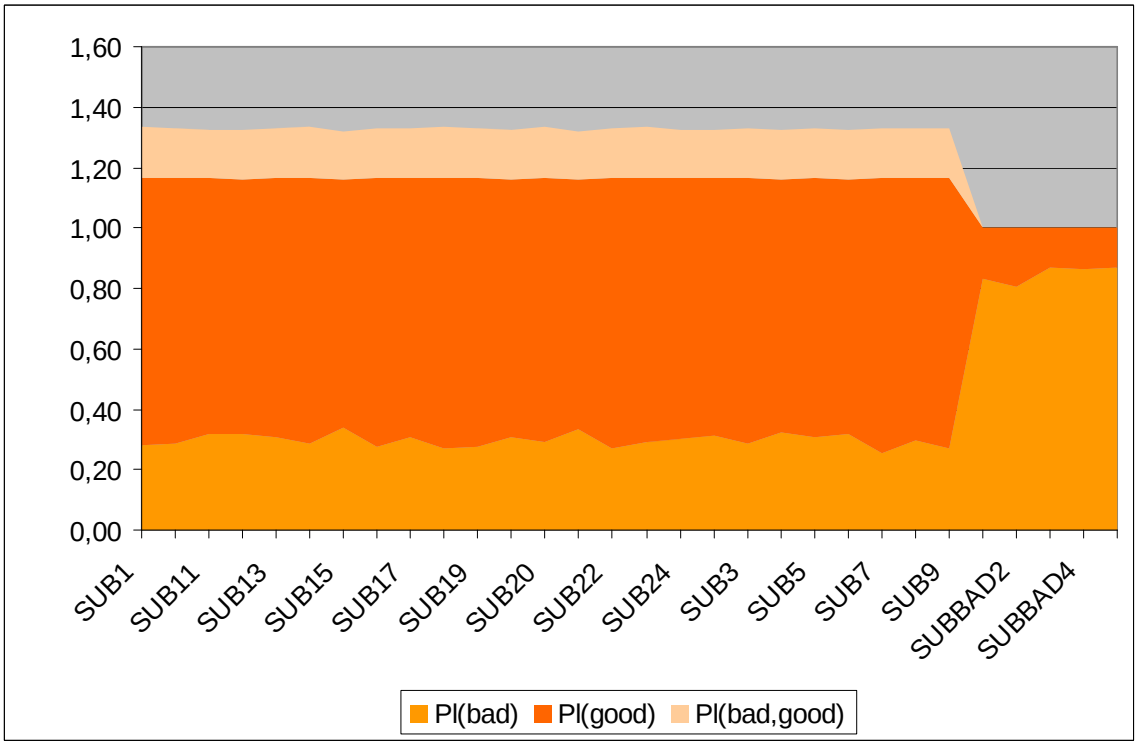
Διαβάζοντας και αναλύοντας τον παραπάνω πίνακα θα μπορούσε κανείς να καταλήξει στα εξής συμπεράσματα:

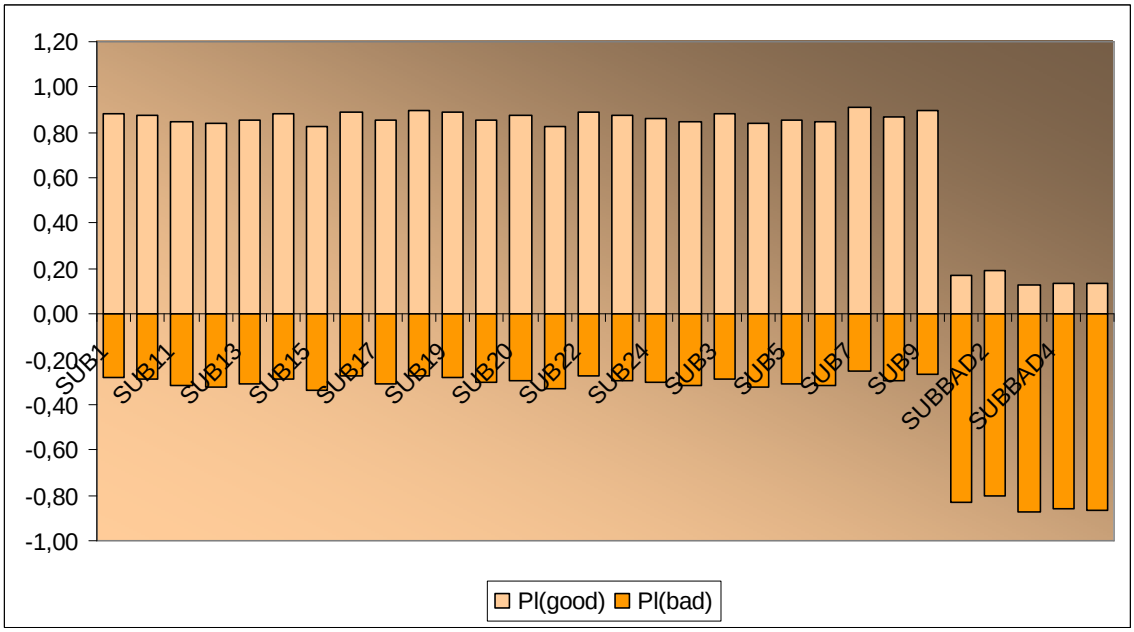
- Για ευλογοφάνεια μεγαλύτερη από 0.80 το σύστημα αποτιμά σωστά τόσο τους καλούς όσο και τους κακούς. Όπως φαίνεται από το παράδειγμα η διαφορά στην ευλογοφάνεια των καλών και των κακών είναι τόσο μεγάλη που μας επιτρέπει να επιλέξουμε ένα πολύ μικρότερο threshold.
- Γενικά παρατηρούμε ότι στους κακούς η αβεβαιότητα είναι μηδέν. Αυτό φαίνεται λογικό αφού έχουν πολύ περισσότερες ενέργειες ύποπτες από τις καλές και όλες οι

ύποπτες ενέργειες είναι συγκεντρωμένες (σε μία για την ακρίβεια)

– Η επιλογή της ευλογοφάνειας των καλών χρηστών σαν μέτρο αποτίμησης ουσιαστικά πριμοδοτεί την πίστη για τους καλούς χρήστες με την αβεβαιότητα. Σε ένα σενάριο που κινείται σε ορθολογική βάση αυτό είναι περισσότερο από ικανοποιητικό.

Γραφικά τα αποτελέσματα φαίνονται στα παρακάτω διαγράμματα:





Σενάριο 2

Στο σενάριο αυτό περιορίζουμε τις καλές ενέργειες των καλών χρηστών στο 50% και αντίστοιχα οι ύποπτες ενέργειες αυξάνονται στο 50%. Οι καλοί χρήστες στο σύστημα αντιστοιχούν στο 80% των συνολικών χρηστών. Η συμπεριφορά των κακών χρηστών παραμένει η ίδια και είναι 85% ύποπτες ενέργειες και 15% καλές ενέργειες.

Τα δεδομένα που κατέγραψε ο broker συγκεντρωτικά είναι:

		ΚΑΛΕΣ	***ΥΠΟΠΤΕΣ***							
SUB	***BROKER***	***X***	***Y1***	***Y2***	***Y3***	***Y4***	***Y5***	***Y6***	***Y7***	
SUB1	Bid000	151	20	13	26	22	34	23	16	305
SUB10	Bid000	148	24	25	27	24	22	24	19	313
SUB11	Bid000	177	20	28	28	17	29	25	18	342
SUB12	Bid000	169	25	25	23	19	32	12	19	324
SUB13	Bid000	174	27	35	22	22	29	25	20	354
SUB14	Bid000	184	27	26	21	20	26	27	24	355
SUB15	Bid000	171	35	21	19	30	20	29	27	352
SUB16	Bid000	173	22	25	22	26	33	25	28	354
SUB17	Bid000	180	26	29	27	24	28	26	16	356
SUB18	Bid000	194	23	18	18	24	27	22	25	351
SUB19	Bid000	173	19	5	29	27	32	24	31	340
SUB2	Bid000	145	22	19	29	20	24	29	21	309
SUB20	Bid000	169	22	26	22	27	34	30	30	360
SUB21	Bid000	173	28	13	28	19	23	24	33	341
SUB22	Bid000	179	21	25	21	25	29	26	23	349
SUB23	Bid000	170	23	18	31	17	31	23	21	334
SUB3	Bid000	157	21	23	19	23	26	18	26	313
SUB4	Bid000	148	23	16	23	28	21	14	28	301
SUB5	Bid000	155	21	28	19	24	24	25	25	321
SUB6	Bid000	165	25	31	21	24	18	29	18	331
SUB7	Bid000	189	19	28	25	27	21	22	20	351
SUB8	Bid000	158	27	18	26	25	22	24	22	322
SUB9	Bid000	178	17	19	17	14	25	17	18	305
SUBBAD1	Bid000	95		226						321
SUBBAD2	Bid000	100	243							343
SUBBAD3	Bid000	93		210						303
SUBBAD4	Bid000	106						260		366
SUBBAD5	Bid000	100			211					311
SUBBAD6	Bid000	99					250			349
SUBBAD7	Bid000	98			226					324
Grand Total:		**4571**	**780**	**950**	**980**	**528**	**860**	**803**	**528**	**10000**

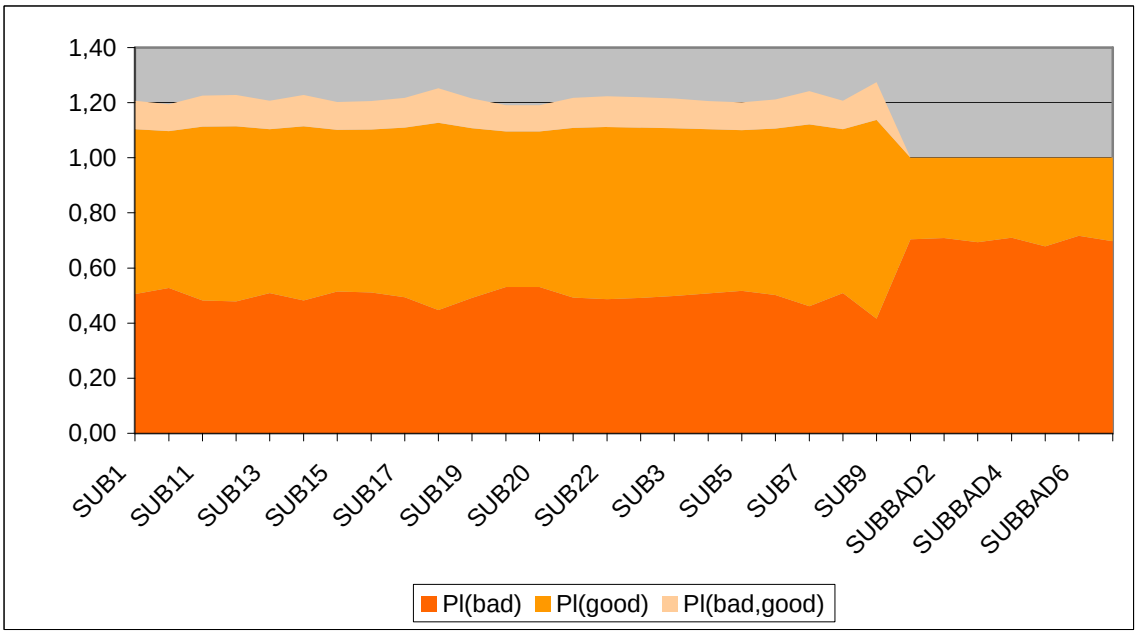
Πίνακας 8 – Δεδομένα Σεναρίου 2

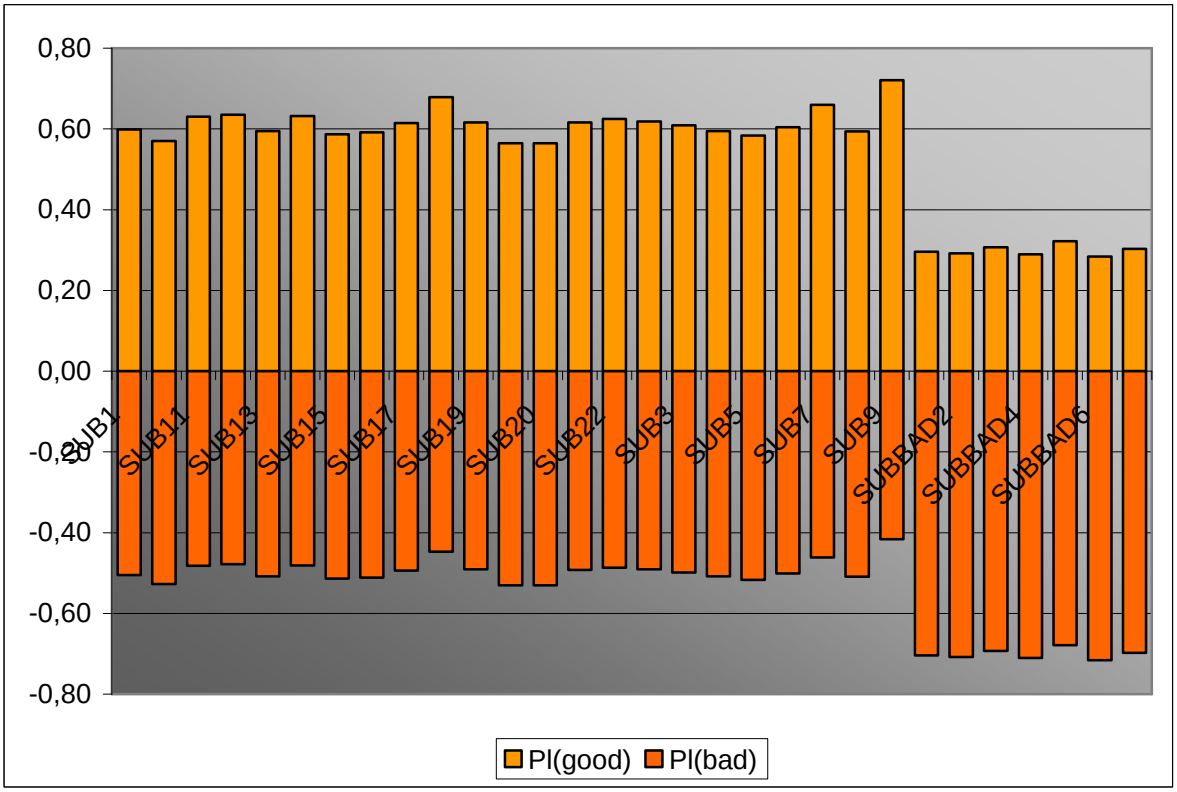
SUB	m(bad)	m(good)	m(bad,good)	bel(bad)	bel(good)	bel(bad,good)	Pl(bad)	Pl(good)	Pl(bad,good)
SUB1	0,40	0,50	0,10	0,40	0,50	1	0,50	0,60	0,10
SUB10	0,43	0,47	0,10	0,43	0,47	1	0,53	0,57	0,10
SUB11	0,37	0,52	0,11	0,37	0,52	1	0,48	0,63	0,11
SUB12	0,36	0,52	0,11	0,36	0,52	1	0,48	0,64	0,11
SUB13	0,41	0,49	0,10	0,41	0,49	1	0,51	0,59	0,10
SUB14	0,37	0,52	0,11	0,37	0,52	1	0,48	0,63	0,11
SUB15	0,41	0,49	0,10	0,41	0,49	1	0,51	0,59	0,10
SUB16	0,41	0,49	0,10	0,41	0,49	1	0,51	0,59	0,10
SUB17	0,39	0,51	0,11	0,39	0,51	1	0,49	0,61	0,11
SUB18	0,32	0,55	0,13	0,32	0,55	1	0,45	0,68	0,13
SUB19	0,38	0,51	0,11	0,38	0,51	1	0,49	0,62	0,11
SUB2	0,44	0,47	0,10	0,44	0,47	1	0,53	0,56	0,10
SUB20	0,44	0,47	0,10	0,44	0,47	1	0,53	0,56	0,10
SUB21	0,38	0,51	0,11	0,38	0,51	1	0,49	0,62	0,11
SUB22	0,38	0,51	0,11	0,38	0,51	1	0,49	0,62	0,11
SUB23	0,38	0,51	0,11	0,38	0,51	1	0,49	0,62	0,11
SUB3	0,39	0,50	0,11	0,39	0,50	1	0,50	0,61	0,11
SUB4	0,41	0,49	0,10	0,41	0,49	1	0,51	0,59	0,10
SUB5	0,42	0,48	0,10	0,42	0,48	1	0,52	0,58	0,10
SUB6	0,40	0,50	0,11	0,40	0,50	1	0,50	0,60	0,11
SUB7	0,34	0,54	0,12	0,34	0,54	1	0,46	0,66	0,12
SUB8	0,41	0,49	0,10	0,41	0,49	1	0,51	0,59	0,10
SUB9	0,28	0,58	0,14	0,28	0,58	1	0,42	0,72	0,14
SUBBAD1	0,70	0,30	0,00	0,70	0,30	1	0,70	0,30	0,00
SUBBAD2	0,71	0,29	0,00	0,71	0,29	1	0,71	0,29	0,00
SUBBAD3	0,69	0,31	0,00	0,69	0,31	1	0,69	0,31	0,00
SUBBAD4	0,71	0,29	0,00	0,71	0,29	1	0,71	0,29	0,00
SUBBAD5	0,68	0,32	0,00	0,68	0,32	1	0,68	0,32	0,00
SUBBAD6	0,72	0,28	0,00	0,72	0,28	1	0,72	0,28	0,00
SUBBAD7	0,70	0,30	0,00	0,70	0,30	1	0,70	0,30	0,00

Πίνακας 9 – Αποτελέσματα Σεναρίου 2

Μελετώντας τον παραπάνω πίνακα μπορούμε να καταλήξουμε στα συμπεράσματα:

- Παρόλο που οι καλές ενέργειες των καλών χρηστών μειώθηκαν αισθητά, η ευλογοφάνεια παραμένει σε ικανοποιητικά επίπεδα και σταθερά πάνω από 0,50. Στο παράδειγμα μας με threshold 0,55 χαρακτηρίζει σωστά τους χρήστες τόσο τους καλούς όσο και τους κακούς. Η κατάσταση είναι όμως φαίνεται οριακή και στατιστικά είναι δυνατό να προκύψουν λανθασμένες εκτιμήσεις (false positive κυρίως σε αυτό το σενάριο)
- Η ευλογοφάνεια για τους κακούς χρήστες άλλαξε αρκετά αφού μειώθηκε επίσης η ύποπτη δραστηριότητα.
- Το χαρακτηριστικό της μηδενικής αβεβαιότητας για τους κακούς παραμένει και ο λόγος είναι ότι εξακολουθούν να εφαρμόζουν την τακτική της μίας συγκεκριμένης ύποπτης ενέργειας


1,40
1,20
1,00
0,80
0,60
0,40
0,20
0,00
SUB1
SUB11
SUB13
SUB15
SUB17
SUB19
SUB20
SUB22
SUB3
SUB5
SUB7
SUB9
SUBBAD2
SUBBAD4
SUBBAD6
Pl(bad)
Pl(good)
Pl(bad,good)


0,80
0,60
0,40
0,20
0,00
-0,20
-0,40
-0,60
-0,80
SUB1
SUB11
SUB13
SUB15
SUB17
SUB19
SUB20
SUB22
SUB3
SUB5
SUB7
SUB9
SUBBAD2
SUBBAD4
SUBBAD6
Pl(good)
Pl(bad)

Σενάριο 3

Στο τρίτο σενάριο έγινε μια προσπάθεια να πιεστεί λίγο το σύστημα. Το ποσοστό των καλών και των κακών χρηστών είναι 80% και 20% αντίστοιχα. Η συμπεριφορά των καλών είναι 80% καλές ενέργειες και 20% ύποπτες ενέργειες. Αυτό που θα αλλάζει αρκετά σε αυτό το σενάριο είναι το ποσοστό των ύποπτων και των καλών ενεργειών των κακών χρηστών το οποίο ανέρχεται στο 50% τόσο για τις ύποπτες όσο και για τις καλές ενέργειες. Αυτό σημαίνει ότι ένας κακόβουλος χρήστης έχει τις μισές του ενέργειες ως ύποπτες και τις άλλες μισές κανονικές και υποθέτουμε ότι αυτό είναι αρκετό για να στρεσάρει το σύστημα.

Τα δεδομένα που κατέγραψε ο broker συγκεντρωτικά είναι:

		ΚΑΛΕΣ	***ΥΠΟΠΤΕΣ***							
SUB	***BROKER***	***X***	***Y1***	***Y2***	***Y3***	***Y4***	***Y5***	***Y6***	***Y7***	
SUB1	Bid000	296	7	13	5	11	16	9	12	369
SUB10	Bid000	272	7	15	11	4	12	11	9	341
SUB11	Bid000	250	5	3	10	12	9	6	6	301
SUB12	Bid000	253	8	19	7	8	5	13	8	321
SUB13	Bid000	257	8	5	10	8	11	11	6	316
SUB14	Bid000	262	12	15	12	4	13	10	12	340
SUB15	Bid000	250	12	13	8	9	10	8	9	319
SUB16	Bid000	263	9	7	7	9	11	7	9	322
SUB17	Bid000	281	14	9	12	10	15	8	10	359
SUB18	Bid000	284	11	11	4	9	10	10	9	348
SUB19	Bid000	263	10	11	7	6	12	11	8	328
SUB2	Bid000	257	5	9	9	6	6	5	6	303
SUB20	Bid000	281	10	8	8	4	6	17	14	348
SUB21	Bid000	268	11	4	9	8	8	11	8	327
SUB3	Bid000	283	8	13	12	11	20	10	14	371
SUB4	Bid000	252	10	12	9	6	9	11	8	317
SUB5	Bid000	281	9	14	12	9	14	5	3	347
SUB6	Bid000	258	11	8	14	14	7	13	5	330
SUB7	Bid000	251	15	7	12	7	11	8	12	323
SUB8	Bid000	273	13	11	12	11	7	14	14	355
SUB9	Bid000	264	10	11	5	6	5	8	11	320
SUBBAD1	Bid000	197							180	377
SUBBAD2	Bid000	179				155				334
SUBBAD3	Bid000	173				163				336
SUBBAD4	Bid000	183			164					347
SUBBAD5	Bid000	155							164	319
SUBBAD6	Bid000	172				144				316
SUBBAD7	Bid000	157						186		343
SUBBAD9	Bid000	155					144			299
SUBBAD8	Bid000	158					166			324
	Grand Total:	**7128**	**205**	**218**	**359**	**634**	**527**	**392**	**537**	**10000**

Πίνακας 10 – Δεδομένα Σεναρίου 3

Στη συνέχεια υπολογίζουμε πάλι τις τιμές βασικής πιθανότητας, πίστης και ευλογοφάνειας.

SUB	m(bad)	m(good)	m(bad,good)	bel(bad)	bel(good	bel(bad,good)	Pl(bad)	Pl(good)	Pl(bad,good)
SUB1	0,04	0,80	0,15	0,04	0,80	1	0,20	0,96	0,15
SUB10	0,05	0,80	0,16	0,05	0,80	1	0,20	0,95	0,16
SUB11	0,03	0,83	0,14	0,03	0,83	1	0,17	0,97	0,14
SUB12	0,05	0,79	0,16	0,05	0,79	1	0,21	0,95	0,16
SUB13	0,04	0,81	0,15	0,04	0,81	1	0,19	0,96	0,15
SUB14	0,07	0,77	0,16	0,07	0,77	1	0,23	0,93	0,16
SUB15	0,05	0,78	0,16	0,05	0,78	1	0,22	0,95	0,16
SUB16	0,03	0,82	0,15	0,03	0,82	1	0,18	0,97	0,15
SUB17	0,06	0,78	0,16	0,06	0,78	1	0,22	0,94	0,16
SUB18	0,03	0,82	0,15	0,03	0,82	1	0,18	0,97	0,15
SUB19	0,04	0,80	0,16	0,04	0,80	1	0,20	0,96	0,16
SUB2	0,02	0,85	0,13	0,02	0,85	1	0,15	0,98	0,13
SUB20	0,04	0,81	0,15	0,04	0,81	1	0,19	0,96	0,15
SUB21	0,03	0,82	0,15	0,03	0,82	1	0,18	0,97	0,15
SUB3	0,07	0,76	0,16	0,07	0,76	1	0,24	0,93	0,16
SUB4	0,05	0,79	0,16	0,05	0,79	1	0,21	0,95	0,16
SUB5	0,04	0,81	0,15	0,04	0,81	1	0,19	0,96	0,15
SUB6	0,06	0,78	0,16	0,06	0,78	1	0,22	0,94	0,16
SUB7	0,06	0,78	0,16	0,06	0,78	1	0,22	0,94	0,16
SUB8	0,07	0,77	0,16	0,07	0,77	1	0,23	0,93	0,16
SUB9	0,03	0,83	0,15	0,03	0,83	1	0,18	0,97	0,15
SUBBAD1	0,48	0,52	0,00	0,48	0,52	1	0,48	0,52	0,00
SUBBAD2	0,46	0,54	0,00	0,46	0,54	1	0,46	0,54	0,00
SUBBAD3	0,49	0,51	0,00	0,49	0,51	1	0,49	0,51	0,00
SUBBAD4	0,47	0,53	0,00	0,47	0,53	1	0,47	0,53	0,00
SUBBAD5	0,51	0,49	0,00	0,51	0,49	1	0,51	0,49	0,00
SUBBAD6	0,46	0,54	0,00	0,46	0,54	1	0,46	0,54	0,00
SUBBAD7	0,54	0,46	0,00	0,54	0,46	1	0,54	0,46	0,00
SUBBAD9	0,48	0,52	0,00	0,48	0,52	1	0,48	0,52	0,00
SUBBAD8	0,51	0,49	0,00	0,51	0,49	1	0,51	0,49	0,00

Πίνακας 11 – Αποτελέσματα Σεναρίου 3

Διαβάζοντας και αναλύοντας τον παραπάνω πίνακα μπορούμε να καταλήξουμε στα εξής συμπεράσματα:

- Η αύξηση από 70% σε 80% των καλών κινήσεων και η ταυτόχρονη μείωση των κακών κινήσεων οδήγησε την ευλογοφάνεια των καλών χρηστών πολύ υψηλότερα.
- Το ίδιο συνέβη και με την ευλογοφάνεια για τους κακούς χρήστες αλλά φαίνεται ότι δεν πριμοδοτήθηκε τόσο σημαντικά ώστε να φτάσει στα προηγούμενα επίπεδα των καλών. Μάλιστα ένα threshold 0.55 όπως και πριν χαρακτηρίζει σωστά τους καλούς και τους κακούς και ισχύει και για τα δύο προηγούμενα σενάρια. Είναι όμως φανερό ότι η κατάσταση είναι πλέον οριακή και στατιστικά είναι δυνατό να προκύψουν λανθασμένες εκτιμήσεις (false negative κυρίως στο συγκεκριμένο σενάριο)
- Το χαρακτηριστικό της μηδενικής αβεβαιότητας για τους κακούς παραμένει φυσικά και εδώ για τους ίδιους λόγους.

Σενάριο 4

Πρόκειται για το πιο «δύσκολο» σενάριο μέχρι τώρα που αφορά στο κεντρικοποιημένο μοντέλο. Εδώ οι καλοί χρήστες δημιουργούν 70% ύποπτες ενέργειες στο σύστημα με τους κακούς να δημιουργούν 85% ύποπτες ενέργειες στο σύστημα. Δηλαδή οι καλοί χρήστες έχουν πολύ υψηλό ποσοστό ύποπτων ενεργειών και ταυτόχρονα το ποσοστό αυτό δεν απέχει πολύ από των κακόβουλων χρηστών. Φυσικά είναι μάλλον μη ρεαλιστική υπόθεση αφού είναι δύσκολο ένα σύστημα να ανέχεται υψηλό ποσοστό «βλαβερών» ενεργειών και να μη θέλει να δέχεται ένα ποσοστό ελαφρώς μεγαλύτερο. Έτσι το σενάριο «παίζει» στα όρια του μοντέλου μας.

		ΚΑΛΕΣ	***ΥΠΟΠΤΕΣ***							
SUB	***BROKER***	***X***	***Y1***	***Y2***	***Y3***	***Y4***	***Y5***	***Y6***	***Y7***	
SUB1	Bid000	127	42	35	23	30	42	32	40	371
SUB10	Bid000	115	18	28	32	30	30	22	19	294
SUB11	Bid000	114	35	25	24	20	37	32	17	304
SUB12	Bid000	152	30	28	29	33	28	28	30	358
SUB13	Bid000	123	36	35	32	23	35	32	20	336
SUB14	Bid000	123	20	36	26	19	22	22	34	302
SUB15	Bid000	124	25	24	37	24	36	22	25	317
SUB16	Bid000	116	34	28	26	39	32	38	26	339
SUB17	Bid000	122	40	35	28	34	21	22	30	332
SUB18	Bid000	116	24	29	30	26	35	31	36	327
SUB19	Bid000	107	33	26	32	28	25	29	30	310
SUB2	Bid000	106	36	28	27	29	47	35	28	336
SUB20	Bid000	97	41	32	31	26	29	29	40	325
SUB21	Bid000	117	28	41	31	45	32	36	35	365
SUB22	Bid000	118	25	30	34	32	34	31	31	335
SUB23	Bid000	126	39	27	27	27	27	27	20	320
SUB24	Bid000	122	33	25	35	26	29	30	25	325
SUB25	Bid000	122	30	32	36	38	28	38	26	350
SUB26	Bid000	152	28	35	34	20	29	33	28	359
SUB3	Bid000	114	26	35	37	35	31	32	44	354
SUB4	Bid000	112	30	28	29	28	36	43	30	336
SUB5	Bid000	116	41	30	30	39	36	26	30	348
SUB6	Bid000	112	32	25	31	33	33	24	30	320
SUB7	Bid000	124	38	30	28	40	37	31	27	355
SUB8	Bid000	124	38	29	25	28	26	33	31	334
SUB9	Bid000	129	31	26	34	33	36	21	26	336
SUBBAD1	Bid000	47						262		309
SUBBAD2	Bid000	51		282						333
SUBBAD3	Bid000	50						288		338
SUBBAD4	Bid000	52				280				332
Grand Total		3330	833	1064	788	1065	833	1329	758	10000

Πίνακας 12 – Δεδομένα Σεναρίου 4

SUB	m(bad)	m(good)	m(bad,good)	bel(bad)	bel(good	bel(bad,good)	Pl(bad)	Pl(good)	Pl(bad,good)
SUB1	0,61	0,34	0,05	0,61	0,34	1	0,66	**0,39**	0,05
SUB10	0,54	0,39	0,07	0,54	0,39	1	0,61	**0,46**	0,07
SUB11	0,56	0,38	0,06	0,56	0,38	1	0,63	**0,44**	0,06
SUB12	0,50	0,42	0,08	0,50	0,42	1	0,58	**0,50**	0,08
SUB13	0,57	0,37	0,06	0,57	0,37	1	0,63	**0,43**	0,06
SUB14	0,52	0,41	0,07	0,52	0,41	1	0,59	**0,48**	0,07
SUB15	0,54	0,39	0,07	0,54	0,39	1	0,61	**0,46**	0,07
SUB16	0,61	0,34	0,05	0,61	0,34	1	0,66	**0,39**	0,05
SUB17	0,57	0,37	0,06	0,57	0,37	1	0,63	**0,43**	0,06
SUB18	0,59	0,35	0,06	0,59	0,35	1	0,65	**0,41**	0,06
SUB19	0,60	0,35	0,05	0,60	0,35	1	0,65	**0,40**	0,05
SUB2	0,64	0,32	0,04	0,64	0,32	1	0,68	**0,36**	0,04
SUB20	0,66	0,30	0,04	0,66	0,30	1	0,70	0,34	0,04
SUB21	0,63	0,32	0,05	0,63	0,32	1	0,68	**0,37**	0,05
SUB22	0,59	0,35	0,06	0,59	0,35	1	0,65	**0,41**	0,06
SUB23	0,54	0,39	0,07	0,54	0,39	1	0,61	**0,46**	0,07
SUB24	0,56	0,38	0,06	0,56	0,38	1	0,62	**0,44**	0,06
SUB25	0,60	0,35	0,05	0,60	0,35	1	0,65	**0,40**	0,05
SUB26	0,50	0,42	0,08	0,50	0,42	1	0,58	**0,50**	0,08
SUB3	0,63	0,32	0,05	0,63	0,32	1	0,68	**0,37**	0,05
SUB4	0,62	0,33	0,05	0,62	0,33	1	0,67	**0,38**	0,05
SUB5	0,62	0,33	0,05	0,62	0,33	1	0,67	**0,38**	0,05
SUB6	0,60	0,35	0,05	0,60	0,35	1	0,65	**0,40**	0,05
SUB7	0,60	0,35	0,05	0,60	0,35	1	0,65	**0,40**	0,05
SUB8	0,57	0,37	0,06	0,57	0,37	1	0,63	**0,43**	0,06
SUB9	0,55	0,38	0,07	0,55	0,38	1	0,62	**0,45**	0,07
SUBBAD1	0,85	0,15	0,00	0,85	0,15	1	0,85	0,15	0,00
SUBBAD2	0,85	0,15	0,00	0,85	0,15	1	0,85	0,15	0,00
SUBBAD3	0,85	0,15	0,00	0,85	0,15	1	0,85	0,15	0,00
SUBBAD4	0,84	0,16	0,00	0,84	0,16	1	0,84	0,16	0,00

Πίνακας 13 – Αποτελέσματα Σεναρίου 4

Τα συμπεράσματα σε αυτό το ακραίο σενάριο είναι τα παρακάτω:

- Η ευλογοφάνεια των καλών χρηστών έχει πέσει κάτω από το 50% ενώ η ευλογοφάνεια (κακής συμπεριφοράς) των καλών επίσης έχει ανέβει αρκετά πάνω από το 50%
- Η ευλογοφάνεια για τους κακούς χρήστες δεν έχει αλλάξει.
- Πάντως ένα threshold 0.35 χαρακτηρίζει σωστά τους καλούς εκτός από μία περίπτωση (false positive) και επίσης χαρακτηρίζει σωστά τους κακούς. Στο σενάριο ήδη είναι εμφανές ότι θα υπάρχουν λανθασμένα αποτελέσματα.
- Το σενάριο αυτό λαμβάνεται υπόψη μόνο στο βαθμό που εξετάζουμε τα όρια του μοντέλου μας. Είναι απίθανο να υπάρχει σύστημα με αυτές τις υποθέσεις.

Σενάριο 5

Στο σενάριο αυτό οι παράμετροι το συστήματος είναι ίδιοι με το σενάριο 1 με την διαφορά ότι υπάρχουν 3 Brokers στο σύστημα. Ουσιαστικά κάνουμε χρήση του κανόνα του Dempster για τον συνδυασμό εκτιμήσεων. Στο σύστημα υπάρχουν 80% καλοί χρήστες στο σύνολο.

Εδώ οι καλοί χρήστες δημιουργούν 70% καλές ενέργειες στο σύστημα με τους κακούς να δημιουργούν 85% ύποπτες ενέργειες στο σύστημα. Πρέπει να επισημανθεί ότι ένας κακός χρήστης επαναλαμβάνει την ίδια ύποπτη ενέργεια ενώ η συμπεριφορά του είναι παρόμοια σε όλους τους brokers που είναι συνδεδεμένος.Δεν κάνει όμως τις ίδιες ακριβώς ενέργειες παρά μόνο κατά μέσο όρο.

		ΚΑΛΕΣ	***ΥΠΟΠΤΕΣ***							
BROKER	***SUB***	***X***	***Y1***	***Y2***	***Y3***	***Y4***	***Y5***	***Y6***	***Y7***	
Bid000	SUB1	247	24	14	8	17	8	14	13	345
	SUB10	225	11	19	18	15	12	11	18	329
	SUB11	232	14	10	18	16	7	11	19	327
	SUB12	244	14	10	16	20	14	16	15	349
	SUB13	245	16	9	16	18	19	9	11	343
	SUB14	247	14	15	13	10	11	21	15	346
	SUB15	221	15	16	16	12	17	17	13	327
	SUB16	218	16	11	15	14	12	10	14	310
	SUB17	247	10	13	15	12	19	10	12	338
	SUB18	221	11	15	13	16	14	18	18	326
	SUB19	231	15	10	24	14	16	22	10	342
	SUB2	236	18	11	13	16	13	7	15	329
	SUB20	239	14	18	13	21	12	15	13	345
	SUB21	240	13	12	16	10	11	17	14	333
	SUB22	239	12	11	18	24	8	13	17	342
	SUB23	253	15	23	25	18	12	14	13	373
	SUB24	234	18	17	11	15	12	16	13	336
	SUB25	232	19	10	11	15	10	15	9	321
	SUB3	227	14	15	18	19	13	10	9	325
	SUB4	239	14	13	13	12	7	21	25	344
	SUB5	244	13	11	22	21	14	11	20	356
	SUB6	229	13	17	11	20	13	13	7	323
	SUB7	228	25	14	12	17	8	16	10	330
	SUB8	270	13	19	14	19	16	13	13	377
	SUB9	203	19	20	16	15	17	10	13	313
	SUBBAD1	59					281			340
	SUBBAD2	44		264						308
	SUBBAD3	50					267			317
	SUBBAD4	52					256			308
	SUBBAD5	39	259							298
Bid000 Total		**6135**	**639**	**617**	**385**	**406**	**1119**	**350**	**349**	**10000**
Bid001	SUB1	219	14	17	18	19	20	18	20	345
	SUB10	234	11	12	17	15	12	7	21	329
	SUB11	231	17	13	18	13	12	11	12	327
	SUB12	245	12	13	20	15	14	12	18	349

	SUB13	237	21	20	10	18	15	12	10	343
	SUB14	243	19	17	10	12	12	15	18	346
	SUB15	231	16	13	11	7	18	16	15	327
	SUB16	219	9	13	13	15	13	12	16	310
	SUB17	246	14	11	14	13	13	10	17	338
	SUB18	239	16	15	5	9	10	15	17	326
	SUB19	242	13	19	21	11	13	12	11	342
	SUB2	226	17	17	15	13	14	12	15	329
	SUB20	232	16	13	13	30	13	13	15	345
	SUB21	241	14	7	22	11	15	12	11	333
	SUB22	235	15	15	10	21	13	20	13	342
	SUB23	253	22	21	19	18	18	14	8	373
	SUB24	240	15	12	10	9	19	10	21	336
	SUB25	221	16	13	15	17	8	15	16	321
	SUB3	232	13	10	16	15	12	16	11	325
	SUB4	229	19	15	12	17	8	19	25	344
	SUB5	255	13	12	23	15	11	10	17	356
	SUB6	222	12	17	18	16	11	12	15	323
	SUB7	231	17	16	6	16	12	16	16	330
	SUB8	262	21	15	16	16	22	16	9	377
	SUB9	221	16	14	12	15	12	8	15	313
	SUBBAD1	55					285			340
	SUBBAD2	48		260						308
	SUBBAD3	48					269			317
	SUBBAD4	49					259			308
	SUBBAD5	41	257							298
Bid001 Total		**6127**	**645**	**620**	**364**	**376**	**1153**	**333**	**382**	**10000**
Bid002	SUB1	251	17	12	11	16	13	15	10	345
	SUB10	231	14	17	16	11	14	13	13	329
	SUB11	234	12	16	15	18	10	6	16	327
	SUB12	253	12	12	13	18	14	12	15	349
	SUB13	241	14	19	10	14	15	13	17	343
	SUB14	231	14	15	12	20	19	22	13	346
	SUB15	226	11	10	15	21	20	12	12	327
	SUB16	215	16	14	13	12	16	9	15	310
	SUB17	223	15	16	16	17	18	18	15	338
	SUB18	238	12	11	8	12	9	16	20	326
	SUB19	239	13	17	17	15	16	12	13	342
	SUB2	235	12	13	15	14	10	9	21	329
	SUB20	222	14	16	23	28	9	15	18	345
	SUB21	221	12	16	25	14	15	16	14	333
	SUB22	265	10	10	14	11	9	13	10	342
	SUB23	275	17	12	17	15	12	19	6	373
	SUB24	235	13	13	15	13	14	7	26	336
	SUB25	223	17	12	12	20	12	17	8	321
	SUB3	233	16	13	15	18	11	8	11	325
	SUB4	256	10	15	11	12	9	17	14	344
	SUB5	265	14	17	16	12	4	11	17	356
	SUB6	220	11	15	13	18	21	12	13	323
	SUB7	238	12	11	5	21	12	18	13	330
	SUB8	275	11	19	15	15	13	16	13	377
	SUB9	219	18	13	12	11	15	10	15	313
	SUBBAD1	45					295			340

	SUBBAD2	43		265						308
	SUBBAD3	37					280			317
	SUBBAD4	49					259			308
	SUBBAD5	49	249							298
Bid002 Total		**6187**	**586**	**619**	**354**	**396**	**1164**	**336**	**358**	**10000**
Grand Total		***18449***	***1870***	***1856***	***1103***	***1178***	***3436***	***1019***	***1089***	***30000***

Πίνακας 14 – Δεδομένα Σεναρίου 5

Στον παρακάτω πίνακα περιγράφουμε τα αποτελέσματα συγκεντρωτικά και επιπλέον για λόγους σύγκρισης εμφανίζουμε και τις τιμές πιθανότητας που θα δίναμε αν το σύστημα ήταν κεντρικοποιημένο και υπήρχε ένας κεντρικός broker που συγκέντρωνε και κατέγραφε όλες τις κινήσεις.

		Distributed			**Non-distributed**					
SUB	***BROK***	***m_d(bad)***	***m_d(good)***	***m_d(bad,good)***	***b***	***g***	***u***	***PI_d(bad)***	***PI_d(good)***	***PI_d(bad,good)***
SUB1	ALL	0,02	0,74	0,24	0,14	0,69	0,17	0,26	0,98	0,24
SUB10	ALL	0,02	0,75	0,24	0,14	0,70	0,17	0,25	0,98	0,24
SUB11	ALL	0,01	0,76	0,23	0,12	0,71	0,17	0,24	0,99	0,23
SUB12	ALL	0,01	0,75	0,23	0,12	0,71	0,17	0,25	0,99	0,23
SUB13	ALL	0,01	0,75	0,24	0,13	0,70	0,17	0,25	0,99	0,24
SUB14	ALL	0,02	0,74	0,24	0,14	0,69	0,17	0,26	0,98	0,24
SUB15	ALL	0,02	0,74	0,24	0,14	0,69	0,16	0,26	0,98	0,24
SUB16	ALL	0,02	0,75	0,24	0,13	0,70	0,17	0,25	0,98	0,24
SUB17	ALL	0,01	0,75	0,24	0,13	0,71	0,17	0,25	0,99	0,24
SUB18	ALL	0,01	0,76	0,23	0,12	0,71	0,17	0,24	0,99	0,23
SUB19	ALL	0,02	0,75	0,24	0,14	0,69	0,16	0,25	0,98	0,24
SUB2	ALL	0,01	0,75	0,23	0,13	0,71	0,17	0,25	0,99	0,23
SUB20	ALL	0,02	0,72	0,26	0,17	0,67	0,16	0,28	0,98	0,26
SUB21	ALL	0,01	0,75	0,24	0,13	0,70	0,17	0,25	0,99	0,24
SUB22	ALL	0,01	0,76	0,23	0,11	0,72	0,17	0,24	0,99	0,23
SUB23	ALL	0,02	0,75	0,24	0,14	0,70	0,16	0,25	0,98	0,24
SUB24	ALL	0,02	0,75	0,24	0,13	0,70	0,17	0,25	0,98	0,24
SUB25	ALL	0,01	0,75	0,24	0,13	0,70	0,17	0,25	0,99	0,24
SUB3	ALL	0,01	0,76	0,23	0,12	0,71	0,17	0,24	0,99	0,23
SUB4	ALL	0,02	0,75	0,24	0,13	0,70	0,16	0,25	0,98	0,24
SUB5	ALL	0,01	0,76	0,23	0,12	0,72	0,17	0,24	0,99	0,23
SUB6	ALL	0,02	0,74	0,24	0,14	0,69	0,16	0,26	0,98	0,24
SUB7	ALL	0,02	0,75	0,23	0,13	0,70	0,16	0,25	0,98	0,23
SUB8	ALL	0,01	0,76	0,23	0,12	0,71	0,17	0,24	0,99	0,23
SUB9	ALL	0,02	0,74	0,24	0,15	0,68	0,16	0,26	0,98	0,24
SUBBAD1	ALL	0,99	0,01	0,00	0,84	0,16	0,00	0,99	0,01	0,00
SUBBAD2	ALL	1,00	0,00	0,00	0,85	0,15	0,00	1,00	0,00	0,00
SUBBAD3	ALL	1,00	0,00	0,00	0,86	0,14	0,00	1,00	0,00	0,00
SUBBAD4	ALL	0,99	0,01	0,00	0,84	0,16	0,00	0,99	0,01	0,00
SUBBAD5	ALL	1,00	0,00	0,00	0,86	0,14	0,00	1,00	0,00	0,00

Πίνακας 15 – Αποτελέσματα Σεναρίου 5

Γενικά τα αποτελέσματα φαίνονται πολύ ικανοποιητικά στο κατανεμημένο μοντέλο. Πιο συγκεκριμένα έχουμε τις παρακάτω παρατηρήσεις:

- Η ευλογοφάνεια των καλών (για την υπόθεση Καλός) και των κακών (για την υπόθεση Κακός) ενισχύεται σημαντικά με τον συνδυασμό των εκτιμήσεων
- Η αβεβαιότητα ενισχύεται επίσης αλλά όχι τόσο πολύ
- Η καταχώρηση πιθανότητας m(bad) για τους καλούς και m(good) για τους κακούς σχεδόν μηδενίζεται. Αυτό δείχνει πόσο σημαντικό ρόλο στην θεωρία Dempster-Shafer παίζει ο κανόνας του Dempster για τον συνδυασμό εκτιμήσεων
- Αν το μέτρο αποτίμησης είναι η ευλογοφάνεια (για την υπόθεση Καλός Pl(good)) τότε αποδίδει πολύ αποτελεσματικά την συμπεριφορά τόσο των καλών όσο και των κακών (πχ με ένα threshold 0.80)
- Το γεγονός ότι ο κάθε broker έχει μια ελαφρώς διαφοροποιημένη καταγραφή λόγω της τυχαιότητας που παράγει ενέργειες ο κάθε subscriber καθιστά το μοντέλο πιο αξιόπιστο σε πραγματικές συνθήκες.

Σενάριο 6

Στο σενάριο αυτό οι παράμετροι το συστήματος είναι ίδιοι με το σενάριο 2 με την διαφορά ότι υπάρχουν 3 Brokers στο σύστημα. Όπως παραπάνω κάνουμε χρήση του κανόνα του Dempster για τον συνδυασμό εκτιμήσεων. Στο σύστημα υπάρχουν 80% καλοί χρήστες στο σύνολο.

Οι καλοί χρήστες δημιουργούν 50% καλές ενέργειες (και 50% ύποπτες) στο σύστημα με τους κακούς να δημιουργούν 85% ύποπτες ενέργειες στο σύστημα. Ένας κακός χρήστης επαναλαμβάνει την ίδια ύποπτη ενέργεια ενώ η συμπεριφορά του είναι παρόμοια σε όλους τους brokers που είναι συνδεδεμένος ενώ δεν κάνει τις ίδιες ακριβώς ενέργειες παρά μόνο κατά μέσο όρο.

Ο παρακάτω πίνακας περιγράφει τα αποτελέσματα συγκεντρωτικά και επιπλέον για λόγους σύγκρισης εμφανίζουμε και τις τιμές πιθανότητας που θα δίναμε αν το σύστημα ήταν κεντρικοποιημένο και υπήρχε ένας κεντρικός broker που συγκέντρωνε και κατέγραφε όλες τις κινήσεις.

		Distributed			Non-distributed					
SUB	***BROKER***	m_d***(bad)***	m_d***(good)***	m_d***(bad,good0***	***b***	***g***	***u***	Pl_d***(bad)***	Pl_d***(good)***	Pl_d***(bad,good)***
SUB1	ALL	0,24	0,47	0,29	0,40	0,50	0,11	0,53	0,76	0,29
SUB10	ALL	0,22	0,47	0,31	0,39	0,50	0,11	0,53	0,78	0,31
SUB11	ALL	0,25	0,45	0,30	0,41	0,49	0,10	0,55	0,75	0,30
SUB12	ALL	0,24	0,45	0,31	0,40	0,49	0,10	0,55	0,76	0,31
SUB13	ALL	0,24	0,48	0,28	0,39	0,50	0,11	0,52	0,76	0,28
SUB14	ALL	0,32	0,41	0,28	0,44	0,47	0,10	0,59	0,68	0,28
SUB15	ALL	0,26	0,45	0,30	0,41	0,49	0,10	0,55	0,74	0,30
SUB16	ALL	0,20	0,49	0,31	0,38	0,51	0,11	0,51	0,80	0,31
SUB17	ALL	0,19	0,49	0,32	0,38	0,51	0,11	0,51	0,81	0,32
SUB18	ALL	0,22	0,47	0,31	0,39	0,50	0,11	0,53	0,78	0,31
SUB19	ALL	0,19	0,49	0,32	0,38	0,51	0,11	0,51	0,81	0,32
SUB2	ALL	0,19	0,52	0,29	0,37	0,52	0,11	0,48	0,81	0,29
SUB20	ALL	0,27	0,44	0,29	0,42	0,48	0,10	0,56	0,73	0,29
SUB21	ALL	0,31	0,40	0,29	0,44	0,47	0,09	0,60	0,69	0,29
SUB22	ALL	0,18	0,51	0,30	0,36	0,52	0,11	0,49	0,82	0,30
SUB23	ALL	0,26	0,44	0,30	0,41	0,49	0,10	0,56	0,74	0,30
SUB24	ALL	0,22	0,49	0,29	0,38	0,51	0,11	0,51	0,78	0,29
SUB3	ALL	0,29	0,41	0,30	0,43	0,47	0,10	0,59	0,71	0,30
SUB4	ALL	0,24	0,47	0,30	0,40	0,50	0,11	0,53	0,76	0,30
SUB5	ALL	0,23	0,45	0,32	0,40	0,49	0,10	0,55	0,77	0,32
SUB6	ALL	0,25	0,43	0,32	0,41	0,48	0,10	0,57	0,75	0,32
SUB7	ALL	0,26	0,45	0,29	0,41	0,49	0,10	0,55	0,74	0,29
SUB8	ALL	0,26	0,45	0,29	0,41	0,49	0,10	0,55	0,74	0,29
SUB9	ALL	0,20	0,50	0,30	0,37	0,52	0,11	0,50	0,80	0,30
SUBBAD1	ALL	0,93	0,07	0,00	0,70	0,30	0,00	0,93	0,07	0,00
SUBBAD2	ALL	0,94	0,06	0,00	0,71	0,29	0,00	0,94	0,06	0,00
SUBBAD3	ALL	0,93	0,07	0,00	0,70	0,30	0,00	0,93	0,07	0,00
SUBBAD4	ALL	0,92	0,08	0,00	0,70	0,30	0,00	0,92	0,08	0,00

SUBBAD5 ALL	0,93	0,07	0,00	0,70	0,30	0,00	0,93	0,07	0,00
SUBBAD6 ALL	0,95	0,05	0,00	0,72	0,28	0,00	0,95	0,05	0,00

Πίνακας 16 – Αποτελέσματα Σεναρίου 6

Γενικά τα αποτελέσματα φαίνονται ικανοποιητικά και σε αυτό το σενάριο για κατανεμημένο μοντέλο. Πιο συγκεκριμένα έχουμε τις παρακάτω παρατηρήσεις:

- Η ευλογοφάνεια των καλών (για την υπόθεση Καλός) και των κακών (για την υπόθεση Κακός) ενισχύεται με τον συνδυασμό των εκτιμήσεων όπως και προηγουμένως
- Η αβεβαιότητα όμως ενισχύεται κατά πολύ αυτή τη φορά. Για την ακρίβεια είναι δύο με τρεις φορές μεγαλύτερη από ότι θα ήταν στο κεντρικοποιημένο μοντέλο.
- Η καταχώρηση πιθανότητας m(bad) για τους καλούς και m(good) για τους κακούς μειώνεται αλλά όχι πολύ σε σχέση με το σενάριο 5.
- Αν το μέτρο αποτίμησης είναι η ευλογοφάνεια (για την υπόθεση Καλός Pl(good)) τότε αποδίδει πολύ αποτελεσματικά την συμπεριφορά τόσο των καλών όσο και των κακών. Με ένα threshold 0.55 καλύπτει και την περίπτωση του κεντρικοποιημένου μοντέλου με τις ίδιες παραμέτρους
- Μπορούμε επίσης να παρατηρήσουμε ότι σε σχέση με το κεντρικοποιημένο μοντέλο υπάρχει διαφοροποίηση και στην πιθανότητα m(good) η οποία είναι ανεπαίσθητα χαμηλότερη

Σενάριο 7

Πρόκειται για ένα πολύ τραβηγμένο σενάριο στο οποίο οι κακοί χρήστες έχουν ελάχιστα χειρότερη συμπεριφορά σε σχέση με τους καλούς χρήστες. Πιο συγκεκριμένα έχουν 35% ύποπτες ενέργειες σε σχέση με το 20% ύποπτων ενεργειών των καλών χρηστών.

		Distributed			Non-distributed					
SUB	***BROKER***	m_d***(bad)***	m_d***(good)***	m_d***(bad,good)***	***b***	***g***	***u***	Pl_d***(bad)***	Pl_d***(good)***	Pl_d***(bad,good)***
SUB1	ALL	0,00	0,83	0,17	0,03	0,82	0,15	0,17	**1,00**	0,17
SUB10	ALL	0,00	0,81	0,19	0,04	0,80	0,16	0,19	**1,00**	0,19
SUB11	ALL	0,00	0,81	0,19	0,05	0,79	0,16	0,19	**1,00**	0,19
SUB12	ALL	0,00	0,81	0,19	0,06	0,78	0,16	0,19	**1,00**	0,19
SUB13	ALL	0,00	0,83	0,17	0,04	0,81	0,15	0,17	**1,00**	0,17
SUB14	ALL	0,00	0,81	0,18	0,04	0,81	0,15	0,19	**1,00**	0,18
SUB15	ALL	0,00	0,83	0,16	0,04	0,80	0,16	0,17	**1,00**	0,16
SUB16	ALL	0,00	0,82	0,18	0,04	0,80	0,16	0,18	**1,00**	0,18
SUB17	ALL	0,00	0,82	0,18	0,04	0,80	0,16	0,18	**1,00**	0,18
SUB18	ALL	0,00	0,83	0,17	0,04	0,81	0,15	0,17	**1,00**	0,17
SUB19	ALL	0,00	0,82	0,18	0,04	0,80	0,16	0,18	**1,00**	0,18
SUB2	ALL	0,00	0,84	0,16	0,04	0,81	0,16	0,16	**1,00**	0,16
SUB20	ALL	0,00	0,83	0,16	0,04	0,81	0,15	0,17	**1,00**	0,16
SUB21	ALL	0,00	0,82	0,18	0,04	0,81	0,15	0,18	**1,00**	0,18
SUB22	ALL	0,00	0,83	0,17	0,03	0,82	0,15	0,17	**1,00**	0,17
SUB23	ALL	0,00	0,83	0,17	0,04	0,81	0,15	0,17	**1,00**	0,17
SUB24	ALL	0,00	0,84	0,16	0,03	0,82	0,15	0,16	**1,00**	0,16
SUB25	ALL	0,00	0,84	0,16	0,03	0,83	0,15	0,16	**1,00**	0,16
SUB3	ALL	0,00	0,82	0,18	0,06	0,78	0,16	0,18	**1,00**	0,18
SUB4	ALL	0,00	0,82	0,18	0,04	0,80	0,16	0,18	**1,00**	0,18
SUB5	ALL	0,00	0,81	0,19	0,06	0,78	0,16	0,19	**1,00**	0,19
SUB6	ALL	0,00	0,83	0,17	0,05	0,79	0,16	0,17	**1,00**	0,17
SUB7	ALL	0,00	0,83	0,16	0,04	0,80	0,16	0,17	**1,00**	0,16
SUB8	ALL	0,00	0,82	0,18	0,05	0,79	0,16	0,18	**1,00**	0,18
SUB9	ALL	0,00	0,85	0,15	0,04	0,81	0,15	0,15	**1,00**	0,15
SUBBAD1	ALL	0,16	0,84	0,00	0,36	0,64	0,00	0,16	**0,84**	0,00
SUBBAD2	ALL	0,11	0,89	0,00	0,33	0,67	0,00	0,11	**0,89**	0,00
SUBBAD3	ALL	0,14	0,86	0,00	0,35	0,65	0,00	0,14	**0,86**	0,00
SUBBAD4	ALL	0,13	0,87	0,00	0,35	0,65	0,00	0,13	**0,87**	0,00
SUBBAD5	ALL	0,20	0,80	0,00	0,39	0,61	0,00	0,20	**0,80**	0,00

Πίνακας 17 – Αποτελέσματα Σεναρίου 7

Γενικά στα αποτελέσματα φαίνεται ότι το μοντέλο μας δεν μπορεί να διακρίνει μεταξύ των καλών και κακών χρηστών αφού και οι δύο εμφανίζουν πολύ υψηλή ευλογοφάνεια.

– Η ευλογοφάνεια των καλών (για την υπόθεση Καλός) και των κακών (για την υπόθεση Κακός) ενισχύεται σημαντικά με τον συνδυασμό των εκτιμήσεων

– Αυτό βέβαια προϋποθέτει ότι το σύστημα μπορεί να ανεχτεί ένα χαμηλό ποσοστό ύποπτων ενεργειών πριν αυτές γίνουν βλαβερές για το σύστημα και ταυτόχρονα οι κακοί χρήστες το εκμεταλλεύονται χωρίς να χρειάζεται μεγάλη προσπάθεια. Δεν είναι πολύ ρεαλιστικό σενάριο καθώς το ποσοστό των κακών χρηστών μπορούν άνετα να το καλύψουν οι καλοί χρήστες συγκεντρωτικά. Για παράδειγμα κάθε κακός χρήστης αντιστοιχεί σε λιγότερο από δύο κακούς σε σχέση με τις συνολικές ύποπτες ενέργειες που κάνει.

Συμπεράσματα

Η έρευνα για το συγκεκριμένο μοντέλο σε συστήματα publish/subscribe έγινε γιατί υπάρχει η ανάγκη περιγραφής γενικών trust μοντέλων που ανταποκρίνονται στις απαιτήσεις των συστημάτων αυτών πολλά από τα οποία είναι σε ερευνητικό στάδιο. Έγινε προσπάθεια να προσεγγιστεί το θέμα από μια άλλη σκοπιά χωρίς να παραβλέπεται η υπάρχουσα γνώση και εμπειρία.

Με βάση τα αποτελέσματα μπορούμε να πούμε ότι η συμπεριφορά του μοντέλου κρίνεται ικανοποιητική. Αποτιμά σωστά την συμπεριφορά των χρηστών στις περισσότερες περιπτώσεις και εμφανίζει αποκλίσεις μόνο σε τραβηγμένα σενάρια. Επιπλέον λειτουργεί ικανοποιητικά τόσο σε κεντρικοποιημένο σύστημα όσο και σε κατανεμημένο σύστημα. Επιπλέον μπορούμε να ξεχωρίσουμε να ακόλουθα πλεονεκτήματα.

Για την υλοποίηση του μοντέλου δεν υπάρχουν υψηλές απαιτήσεις σε πόρους τους συστήματος. Η πολυπλοκότητα των υπολογισμών είναι χαμηλή, O(n) σε σχέση με τον αριθμό ενεργειών που επιλέγουμε ως ύποπτες και τα στοιχεία που χρειάζονται να αποθηκεύονται και να ανανεώνονται για κάθε συνδρομητή είναι ένα πίνακας 10 θέσεων. Πολλά συστήματα publish/subscribe κρατάνε ήδη κάποια κατάσταση ανά συνδρομητή άρα το να προσθέσουν μερικά στοιχεία ακόμα είναι σχετικά εύκολο. Επίσης οι πράξεις υλοποίησης είναι απλές.

Η αποτίμηση επίσης ακολουθεί μια απλή προσέγγιση και η χρήση της τιμής της ευλογοφάνειας είναι εύκολη αφού υπολογίζεται και ανανεώνεται για κάθε συνδρομητή όταν αυτός αλληλεπιδρά με το σύστημα.

Πολύ σημαντικό είναι επίσης ότι κάθε συνδρομητής χαρακτηρίζεται σαν κακός ή καλός με βάση μόνο την δική του συμπεριφορά ανεξάρτητα τι κάνουν οι άλλοι και τι συμβαίνει στο σύστημα κάθε στιγμή. Αυτό έχει επίσης σαν αποτέλεσμα να μπορούν οι κανόνες για το evidence που προκύπτει από τις ενέργειες να διαφοροποιούνται χωρίς να επηρεάζεται όλο το σύστημα. Για παράδειγμα θα μπορούσε να γίνει στάθμιση των βασικών πιθανοτήτων και της αβεβαιότητας. Στα σενάρια που περιγράψαμε αυτό είναι εφικτό αφού πριμοδοτούμε με την αβεβαιότητα τους καλούς χρήστες άρα οποιαδήποτε στάθμιση αυτής μπορεί να αμβλύνει ή να οξύνει το τελικό αποτέλεσμα υπέρ των καλών ή των κακών αντίστοιχα.

Επίσης το μοντέλο είναι σχετικά ανοιχτό στο τρόπο που λειτουργεί το mass function. Θα μπορούσαν να ληφθούν υπόψη επιπλέον στοιχεία για μεγαλύτερη ακρίβεια στις καταχώρηση των πιθανοτήτων, στοιχεία στατιστικής φύσεως ή άλλα (πχ. το γεγονός ότι οι κακοί χρήστες έχουν 0 αβεβαιότητα δεν έχει αξιοποιηθεί). Χρησιμοποιώντας στατιστικά στοιχεία του συστήματος είναι δυνατό να κάνει το μοντέλο πιο δυναμικό. Αυτό όμως θέλει προσοχή καθώς η συμπεριφορά κάθε χρήστη θα επηρεάζεται πλέον και από την συμπεριφορά των άλλων. Επιπλέον σαν επέκταση του μοντέλου θα μπορούσε οι υποθέσεις να είναι περισσότερες από δύο (αυξάνοντας κάπως την πολυπλοκότητα των υπολογισμών). Για παράδειγμα εκτός από τις υποθέσεις Bad, Good θα μπορούσε να υπάρχει και υπόθεση TrustedFor όπου θα χαρακτήριζε Ναι ή Όχι για συγκεκριμένο σετ ενεργειών ή ακόμα και για διαφορετικό σετ ενεργειών.

Εν αντιθέσει με τα θετικά στοιχεία που περιγράψαμε υπάρχουν και τα αρνητικά στοιχεία τα οποία αναλύονται παρακάτω.

Το πιο βασικό πρόβλημα της συγκεκριμένης πρότασης γίνεται αντιληπτό όταν το σύστημα μπορεί και θέλει να ανέχεται υψηλό ποσοστό ύποπτων ενεργειών. Σε αυτή την περίπτωση δεν μπορεί να διακρίνει με ασφάλεια τους κακούς συνδρομητές από τους καλούς. Αυτό συμβαίνει γιατί όσο υψηλότερα είναι τα ποσοστά των ύποπτων ενεργειών τόσο περισσότερη ενισχύεται η πεποίθηση ότι αυτός ο χρήστης είναι κακός. Αυτό αμβλύνεται κάπως όταν υπάρχει καταμερισμός στις ύποπτες ενέργειες αλλά όχι αρκετά στην περίπτωση που είναι υψηλά τα ποσοστά τους.

Παρόμοιο πρόβλημα δημιουργείται και όταν είναι πολύ μικρή η διαφορά των ύποπτων ενεργειών μεταξύ καλών και κακών χρηστών. Πάλι το μοντέλο μας δεν μπορεί να διακρίνει χωρίς λάθη τους κακούς και τους καλούς και αυτό συμβαίνει γιατί η συμπεριφορές είναι παρόμοιες και η αβεβαιότητα που συνήθως υπάρχει λόγω καταμερισμού των ύποπτων ενεργειών δεν ενισχύει τόσο σημαντικά τους καλούς.

Βέβαια και οι δύο περιπτώσει είναι δύσκολο να υπάρξουν σε πραγματικά συστήματα όπου πιστεύουμε ότι οι κακόβουλοι συνδρομητές θα παράγουν αρκετά υψηλό ποσοστό ύποπτων ενεργειών. Επιπλέον κάποιες ύποπτες ενέργειες θα μπορούσαν με κάποια στάθμιση να συνεισφέρουν περισσότερο στην περίπτωση του κακός ιδιαίτερα όταν δεν είναι απαραίτητος μεγάλος όγκος από αυτές τις ενέργειες για να προκαλέσουν πρόβλημα.

Σε σενάρια που δεν θεωρούνται εύκολα για το μοντέλο μας παρατηρείται πιο έντονα το φαινόμενο να έχουμε αποκλίσεις στην κατανεμημένη περίπτωση. Ο λόγος είναι ότι υπάρχει μεγάλη ενίσχυση μίας υπόθεσης όταν συνδυάζονται οι εκτιμήσεις. Άρα αν εμφανίζουν μια μικρή απόκλιση οι επιμέρους brokers, η απόκλιση αυτή θα είναι πιο έντονη όταν συνδυάζουν τις εκτιμήσεις τους με τον κανόνα Dempster-Shafer.

Παρόλο που είναι θετικό ότι η αποτίμηση της συμπεριφοράς κάθε συνδρομητή γίνεται αποκλειστικά από τις δικές του ενέργειες, το γεγονός ότι δεν λαμβάνεται καθόλου υπόψη το σύστημα μπορεί να δημιουργήσει προβλήματα. Είναι φανερό ότι υπάρχουν περιπτώσεις που λόγω πχ. μεγαλύτερης ισχύος, ένα σύστημα μπορεί παροδικά να επιτρέψει υψηλότερο ποσοστό ύποπτων ενεργειών ή λόγω φόρτου και πριν από κάποια αναβάθμιση ο μέσος φόρτος να έχει ανέβει σε ήδη μεγάλα επίπεδα. Σε αυτές τις περιπτώσεις το μοντέλο μας λειτουργεί περισσότερο ντετερμινιστικά και δεν επιτρέπει βελτίωση του συστήματος ενώ αυξάνει και τα λάθος αποτελέσματα. Αυτό μπορεί να διορθωθεί προσωρινά αν υπάρχει παρακολούθηση των συστημάτων από τους διαχειριστές και προσαρμογή των thresholds μέχρι το επίπεδο φυσικά που περιγράψαμε πιο πάνω.

Τέλος, είναι γενικά αποδεκτό ότι δεν έχουν όλες οι ύποπτες ενέργειες την ίδια επίπτωση στο σύστημα και όπως περιγράψαμε πιο πάνω δεν είναι απαραίτητος να υπάρχει για όλες ο ίδιος όγκος μηνυμάτων για να προκαλέσουν πρόβλημα. Έτσι το ότι δεν υπάρχει στάθμιση που να ορίζει την βαρύτητα κάθε ενέργειας οδηγεί σε μη βέλτιστη απόδοση του μοντέλου μας και ενδεχομένως σε κάποιες περιπτώσεις και προβληματική συμπεριφορά.

Γίνεται εύκολα αντιληπτό ότι το μοντέλο μας δουλεύει πολύ ικανοποιητικά όταν γίνονται ορθολογικές υποθέσεις για το σύστημα και όταν το σύστημα δεν έχει προβλήματα φόρτου. Αυτό δεν σημαίνει ότι δεν χρήζει περαιτέρω διερεύνησης. Υπάρχουν περιθώρια βελτίωσης και ανάλυσης που θα μπορούσαν να ελαττώσουν τα προβλήματα σε τραβηγμένες καταστάσεις. Συνοψίζοντας αυτά μπορούμε να πούμε ότι θα ήταν χρήσιμα τα παρακάτω

- Μελέτη για την δυναμική προσαρμογή των threshold με βάση στατιστική ανάλυση της ευλογοφάνειας ή της πίστης/πεποίθησης ή βασικής πιθανότητας. Με βάση αυτό τα threshold θα μπορούσαν να προσαρμόζονται δυναμικά σε διακριτές ή συνεχείς τιμές μεταξύ κάποιων ορίων ή περιορισμών ή συνθηκών. Αυτό είναι διαφορετικό από την γενική στατιστική ανάλυση που περιγράφεται πιο κάτω.

- Στάθμιση της σοβαρότητας κάθε ύποπτης ενέργειας η οποία μπορεί να είναι ανεξάρτητη ανά σύστημα ή ανά broker ή κοινή για όλους.
- Στατιστική ανάλυση της συμπεριφοράς του συστήματος και ανάλογη προσαρμογή των thresholds και των παραμέτρων. Αυτό οδηγεί σε απόλυτα δυναμικό και αυτο-προσαρμοζόμενο μοντέλο. Ιδιαίτερη προσοχή πρέπει να δοθεί και σε ακραίες περιπτώσεις όπου δεν υπάρχει κανείς κακόβουλος χρήστης στο σύστημα (είτε γιατί δεν έχει εισέλθει ακόμα είτε γιατί το ίδιο το σύστημα τον έχει αποκλείσει).
- Μια πιο δύσκολη προσέγγιση θα ήταν η μελέτη του μοντέλου όταν το σύστημα δεν μπορεί να αναγνωρίσει με 100% πιθανότητα τις ύποπτες ενέργειες. Η εισαγωγή της πιθανότητας αναγνώρισης θα είχε ενδιαφέρον να μελετηθεί επίσης.

Μελλοντικά θα μπορούσε να γίνει μελέτη για την συμπεριφορά του συστήματος όταν οι κακόβουλοι συνδρομητές έχουν την ίδια (κακή) συμπεριφορά στην πλειοψηφία των brokers που είναι συνδεδεμένοι (πχ ίδια συμπεριφορά στους 2 από τους 3 brokers). Το αναμενόμενο αποτέλεσμα με βάση τη θεωρία είναι ότι αν υπάρχει πλειοψηφία σε μία άποψη αυτή η άποψη θα υπερισχύσει.

Συνοψίζοντας θα μπορούσαμε να πούμε ότι δημιουργήσαμε ένα μοντέλο που λειτουργεί πολύ ικανοποιητικά κάτω από κανονικές συνθήκες, είναι εύκολο στην υλοποίηση του και υπάρχει δυνατότητα επέκτασης του.

Βιβλιογραφία και πηγές

Zadeh, L A. (1984) Review of Shafer's a mathematical theory of evidence AI Magazine 5(3): 81-83

S.D., Kamvar, M.T. Schlosser and H. Garcia-Molina, "The EigenTrust Algorithm for Reputation Management in P2P Networks", ACM Press, Proc. 12th International Conference on World Wide Web, pp. 640–651, 2003.

Alex Wun, Alex Cheung, Hans-Arno Jacobsen, A Taxonomy for Denial of Service Attacks in Content-based Publish/Subscribe Systems (2007) ACM, Vol. 233, pp 116 - 127

Wun, Alex , Denial of service in content-based publish/scribe systems. Masters Abstracts International. Vol. 46, no. 6. 2007

J. Mirkovic and P. Reiher. A Taxonomy of DDoS Attack and DDoS Defense Mechanisms. ACM SIGCOMM, Computer Communications Review, 34 (2):39{54, 2004.

G. Banavar, T. Chandra, B. Mukherjee, J. Nagarajaro, R. Strom and D. Sturman, "An Efficient Multicast Protocol for Content-Based Publish-Subscribe Systems," IEEE Computer, Proc. of the 19th International Conference on Distributed Computing Systems, vol 19, pp 262-272, 1999.

C. Wang, A Carzaniga, D. Evans and A.L. Wolf, "Security Issues and Requirements for Internet-Scale Publish-Subscribe Systems," Proc. of the HICSS'02, pp. 3940-3947,2002.

L. Fiege, A. Zeidler, A. Buchmann, R. Kilian-Kehr and G. Muh, "Security Aspects in Publish/Subscribe Systems," Proc. Intl. Workshop on Distributed Event-based Systems (DEBS), 2004.

G. Danezis, C. Lesniewski-Laas, M.F. Kaashoek and R.J. Anderson, "Sybil-Resistant DHT Routing," Springer-Verlag, Lecture Notes in Computer Science, no. 3679, pp. 305-318, 2005

Reasoning with Uncertain and Conflicting Opinions in Open Reputation Systems Andreas Gutscher1, Electronic Notes in Theoretical Computer Science 244 (2009) 67–79, www.elsevier.com/locate/entcs

Preventing Spam in Publish/Subscribe Sasu Tarkoma, Proceedings of the 26th IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW'06)

State-of-the-Art Report and Technical Requirements, (D2.1), 2008, http://psirp.org/publications

Conceptual architecture of PSIRP including subcomponent descriptions (D2.2), June 2008. http://psirp.org/publications.

Architecture Definition, Component Descriptions,and Requirements (D2.3), 2008, http://psirp.org/publications

Zie Zhang,Robin Cohen, Evaluating the trustworthiness of advice about seller agents in e-marketplaces: A personilized approach, Electronic Commerce Research and Applications 7 (2008) 330-340

Christian Thiel, Friedhelm Schwenker, Gunther Palm , Using Dempster-Shafer Theory in MCF Systems to Reject Samples, MCS 2005, LNCS 3541, pp 118-127, 2005, Springer-Verlag Berlin Heidelberg 2005

Daniel Cvrcek, Vaclav Matyas Jr., Ahmed Patel, Evidence processing and privacy issues in evidence-based reputation systems, 2005 Elsevier B.V. doi:10.1016/j.csi.2005.01.011

Himanshu Khurana, Scalable Security and Accounting Services for Content-based Publish/Subscribe Systems, 2005 ACM Symposium on Applied Computing, pp: 801 - 807, 2005

Scott Ferson, Vladik Kreinovich, Lev Ginzburg, Davis S. Myers, and Kari Sentz, Constructing Probability Boxes and Dempster-Shafer Structures, SAND REPORT, SAND2002-4015, 2003

Qi Chen and Uwe Aickelin, Anomaly Detection Using the Dempster-Shafer Method Neural and Evolutionary Computing (cs.NE); Artificial Intelligence (cs.AI); Cryptography and Security (cs.CR), Proceedings of the International Conference on Data Mining (DMIN 2006), pp 232-238, Las Vegas, USA 2006

Thomas M. Chen, Varadharajan Venkataramanan, Dempster-Shafer Theory for Intrusion Detection in Ad Hoc Networks, IEEE Internet Computing, Volume 9 , Issue 6 (November 2005), Pages: 35 – 41, 2005

A.K. Datta, M. Gradinariu, M. Raynal, G. Simon, Anonymous Publish/Subscribe in P2P Networks, 2003, Proceedings of IPDPS, Conference, pp 22--26

Van Jacobson, Diana K.Smetters, James D. Thornton, Michael F. Plass, Networking Named Content CoNEXT'09, December1-4, 2009, pp: 1-12

John R. Douceur, The Sybil Attack, Lecture Notes In Computer Science; Vol. 2429, pp : 251 – 260, 2002

Guido Urdaneta, A Survey on DHT Security Techniques, Accepted for publication, ACM Computing Surveys, 2009

Jean Gordon and Edward H. Shortliffe, Classic Works of the Dempster-Shafer Theory of Belief Functions, Chapter 13 - The Dempster-Shafer Theory of Evidence, volume : 219, 2008

Nic Wilson, Algorithms for Dempster-Shafer Theory, School of Computing and Mathematical Sciences, Conference, 2000, pp.421-475

L. Opyrchal and A. Prakash. Secure distribution of events in content-based publish subscribe systems. In 10th USENIX Security Symposium, Aug. 2001.

C. Wang, A. Carzaniga, D. Evans, and A. L. Wolf. Security issues and requirements for Internet-scale publish-subscribe systems. In Proc. of the HICSS-35, Big Island, Hawaii, Jan. 2002.

N. Fotiou, G.C. Polyzos, D. Trossen, "Illustrating a Publish-Subscribe Internet Architecture," Future Internet Architectures: New Trends in Service Architectures (2nd Euro-NF Workshop), June 2009

Dempster–Shafer theory, From Wikipedia, the free encyclopedia

PSIRP Project Website http://www.psirp.org, 2008.