



Πανεπιστήμιο Θεσσαλίας

Τμήμα Μηχανικών Η/Υ, Τηλεπικοινωνιών και Δικτύων

Πρόγραμμα Μεταπτυχιακών Σπουδών

---

## **Αναπαράσταση και Διαχείριση Εμπιστοσύνης σε Ανοιχτά Πολυπρακτορικά Συστήματα**

*Μια πλατφόρμα για τη μελέτη και την αξιολόγηση  
μοντέλων εμπιστοσύνης*

---

*Μεταπτυχιακή Διατριβή*

**Μπαλταγιάννης Μιχάλης**

**Ιούνιος 2006**



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ  
ΒΙΒΛΙΟΘΗΚΗ & ΚΕΝΤΡΟ ΠΛΗΡΟΦΟΡΗΣΗΣ  
ΕΙΔΙΚΗ ΣΥΛΛΟΓΗ «ΓΚΡΙΖΑ ΒΙΒΛΙΟΓΡΑΦΙΑ»**

Αριθ. Εισ.: 5049/1  
Ημερ. Εισ.: 24-09-2007  
Δωρεά: Συγγραφέα  
Ταξιθετικός Κωδικός: Δ  
005.43  
ΜΠΑ

*στη μνήμη του Ευθύμη*

## Ευχαριστίες

Η μεταπτυχιακή αυτή διατριβή εκπονήθηκε στο Τμήμα Μηχανικών Η/Υ Τηλεπικοινωνιών και Δικτύων του Πανεπιστημίου Θεσσαλίας κατά το χρονικό διάστημα από το Μάρτιο του 2005 έως και τον Ιούλιο του 2006. Για την ολοκλήρωσή της δαπανήθηκαν πάρα πολλές ώρες ερευνητικής δουλειάς που όμως από μόνες τους δεν θα επαρκούσαν εάν δεν υποστηριζόταν και από τη μικρή ή μεγάλη βοήθεια κάποιων ανθρώπων.

Η επιβλέπουσα της διατριβής, Δρ Ασπασία Δασκαλοπούλου, αποτέλεσε μια ανεξάντλητη πηγή επιστημονικής καθοδήγησης, έμπνευσης και διάθεσης για δημιουργική δουλειά. Καθ' όλη τη διάρκεια αυτής της προσπάθειας ήταν δίπλα, προτείνοντας λύσεις και δίνοντας κατευθύνσεις για την αντιμετώπιση των δύσκολων ζητημάτων που ανέκυπταν. Περισσότερο απ' όλα όμως υπήρξε μια πολύτιμη φίλη, που όταν οι δυσκολίες έδειχναν να είναι ανυπέρβλητες έβρισκε τον τρόπο να μεταδίδει ηρεμία και αισιοδοξία.

Ο Δρ Σπύρος Λάλης, μέλος της τριμελούς επιτροπής αξιολόγησης, βοήθησε στην βελτίωση της αρτιότητας της διατριβής με τις εύστοχες παρατηρήσεις και τις πολύ ενδιαφέρουσες προτάσεις του. Επίσης, αρκετές από τις κατευθύνσεις μελλοντικής δουλειάς βασίζονται σε δικές του ιδέες.

Ο Δρ Χρήστος Ηλιούδης συνεισέφερε ως δάσκαλος αλλά και ως φίλος. Η εμπειρία του και οι συμβουλές του βοήθησαν στην αποσαφήνιση αρκετών συγκεχυμένων σημείων και συνέβαλαν στον γρηγορότερο και ακριβέστερο εντοπισμό της ουσίας των ζητημάτων με τα οποία ασχολήθηκε η διατριβή.

Ο Δημοσθένης Μπέκας, τελειόφοιτος του Τμήματος Μηχανικών Η/Υ Τηλεπικοινωνιών και Δικτύων και φίλος, βοήθησε στο κομμάτι της υλοποίησης, διαθέτοντας τον πολύτιμο χρόνο του για να προσφέρει υλικό και γνώσεις γύρω από το περιβάλλον JADE.

Αφανείς υποστηρικτές αυτής της προσπάθειας ήταν ακόμη ο μικρός Νικόλας με τη νεογέννητη αδελφούλα του. Το χαμόγελό τους και η μικρή αγκαλιά τους αποτέλεσαν μια όαση ξεκούρασης και ψυχολογικής τόνωσης, ιδιαίτερα όταν η κόπωση έκανε την εναπομένουσα δουλειά να φαντάζει ακατόρθωτη. Η συνεισφορά τους αποκτά ιδιαίτερη σπουδαιότητα αν συνυπολογίσει κανείς και όλες εκείνες τις ώρες που αδιαμαρτύρητα στερήθηκαν τον πατέρα τους.

Τέλος, χωρίς αμφιβολία, τη σημαντικότερη βοήθεια την προσέφερε η Μαρία. Σύζυγος, φίλη αλλά και συνεργάτης, έκανε εφικτή την ολοκλήρωση αυτής της δύσκολης προσπάθειας. Με τη φροντίδα, τη κατανόηση και τη συμπαράστασή της λειτούργησε ως στυλοβάτης και κινητήρια δύναμη μαζί. Η προσφορά της είναι τόσο μεγάλη και σημαντική που για την περιγραφή της θα δικαιούταν ίσως περισσότερες σελίδες και απ' αυτές της διατριβής.

Σε όλους τους παραπάνω, σημαντικούς για μένα ανθρώπους, αλλά και σε όσους άλλους συνέβαλαν με τον ένα ή τον άλλο τρόπο στην ολοκλήρωση αυτού του έργου και τους οποίους μπορεί να ξεχνώ αυτή τη στιγμή, θα ήθελα να εκφράσω, μέσα από την καρδιά μου, ένα μεγάλο ευχαριστώ.

Μιχάλης Μπαλταγιάννης

# Περιεχόμενα

<b>1. Εισαγωγή.....</b>	<b>1</b>
1.1 Εμπιστοσύνη : Ο θεμέλιος λίθος .....	1
1.2 Διαχείριση και μοντέλα εμπιστοσύνης.....	3
1.3 Εφαρμογές των μοντέλων εμπιστοσύνης.....	7
1.4 Αντικείμενο της διατριβής.....	9
1.5 Συναφείς εργασίες .....	9
1.6 Παραδοχές - Συμβάσεις.....	10
1.7 Διάρθρωση της διατριβής.....	11
<b>2. Μοντέλα Εμπιστοσύνης.....</b>	<b>13</b>
<b>2.1 Η περίπτωση eBay.....</b>	<b>13</b>
2.1.1 Τρόπος λειτουργίας .....	13
2.1.2 Το σύστημα ανατροφοδότησης .....	14
2.1.3 Σχολιασμός - κριτική .....	15
<b>2.2 Το μοντέλο του Marsh.....</b>	<b>16</b>
2.2.1 Παραδοχές- Χαρακτηριστικά .....	16
2.2.2 Αναπαράσταση της εμπιστοσύνης.....	17
2.2.3 Υπολογισμός της εμπιστοσύνης .....	19
2.2.4 Το κατώφλι συνεργασίας.....	21
2.2.5 Δομές .....	24
2.2.6 Σχολιασμός - κριτική.....	24
<b>2.3 Το μοντέλο PeerTrust.....</b>	<b>25</b>
2.3.1 Παραδοχές- Χαρακτηριστικά .....	25
2.3.2 Αναπαράσταση της εμπιστοσύνης.....	27
2.3.3 Υπολογισμός της εμπιστοσύνης .....	27
2.3.4 Δομές .....	30
2.3.5 Σχολιασμός - κριτική .....	31
<b>2.4 Το μοντέλο των Mui - Mohtashemi - Halberstadt.....</b>	<b>32</b>
2.4.1 Παραδοχές- Χαρακτηριστικά .....	32
2.4.2 Αναπαράσταση της εμπιστοσύνης.....	34
2.4.3 Αξιολόγηση της εμπιστοσύνης.....	35
2.4.4 Μηχανισμός διάδοσης της φήμης.....	38
2.4.5 Δομές .....	40
2.4.6 Σχολιασμός - κριτική .....	40
<b>2.5 Το μοντέλο των Abdul-Rahman &amp; Hailes.....</b>	<b>40</b>
2.5.1 Παραδοχές- Χαρακτηριστικά .....	41
2.5.2 Αναπαράσταση εμπιστοσύνης.....	42
2.5.3 Δομές Δεδομένων .....	43
2.5.4 Υπολογισμοί .....	44
2.5.5 Σχολιασμός - κριτική.....	47
<b>2.6 Το μοντέλο των F. Azzerdin and M. Maheswaran .....</b>	<b>48</b>
2.6.1 Παραδοχές- Χαρακτηριστικά .....	48
2.6.2 Αναπαράσταση εμπιστοσύνης.....	50
2.6.3 Δομές .....	51
2.6.4 Υπολογισμός της εμπιστοσύνης .....	52
2.6.5 Ενημέρωση των δομών.....	53
2.6.6 Νέα μέλη & κληρονομικότητα .....	54
2.6.7 Σχολιασμός - κριτική.....	54
<b>2.7 Το μοντέλο των Bin Yu &amp; Munindar P. Singh .....</b>	<b>55</b>
2.7.1 Παραδοχές - Υπόβαθρο .....	56

2.7.2	Αναπαράσταση της εμπιστοσύνης.....	57
2.7.3	Υπολογισμοί.....	57
2.7.4	Απεικόνιση των ιδιοτήτων της εμπιστοσύνης.....	61
2.7.5	Σχολιασμός - κριτική.....	62
<b>3.</b>	<b>Τα Χαρακτηριστικά των Μοντέλων Εμπιστοσύνης.....</b>	<b>65</b>
3.1	Βασικά χαρακτηριστικά των μοντέλων εμπιστοσύνης.....	65
3.2	Η λειτουργικότητα των πρακτόρων εμπιστοσύνης.....	69
3.2.1	Ένα υποθετικό σενάριο από τον πραγματικό κόσμο.....	69
3.2.2	Οι βασικές συνιστώσες κοινωνικής συμπεριφοράς.....	71
3.2.3	Οι βασικοί μηχανισμοί συμπεριφοράς ενός πράκτορα εμπιστοσύνης.....	73
<b>4.</b>	<b>Ο Προσομοιωτής.....</b>	<b>79</b>
4.1	Βασικές Απαιτήσεις.....	79
4.2	Σχεδιαστικές επιλογές.....	80
4.2.1	Η πλατφόρμα.....	80
4.2.2	Το JADE.....	80
4.2.3	Τα κυριότερα χαρακτηριστικά του JADE.....	82
4.2.4	Η γλώσσα ανάπτυξης.....	84
4.2.5	Η λειτουργικότητα των πρακτόρων στο JADE.....	85
4.3	Η Αρχιτεκτονική του προσομοιωτή.....	86
4.4	Οι βασικές λειτουργίες του TrustSIM.....	91
4.5	Το βασικό μοντέλο του TrustSIM.....	93
4.5.1	Συμβάσεις.....	94
4.5.2	Οι συμπεριφορές του βασικού μοντέλου.....	95
4.5.3	Σύνοψη.....	114
4.6	Οι κυριότερες δομές δεδομένων.....	115
<b>5.</b>	<b>Μετρικές Αξιολόγησης Μοντέλων Εμπιστοσύνης.....</b>	<b>121</b>
5.1	Trusting index.....	121
5.2	Trusted index.....	122
5.3	Trustworthiness evaluation accuracy.....	123
5.4	Malicious activity index.....	124
5.5	Malicious activity success.....	125
5.6	Malicious activity efficiency.....	125
5.7	Interaction prevention index.....	126
5.8	Interaction prevention efficiency.....	127
<b>6.</b>	<b>Αξιολόγηση Μοντέλων Εμπιστοσύνης με τον Προσομοιωτή.....</b>	<b>129</b>
6.1	Τα σενάρια των προσομοιώσεων.....	130
6.2	Προσομοιώσεις του Βασικού Μοντέλου.....	134
6.3	Προσομοιώσεις του μοντέλου PeerTrust.....	141
6.4	Συμπεράσματα.....	148
<b>7.</b>	<b>Συμπεράσματα.....</b>	<b>151</b>
7.1	Συνεισφορά.....	151
7.2	Δυσκολίες.....	152
7.3	Μελλοντική δουλειά.....	153

7.4 Γενικά συμπεράσματα.....	155
<i>Βιβλιογραφία</i> .....	157
<i>Παράρτημα Α : Οι παράμετροι προσομοίωσης του TrustSIM</i> .....	161
<i>Παράρτημα Β : Τα αρχεία ρύθμισης μοντέλου (model configuration files)</i> .....	167
<i>Παράρτημα Γ : Τα μηνύματα επικοινωνίας των πρακτόρων του TrustSIM</i> .....	169

# ΚΕΦΑΛΑΙΟ

## 1. Εισαγωγή

### 1.1 Εμπιστοσύνη : Ο θεμέλιος λίθος

*Μια μέρα, καθώς βγαίνετε από το super market έχοντας ολοκληρώσει τα καθημερινά σας ψώνια, σας σταματά ένας νεαρός ζητώντας να του αφιερώσετε λίγα λεπτά από το χρόνο σας. Σας προσφέρει μια "καταπληκτική" ασφάλεια αυτοκινήτου, με χαμηλά ετήσια ασφαλιστρα, εξαιρετικές καλύψεις και πολλά bonuses. Δυστυχώς για εκείνον, είστε ένας σχετικά επιφυλακτικός άνθρωπος και δεν εμπιστεύεστε εύκολα αγνώστους. Έτσι, απορρίπτετε την προσφορά του, αν και αναγνωρίζετε ότι είναι αρκετά δελεαστική. Μετά από μερικές μέρες, επισκέπτεστε τον καλύτερό σας φίλο, στη γιορτή του. Ανάμεσα στους παρευρισκόμενους αναγνωρίζετε τον πωλητή του super market. Από το φίλο σας μαθαίνετε ότι πρόκειται για ένα πολύ καλό παιδί που τα τελευταία χρόνια εργάζεται ως ασφαλιστής σε μια γνωστή ασφαλιστική εταιρία. Μάλιστα, ο φίλος σας έχει ασφαλίσει και τα δύο του αυτοκίνητα στο συγκεκριμένο ασφαλιστή και είναι ιδιαίτερα ευχαριστημένος. Μετά τις πολύ καλές συστάσεις και επειδή η δική σας ασφάλεια λήγει τον επόμενο μήνα, αποφασίζεται να ασφαλίσετε και εσείς το αυτοκίνητό σας στον νεαρό ασφαλιστή. Κάποια στιγμή, μερικούς μήνες αφότου αλλάξατε την ασφάλεια του αυτοκινήτου σας, ο νεαρός ασφαλιστής επικοινωνεί μαζί σας και σας ενημερώνει ότι η εταιρία του διαθέτει και επενδυτικά προγράμματα. Σας προτείνει λοιπόν να επενδύσετε, μέσω του ιδίου και της εταιρίας του, σε συγκεκριμένο χαρτοφυλάκιο μετοχών, με δυναμική (όπως σας διαβεβαιώνει) πορεία, υποσχόμενος ετήσιο διπλασιασμό της επένδυσής. Όμως, είχατε στο παρελθόν μια άσχημη εμπειρία στο χρηματιστήριο, κατά την οποία χάσατε ένα σημαντικό χρηματικό ποσό, κι έτσι αποφασίζετε να απορρίψετε την πρόταση, θεωρώντας ότι εμπεριέχει πολύ μεγάλο ρίσκο.*

Κεντρικό στοιχείο σε όλες τις φάσεις της παραπάνω ιστορίας είναι η διακύμανση της εμπιστοσύνης που έχετε προς το πρόσωπο του νεαρού ασφαλιστή. Ο υφιστάμενος βαθμός εμπιστοσύνης είναι ο κυρίαρχος παράγοντας που καθοδηγεί τις επιμέρους επιλογές και αποφάσεις σας. Παρόμοιες καταστάσεις, στις οποίες καλούμαστε να πάρουμε αποφάσεις με κριτήριο την εμπιστοσύνη που έχουμε διαμορφώσει, συνθέτουν το ψηφιδωτό της καθημερινότητας του κάθε ανθρώπου. Από το κρεοπωλείο στο οποίο ψωνίζουμε το κρέας για την καθημερινή μας διατροφή, μέχρι το γιατρό στον οποίο εμπιστευόμαστε την υγεία μας, οι επιλογές μας εμπεριέχουν κάποιου είδους αξιολόγηση εμπιστοσύνης. Ακόμη και το απλό γεγονός ότι διασχίζουμε μια διασταύρωση όταν έχει ανάψει πράσινο, είναι μια αυτοματοποιημένη και σχεδόν υποσυνείδητη επιλογή που βασίζεται στην εμπιστοσύνη που έχουμε προς τους άλλους ότι δεν θα παραβιάσουν το κόκκινο.

Η σπουδαιότητα της εμπιστοσύνης στις ανθρώπινες κοινωνίες έχει αναλυθεί από ένα πλήθος κοινωνιολογικών και άλλων ερευνών. Η διαπίστωση ότι οπουδήποτε υπάρχει κοινωνία υπάρχει και η έννοια της εμπιστοσύνης [1], καθώς και ότι η απουσία της αποτελεί ένα σημαντικό ανασταλτικό παράγοντα στην ανάπτυξη μιας κοινωνίας [5], οδήγησε τα τελευταία χρόνια όλο και περισσότερους



ερευνητές να στραφούν προς την εμπιστοσύνη και να προσπαθήσουν να ενσωματώσουν τα χαρακτηριστικά της και στις ψηφιακές κοινωνίες.

Η εμπιστοσύνη είναι μια κατευθυνόμενη σχέση ανάμεσα σε δύο μέρη, στην οποία το ένα (trustor) εμπιστεύεται το άλλο (trustee). Στη βιβλιογραφία έχει εμφανιστεί μια πολύ μεγάλη ποικιλία ορισμών για την εμπιστοσύνη, ο καθένας από τους οποίους φαίνεται να εστιάζει σε διαφορετικές πτυχές της σημασίας της [3]. Ο ορισμός που φαίνεται να κυριαρχεί προέρχεται από τον Gambetta [2] :

*Εμπιστοσύνη είναι η υποκειμενική πιθανότητα με την οποία ένα άτομο A αναμένει ότι ένα άλλο άτομο B θα εκτελέσει μια ενέργεια από την οποία εξαρτάται η ευημερία του A .*

Ο ορισμός του Gambetta, αν και γενικά αποδεκτός, έχει δεχτεί κριτική [5, 7, 8]. Ο λόγος είναι ότι, δεν αναφέρεται ξεκάθαρα στην διάθεση του εμπιστευόντος να εξαρτηθεί από την ικανότητα του εμπιστευόμενου, αλλά και στη σχέση της διάθεσης αυτής με τις συγκεκριμένες συνθήκες (περίσταση) κάτω από τις οποίες λαμβάνεται. Ένας πιο ολοκληρωμένος ορισμός δόθηκε από τους T.Grandison και M.Sloman [7]. Σύμφωνα με τον ορισμό αυτό εμπιστοσύνη είναι <sup>1</sup> :

*"η ισχυρή πεποίθηση στην ικανότητα μιας οντότητας ότι μπορεί να ενεργήσει αξιόπιστα, με ασφάλεια και με ένα τρόπο που να μπορείς να εξαρτηθείς απ' αυτήν, σε μια συγκεκριμένη περίπτωση"*

Και στους δύο παραπάνω ορισμούς (αλλά και σε όσους άλλους εμφανίζονται στη βιβλιογραφία) κυρίαρχο χαρακτηριστικό είναι ότι η εμπιστοσύνη εμφανίζεται ως μια πεποίθηση και υπό την έννοια αυτή μπορεί να χαρακτηρίζει μόνο σκεπτόμενες οντότητες. Τα πολυπρακτορικά συστήματα, (δηλ., η κατανεμημένη τεχνητή νοημοσύνη) και το ηλεκτρονικό εμπόριο αποτελούν μερικές από τις πιο προφανείς περιοχές εφαρμογής των αρχών και των χαρακτηριστικών της εμπιστοσύνης.

Γενικά, η έννοια της εμπιστοσύνης έχει νόημα σε περιβάλλοντα όπου υπάρχει αβεβαιότητα σχετικά με το ενδεχόμενο κακόβουλης ή ανεπιθύμητης συμπεριφοράς. Ιδιαίτερα σε εφαρμογές με online συναλλαγές, υπάρχει πάντα ο εγγενής κίνδυνος της εξαπάτησης, καθώς δεν είναι δυνατό τα συναλλασσόμενα μέρη να γνωρίζουν με βεβαιότητα τις προσθέσεις και την αξιοπιστία της άλλης πλευράς. Η εμπιστοσύνη μπορεί να παίζει σημαντικό ρόλο στη διαχείριση του ρίσκου που εμπεριέχουν τα online περιβάλλοντα, βοηθώντας με τον τρόπο αυτό στην βελτίωση της ασφάλειας που προσφέρουν και κατ' επέκταση στην ενίσχυση της αποδοχής τους από τους πολίτες.

Οι πολλές και σημαντικές, για την ανθρώπινη δραστηριότητα, εφαρμογές της εμπιστοσύνης οδήγησαν τους ερευνητές στο να προτείνουν τα τελευταία χρόνια μια σειρά μοντέλων και τεχνικών για τη διαχείρισή της. Η εργασία αυτή ασχολείται με την έννοια της εμπιστοσύνης στον ψηφιακό κόσμο, εστιάζοντας κυρίως στον τομέα της αξιολόγησης των μοντέλων που αναπτύχθηκαν για τη διαχείρισή

<sup>1</sup> Ελεύθερη μετάφραση από το αγγλικό κείμενο.

της. Στόχος της είναι η δημιουργία μιας γενικής πλατφόρμας για τη μελέτη διαφόρων μοντέλων εμπιστοσύνης και η ανάπτυξη ενός συνόλου μετρικών για την αξιολόγησή τους.

## 1.2 Διαχείριση και μοντέλα εμπιστοσύνης

Οι τεράστιες δυνατότητες επικοινωνίας που δημιούργησε η ραγδαία ανάπτυξη των δικτύων και των υπολογιστών έχουν αλλάξει ριζικά τη φυσιογνωμία των ανθρώπινων αλληλεπιδράσεων. Συναλλαγές και συνεργασίες μπορούν να λάβουν χώρα ηλεκτρονικά, χωρίς να απαιτείται η φυσική παρουσία των συμμετεχόντων. Σ' αυτό το ταχύτατα μεταβαλλόμενο περιβάλλον είναι εξαιρετικά σημαντικό να μπορούν να δημιουργηθούν on-line συστήματα που να εμπνέουν στους χρήστες τους εμπιστοσύνη. Έτσι, προέκυψε η ανάγκη για την ανάπτυξη τεχνικών που να επιτρέπουν στα μέλη μιας κοινωνίας, αφενός να εκτιμούν την αξιοπιστία των άλλων μελών και, αφετέρου να διαμορφώνουν τη δική τους εικόνα αξιοπιστίας που αντιλαμβάνονται τα άλλα μέλη. Προέκυψε η ανάγκη για τη δημιουργία συστημάτων διαχείρισης εμπιστοσύνης.

Οι Blaze, Feigenbaum και Lacy [9], το 1996, εισήγαγαν για πρώτη φορά την έννοια της διαχείρισης εμπιστοσύνης ορίζοντάς την ως

*"μια ενοποιημένη προσέγγιση για τον προσδιορισμό και την ερμηνεία πολιτικών ασφάλειας, διαπιστευτηρίων και σχέσεων που επιτρέπουν την απ' ευθείας εξουσιοδότηση κρίσιμων ως προς την ασφάλεια ενεργειών".*

Σ' ένα πιο τυπικό ορισμό οι Josang, Keser και Dimitrakos [5] θεωρούν ότι διαχείριση εμπιστοσύνης είναι :

*"Η δραστηριότητα της δημιουργίας συστημάτων και μεθόδων που επιτρέπουν στα εμπλεκόμενα μέρη να κάνουν εκτιμήσεις και να παίρνουν αποφάσεις σχετικά με την εξάρτησή τους από εν δυνάμει συναλλαγές που εμπεριέχουν ρίσκο, και που επιτρέπουν στους παίκτες και τους ιδιοκτήτες συστημάτων να αναπαριστούν σωστά και να αυξάνουν την αξιοπιστία των ιδίων και των συστημάτων τους."*

Με απλά λόγια, διαχείριση εμπιστοσύνης είναι ένα σύνολο μηχανισμών και διαδικασιών που επιτρέπουν την αναπαράσταση, τον υπολογισμό, την ανταλλαγή και την εξέλιξη πληροφοριών που αφορούν την εμπιστοσύνη των μελών μιας κοινωνίας.

Ένα σύστημα διαχείρισης εμπιστοσύνης για να είναι εφαρμόσιμο θα πρέπει να έχει τουλάχιστο τα παρακάτω χαρακτηριστικά :

- Να αναπαριστά την εμπιστοσύνη σε μια μορφή που να επιτρέπει την εκτέλεση υπολογισμών
- Να διαθέτει κάποιο μηχανισμό αναγνώρισης ταυτότητας μέσω του οποίου να μπορεί να αντιστοιχεί μια αξιολόγηση εμπιστοσύνης σε μια αναγνωρίσιμη οντότητα.
- Να διαθέτει μηχανισμούς υπολογισμού και δυναμικής ενημέρωσης της εμπιστοσύνης

- Να επιτρέπει τη διάδοση της εμπιστοσύνης μεταξύ των μελών μιας κοινωνίας. Ο μηχανισμός αυτός είναι γνωστός και ως φήμη.

Η χρησιμότητα της διαχείρισης εμπιστοσύνης στις εφαρμογές του ψηφιακού κόσμου μπορεί να συνοψισθεί στους εξής παράγοντες :

- **Αυξάνει την ασφάλεια των συστημάτων**

Η διαχείριση εμπιστοσύνης προσφέρει στα συστήματα ασφάλειας μια επιπλέον διάσταση, της παρακολούθησης και αξιολόγησης της συμπεριφοράς των χρηστών. Οι παραδοσιακές τεχνικές ασφάλειας βασίζονται στην αυθεντικοποίηση (authentication) των χρηστών και στην εξουσιοδότηση τους (authorization) με τα επιτρεπόμενα δικαιώματα πρόσβασης. Η διαχείριση εμπιστοσύνης προχωρά ένα βήμα πιο πέρα και επιτρέπει τη δυναμική προσαρμογή των δικαιωμάτων πρόσβασης ανάλογα με την συμπεριφορά του χρήστη. Η διαφορά των δύο προσεγγίσεων επισημάνθηκε για πρώτη φορά από τους Rasmussen και Jasson [10], οι οποίοι χρησιμοποίησαν τους όρους *hard security* για την παραδοσιακή ασφάλεια και *soft security* για την ασφάλεια μέσω διαχείρισης εμπιστοσύνης.

- **Βοηθά στη διαχείριση του ρίσκου των συναλλαγών**

Η εμπιστοσύνη στον ψηφιακό κόσμο βασίζεται στη γνώση<sup>2</sup> [11]. Υπό την έννοια αυτή μπορεί να ειπωθεί και ως ένας τρόπος αναπαράστασης της διαθέσιμης γνώσης σχετικά με την αξιοπιστία μιας οντότητας. Μέσα από τη διαχείριση εμπιστοσύνης δημιουργείται καλύτερη αίσθηση της διαθέσιμης γνώσης με άμεσο αποτέλεσμα την αποτελεσματικότερη διαχείριση του ρίσκου των online συναλλαγών. Για παράδειγμα η απόφαση για την συμμετοχή ή όχι σε μια επικίνδυνη συναλλαγή μπορεί να αναχθεί στη σύγκριση της εμπιστοσύνης της άλλης πλευράς με ένα προκαθορισμένο κατώφλι που σχετίζεται με την επικινδυνότητα της συναλλαγής. Επιπλέον, μέσω της διάδοσης της εμπιστοσύνης μπορεί να αναπληρωθεί η έλλειψη άμεσης γνώσης και να επιτευχθεί ακριβέστερη εκτίμηση της εμπιστοσύνης.

- **Αποθαρρύνει τις κακόβουλες ή ανεπιθύμητες συμπεριφορές**

Τα μέλη μιας ψηφιακής κοινωνίας, όταν γνωρίζουν ότι η αξιοπιστίας τους αξιολογείται διαρκώς, με βάση τη συμπεριφορά τους, επιδιώκουν να συμπεριφέρονται με τον προβλεπόμενο από την κοινωνία τρόπο ώστε να διατηρούν ή και να βελτιώνουν την αξιοπιστία τους. Επιπλέον η δυνατότητα διάδοσης της εμπιστοσύνης σε όλη την κοινωνία λειτουργεί ως αντικίνητρο για την εκδήλωση παραβατικών συμπεριφορών, αφού μέσω αυτών των συμπεριφορών οι παραβάτες μπορούν να οδηγηθούν στην κοινωνική απομόνωση.

- **Ενθαρρύνει τις υγιείς συμπεριφορές**

<sup>2</sup> Στον πραγματικό κόσμο η εμπιστοσύνη πέραν της γνώσης μπορεί να βασίζεται και στην πίστη. Μάλιστα κάποιες φορές μπορεί να έρχεται και σε αντίθεση με τα όσα υποδεικνύει η γνώση. Η εμπιστοσύνη αυτή ονομάζεται "παράλογη".

Ένα βασικό χαρακτηριστικό όλων των κοινωνιών είναι η συνεργασία. Μέσα από τη συνεργασία τα μέλη μιας κοινωνίας προσπαθούν να πετύχουν τους προσωπικούς τους στόχους και να ικανοποιήσουν τις ανάγκες τους. Μάλιστα, η τάση για συνεργασία συχνά οδηγεί και στη δημιουργία κοινωνικών ομάδων, των οποίων τα μέλη συνδέονται με κοινές επιδιώξεις και αρχές. Η εμπιστοσύνη παίζει έναν πολύ σημαντικό ρόλο στην αναζήτηση συνεργασιών, καθώς τα άτομα διαλέγουν να συνεργαστούν με την πιο αξιόπιστη απ' τις διαθέσιμες επιλογές [1]. Υπό την έννοια αυτή, τα μέλη μιας κοινωνίας επιδιώκουν να βελτιώσουν την αξιοπιστία τους, ώστε να μεγιστοποιήσουν τις επιλογές τους για συνεργασία και τελικά να πετύχουν την ευημερία τους. Ακόμη και τα κακόβουλα μέλη μιας κοινωνίας θα προσπαθήσουν να κτίσουν ένα σημαντικό επίπεδο εμπιστοσύνης πριν επιχειρήσουν να εξαπατήσουν. Σε εκείνο το σημείο όμως θα αντιμετωπίσουν το δίλημμα, αν αξίζει το κόπο να θυσιάσουν (ίσως και μόνιμα) την αξιοπιστία τους για τα πρόσκαιρα οφέλη μιας συναλλαγής.

Για τη διαχείριση εμπιστοσύνης έχουν προταθεί πάρα πολλά μοντέλα εμπιστοσύνης. Κεντρικό σημείο όλων των μοντέλων είναι ο μηχανισμός υπολογισμού της εμπιστοσύνης, ο οποίος συνήθως περιλαμβάνει μια συνάρτηση, η οποία είναι γνωστή με τον όρο συνάρτηση εμπιστοσύνης (trust function). Με βάση τα ιδιαίτερα χαρακτηριστικά του υπολογισμού της εμπιστοσύνης τα μοντέλα εμπιστοσύνης μπορούν να ταξινομηθούν στις παρακάτω κατηγορίες [12]:

#### 1. Με βάση τη φύση της αξιολόγησης

- **Μοντέλα ταυτότητας (Identity-based)**

Στην κατηγορία αυτή εντάσσονται τα μοντέλα που υπολογίζουν την εμπιστοσύνη ελέγχοντας την ταυτότητα της αξιολογούμενης οντότητας και των διαπιστευτηρίων που διαθέτει, χρησιμοποιώντας τις παραδοσιακές τεχνικές της αυθεντικοποίησης και της εξουσιοδότησης.

- **Μοντέλα συμπεριφοράς (Behaviour-based)**

Στα μοντέλα αυτά η εμπιστοσύνη διαμορφώνεται λαμβάνοντας υπόψη τη συμπεριφορά της αξιολογούμενης οντότητας χρησιμοποιώντας τις εμπειρίες από προηγούμενες αλληλεπιδράσεις μαζί της. Όλα τα σύγχρονα μοντέλα εμπιστοσύνης εμπίπτουν σ' αυτή την κατηγορία.

#### 2. Με βάση τον τύπο διαχείρισης

- **Μοντέλα κεντρικής διαχείρισης (Centralized)**

Στα μοντέλα κεντρικής διαχείρισης η εμπιστοσύνη υπολογίζεται σε ένα κεντρικό κόμβο χρησιμοποιώντας τα δεδομένα που αποστέλλουν τα μέλη της κοινωνίας από τις προσωπικές τους εμπειρίες (αλληλεπιδράσεις). Όταν ένα μέλος χρειάζεται μια εκτίμηση εμπιστοσύνης (για κάποιο άλλο μέλος) τη ζητά από τον κεντρικό κόμβο.

- **Μοντέλα καταναμημένης διαχείρισης (Distributed)**

Στα μοντέλα κατανεμημένης ή αποκεντρωμένης διαχείρισης κάθε μέλος της κοινωνίας υπολογίζει μόνο του την εμπιστοσύνη, χρησιμοποιώντας δεδομένα κυρίως από τις δικές του εμπειρίες και ίσως και από πληροφορίες που φτάνουν σ' αυτό από άλλα μέλη της κοινωνίας (φήμη).

### 3. Με βάση την υποκειμενικότητα της αξιολόγησης

- **Μοντέλα αντικειμενικής αξιολόγησης (*Objective*)**

Στα μοντέλα αντικειμενικής αξιολόγησης, η εκτίμηση της εμπιστοσύνης μιας οντότητας είναι ίδια για όλα τα μέλη της κοινωνίας. Έτσι όσα μέλη και αν υπολογίσουν την εμπιστοσύνη μιας οντότητας, θα καταλήξουν στο ίδιο αποτέλεσμα. Η κατηγορία αυτή είναι συγγενική με την κατηγορία της κεντρικής διαχείρισης, καθώς στην κεντρική διαχείριση έχουμε μια κεντρική αξιολόγηση της εμπιστοσύνης την οποία χρησιμοποιούν όλα τα μέλη της κοινωνίας.

- **Μοντέλα υποκειμενικής αξιολόγησης (*Subjective*)**

Στα μοντέλα υποκειμενικής αξιολόγησης, το κάθε μέλος της κοινωνίας εκτιμά με το δικό του υποκειμενικό τρόπο, βάσει των αρχών και των πεποιθήσεών του, την αξιοπιστία κάποιου άλλου μέλους. Τα μοντέλα αυτής της κατηγορίας είναι πιο κοντά στην πραγματικότητα, αφού η εμπιστοσύνη εκ φύσεως (ως πεποίθηση) είναι υποκειμενική.

### 4. Με βάση τα δεδομένα που χρησιμοποιούνται στην αξιολόγηση

- **Μοντέλα συναλλαγών (*Transaction-based*)**

Τα μοντέλα αυτής της κατηγορίας υπολογίζουν την εμπιστοσύνη χρησιμοποιώντας δεδομένα μόνο από τις προσωπικές συναλλαγές του κάθε μέλους, χωρίς να λαμβάνονται υπόψη οι αντίστοιχες εμπειρίες άλλων μελών.

- **Μοντέλα φήμης (*Reputation-based*)**

Στην κατηγορία αυτή η εμπιστοσύνη μιας οντότητας υπολογίζεται λαμβάνοντας υπόψη τη προσωπική γνώμη που έχουν γι' αυτή τα άλλα μέλη της κοινωνίας. Βέβαια, το κάθε μέλος μπορεί να διαμορφώνει την προσωπική του γνώμη από τα δεδομένα των επιμέρους συναλλαγών, όμως στα άλλα μέλη μεταφέρει μια συνολική εκτίμηση της αξιολογούμενης οντότητας και όχι τις λεπτομέρειες των συναλλαγών.

### 5. Με βάση την προέλευση των δεδομένων που χρησιμοποιούνται στην αξιολόγηση

- **Μοντέλα τοπικών δεδομένων (*Localized*)**

Στα μοντέλα αυτής της κατηγορίας τα δεδομένα που χρησιμοποιούνται στον υπολογισμό της εμπιστοσύνης είναι τοπικά, δηλ. προέρχονται είτε από το ίδιο το μέλος που διενεργεί τον αξιολόγηση (προσωπικές εμπειρίες συναλλαγών), είτε από γειτονικά μέλη με τα οποία διατηρεί σχέσεις συνεργασίας. Επειδή τα δεδομένα που χρησιμοποιεί κάθε μέλος για τον υπολογισμό της

εμπιστοσύνης μιας οντότητας είναι διαφορετικά, ο υπολογισμός δεν μπορεί παρά να είναι υποκειμενικός.

- **Μοντέλα καθολικών δεδομένων (Global)**

Τα μοντέλα καθολικών δεδομένων χρησιμοποιούν το σύνολο των διαθέσιμων δεδομένων από όλα τα μέλη της κοινωνίας. Αυτό υλοποιείται είτε με κεντρική διαχείριση, με τα δεδομένα να συγκεντρώνονται σε ένα κεντρικό κόμβο, είτε με συλλογή δεδομένων από ολόκληρη την κοινωνία. Η δεύτερη εκδοχή προφανώς έχει πολύ μεγάλο κόστος επικοινωνίας. Η καθολικότητα των δεδομένων καθιστά την αξιολόγηση της εμπιστοσύνης αντικειμενική, αφού όλα τα μέλη χρησιμοποιούν τα ίδια δεδομένα.

#### 6. Με βάση τον τρόπο χρησιμοποίησης του αποτελέσματος της αξιολόγησης

- **Μοντέλα κατάταξης (Rank-based)**

Στα μοντέλα αυτής της κατηγορίας η αξιολόγηση της εμπιστοσύνης μιας οντότητας από μόνη της δεν προσφέρει σημαντική πληροφόρησης. Μπορεί όμως να χρησιμοποιηθεί για τη σύγκριση της οντότητας με άλλες οντότητες και την κατάταξή της σε μια κλίμακα αξιοπιστίας.

- **Μοντέλα κατωφλίου (Threshold-based)**

Στα μοντέλα κατωφλίου υπάρχει μια προκαθορισμένη τιμή κατωφλίου ως προς την οποία συγκρίνεται η αξιολόγηση της εμπιστοσύνης μιας οντότητας. Το αποτέλεσμα της σύγκρισης καθορίζει το αν θα πραγματοποιηθεί κάποια αλληλεπίδραση με την αξιολογούμενη οντότητα. Τα περισσότερα μοντέλα εμπίπτουν σ' αυτή την κατηγορία.

Πρέπει να τονισθεί ότι οι παραπάνω κατηγορίες δεν είναι αμοιβαία αποκλειόμενες. Έτσι ένα μοντέλο μπορεί να εμπίπτει ταυτόχρονα σε περισσότερες από μια κατηγορίες. Στατιστικά, κυρίως λόγω της πληθώρας των εφαρμογών κατανεμημένης τεχνητής νοημοσύνης, τα σύγχρονα μοντέλα εμπιστοσύνης, στη συντριπτική τους πλειοψηφία, εμπίπτουν στις κατηγορίες της κατανεμημένης διαχείρισης, της υποκειμενικής αξιολόγησης, της φήμης, των τοπικών δεδομένων και του κατωφλίου.

### 1.3 Εφαρμογές των μοντέλων εμπιστοσύνης

Το πεδίο εφαρμογής των μοντέλων εμπιστοσύνης είναι αρκετά ευρύ και συνεχώς διευρύνεται. Οι πιο προφανείς περιοχές εφαρμογών είναι τα περιβάλλοντα στα οποία αναπτύσσονται ψηφιακές κοινωνίες με online συναλλαγές. Ας δούμε, σύντομα, μερικές χαρακτηριστικές περιοχές.

#### **E-commerce**

Το ηλεκτρονικό εμπόριο είναι ίσως η πιο σημαντική εφαρμογή του ψηφιακού κόσμου (κυρίως λόγω του οικονομικού παράγοντα), στην οποία η εμπιστοσύνη παίζει καταλυτικό ρόλο. Η στροφή της ανθρώπινης επιχειρηματικότητας προς το ηλεκτρονικό εμπόριο έχει οδηγήσει στην ανάπτυξη τεχνολογιών για την αξιοποίηση των αρχών διαχείρισης εμπιστοσύνης στις online συναλλαγές [14]. Οι τεχνολογίες αυτές

αποτελούν το οξυγόνο του νέου περιβάλλοντος καθώς διαμορφώνουν ένα κλίμα εμπιστοσύνης και ασφάλειας ανάμεσα στους καταναλωτές και στις επιχειρήσεις. Δύο πολύ γνωστά παραδείγματα εφαρμογών ηλεκτρονικού εμπορίου που χρησιμοποιούν μηχανισμούς διαχείρισης φήμης για τη υποστήριξη της λειτουργικότητάς τους είναι το [18, 24] και το Bizrate [19].

### **Peer-to-Peer Resource Sharing**

Οι εφαρμογές διαμοιρασμού πόρων σε P2P δίκτυα είναι ένας ακόμη χώρος στον οποίο η διαχείριση εμπιστοσύνης μπορεί να προσφέρει σημαντική βοήθεια. Η αυτονομία των κόμβων, η δυνατότητα διατήρησης της ανωνυμίας, η δυναμική φύση των συνθηκών του δικτύου και η απουσία αποτελεσματικού μηχανισμού ασφάλειας καθιστούν τις συναλλαγές σ' ένα P2P περιβάλλον εξαιρετικά επικίνδυνες. Αρκετά μοντέλα εμπιστοσύνης έχουν προσφέρει αξιόπιστες λύσεις στην περιοχή αυτή, στην κατεύθυνση της προφύλαξης των κόμβων από κακόβουλες συμπεριφορές [15,28].

### **Mobile Systems**

Οι εξελίξεις στον τομέα του hardware έχουν επιτρέψει τη δημιουργία φορητών συσκευών που έχουν διεισδύσει στην καθημερινή μας ζωή. Οι εντυπωσιακές δυνατότητες αυτών των συσκευών ενισχύονται από την ανάπτυξη των ασύρματων δικτύων, επιτρέποντας στους χρήστες τους να έχουν πρόσβαση σε πόρους και υπηρεσίες οποιαδήποτε στιγμή και από οπουδήποτε. Τα ζητήματα που προκύπτουν σ' αυτό το περίπλοκο περιβάλλον σχετίζονται με την επικινδυνότητα που δημιουργεί η παραχώρηση ελέγχου σε άγνωστες οντότητες κατά τη διάρκεια μιας συναλλαγής. Η ενίσχυση των φορητών συσκευών με μηχανισμούς διαχείρισης εμπιστοσύνης επιτρέπει τη μείωση του ρίσκου των συναλλαγών καθώς τους δίνει τη δυνατότητα να επιλέγουν τους πιο αξιόπιστους κόμβους για τις συναλλαγές τους [6].

### **Grid**

Το Grid μπορεί να ειπωθεί ως ένα πολυπρακτορικό σύστημα που παρέχει υπηρεσίες διαμοιρασμού υπολογιστικών πόρων, με τους χρήστες και τους πόρους να λειτουργούν ως αυτόνομοι πράκτορες που έχουν τον έλεγχο της συμπεριφοράς τους. Η ύπαρξη της αυτονομίας δημιουργεί αβεβαιότητα καθώς δεν μπορούμε να γνωρίζουμε με ακρίβεια το πως θα συμπεριφερθεί μια οντότητα σε μια αλληλεπίδραση. Οι συχνές μεταβολές που λαμβάνουν χώρα στο ανοικτό περιβάλλον του Grid με την είσοδο και την αποχώρηση κόμβων επιτείνουν το πρόβλημα. Η αβεβαιότητα και κατ' επέκταση το ρίσκο που εμφανίζουν οι αλληλεπιδράσεις σε περιβάλλον Grid μπορεί να αντιμετωπισθεί με την ενσωμάτωση της εμπιστοσύνης στη λογική των πρακτόρων που συμμετέχουν στο Grid [16], κατά τρόπο που να μπορούν να ξεχωρίσουν και να επιλέξουν τις πιο αξιόπιστες οντότητες για τις συναλλαγές τους.

### **Ασφάλεια κατανεμημένων συστημάτων**

Η παραδοσιακή προσέγγιση για την ασφάλεια των συστημάτων χειρίζεται την ασφάλεια με ένα συνδυασμό μηχανισμών αυθεντικοποίησης και ελέγχου πρόσβασης. Πρώτα γίνεται η αναγνώριση της οντότητας που ζητά πρόσβαση στο σύστημα και στη συνέχεια αναζητούνται σε κάποια εσωτερική βάση δεδομένων με τα δικαιώματα πρόσβασης, τα δικαιώματα της συγκεκριμένης οντότητας. Η προσέγγιση

αυτή όμως είναι ανεπαρκής στο δυναμικό και ετερογενές περιβάλλον των κατανεμημένων συστημάτων, αφενός γιατί οι συμμετέχουσες οντότητες δεν είναι εκ των προτέρων γνωστές και αφετέρου διότι τα δικαιώματα πρόσβασης μεταβάλλονται δυναμικά με το χρόνο και τις εκάστοτε συνθήκες. Η χρήση της διαχείρισης εμπιστοσύνης, με τον έλεγχο της συμπεριφοράς, δίπλα στους παραδοσιακούς μηχανισμούς ασφάλειας, ενισχύει την αποτελεσματικότητά τους επιτρέποντας τους να αντεπεξέλθουν στις προκλήσεις που εισάγει η δυναμική φύση των κατανεμημένων συστημάτων [17].

## 1.4 Αντικείμενο της διατριβής

Παρά την πολύ μεγάλη προσπάθεια που γίνεται για την ανάπτυξη νέων μοντέλων εμπιστοσύνης και την ενσωμάτωσή τους στις υπάρχουσες εφαρμογές του ψηφιακού κόσμου, ελάχιστη δουλειά έχει γίνει στην κατεύθυνση της αξιολόγησης των μοντέλων εμπιστοσύνης. Η εργασία αυτή επιχειρεί να καλύψει το κενό που παρατηρείται στον τομέα αυτό, προτείνοντας μια ολοκληρωμένη πλατφόρμα για την μελέτη και αξιολόγηση μοντέλων εμπιστοσύνης. Η πλατφόρμα περιλαμβάνει :

- α) Ένα παραμετροποιημένο περιβάλλον προσομοίωσης μοντέλων εμπιστοσύνης με δυνατότητα real time γραφικής απεικόνισης της εξέλιξης των προσομοιώσεων
- β) Ένα interface για την κωδικοποίηση μοντέλων εμπιστοσύνης και την εισαγωγή τους στον προσομοιωτή
- γ) Ένα σύνολο γενικών μετρικών αξιολόγησης μοντέλων εμπιστοσύνης. Οι μετρικές είναι σχεδιασμένες έτσι ώστε να επιτρέπουν και τη σύγκριση των προσομοιούμενων μοντέλων

Για να γίνει εφικτή η υλοποίηση της πλατφόρμας χρειάστηκε αρχικά να πραγματοποιηθεί μια ευρεία και συστηματική μελέτη του χώρου των μοντέλων εμπιστοσύνης με στόχο :

- να διαπιστωθεί εάν υπάρχει κάποια κοινή βάση προσέγγισης, πάνω την οποία θα μπορούσε να οικοδομηθεί μια πλατφόρμα αξιολόγησης μοντέλων εμπιστοσύνης
- να προσδιορισθεί ο χώρος των διαφορετικών χαρακτηριστικών που εμφανίζουν τα υπάρχοντα μοντέλα εμπιστοσύνης, τα οποία θα πρέπει να υποστηρίξει μια πλατφόρμα αξιολόγησης μοντέλων εμπιστοσύνης

Η προτεινόμενη πλατφόρμα δοκιμάστηκε στην πράξη με την προσομοίωση κάποιων μοντέλων εμπιστοσύνης και τη συγκριτική αξιολόγησή τους. Τα αποτελέσματα των προσομοιώσεων, που παρουσιάζονται στο τελευταίο κομμάτι της εργασίας, δείχνουν πως η ενιαία αξιολόγηση μοντέλων εμπιστοσύνης είναι εφικτή μέσα από τη συγκεκριμένη πλατφόρμα. Ελπίζουμε πως η εργασία αυτή θα αποτελέσει ένα ερέθισμα για να διεξαχθεί περισσότερη έρευνα προς στην κατεύθυνση αυτή.

## 1.5 Συναφείς εργασίες

Καθ' όλη τη διάρκεια εκπόνησης της παρούσας διατριβής δεν μπορέσαμε να εντοπίσαμε κάποια άλλη ερευνητική δουλειά που να είναι συναφής με το αντικείμενο της δικής μας δουλειάς. Βρέθηκαν δύο μόνο εν μέρει σχετικές δουλειές, οι οποίες όμως αφορούσαν την αξιολόγηση των μηχανισμών



διαμόρφωσης φήμης κάποιων μοντέλων εμπιστοσύνης. Κατά το χρόνο ολοκλήρωσης της διατριβής (Ιούνιος 2006), εμφανίστηκαν αναφορές για το project *Agent Reputation and Trust (ART) Testbed* [20]. Πρόκειται για μια διεθνή ερευνητική προσπάθεια που ξεκίνησε το 2005 με στόχο τη δημιουργία μιας πλατφόρμας για την αξιολόγηση μοντέλων εμπιστοσύνης με ένα αρκετά πρωτότυπο τρόπο: βάζοντάς τα να διαγωνιστούν πάνω στην αντιμετώπιση ενός συγκεκριμένου προβλήματος. Ο διαγωνισμός έχει τα χαρακτηριστικά παιχνιδιού στο οποίο οι παίκτες είναι πράκτορες που ενσωματώνουν τη λογική των αξιολογούμενων μοντέλων (ένας πράκτορας για κάθε μοντέλο) και δραστηριοποιούνται προσπαθώντας να κερδίσουν ακολουθώντας τους κανόνες του παιχνιδιού. Φυσικά, οι κανόνες και η θεματολογία του παιχνιδιού είναι επιλεγμένοι έτσι ώστε να απαιτούν τη χρήση μηχανισμών διαχείρισης εμπιστοσύνης από τους πράκτορες-παίκτες. Εκτός από τη λειτουργία σε κατάσταση διαγωνισμού (competition mode), η πλατφόρμα μπορεί να λειτουργήσει και σε κατάσταση πειραματισμού (experimentation mode) στην οποία ο χρήστης μπορεί να ρυθμίσει τις παραμέτρους του παιχνιδιού και να παρακολουθήσει την απόδοση των συμμετεχόντων πρακτόρων. Ο πρώτος διαγωνισμός πραγματοποιήθηκε το Μάιο του 2006 και μέχρι την ολοκλήρωση αυτής της διατριβής, δεν είχαν ανακοινωθεί κάποια σχετικά συμπεράσματα. Δυστυχώς, πέραν των γενικών προδιαγραφών και στόχων της πλατφόρμας [21] δεν έχουμε εντοπίσει κάποια άλλη αναφορά της πλατφόρμας με αποτελέσματα και συμπεράσματα από αξιολογήσεις κάποιων γνωστών μοντέλων εμπιστοσύνης ώστε να μπορούμε να διατυπώσουμε μια εμπειριστατωμένη κριτική. Ως μια πρώτη εντύπωση, φαίνεται να αποτελεί μια όχι ιδιαίτερα ευέλικτη πρόταση που δίνει μεγαλύτερη έμφαση στην αξιολόγηση της ικανότητας των μοντέλων να διαγωνίζονται πάνω στην επίλυση ενός συγκεκριμένου προβλήματος, παρά στην μελέτη και αξιολόγηση των λειτουργικών χαρακτηριστικών τους. Πάντως αναμένουμε με ενδιαφέρον τις πρώτες τεχνικές αναφορές για να διαπιστώσουμε πιθανά σημεία σύγκλισης με τη δική μας πρόταση.

Οι δύο άλλες, λιγότερο συναφείς, δουλειές [22,23] ασχολούνται με την αξιολόγηση του μηχανισμού εκτίμησης (rating) της αξιοπιστίας μέσω της φήμης. Τέτοιο μηχανισμό διαθέτουν τα μοντέλα εμπιστοσύνης που βασίζονται στη φήμη (reputation-based). Και στις δύο δουλειές υπάρχουν πειραματικά αποτελέσματα από προσομοιώσεις που έγιναν πάνω στους μηχανισμούς φήμης κάποιων γνωστών reputation-based μοντέλων εμπιστοσύνης. Ωστόσο μόνο στο [22] το περιβάλλον προσομοίωσης διατίθεται ως αυτόνομο λογισμικό που μπορεί να χρησιμοποιηθεί από τρίτους ως πλατφόρμα αξιολόγησης για την αξιολόγηση μηχανισμών φήμης.

## 1.6 Παραδοχές - Συμβάσεις

Η εμπιστοσύνη είναι ένα κοινωνικό φαινόμενο και ως τέτοιο έχει ενδιαφέρον να το μελετούμε κυρίως εντός των ορίων κάποιας κοινωνίας. Στην εργασία αυτή επικεντρώνουμε το ενδιαφέρον μας στις κοινωνίες πρακτόρων θεωρώντας ότι στα πολυπρακτορικά περιβάλλοντα συναντάμε τις σημαντικότερες εφαρμογές της διαχείρισης εμπιστοσύνης. Στο υπόλοιπο της εργασίας υιοθετούμε τους παρακάτω ορισμούς :

**Μοντέλο εμπιστοσύνης** : Σύνολο κανόνων, μηχανισμών και διαδικασιών που επιτρέπουν την αναπαράσταση και διαχείριση της εμπιστοσύνης.

**Πράκτορας εμπιστοσύνης** : πράκτορας που έχει την ικανότητα να χρησιμοποιεί και να διαχειρίζεται την έννοια της εμπιστοσύνης

Επιπλέον κάνουμε τις παρακάτω παραδοχές :

- Οι όροι *συνεργασία* και *συναλλαγή* θεωρούμε ότι έχουν παραπλήσια σημασία, υπό την έννοια ότι αντιπροσωπεύουν μια κατάσταση κατά την οποία δύο μέρη συνεργάζονται (συμφωνούν να έρθουν σε επαφή) για να ανταλλάξουν κάτι (υπηρεσίες, αγαθά, πληροφορίες κλπ)
- Ασχολούμαστε με καταστάσεις στις οποίες η εμπιστοσύνη δεν είναι απόλυτη, καθώς από τη φύση της η εμπιστοσύνη έχει νόημα εκεί όπου υπάρχει αμφιβολία (έστω και ελάχιστη) και ρίσκο. Γι' αυτό και μελετούμε κοινωνίες που δεν είναι ιδανικές και μπορούν να περιλαμβάνουν και μη υγιείς συμπεριφορές.
- Ως μη υγιή ή κακόβουλη συμπεριφορά θεωρούμε οποιαδήποτε εκδήλωση μιας δραστηριότητα που έχει ένα ανεπιθύμητο αποτέλεσμα, ανεξάρτητα απ' το εάν προέρχεται από κακή πρόθεση ή αν το αποτέλεσμα είναι πραγματικά επιζήμιο<sup>3</sup>.

## 1.7 Διάρθρωση της διατριβής

Η εργασία είναι οργανωμένη σε έξι κεφάλαια. Τα περιεχόμενο των κεφαλαίων που ακολουθούν εν συντομία είναι :

Στο **κεφάλαιο 2** παρουσιάζονται τα σημαντικότερα σημεία της μελέτης κάποιων μοντέλων εμπιστοσύνης. Από τα μοντέλα που μελετήθηκαν, επιλέχθηκαν και παρουσιάζονται τα πιο χαρακτηριστικά, καλύπτοντας τις περισσότερες από τις κατηγορίες των μοντέλων εμπιστοσύνης που καταγράφηκαν παραπάνω.

Στο **κεφάλαιο 3** καταγράφονται τα κοινά χαρακτηριστικά και οι διαφορετικοί μηχανισμοί που εμφανίζουν τα μοντέλα που μελετήθηκαν. Τα στοιχεία αυτά χρησιμοποιούνται για τον προσδιορισμό των προδιαγραφών του περιβάλλοντος προσομοίωσης και τη σχεδίαση της αρχιτεκτονικής του.

Στο **κεφάλαιο 4** παρουσιάζονται οι προδιαγραφές, οι σχεδιαστικές επιλογές και η αρχιτεκτονική του περιβάλλοντος προσομοίωσης καθώς και τα σημαντικότερα λειτουργικά του χαρακτηριστικά.

Στο **κεφάλαιο 5** προτείνονται κάποιες μετρικές αξιολόγησης μοντέλων εμπιστοσύνης. Οι μετρικές αυτές ενσωματώθηκαν στο περιβάλλον προσομοίωσης και επιτρέπουν την συγκριτική αξιολόγηση των προσομοιούμενων μοντέλων.

---

<sup>3</sup> Ανεπιθύμητη μπορεί να θεωρηθεί και μια κακής ποιότητας υπηρεσία, που όμως δεν προκαλεί κάποια συγκεκριμένη ζημιά.

Στο κεφάλαιο 6 παρουσιάζονται τα πειραματικά αποτελέσματα που προέκυψαν από τις προσομοιώσεις που διενεργήθηκαν σε μοντέλα εμπιστοσύνης, χρησιμοποιώντας τις μετρικές του κεφαλαίου 5.

Στο κεφάλαιο 7 συνοψίζονται τα αποτελέσματα και η συνεισφορά της εργασίας ενώ γίνεται και μια σύντομη αναφορά στις δυσκολίες που αντιμετωπίσαμε κατά την εκπόνησή της, κυρίως σε ότι αφορά την ανάπτυξη του περιβάλλοντος προσομοίωσης. Παράλληλα καταγράφονται και οι πολλές και ενδιαφέρουσες δυνατότητες που υπάρχουν για μελλοντική δουλειά.

# ΚΕΦΑΛΑΙΟ

## 2. Μοντέλα Εμπιστοσύνης

Το κεφάλαιο αυτό παρουσιάζει συνοπτικά μερικά από τα πλέον γνωστά μοντέλα εμπιστοσύνης που έχουν προταθεί. Τα μοντέλα που επιλέξαμε να μελετήσουμε σε περισσότερο βάθος είναι κατά το δυνατόν τα περισσότερο αντιπροσωπευτικά, σύμφωνα με τις κατηγορίες που παρουσιάστηκαν στο κεφάλαιο 1. Η συζήτηση που ακολουθεί επικεντρώνεται στα σημαντικότερα από τα λειτουργικά χαρακτηριστικά τους, τα οποία χρησιμοποιήθηκαν ως σημείο αναφοράς για το σχεδιασμό της αρχιτεκτονικής και τον καθορισμό των τεχνικών προδιαγραφών του προσομοιωτή.

### 2.1 Η περίπτωση eBay

Το eBay είναι ίσως μια από τις χαρακτηριστικότερες και πιο ενδιαφέρουσες περιπτώσεις εφαρμογής μηχανισμών διαχείρισης εμπιστοσύνης για την υποστήριξη διαδικασιών ηλεκτρονικού εμπορίου. Εμπίπτει στην κατηγορία των κεντροποιημένων (centralized) μοντέλων με διαχείριση φήμης (reputation management): η εμπιστοσύνη υπολογίζεται κεντρικά μέσα από τις αναφορές των μελών της κοινωνίας. Το μοντέλο εμπιστοσύνης του eBay χαρακτηρίζεται και ως μοντέλο "εμπιστοσύνης κοινότητας" (community trust model) διότι διαθέτει μηχανισμούς που οικοδομούν μια κοινότητα που εξασφαλίζει στα μέλη της ένα περιβάλλον ασφάλειας και εμπιστοσύνης, βασισμένο όχι σε ελέγχους και περιορισμούς αλλά στην καλή διάθεση και τις θετικές προσδοκίες των ίδιων των μελών της [24].

#### 2.1.1 Τρόπος λειτουργίας

Λειτουργικά, το eBay παρέχει μια υπηρεσία διεκπεραίωσης ηλεκτρονικών δημοπρασιών (e-auctions). Γενικά, βασίζεται πάνω στις ίδιες αρχές που διέπουν και τους συμβατικούς (μη ηλεκτρονικούς) οίκους δημοπρασιών με μια σημαντική διαφορά όμως: δεν ελέγχει τα είδη που προσφέρονται. Το eBay λειτουργεί ουσιαστικά ως διαμεσολαβητής, προσφέροντας απλά ένα ψηφιακό χώρο και κάποιες διαδικασίες για την επικοινωνία όσων επιθυμούν να πουλήσουν κάτι, μ' αυτούς που θέλουν να το αγοράσουν.

Οι χρήστες του eBay, πριν συμμετάσχουν σε μια δημοπρασία, είτε ως πωλητές είτε ως υποψήφιοι αγοραστές, εγγράφονται στο σύστημα δίνοντας κάποιες προσωπικές τους πληροφορίες όπως το όνομά τους, μια διεύθυνση επικοινωνίας και ένα ψευδώνυμο με το οποίο εμφανίζονται στις δημοπρασίες. Η μόνη πληροφορία που επαληθεύει το eBay (τουλάχιστο για τους πωλητές) είναι το e-mail. Επιπλέον, το σύστημα δεν αποκαλύπτει τα πραγματικά ονόματα και τις φυσικές διευθύνσεις των χρηστών. Το γεγονός αυτό, σε συνδυασμό με τη δυνατότητα εύκολης απόκτησης e-mail από διάφορες πηγές (Yahoo, HotMail κλπ) σημαίνει ότι οι χρήστες του eBay μπορούν να διατηρήσουν την ανωνυμία τους.

Όταν ένας χρήστης επιθυμεί να πουλήσει ένα αντικείμενο, το δηλώνει, μέσω σχετικής φόρμας, στο site του eBay, δίνοντας μια μικρή περιγραφή, κάποια ψηφιακή φωτογραφία και κάποια

διαδικαστικά στοιχεία, όπως κόστος αποστολής, τρόποι πληρωμής κλπ. Παράλληλα καθορίζει και τα χαρακτηριστικά της επικείμενης δημοπρασίας, όπως την ελάχιστη τιμή πώλησης (τιμή εκκίνησης), τη διάρκεια της δημοπρασίας κλπ.

Όσοι επιθυμούν να αγοράσουν κάτι από το eBay, αφού εντοπίσουν (μέσω των σχετικών εργαλείων αναζήτησης που διαθέτει το site) το αντικείμενο που τους ενδιαφέρει, υποβάλλουν τη μέγιστη προσφορά τους. Μετά το τέλος του χρόνου της δημοπρασίας, το σύστημα συγκρίνει τις προσφορές που έχουν υποβληθεί για το αντικείμενο, βρίσκει την υψηλότερη και στέλνει ένα ενημερωτικό μήνυμα στον πλειοδότη και στον πωλητή. Οι δυο τους θα πρέπει στη συνέχεια, εντός τριών ημερών, να επικοινωνήσουν μεταξύ τους μέσω e-mail και να ανταλλάξουν διευθύνσεις. Ο πλειοδότης αποστέλλει το ποσό της προσφοράς που έκανε και ο πωλητής αποστέλλει το αντικείμενο.

Όλα τα παραπάνω είναι λίγο-πολύ αναμενόμενα και συνηθισμένα για ένα σύστημα ηλεκτρονικών δημοπρασιών. Το ενδιαφέρον κομμάτι της περίπτωσης eBay είναι ότι, μετά την ολοκλήρωση της διαδικασίας, η κάθε πλευρά μπορεί να καταθέσει στο λεγόμενο σύστημα ανατροφοδότησης του eBay (Feedback system) μια αξιολόγηση της άλλης πλευράς, η οποία κυμαίνεται ανάμεσα σε τρεις δυνατές διαβαθμίσεις : θετική, ουδέτερη και αρνητική. Η αξιολόγηση αυτή καταγράφεται στο ιστορικό του χρήστη για να συμπεριληφθεί στον υπολογισμό ενός ατομικού δείκτη που συνοδεύει τον χρήστη σε κάθε του συναλλαγή.

### 2.1.2 Το σύστημα ανατροφοδότησης

Το σύστημα ανατροφοδότησης είναι ουσιαστικά ένα σύστημα φήμης, κεντρικής διαχείρισης, που διέπεται από εξαιρετικά απλούς κανόνες. Μέσω του συστήματος ανατροφοδότησης υπολογίζεται για κάθε χρήστη ένας αριθμός που λειτουργεί ως δείκτης αξιοπιστίας για το χρήστη και ο οποίος αποτελεί το θεμέλιο λίθο και την κινητήρια δύναμη του μοντέλου. Αυτό γίνεται ως εξής:

Μετά από μια συναλλαγή, ο πωλητής και ο αγοραστής μπορούν να αφήσουν ένα σχόλιο ο ένας για τον άλλο. Το σχόλιο αποτελείται από μια γραμμή κειμένου και μια αριθμητική αξιολόγηση η οποία μπορεί να πάρει τις τιμές +1, 0, -1 υποδηλώνοντας αντίστοιχα θετική, ουδέτερη ή αρνητική γνώμη. Αρχικά (μέχρι το 1999) κάθε χρήστης μπορούσε να στείλει ένα σχόλιο για οποιοδήποτε άλλο χρήστη. Από το 2000, τα σχόλια συνδέθηκαν με τις συναλλαγές. Έτσι, πλέον μόνο ο πωλητής και ο αγοραστής μπορούν να στείλουν σχόλιο ο ένας για τον άλλο και το σχόλιο θα αφορά τη συγκεκριμένη συναλλαγή που είχαν μεταξύ τους. Οι αξιολογήσεις που συγκεντρώνει ο κάθε χρήστης χρησιμοποιούνται για τον υπολογισμό του ατομικού δείκτη του χρήστη. Συγκεκριμένα, οι αξιολογήσεις κάθε χρήστη ομαδοποιούνται κατά αποστολέα και για κάθε αποστολέα υπολογίζεται το άθροισμα των τιμών αξιολόγησης που έστειλε για να εκτιμηθεί εάν έχει θετική, αρνητική, ή ουδέτερη γνώμη για το χρήστη. Για παράδειγμα, εάν ένας αποστολέας έστειλε για ένα συγκεκριμένο χρήστη μια θετική (+1) και δύο αρνητικές (-2) αξιολογήσεις, προκύπτει άθροισμα -1 που σημαίνει ότι ο αποστολέας θεωρείται ότι έχει

αρνητική γνώμη για το χρήστη. Το άθροισμα των διακριτών γνωμών που συγκεντρώνει ένας χρήστης με τον τρόπο αυτό αποτελεί και τον δείκτη αξιοπιστίας του. Για παράδειγμα εάν ένας χρήστης συγκεντρώνει δέκα θετικές, δύο αρνητικές και τέσσερις ουδέτερες γνώμες, ο χρήστης θα έχει δείκτη αξιοπιστίας 8.

Αξίζει να τονιστεί ότι το σύστημα ανατροφοδότησης για κάθε χρήστη, λαμβάνει υπόψη κάθε διαφορετικό αποστολέα αξιολόγησης μια μόνο φορά, δηλαδή σαν να έχει στείλει ο χρήστης μια μόνο αξιολόγηση. Αυτό γίνεται για να μην μπορεί να δημιουργηθεί για κάποιο χρήστη παραπλανητική εικόνα, μέσω εικονικών συναλλαγών με κάποιους γνωστούς του.

### 2.1.3 Σχολιασμός - κριτική

Το μοντέλο eBay δεν προτείνει κάποιο συγκεκριμένο τρόπο χρησιμοποίησης του δείκτη ανατροφοδότησης των χρηστών. Ο κάθε χρήστης είναι ελεύθερος να τον αξιολογήσει και αν εκτιμήσει το τι ακριβώς αντιπροσωπεύει αλλά και το πως μπορεί να τον αξιοποιήσει προκειμένου να διαφυλάξει το συμφέρον του στις επικείμενες συναλλαγές του. Βέβαια μέσω της κοινής λογικής αυτός ο δείκτης, υπαγορεύει στους χρήστες κάποιες συμπεριφορές [25] που θα μπορούσαν να εκληφθούν και ως μηχανισμοί του μοντέλου. Συγκεκριμένα :

- Οι χρήστες τείνουν να εμπιστεύονται περισσότερο αυτούς που έχουν διαμορφώσει ένα μεγαλύτερο δείκτη ανατροφοδότησης.
- Οι χρήστες επιδιώκουν να αυξήσουν όσο περισσότερο μπορούν τον δείκτη τους ελπίζοντας σε μεγαλύτερη εμπιστοσύνη από τα άλλα μέλη της κοινότητας.
- Η απειλή είσπραξης αρνητικών αξιολογήσεων αποτρέπει τους χρήστες από κακόβουλες συμπεριφορές.
- Η ύπαρξη του δείκτη δημιουργεί στους χρήστες ένα αίσθημα ασφάλειας διευκολύνοντας την διενέργεια συναλλαγών μεταξύ αγνώστων.

Η απλότητα του μοντέλου eBay εγείρει έντονη αμφισβήτηση και αρνητική κριτική από πολλούς ερευνητές που ασχολούνται με τη μοντελοποίηση και τη διαχείριση εμπιστοσύνης. Τα κυριότερα σημεία αυτής της κριτικής είναι :

- Οι αξιολογήσεις που δίνονται από τους χρήστες δεν μπορεί να είναι αντικειμενικές διότι :
  - Η κλίμακα αξιολόγησης δεν επιτρέπει την απεικόνιση επαρκών αποχρώσεων ικανοποίησης ή δυσαρέσκειας με αποτέλεσμα οι χρήστες να μην μπορούν να εκφράσουν αντικειμενικά την εκτίμησή τους.
  - Υπάρχει η υποκειμενικότητα της εκτίμησης: καταστάσεις που εκλαμβάνονται ως πρόβλημα από ένα χρήστη μπορεί να μην αποτελούν πρόβλημα για κάποιον άλλο (π.χ. καθυστερήσεις).
  - Υπάρχουν χρήστες που στέλνουν ψευδείς αρνητικές αξιολογήσεις ως αντίποινα για την είσπραξη από την άλλη πλευρά αρνητικών αξιολογήσεων.

- Οι χρήστες δεν είναι υποχρεωμένοι να στέλνουν πάντα μια αξιολόγηση.
- Ο μηχανισμός υπολογισμού του δείκτη δεν είναι αξιόπιστος. Κάποιος χρήστης μπορεί να εξαπατά στο ¼ των συναλλαγών του και ο δείκτης του συνεχώς να αυξάνεται.
- Οι χρήστες μπορούν εύκολα να αλλάξουν ταυτότητα και να απαλλαγούν έτσι από την όποια κακή φήμη τους.

Παρά τις αντιδράσεις πάντως, το σύστημα στην πράξη φαίνεται να δουλεύει και μάλιστα ικανοποιητικά. Το μοντέλο "εμπιστοσύνης κοινότητας" του eBay καταφέρνει με ένα πολύ απλό τρόπο να κτίζει εμπιστοσύνη παρά την έλλειψη εξασφαλίσεων, περιορισμών και μέτρων πρόληψης, όπως φαίνεται από το πλήθος των χρηστών του eBay και των συναλλαγών που διεξάγονται μεταξύ τους.

## 2.2 Το μοντέλο του Marsh

Ο Marsh ήταν από τους πρώτους ερευνητές που μελέτησαν το πρόβλημα της αναπαράστασης της εμπιστοσύνης και επικεντρώθηκαν σε τυπικές, μαθηματικές αναπαραστάσεις. [1]. Στόχος του ήταν να αποσαφηνίσει την, μέχρι τότε, θεωρητική έννοια της εμπιστοσύνης, και να διαμορφώσει έτσι μια κοινή βάση συζήτησης για την επιστημονική έρευνα γύρω από το θέμα αυτό. Όπως εξηγεί ο ίδιος, για το λόγο αυτό προσπάθησε να διατηρήσει την πολυπλοκότητα του μοντέλου του σε όσο το δυνατόν χαμηλό επίπεδο, χωρίς να θυσιάσει την εκφραστική δύναμή του. Η τυπική αναπαράσταση της εμπιστοσύνης που πρότεινε, αν και ελλιπής, αποτέλεσε σημείο αναφοράς για την πλειοψηφία των ερευνητών στα χρόνια που ακολούθησαν τη δημοσίευσή της.

### 2.2.1 Παραδοχές- Χαρακτηριστικά

Το μοντέλο παρέχει υποστήριξη σε έναν πράκτορα  $x$  που καλείται να αποφασίσει αν θα προχωρήσει σε αλληλεπίδραση με κάποιον άλλο πράκτορα  $y$ .

Διακρίνονται τρία είδη εμπιστοσύνης : η βασική εμπιστοσύνη (basic trust), η γενική εμπιστοσύνη (general trust) και η εμπιστοσύνη περίπτωσης (situational trust). Η πρώτη αφορά τη γενικότερη τάση που έχει ένας πράκτορας να εμπιστεύεται ή όχι, και μπορεί να καθορίζεται κατά το σχεδιασμό του πράκτορα (να είναι δηλαδή στοιχείο της προσωπικότητάς του) ή να διαμορφώνεται από τις εμπειρίες που αποκτά ο πράκτορας στη διάρκεια της ζωής του. Τα άλλα δύο είδη αφορούν την εμπιστοσύνη που έχει ένας πράκτορας προς κάποιον άλλο, είτε συνολικά, είτε για συγκεκριμένη περίπτωση αλληλεπίδρασης.

Το μοντέλο βασίζεται στις εξής παραδοχές :

- Ένας πράκτορας έχει τη δυνατότητα επιλογής (δηλ. μπορεί να αρνηθεί μια αλληλεπίδραση). Αυτό είναι απόλυτα σύμφωνο με το χαρακτηριστικό της αυτονομίας που, κατά γενική παραδοχή, οφείλει να διαθέτει λογισμικό για να χαρακτηρίζεται πράκτορας.

- Οι πράκτορες δεν οφείλουν εξ' αρχής κάτι ο ένας στον άλλο, έτσι ώστε να λειτουργούν κάτω από μια αίσθηση υποχρέωσης
- Οι πράκτορες έχουν επίγνωση της κατάστασης που χαρακτηρίζει την όποια επικείμενη αλληλεπίδραση, δηλ. μπορούν να διακρίνουν ομοιότητες και διαφορές με άλλες περιστάσεις που έχουν βιώσει στο παρελθόν.
- Μια αλληλεπίδραση μεταξύ δύο πρακτόρων μπορεί να έχει μόνο δύο αποτελέσματα : συνεργασία ή μη συνεργασία.
- Ένας πράκτορας δεν μπορεί ποτέ να είναι απόλυτα σίγουρος για την αξιοπιστία κάποιου άλλου πράκτορα

Η τελευταία παραδοχή βασίζεται στην αντίληψη ότι, όταν μιλάμε για εμπιστοσύνη υπάρχει (μικρή ή μεγάλη) αβεβαιότητα σχετικά με τις προθέσεις της άλλης πλευράς. Αυτό σημαίνει πως απαιτείται κάποια διερεύνηση και σκέψη για το κατά πόσο μπορούμε να βασιστούμε και να ελπίζουμε σε μια συγκεκριμένη συμπεριφορά. Υπό την έννοια αυτή, ο Marsh θεωρεί ότι η κατάσταση της απόλυτης ή τυφλής εμπιστοσύνης δεν πρέπει να αποτελεί τμήμα ενός μηχανισμού διαχείρισης εμπιστοσύνης, αφού είναι ταυτόσημη με την απόλυτη σιγουριά και συνεπώς, πέρα από οποιοδήποτε σκεπτικό διερεύνησης. Για το λόγο αυτό, το πεδίο τιμών που προτείνει για την αναπαράσταση της εμπιστοσύνης είναι ανοικτό στο πάνω άκρο του, με την οριακή (μη επιτρεπτή) τιμή του πάνω άκρου να υποδηλώνει την απόλυτη (τυφλή) εμπιστοσύνη.

Ένα σημαντικό χαρακτηριστικό του μοντέλου του Marsh, είναι η διάκριση ανάμεσα στην ανυπαρξία εμπιστοσύνης και την έλλειψή της. Οι δύο καταστάσεις αν και μοιάζουν ίδιες, διαφέρουν σημαντικά μεταξύ τους. Ως ανυπαρξία εμπιστοσύνης θεωρείται η κατάσταση κατά την οποία δεν μπορεί να εκτιμηθεί η αξιοπιστία ενός πράκτορα (πιθανότατα λόγω έλλειψης στοιχείων). Η αδυναμία εκτίμησης της εμπιστοσύνης αναπαρίσταται με μια τιμή που βρίσκεται στο μέσο του πεδίου τιμών της εμπιστοσύνης. Ως έλλειψη εμπιστοσύνης θεωρείται η κατάσταση όπου ο πράκτορας αξιολογείται ως αναξιόπιστος, κατόπιν όμως διερεύνησης με επαρκή στοιχεία.

### 2.2.2 Αναπαράσταση της εμπιστοσύνης

Για την αναπαράσταση της εμπιστοσύνης ο Marsh χρησιμοποίησε στο μοντέλο του τους παρακάτω συμβολισμούς :

**Γνώση (Knowledge) :**  $K_x(y)$

Αναπαριστά το εάν ένας πράκτορας  $x$  γνωρίζει έναν άλλο πράκτορα  $y$ . Παίρνει την τιμή '1' (ή 'true') εάν ο  $x$  γνωρίζει τον  $y$  και την τιμή '0' (ή 'false') εάν δεν τον γνωρίζει. Το μοντέλο, επιτρέπει (όπου αυτό είναι αναγκαίο) και τη χρήση ενδιάμεσων τιμών στο διάστημα  $[0,1]$  για την αναπαράσταση διαφόρων βαθμίδων γνώσης (λίγο, αρκετά, πολύ κλπ).



**Βασική εμπιστοσύνη (Basic trust) :**  $T_x \in [-1,+1]$ 

Αναπαριστά τη γενική τάση-προδιάθεση που έχει ένας πράκτορας  $x$  να εμπιστεύεται τους άλλους (trusting disposition). Συνήθως επηρεάζεται και διαμορφώνεται από το σύνολο των εμπειριών που έχει αποκτήσει ο πράκτορας στη διάρκεια της ζωής του. Όσο υψηλότερη είναι η τιμή της βασικής εμπιστοσύνης, τόσο μεγαλύτερη είναι τάση που έχει ο πράκτορας να εμπιστεύεται τους άλλους.

**Γενική εμπιστοσύνη (General trust) :**  $T_x(y) \in [-1,+1]$ 

Αναπαριστά την εμπιστοσύνη που έχει ένας πράκτορας  $x$  προς έναν άλλο πράκτορα  $y$ . Δεν σχετίζεται με κάποια συγκεκριμένη κατάσταση αλλά συμβολίζει τη γενικότερη εμπιστοσύνη του  $x$  προς τον  $y$ . Η παράμετρος αυτή εκφράζει την προσδοκία που έχει ο  $x$  ότι ο  $y$  θα συμπεριφερθεί σύμφωνα με τα ενδιαφέροντά του (του  $x$ ) και δεν θα τον βλάψει.

**Εμπιστοσύνη περίπτωσης (Situational trust) :**  $T_x(y, \alpha) \in [-1,+1]$ 

Αναπαριστά την εμπιστοσύνη που έχει ένας πράκτορας  $x$  προς έναν άλλο πράκτορα  $y$ , αλλά για μια συγκεκριμένη περίπτωση  $\alpha$ . Ουσιαστικά, πρόκειται για επέκταση του ορισμού της γενικής εμπιστοσύνης ώστε να λαμβάνονται υπόψη τα ιδιαίτερα χαρακτηριστικά της εκάστοτε αλληλεπίδρασης, αφού τα κριτήρια με τα οποία γίνεται η αξιολόγηση της εμπιστοσύνης διαφοροποιούνται ανάλογα με την περίπτωση.

**Ωφελιμότητα (Utility) :**  $U_x(\alpha) \in [-1,+1]$ 

Αναπαριστά την ωφέλεια που έχει ένας πράκτορας  $x$  από την έκβαση μιας περίπτωσης  $\alpha$ . Ο Marsh υιοθετεί τις αρχές της θεωρίας αναμενόμενης χρησιμότητας [26], σύμφωνα με την οποία, η εκτίμηση της γενικής ωφέλειας από μια συγκεκριμένη περίπτωση θα πρέπει να λαμβάνει υπόψη και να συνυπολογίζει την ωφέλεια που αποφέρουν όλες οι πιθανές, διαφορετικές εκβάσεις αυτής της περίπτωσης.

**Σπουδαιότητα (Importance) :**  $I_x(\alpha) \in [0,+1]$ 

Αναπαριστά την υποκειμενική εκτίμηση του κέρδους που θα έχει ένας πράκτορας  $x$  από την έκβαση μιας περίπτωσης  $\alpha$ . Αν και μοιάζει με την ωφελιμότητα, διαφέρει στο γεγονός ότι η ωφελιμότητα είναι ένας αντικειμενικός δείκτης που βασίζεται στην κατά το δυνατόν αντικειμενική αξιολόγηση του αποτελέσματος κάθε πιθανής έκβασης μιας περίπτωσης. Ο δείκτης αυτός είναι πάντα ίδιος, ανεξάρτητα από το πότε γίνεται η αξιολόγηση. Από την άλλη, ο δείκτης της σπουδαιότητας ενσωματώνει μέσα του την υποκειμενικότητα της κρίσης του πράκτορα αλλά και τις ιδιαίτερες συνθήκες εμφάνισης της περίπτωσης. Έτσι, μπορεί να προκύψει, για την ίδια

περίσταση, διαφορετική εκτίμηση της σπουδαιότητας, ανάλογα με το ποιος κάνει την εκτίμηση και σε ποια χρονική περίοδο<sup>4</sup>.

Όλοι οι παραπάνω συμβολισμοί μπορούν να εκφράσουν και τη χρονική διαφοροποίηση των χαρακτηριστικών που αναπαριστούν εάν εμπλουτιστούν με τη χρήση ενός χρονικού εκθέτη  $t$ , που υποδηλώνει τη χρονική στιγμή για την οποία αναπαρίσταται το κάθε χαρακτηριστικό. Έτσι προκύπτει η "ενισχυμένη" έκδοσή τους :

$K_x(y)^t$ :	Η γνώση του $x$ για τον $y$ τη χρονική στιγμή $t$
$T_x^t$ :	Η γενική διάθεση εμπιστοσύνης του $x$ τη χρονική στιγμή $t$
$T_x(y)^t$ :	Η εμπιστοσύνη του $x$ προς τον $y$ τη χρονική στιγμή $t$
$T_x(y, \alpha)^t$ :	Η εμπιστοσύνη του $x$ προς τον $y$ , για την περίπτωση $\alpha$ , τη χρονική στιγμή $t$
$I_x(\alpha)^t$ :	Η σπουδαιότητα της περίπτωσης $\alpha$ τη χρονική στιγμή $t$

### 2.2.3 Υπολογισμός της εμπιστοσύνης

Από τα τρία είδη εμπιστοσύνης που διακρίνει ο Marsh, το σημαντικότερο είναι η εμπιστοσύνη περίπτωσης, αφού αυτή είναι που επηρεάζει, κατά κύριο λόγο, τη διάθεση των πρακτόρων για συνεργασία. Ο τύπος που προτείνει για τον υπολογισμό της εμπιστοσύνης περίπτωσης είναι :

$$T_x(y, a) = U_x(a) \times I_x(a) \times \overline{T_x(y)} \quad (1)$$

όπου

$\overline{T_x(y)}$  : Η εκτίμηση της γενικής εμπιστοσύνης του πράκτορα  $x$  προς τον πράκτορα  $y$ .

Για τις παραμέτρους της ωφελιμότητας και της σπουδαιότητας ο Marsh δεν προτείνει κάποιο συγκεκριμένο τρόπο υπολογισμού. Αφήνει να εννοηθεί ότι είναι υποκειμενικές εκτιμήσεις του πράκτορα, που εξαρτώνται από τις πεποιθήσεις και επιδιώξεις του.

Για την εκτίμηση της γενικής εμπιστοσύνης, το μοντέλο προτείνει ότι θα πρέπει να λαμβάνονται υπόψη προηγούμενες αξιολογήσεις της εμπιστοσύνης. Οι αξιολογήσεις αυτές αφορούν είτε τις περιστάσεις που είναι παρόμοιες<sup>5</sup> με αυτή που εξετάζεται, είτε όλες τις προηγούμενες περιστάσεις

<sup>4</sup> Για παράδειγμα, το να στοιχηματίσεις σε ποδοσφαιρική νίκη της Ελλάδας επί της Βραζιλίας σε αγώνα του Mundial, έχει πολύ μεγάλη ωφελιμότητα (αφού θα αποφέρει μεγάλο ποσό κερδών), αλλά πολύ μικρή σπουδαιότητα, αφού είναι εξαιρετικά απίθανο ένα τέτοιο ενδεχόμενο. Εάν όμως, με κάποιο τρόπο, γνωρίζεις ότι το αποτέλεσμα θα είναι τελικά νίκη της Ελλάδας (π.χ. επειδή έχεις πληροφορίες ότι το παιχνίδι είναι «στημένο»), το να στοιχηματίσεις αποκτά τρομερή σπουδαιότητα.

<sup>5</sup> Ο Marsh υιοθετεί τη δουλειά των Dechter & Michie [27] αναφορικά με το πότε δύο περιστάσεις είναι παρόμοιες.

μεταξύ  $x$  και  $y$ . Τρεις είναι οι βασικές επιλογές που υποδεικνύει το μοντέλο για το πως μπορούν να χρησιμοποιηθούν αυτές οι αξιολογήσεις :

α) Αισιόδοξη προσέγγιση

Λαμβάνεται υπόψη η μεγαλύτερη από τις προηγούμενες αξιολογήσεις της εμπιστοσύνης. Αντιστοιχεί στη στάση ζωής των αισιόδοξων ανθρώπων που πάντα αναζητούν και συγκρατούν τις καλύτερες στιγμές των άλλων προσδοκώντας ότι θα επαναληφθούν.

β) Απαισιόδοξη προσέγγιση

Λαμβάνεται υπόψη η μικρότερη από τις προηγούμενες αξιολογήσεις της εμπιστοσύνης. Αντιστοιχεί στη στάση ζωής των απαισιόδοξων ανθρώπων που πάντα περιμένουν τα χειρότερα.

γ) Ρεαλιστική προσέγγιση

Λαμβάνεται υπόψη η μέση τιμή των προηγούμενων αξιολογήσεων της εμπιστοσύνης. Δηλ.

$$\overline{T_x(y)} = \frac{1}{|A|} \sum_{a \in A} T_x(y)$$

όπου

$A$  : το σύνολο των περιστάσεων που είναι παρόμοιες με την υπό διερεύνηση περίπτωση.

Η προσέγγιση αυτή αντιστοιχεί στη αντίληψη που έχουν οι περισσότεροι άνθρωποι, ότι τα πάντα κινούνται γύρω από ένα μέσο όρο.

Ανεξάρτητα από την προσέγγιση που θα ακολουθηθεί, το πλήθος των προηγούμενων αξιολογήσεων που θα χρησιμοποιηθούν περιορίζεται από το μέγεθος της μνήμης του πράκτορα. Η χρήση μεγαλύτερης μνήμης, αν και θεωρητικά είναι επιθυμητή, στην πράξη αποδεικνύεται ότι, ενώ αυξάνει δραματικά την υπολογιστική πολυπλοκότητα, προσφέρει μικρά μόνο οφέλη στην τελική εκτίμηση της γενικής εμπιστοσύνης.

Η τεκμηρίωση του τύπου (1) βασίζεται στην απλή παρατήρηση ότι, ένας πράκτορας εμπλέκεται σε μια αλληλεπίδραση όταν ισχύει ένα ή περισσότερα από τα παρακάτω :

- α) με σκοπό την αύξηση του προσωπικού του οφέλους
- β) εφόσον διαπιστώνει κάποια σπουδαιότητα στη συναλλαγή
- γ) εφόσον η «άλλη» πλευρά είναι επαρκώς αξιόπιστη

Έτσι μια συναλλαγή με μικρή σπουδαιότητα και μικρό όφελος θα έχει πολύ μικρότερες πιθανότητες να επιλεγεί, απ' ό,τι μια συναλλαγή με μεγάλη σπουδαιότητα και μεγάλο όφελος. Φυσικά, σε κάθε περίπτωση σημαντικό ρόλο παίζει η γενική εμπιστοσύνη προς τον πράκτορα με τον οποίο πρόκειται να γίνει η συναλλαγή.

Για την ενημέρωση της γενικής εμπιστοσύνης  $T_x(y)$ , ο Marsh προτείνει έναν ad-hoc υπολογισμό που βασίζεται στη συμπεριφορά του  $y$  στην τελευταία αλληλεπίδρασή του με τον  $x$ . Έτσι η

εμπιστοσύνη του  $x$  προς τον  $y$  αυξάνεται εάν ο  $y$  συνεργάστηκε και μειώνεται σε αντίθετη περίπτωση. Το ύψος της αύξησης ή της μείωσης δεν είναι πάντα το ίδιο, αλλά εξαρτάται από το κόστος ή το όφελος της εκάστοτε περίπτωσης  $a$ . Συγκεκριμένα, προτείνεται ο παρακάτω απλός κανόνας :

α) Εάν ο  $y$  συνεργάστηκε, η εμπιστοσύνη του  $x$  προς τον  $y$  θα αυξηθεί

$$\text{Cooperated}(y)^a \rightarrow T_x(y)^{a+1} = T_x(y)^a + C$$

β) Εάν ο  $y$  δεν συνεργάστηκε, η εμπιστοσύνη του  $x$  προς τον  $y$  θα μειωθεί

$$\text{Defected}(y)^a \rightarrow T_x(y)^{a+1} = T_x(y)^a - C$$

όπου  $C$  : το κόστος (ή το όφελος) της περίπτωσης  $a$

Ο Marsh, επιπλέον, δίνει και μια άλλη παραλλαγή του παραπάνω κανόνα ενημέρωσης, στην οποία λαμβάνεται υπόψη και η ανταποδοτικότητα (reciprocity), Συγκεκριμένα :

α) Εάν ο  $x$  είχε βοηθήσει στο παρελθόν τον  $y$  και ο  $y$  τώρα συνεργάστηκε, η εμπιστοσύνη του  $x$  προς τον  $y$  αυξάνεται ή παραμένει ίδια

$$\text{Helped}(x,y,\alpha)^{t-\delta} \wedge \text{Cooperated}(y, \beta)^t \rightarrow T_x(y)^{t+1} \geq T_x(y)^t$$

β) Εάν ο  $x$  είχε βοηθήσει στο παρελθόν τον  $y$  και ο  $y$  τώρα δεν συνεργάστηκε, η εμπιστοσύνη του  $x$  προς τον  $y$  μειώνεται σημαντικά

$$\text{Helped}(x,y,\alpha)^{t-\delta} \wedge \text{Defected}(y,\alpha)^t \rightarrow T_x(y)^{t+1} \ll T_x(y)^t$$

## 2.2.4 Το κατώφλι συνεργασίας

Προκειμένου να αποφασίσει ένας πράκτορας  $x$  αν θα αλληλεπιδράσει με έναν πράκτορα  $y$  λαμβάνει χώρα η ακόλουθη διαδικασία (η οποία υιοθετείται σχεδόν από όλα τα μοντέλα εμπιστοσύνης στη βιβλιογραφία):

- Ο  $x$  υπολογίζει την εμπιστοσύνη του προς τον  $y$  (λαμβάνοντας υπόψη την συγκεκριμένη περίπτωση) και
- συγκρίνει την τιμή της εμπιστοσύνης του προς τον  $y$  με την τιμή κατωφλίου συνεργασίας (cooperation threshold) που διαθέτει.

Εάν η εμπιστοσύνη του  $x$  προς τον  $y$  είναι μεγαλύτερη από το κατώφλι συνεργασίας, ο  $x$  θα προχωρήσει στην αλληλεπίδραση με τον  $y$ , διαφορετικά θα την αποφύγει.

Είναι προφανές, πως η επιλογή κατάλληλης τιμής για το κατώφλι συνεργασίας είναι πολύ σημαντική. Ο Marsh προτείνει τον υπολογισμό του κατωφλίου συνεργασίας σύμφωνα με τον τύπο :

$$\text{Cooperation\_Threshold}_x(a) = \frac{\text{Perceived\_Risk}_x(a)}{\text{Perceived\_Competence}_x(y,a)} \times I_x(a) \quad (2)$$

όπου

$Perceived\_Risk_x(a)$ : Ένα μέτρο του πώς αντιλαμβάνεται ο  $x$  την επικινδυνότητα της συγκεκριμένης περίπτωσης  $a$ , με τιμές στο διάστημα  $[0, \infty)$ .

$Perceived\_Competence_x(y, a)$ : Ένα μέτρο του πώς αντιλαμβάνεται ο  $x$  την ικανότητα του  $y$  στη συγκεκριμένη περίπτωση  $a$ , με τιμές στο διάστημα  $[0, \infty)$ .

Η αιτιολόγηση του τύπου (2) πηγάζει από το λογικό συμπέρασμα ότι, το κατώφλι συνεργασίας θα πρέπει να είναι :

- ανάλογο της επικινδυνότητας και της σπουδαιότητας μιας συναλλαγής
- αντιστρόφως ανάλογο της ικανότητας του πράκτορα  $y$  με τον οποίο θα γίνει η συναλλαγή

Για την επικινδυνότητα, ο Marsh δεν δίνει κάποιο συγκεκριμένο τρόπο αξιολόγησης. Αναφέρει απλά ότι και αυτή η παράμετρος είναι μια υποκειμενική εκτίμηση του πράκτορα αξιολογητή, η οποία εξαρτάται από το κατά πόσο ο πράκτορας εμπιστεύεται τον εαυτό του στην εκτίμηση της επικινδυνότητας των περιπτώσεων. Διακρίνονται τρεις περιπτώσεις :

- α) ο  $x$  δεν έχει κανενός είδους εμπειρία ή γνώση της συγκεκριμένης περίπτωσης  $a$
- β) ο  $x$  έχει ελλιπή εμπειρία ή γνώση της συγκεκριμένης περίπτωσης  $a$
- γ) ο  $x$  έχει σημαντική εμπειρία ή γνώση της συγκεκριμένης περίπτωσης  $a$

Στις δύο πρώτες περιπτώσεις ο  $x$  μπορεί να εκτιμήσει μόνο κατά προσέγγιση (και με μεγάλη πιθανότητα λάθους) την επικινδυνότητα, λαμβάνοντας ίσως υπόψη τις επιδόσεις του σε προηγούμενες εκτιμήσεις άλλων άγνωστων περιπτώσεων. Στην τρίτη περίπτωση, όπου υπάρχει κάποια γνώση, προτείνεται η χρήση κάποιας θεωρίας εκτίμησης κινδύνου (όπως π.χ. η Bayesian θεωρία [26]), που μπορεί να δώσει μια χρησιμοποιήσιμη μετρική. Με οποιοδήποτε τρόπο και αν προκύψει η τιμή της αντιλαμβανόμενης επικινδυνότητας, αυτή θα πρέπει να κανονικοποιηθεί στο διάστημα  $[0, \infty)$ .

Για την εκτίμηση της ικανότητας, τα πράγματα είναι λίγο πιο ξεκάθαρα. Και εδώ, διακρίνονται τρεις περιπτώσεις :

- α) Ο  $x$  δεν γνωρίζει τίποτε για τον  $y$ . Στην περίπτωση αυτή, το μόνο που μπορεί να χρησιμοποιηθεί είναι η γενική διάθεση εμπιστοσύνης του  $x$ , λαμβάνοντας όμως υπόψη και τη σπουδαιότητα της υπό εξέταση περίπτωσης. Ο τύπος που προτείνεται είναι :

$$Perceived\_Competence_x(y, a) = T_x \times I_x(a)$$

- β) Ο  $x$  γνωρίζει τον  $y$  (έχουν προηγούμενες αλληλεπιδράσεις μεταξύ τους), αλλά όχι σε παρόμοια περίπτωση. Στην περίπτωση αυτή, υπάρχουν στοιχεία, από άλλες περιπτώσεις, που μπορούν να χρησιμοποιηθούν για την αξιολόγηση της ικανότητας. Ο τύπος που προτείνεται είναι :

$$Perceived\_Competence_x(y, a) = \frac{1}{|B|} \sum_{\beta \in B} Experienced\_Competence_x(y, \beta) \times \overline{T_x(y)}$$

όπου

$B$  : το σύνολο των περιστάσεων που είναι διαφορετικές από την  $a$ .

$\overline{T_x(y)}$  : όπως στην παράγραφο **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**

β) Ο  $x$  γνωρίζει τον  $y$  και τον εμπιστεύεται. Στην περίπτωση αυτή, υπάρχουν στοιχεία και από παρόμοιες περιστάσεις. Ο τύπος που προτείνεται είναι παραπλήσιος μ' αυτόν της περίπτωσης β), αλλά με εφαρμογή στο σύνολο των παρόμοιων περιστάσεων :

$$Perceived\_Competence_x(y, a) = \frac{1}{|A|} \sum_{\beta \in A} Experienced\_Competence_x(y, \beta) \times \overline{T_x(y)}$$

όπου

$A$  : το σύνολο των περιστάσεων που είναι παρόμοιες με την  $a$ .

Επιστρέφοντας στον υπολογισμό του κατώφλιου συνεργασίας, ο Marsh κάνει τη λογική παρατήρηση ότι, το κατώφλι συνεργασίας θα πρέπει, επιπλέον, να είναι αντιστρόφως ανάλογο της εμπιστοσύνης που έχει ο  $x$  στον  $y$ . Δηλαδή όσο λιγότερο εμπιστεύεται ο  $x$  τον  $y$  τόσο υψηλότερο θα πρέπει να είναι το κατώφλι. Αυτό επιτυγχάνεται με τον παρακάτω πιο σύνθετο τύπο :

$$Cooperation\_Threshold_x(a) = \frac{Perceived\_Risk_x(a)}{Perceived\_Competence_x(y, a) + \overline{T_x(y)}} \times I_x(a) \quad (3)$$

Δυστυχώς ο τύπος (3), παρουσιάζει αρκετά προβλήματα, εξαιτίας των αρνητικών τιμών που μπορεί να πάρει η εκτίμηση εμπιστοσύνης  $\overline{T_x(y)}$ . Έτσι, στην περίπτωση που η εκτίμηση εμπιστοσύνης είναι αρνητική και ίση σε μέγεθος με την αντιλαμβανόμενη ικανότητα ( $Perceived\_Competence_x(y, a) = -\overline{T_x(y)}$ ), ο παρονομαστής μηδενίζεται και το κατώφλι τείνει στο άπειρο, κατάσταση μη αποδεκτή. Επίσης, εάν η εκτίμηση εμπιστοσύνης είναι αρνητική και μεγαλύτερη σε μέγεθος από την αντιλαμβανόμενη ικανότητα θα προκύψει αρνητική τιμή κατώφλιου, γεγονός που έρχεται σε αντίθεση με την απαίτηση το κατώφλι να μεγαλώνει όσο η εμπιστοσύνη μικραίνει. Τα προβλήματα αυτά επισημαίνονται και από τον ίδιο τον Marsh, ο οποίος τελικά παραπέμπει σε μελλοντική εργασία την ορθότερη χρήση της εμπιστοσύνης στον υπολογισμό του κατώφλιου συνεργασίας.

Τέλος, στο μοντέλο παρουσιάζεται και μια ακόμη προσέγγιση για τον υπολογισμό του κατώφλιου συνεργασίας. Σύμφωνα με την προσέγγιση αυτή όσο σημαντικότερη είναι η εκτέλεση μιας συναλλαγής, τόσο χαμηλότερο θα πρέπει να είναι το κατώφλι συνεργασίας, ώστε να την καθιστά πιο εφικτή. Στην περίπτωση αυτή, ο παράγοντας της σπουδαιότητας μετακινείται στον παρονομαστή και ο τύπος γίνεται :

$$Cooperation\_Threshold_x(a) = \frac{Perceived\_Risk_x(a)}{(Perceived\_Competence_x(y,a) + T_x(y)) \times I_x(a)} \quad (4)$$

Εδώ, παρατηρείται το παράδοξο, κατά το οποίο για την ίδια σπουδαιότητα μιας περίπτωσης  $a$ , ο τύπος (4) δίνει μεγαλύτερη τιμή καταφλίου απ' ό τι ο τύπος (3).

### 2.2.5 Δομές

Το μοντέλο που προτείνει ο Marsh είναι θεωρητικό, δηλαδή δεν προβλέπει συγκεκριμένες δομές δεδομένων για ενδεχόμενη υπολογιστική υλοποίησή του. Από την ανάλυση των υπολογισμών που περιλαμβάνει το μοντέλο συμπεράναμε ότι κάθε πράκτορας χρειάζεται να διατηρεί τις παρακάτω δομές:

- Έναν πίνακα εμπιστοσύνης  $T$ , δύο διαστάσεων. Κάθε στοιχείο  $T[i,j]$  του πίνακα αντιπροσωπεύει την τρέχουσα εμπιστοσύνη του πράκτορα προς έναν πράκτορα  $i$  αναφορικά με μια περίπτωση  $j$ .
- Έναν πίνακα ιστορικού εμπιστοσύνης  $H_T$ , δύο διαστάσεων. Κάθε στοιχείο  $H_T[i,j]$  του πίνακα είναι μια λίστα με τις προηγούμενες τιμές εμπιστοσύνης του πράκτορα προς έναν πράκτορα  $i$  αναφορικά με μια περίπτωση  $j$ .
- Έναν πίνακα ιστορικού εμπειριών αλληλεπίδρασης  $H_I$  με στοιχεία τις αξιολογήσεις της ικανότητας των πρακτόρων με τους οποίους έλαβαν χώρα οι αλληλεπιδράσεις. Ο πίνακας είναι δύο διαστάσεων και κάθε στοιχείο  $H_I[i,j]$  είναι μια λίστα με τις καταγραφείσες ικανότητες ενός πράκτορα  $i$  σε αλληλεπιδράσεις περίπτωσης  $j$ .

### 2.2.6 Σχολιασμός - κριτική

Το μοντέλο του Marsh ήταν πρωτοπόρο γιατί επιχείρησε να μεταφέρει μια δύσκολη και πολυδιάστατη κοινωνιολογική έννοια στον κόσμο των μαθηματικών. Το γεγονός ότι επηρέασε την πλειονότητα των ερευνητών στην συγκεκριμένη επιστημονική κοινότητα είναι ενδεικτικό της σπουδαιότητάς του. Μελετώντας το συμπεράναμε ότι διαθέτει τόσο θετικά όσο και αρνητικά σημεία, τα οποία συνοψίζουμε εδώ:

#### Θετικά

- + Λαμβάνει υπόψη πολλά από τα κοινωνιολογικά χαρακτηριστικά της έννοιας της εμπιστοσύνης (ωφελιμότητα, η σπουδαιότητα και η αμοιβαιότητα κλπ) οπότε η τυπική αναπαράστασή της είναι πολύ ρεαλιστική.
- + Προτείνει ένα τρόπο υπολογισμού της εμπιστοσύνης που λαμβάνει υπόψη τα χαρακτηριστικά της εκάστοτε περίπτωσης.
- + Διακρίνει ανάμεσα στην έλλειψη εμπιστοσύνης (αρνητικές τιμές αναπαράστασης) και την αδυναμία αξιολόγησής της (τιμή αναπαράστασης : 0)

- + Προτείνει συγκεκριμένο τρόπο υπολογισμού του κατωφλίου συνεργασίας, επιτρέποντας την προσαρμογή του στα χαρακτηριστικά της εκάστοτε περίπτωσης

### Αρνητικά

- Στην προσπάθεια για μέγιστο ρεαλισμό, το μοντέλο ενσωματώνει πάρα πολλές μεταβλητές, οι οποίες αυξάνουν δραματικά την πολυπλοκότητα και προσθέτουν εστίες διφορούμενης λογικής.
- Οι υπολογισμοί που προτείνει είναι ad hoc και περιλαμβάνουν αρκετά υποκειμενικά στοιχεία που άπτονται των προσωπικών εκτιμήσεων των πρακτόρων.
- Οι αρνητικές τιμές που προτείνει για την αναπαράσταση της έλλειψης εμπιστοσύνης παράγουν σε αρκετές περιπτώσεις "περίεργα" αποτελέσματα
- Λαμβάνει υπόψη μόνο δύο ενδεχόμενα ως αποτέλεσμα μιας αλληλεπίδρασης: τη συνεργασία και τη μη συνεργασία
- Δεν λαμβάνει υπόψη την φήμη. Έτσι οι πράκτορες δεν μπορούν να επωφεληθούν από τις εμπειρίες και τις συνεπαγόμενες γνώμες άλλων πρακτόρων που ανήκουν στην κοινωνία, οπότε αν και είναι μέλη ενός κοινωνικού συνόλου, η λήψη απόφασής τους διεξάγεται σαν να ήταν απομονωμένοι.

## **2.3 Το μοντέλο PeerTrust**

Οι L.Xiong και L.Liu, ανέπτυξαν, σχετικά πρόσφατα, το 2002, το μοντέλο PeerTrust [28], το οποίο μελετά το πρόβλημα της εμπιστοσύνης μέσα σε κοινότητες ομότιμων μελών (peer-to-peer communities). Κύριος στόχος τους ήταν, μέσω του μοντέλου να επισημανθούν οι διαφορετικοί παράγοντες που εμπλέκονται στην αξιολόγηση της εμπιστοσύνης καθώς και ο τρόπος με τον οποίο αυτοί οι παράγοντες την επηρεάζουν.

### **2.3.1 Παραδοχές- Χαρακτηριστικά**

Το PeerTrust ανήκει στην κατηγορία των μοντέλων που βασίζονται στη φήμη (reputation based). Εστιάζοντας την εφαρμογή του στο χώρο των κοινοτήτων με ομότιμα μέλη (peer-to-peer communities) προτείνει ένα μηχανισμό αξιολόγησης της εμπιστοσύνης που είναι συνάρτηση πέντε διαφορετικών παραγόντων. Ο σημαντικότερος από αυτούς είναι η φήμη που ένα μέλος δημιουργεί καθώς παρέχει υπηρεσίες σε άλλα μέλη της κοινότητας.

Κάθε πράκτορας  $w$  που ξεκινά με δική του πρωτοβουλία μια συναλλαγή με έναν πράκτορα  $u$ , δίνει μετά το πέρας της συναλλαγής μια ανατροφοδότηση (feedback) για το μέγεθος της ικανοποίησης που του προσέφερε η συναλλαγή του με τον  $u$ . Αυτή η ανατροφοδότηση μπορεί να είναι είτε αξιόπιστη



είτε αναξιόπιστη, ανάλογα με το εάν ο  $w$  είναι αξιόπιστος ή όχι. Το μοντέλο θεωρεί ότι οι ανατροφοδοτήσεις συλλέγονται από ένα σύστημα ανατροφοδοτήσεων (feedback system) το οποίο βασίζεται στις συναλλαγές (transaction based), δηλαδή συλλέγει στοιχεία αμέσως μετά από κάθε συναλλαγή και όχι σε βάθος χρόνου.

Οι πέντε κύριοι παράγοντες που το μοντέλο θεωρεί ότι επηρεάζουν τη διαμόρφωση της εμπιστοσύνης είναι :

- **Η ανατροφοδότηση σχετικά με μια συναλλαγή**

Είναι η αξιολόγηση που κάνει το μέλος που ζήτησε μια συναλλαγή, για το πόσο καλά συμπεριφέρθηκε η άλλη πλευρά. Στο μοντέλο η συμπεριφορά ενός μέλους συνδέεται άμεσα με την παροχή ικανοποίησης. Έτσι μια «καλή» συμπεριφορά (που τηρεί δηλ. τα συμφωνηθέντα) θα δώσει υψηλή ικανοποίηση, ενώ αντίθετα μια «κακή» συμπεριφορά θα έχει ως αποτέλεσμα χαμηλή ικανοποίηση. Η κατάθεση-δημοσιοποίηση της αξιολόγησης μιας συναλλαγής από τα εμπλεκόμενα μέρη είναι ένας μηχανισμός ζωτικής σημασίας για τα μοντέλα PeerTrust, αλλά και γενικότερα για τα μοντέλα που βασίζονται στη φήμη.

- **Η αξιοπιστία της ανατροφοδότησης**

Οι ανατροφοδοτήσεις δεν είναι πάντα αληθείς. Κι αυτό γιατί, εκτός από την υποκειμενικότητα που εμπεριέχεται στις περισσότερες διαδικασίες αξιολόγησης, υπάρχει και η πιθανότητα κακόβουλης διάχυσης ψευδών αξιολογήσεων. Είναι απαραίτητο λοιπόν, κάθε μοντέλο που βασίζεται στη φήμη να έχει κάποιο μηχανισμό εκτίμησης της εγκυρότητας της πηγής κάθε ανατροφοδότησης.

- **Το πλήθος των συναλλαγών**

Είναι το πλήθος των συναλλαγών που έχει διενεργήσει ένα μέλος μέσα στην κοινωνία.. Το μοντέλο επισημαίνει ότι σε κοινωνίες με ανομοιόμορφη κατανομή συναλλαγών μεταξύ των μελών τους, η μη χρήση αυτού του δεδομένου μπορεί να οδηγήσει σε λανθασμένη εκτίμηση της αξιοπιστίας των μελών. Για παράδειγμα, μια εκτίμηση της αξιοπιστίας που αρκείται στην απλή άθροιση της ικανοποίησης που προσέφερε ένα μέλος στις συναλλαγές του, χωρίς να λαμβάνει υπόψη το πλήθος των συναλλαγών, θα εμφανίσει κάποιον, που έχει 100 «καλές» συναλλαγές και 20 «κακές», ως περισσότερο αξιόπιστο από κάποιον άλλο που έχει μόνο 50 «καλές» συναλλαγές, έστω και αν ο δεύτερος συμπεριφέρθηκε άψογα σε όλες του τις συναλλαγές. Συνήθως ο παράγοντας του πλήθους των συναλλαγών αξιοποιείται στον υπολογισμό του μέσου όρου της ικανοποίησης που προσέφερε ένα μέλος στο σύνολο των συναλλαγών του.

- **Ο παράγοντας πλαισίου συναλλαγής (transaction context factor)**

Ο παράγοντας αυτός περιλαμβάνει τα ιδιαίτερα χαρακτηριστικά της κάθε συναλλαγής και επιτρέπει σε στοιχεία όπως το μέγεθος, η σπουδαιότητα, η κρισιμότητα ή η επικινδυνότητα

μιας συναλλαγής να συμμετέχουν στην αξιολόγηση της εμπιστοσύνης. Σύμφωνα με το μοντέλο, μπορεί να λειτουργήσει ως ένας σημαντικός αμυντικός μηχανισμός ενάντια σε κακόβουλες συμπεριφορές κατά τις οποίες ένα μέλος συμπεριφέρεται «καλά» στις συναλλαγές με μικρή αξία ή σπουδαιότητα, ώστε να κτίσει μια εικόνα αξιοπιστίας και αργότερα, στις συναλλαγές υψηλής αξίας, να μπορέσει να εξαπατήσει.

- **Ο παράγοντας πλαισίου κοινότητας (community context factor)**

Ο παράγοντας αυτός, κατ' αντιστοιχία με τον προηγούμενο παράγοντα, περιλαμβάνει τα ιδιαίτερα χαρακτηριστικά της κοινότητας. Δίνει τη δυνατότητα στο μηχανισμό αξιολόγησης της εμπιστοσύνης να προσαρμόζεται στις ιδιαιτερότητες της εκάστοτε κοινότητας ενισχύοντας ή αποτρέποντας συμπεριφορές ανάλογα με τις απαιτήσεις και τα πρότυπα που έχουν υιοθετηθεί. Για παράδειγμα σε μια κοινότητα διαμοίρασης πόρων, ο παράγοντας αυτός θα μπορούσε να αντιπροσωπεύει ένα μέτρο της συνεισφοράς (με τη μορφή της παροχής κοινόχρηστων πόρων) του κάθε μέλους προς την κοινότητα. Έτσι, δίνοντας μεγαλύτερη τιμή εμπιστοσύνης στα μέλη με τις περισσότερες παροχές πόρων, η κοινότητα μπορεί να δημιουργήσει ένα κίνητρο για την προσφορά πόρων, που είναι παράμετρος ιδιαίτερα σημαντική για την βιωσιμότητα και την ευημερία της.

## 2.3.2 Αναπαράσταση της εμπιστοσύνης

Οι Xiong και Liu διατύπωσαν το μοντέλο τους με πολύ γενικό τρόπο, χωρίς να υιοθετούν κάποια συγκεκριμένη αναπαράσταση για την εμπιστοσύνη. Έτσι είναι δυνατές πολλές διαφορετικές υλοποιήσεις. Ωστόσο, όπως προκύπτει από την εξέταση του τύπου για τον υπολογισμό της εμπιστοσύνης, θα πρέπει οι τιμές που λαμβάνει αυτή να είναι στο διάστημα [0,1].

## 2.3.3 Υπολογισμός της εμπιστοσύνης

Η εμπιστοσύνη ενός πράκτορα προς έναν άλλο πράκτορα  $u$ , υπολογίζεται σύμφωνα με τον τύπο:

$$T(u) = a \times \frac{\sum_{i=1}^{I(u)} S(u,i) \times Cr(p(u,i)) \times TF(u,i)}{I(u)} + \beta \times CF(u) \quad (2.3.1)$$

όπου

- $I(u)$  : Το πλήθος των συναλλαγών που έχει πραγματοποιήσει το μέλος  $u$
- $p(u,i)$  : Το μέλος με το οποίο αλληλεπιδρά το μέλος  $u$  στην  $i$ -οστή συναλλαγή του
- $S(u,i)$  : Το κανονικοποιημένο μέγεθος της ικανοποίησης που προσφέρει το μέλος  $u$  κατά την  $i$ -οστή συναλλαγή του (με το μέλος  $p(u,i)$ )
- $Cr(p(u,i))$  : Η αξιοπιστία της ανατροφοδότησης που παρέχεται από το μέλος  $p(u,i)$  σχετικά με την ικανοποίηση που του πρόσφερε το μέλος  $u$  στην  $i$ -οστή συναλλαγή του
- $TF(u,i)$  : Ο παράγοντας πλαισίου συναλλαγής για την  $i$ -οστή συναλλαγή του μέλους  $u$
- $CF(u,i)$  : Ο παράγοντας πλαισίου κοινότητας του μέλους  $u$

Ο υπολογισμός της εμπιστοσύνης αποτελείται από δύο μέρη :

- Το πρώτο μέρος είναι ο μέσος όρος της ικανοποίησης που προσέφερε το αξιολογούμενο μέλος  $u$  κατά τις συναλλαγές του, όπως αυτή καταγράφηκε μέσα από αξιόπιστες ανατροφοδοτήσεις. Από μια άποψη, αυτός ο παράγοντας, που βασίζεται στο ιστορικό του  $u$ , αποτελεί μια πρόβλεψη για την πιθανότητα που υπάρχει για μια ικανοποιητική συναλλαγή του  $u$  στο μέλλον.
- Το δεύτερο μέρος, που είναι ο παράγοντας του πλαισίου της κοινότητας, δίνει τη δυνατότητα να ενσωματωθούν στην αξιολόγηση κάποια από τα ιδιαίτερα χαρακτηριστικά της κοινότητας

Οι συντελεστές  $\alpha$  και  $\beta$  ρυθμίζουν τη βαρύτητα με την οποία κάθε μέρος συμμετέχει στον υπολογισμό και καθορίζονται με βάση τις απαιτήσεις του εκάστοτε πεδίου εφαρμογής του μοντέλου.

Ο υπολογισμός (2.3.1) είναι αρκετά γενικός ώστε να καλύπτει μια μεγάλη γκάμα αρχιτεκτονικών επιλογών για την αξιολόγηση της εμπιστοσύνης. Περικλείει και τους πέντε παράγοντες διαμόρφωσης της εμπιστοσύνης, επιτρέποντας την επιλεκτική ενεργοποίηση ή απενεργοποίησή τους, ανάλογα με τα χαρακτηριστικά και τις απαιτήσεις της κάθε κοινότητας.

Μερικές από τις διαφορετικές μορφές που μπορεί να πάρει αυτός ο υπολογισμός, έτσι ώστε να συνεισφέρει στην επίλυση γνωστών προβλημάτων ακολουθούν:

- **Η βασική μορφή**

Στην βασική έκδοση του υπολογισμού απενεργοποιούνται οι παράγοντες πλαισίου συναλλαγής ( $TF(u,i)=1$ ) και πλαισίου κοινότητας ( $\alpha=1$  και  $\beta=0$ ) και λαμβάνονται υπόψη μόνο το πλήθος των συναλλαγών ( $I(u)$ ), οι ανατροφοδοτήσεις τους ( $S(u,i)$ ) και η αξιοπιστία των πηγών ανατροφοδότησης ( $Cr(p(u,i))$ ). Έτσι έχουμε :

$$T(u) = \frac{\sum_{i=1}^{I(u)} S(u,i) \times Cr(p(u,i))}{I(u)}$$

Αυτή η μορφή υπολογισμού της εμπιστοσύνης υπάρχει (αυτούσια ή με παραλλαγές) στα περισσότερα μοντέλα εμπιστοσύνης που βασίζονται στη φήμη.

Το μοντέλο αφήνει ανοικτή την επιλογή του τύπου δεδομένων της ανατροφοδότησης. Μπορεί να ποικίλει μεταξύ θετικών, αρνητικών ή και μικτών τιμών, όμως ο όρος  $S(u,i)$  αντιπροσωπεύει την κανονικοποιημένη τιμή της ικανοποίησης (όπως αυτή προκύπτει μετά από κατάλληλη επεξεργασία της ανατροφοδότησης) και κυμαίνεται μεταξύ 0 και 1.

Η αξιοπιστία της πηγής μιας ανατροφοδότησης είναι ένα χαρακτηριστικό που δύσκολα υπολογίζεται με αντικειμενικό τρόπο. Οι Xiong και Liu, αναγνωρίζοντας αυτή τη δυσκολία και

θεωρώντας ότι το πρόβλημα αυτό, από μόνο του, αποτελεί μια σημαντική ερευνητική περιοχή, υιοθετούν μια σχετικά απλοϊκή προσέγγιση, θεωρώντας ότι η αξιοπιστία μιας ανατροφοδότησης μπορεί να ταυτιστεί με την τιμή εμπιστοσύνης της πηγής της. Βέβαια η παραδοχή αυτή έχει αρκετές αδυναμίες, γεγονός που αναγνωρίζουν και οι ίδιοι οι συγγραφείς.

- **Με ενσωμάτωση του πλαισίου συναλλαγής**

Το πλαίσιο συναλλαγής, κατά το μοντέλο, θα πρέπει να εκφράζει κάποια από τα σημαντικότερα χαρακτηριστικά των συναλλαγών που λαμβάνουν χώρα στην κοινότητα. Τέτοια χαρακτηριστικά, όπως αναφέρθηκε και προηγουμένως, θα μπορούσαν να είναι η αξία της συναλλαγής σε χρήματα, εάν μιλάμε για κοινότητα ηλεκτρονικού εμπορίου, ή η διαβάθμιση (π.χ. απόρρητο) των ανταλλασσόμενων δεδομένων, σε περιπτώσεις στρατιωτικών ή ιατρικών εφαρμογών.

Έτσι εάν στη βασική μορφή προσθέσουμε και τον παράγοντα πλαισίου συναλλαγής, με πεδίο εφαρμογής το ηλεκτρονικό εμπόριο και αξιολογούμενο χαρακτηριστικό την αξία της συναλλαγής (χρήματα), ο υπολογισμός διαμορφώνεται ως εξής :

$$T(u) = \frac{\sum_{i=1}^{I(u)} S(u,i) \times Cr(p(u,i)) \times D(u,i)}{I(u)}$$

όπου

$D(u,i)$  : το μέτρο της αξίας της  $i$ -οστής συναλλαγής του μέλους  $u$

Το μοντέλο δεν περιγράφει το πεδίο τιμών του όρου  $D(u,i)$ . Προφανώς όμως, εάν θεωρήσουμε ότι η εμπιστοσύνη παίρνει τιμές στην περιοχή  $[0,1]$ , θα πρέπει και ο όρος  $D(u,i)$  να παίρνει επίσης τιμές στην περιοχή  $[0,1]$ .

- **Με ενσωμάτωση του πλαισίου κοινότητας**

Μέσα στο πλαίσιο κοινότητας, όπως έχουμε πει, μπορούν να ενσωματωθούν στοιχεία που πηγάζουν από τα ιδιαίτερα χαρακτηριστικά και τον προσανατολισμό της κοινότητας. Τέτοια στοιχεία μπορεί να είναι :

- **Η χρονική προσαρμογή :**

Σε πολλές κοινότητες οι παλαιότερες συναλλαγές δεν έχουν την ίδια σπουδαιότητα με τις πιο πρόσφατες. Εάν θεωρήσουμε ότι το πλήθος των παλαιότερων συναλλαγών είναι  $I_h(u)$ , ο υπολογισμός της εμπιστοσύνης διαμορφώνεται ως εξής :

$$T(u) = a \times \frac{\sum_{i=1}^{I(u)} S(u,i) \times Cr(p(u,i))}{I(u)} + \beta \times \frac{\sum_{i=1}^{I_h(u)} S(u,i) \times Cr(p(u,i))}{I_h(u)}$$

με τα  $a$  και  $\beta$  να ρυθμίζουν την αναλογία στη συμμετοχή των παλαιών και των πρόσφατων συναλλαγών.

Το όριο που καθορίζει από ποια χρονική στιγμή και μετά μια συναλλαγή θα θεωρείται παλαιά εξαρτάται από την συγκεκριμένη κοινότητα.

#### – Το κίνητρο παροχής ανατροφοδότησης

Το πρόβλημα της έλλειψης στοιχείων αξιολόγησης των συναλλαγών μπορεί να αντιμετωπισθεί, εάν στον υπολογισμό της εμπιστοσύνης συμπεριληφθεί και το πλήθος των ανατροφοδοτήσεων  $F(u)$ , που δίνει ένα μέλος στο σύστημα ανατροφοδότησης. Στην περίπτωση αυτή ο υπολογισμός θα πάρει τη μορφή :

$$T(u) = a \times \frac{\sum_{i=1}^{J(u)} S(u,i) \times Cr(p(u,i))}{I(u)} + \beta \times \frac{F(u)}{I(u)}$$

Με τη ρύθμιση αυτή, τα μέλη της κοινότητας έχουν ένα σημαντικό κίνητρο ώστε να δημοσιοποιούν την αξιολόγησή τους για την άλλη πλευρά σε κάθε συναλλαγή που κάνουν, και όχι μόνο για την έκφραση παραπόνων σε προβληματικές συναλλαγές.

#### – Το κίνητρο προσφοράς προς την κοινότητα

Ένα άλλο σημαντικό πρόβλημα που αντιμετωπίζουν οι κοινωνίες ομότιμων μελών με χαρακτηριστικά διαμοιρασμού πόρων ή παροχής υπηρεσιών είναι οι συμπεριφορές που έχουν χαρακτηριστεί ως "free riding", δηλ. όταν τα μέλη αρκούνται στο να καταναλώνουν και όχι να προσφέρουν. Το φαινόμενο μπορεί να αντιμετωπισθεί ενσωματώνοντας στον τύπο το πλήθος των συναλλαγών  $U(u)$  κατά τις οποίες ένα μέλος προσέφερε πόρους ή υπηρεσίες σε άλλα μέλη της κοινότητας. Στην περίπτωση αυτή ο υπολογισμός θα πάρει τη μορφή :

$$T(u) = a \times \frac{\sum_{i=1}^{J(u)} S(u,i) \times Cr(p(u,i))}{I(u)} + \beta \times \frac{U(u)}{I(u)}$$

Με τον τρόπο αυτό, το μοντέλο δίνει κίνητρο στα μέλη της κοινότητας να προσφέρουν και όχι μόνο να καταναλώνουν.

### 2.3.4 Δομές

Και στην περίπτωση του μοντέλου PeerTrust δεν περιγράφονται συγκεκριμένες δομές που είναι αναγκαίες για την υλοποίησή του. Από την εξέταση όμως του γενικού τύπου υπολογισμού της εμπιστοσύνης μπορούμε να συμπεράνουμε πως κάθε πράκτορας χρειάζεται τουλάχιστον δύο δομές:

- Ένας πίνακας ιστορικού αλληλεπιδράσεων  $I_h$ , στον οποίο διατηρούνται οι τιμές ικανοποίησης των αλληλεπιδράσεων των πρακτόρων με τους οποίους ήρθε σε επαφή ο συγκεκριμένος πράκτορας. Ο πίνακας είναι μονοδιάστατος και κάθε στοιχείο  $I_h[i]$  είναι μια λίστα με τις τιμές ικανοποίησης που έδωσε ο πράκτορας  $i$  στις διάφορες αλληλεπιδράσεις του με το συγκεκριμένο πράκτορα που διατηρεί τον πίνακα.

- Ο μονοδιάστατος πίνακας αξιοπιστίας  $T$ . Κάθε στοιχείο  $T[i]$  του πίνακα αντιπροσωπεύει την τρέχουσα εμπιστοσύνη του συγκεκριμένου πράκτορα προς τον πράκτορα  $i$ . Ο πίνακας αυτός λειτουργεί ως τοπική cache για την μείωση του φόρτου επικοινωνίας κατά τον υπολογισμό της εμπιστοσύνης ενός πράκτορα.

### 2.3.5 Σχολιασμός - κριτική

Οι Xiong και Liu προσπάθησαν να διαμορφώσουν ένα μοντέλο που να περιλαμβάνει τους σημαντικότερους, κατά τη γνώμη τους, παράγοντες που επιδρούν στη διαμόρφωση της εμπιστοσύνης. Η πρότασή τους καταφέρνει, με ένα απλό τρόπο, να συμπεριλάβει, σε ένα και μόνο τύπο, την ουσία πολλών από τα ήδη υπάρχοντα μοντέλα εμπιστοσύνης. Το σημαντικότερο όμως χαρακτηριστικό του μοντέλου τους είναι, ότι μπορεί εύκολα να προσαρμοστεί στα ιδιαίτερα χαρακτηριστικά μιας κοινότητας και να προσφέρει λύσεις σε γνωστά προβλήματα.

Από τη μελέτη μας εντοπίσαμε τα παρακάτω θετικά και αρνητικά σημεία για το PeerTrust.

#### Θετικά

- + Παρέχει ένα γενικό τύπο που μπορεί να προσαρμοστεί στις εκάστοτε ανάγκες και τα χαρακτηριστικά της περιοχής εφαρμογής. Με τον τρόπο αυτό μελετά τους πέντε σημαντικότερους παράγοντες καθορισμού της εμπιστοσύνης.
- + Εισάγει τους παράγοντες transaction context factor και community context factor
- + Προτείνει την αξιολόγηση της αξιοπιστίας με βάση τα τρέχοντα στοιχεία των συναλλαγών από όλους τους πράκτορες. Αυτό δίνει μεγαλύτερη εγκυρότητα σε σχέση με άλλα μοντέλα όπου η αξιολόγηση βασίζεται στην τελευταία εκτίμηση της εμπιστοσύνης μετά την οποία όμως μπορεί να έχουν λάβει χώρα αρκετές άλλες συναλλαγές.

#### Αρνητικά

- Υιοθετεί την απλοϊκή παραδοχή ότι η αξιοπιστία των πηγών ανατροφοδότησης υποδεικνύεται από την τιμή εμπιστοσύνης τους.
- Ο τύπος υπολογισμού της εμπιστοσύνης δίνει μια εκτίμηση μόνο για το ύψος της εμπιστοσύνης (υψηλή, χαμηλή) χωρίς να καλύπτει την περίπτωση της έλλειψης εμπιστοσύνης.
- Αφήνει αρκετά ανοικτά θέματα αναφορικά με την υλοποίηση του μοντέλου. Για παράδειγμα δεν διευκρινίζει που αποθηκεύονται τα δεδομένα της αξιολόγησης των συναλλαγών, με ποιο τρόπο γίνεται η αναζήτησή τους κλπ)
- Ο υπολογισμός της εμπιστοσύνης χρειάζεται ως δεδομένα το σύνολο των συναλλαγών που έχουν λάβει χώρα. Αυτό όμως εμπεριέχει πολύ μεγάλο φόρτο επικοινωνίας για την συλλογή των δεδομένων, γεγονός που καθιστά το μοντέλο μη αποδοτικό.

- Σε ότι αφορά την εφαρμογή του υπολογισμού του μοντέλου, διαπιστώνεται ότι στην εκτέλεση του υπολογισμού της εμπιστοσύνης ενός μέλους  $u$  (από άλλα μέλη) εμφανίζονται δύο ζητήματα:

α) Δεν διευκρινίζεται πως και από πού αντλούνται οι ανατροφοδοτήσεις-αξιολογήσεις των συναλλαγών που έχει κάνει το μέλος  $u$ . Το πρόβλημα έχει σχέση και με το ότι στο μοντέλο δεν αναφέρεται τίποτε για τις δομές δεδομένων που απαιτούνται. Για παράδειγμα, δεν διευκρινίζεται εάν οι ανατροφοδοτήσεις-αξιολογήσεις των συναλλαγών κρατούνται στην πηγή (δηλ. στο μέλος που έκανε την αξιολόγηση) ή στον προορισμό (δηλ. στο μέλος που αξιολογήθηκε και δέχτηκε την ανατροφοδότηση).

β) Χρειάζεται να υπάρχουν οι τιμές εμπιστοσύνης των μελών που έχουν κάνει συναλλαγές με το  $u$ , για να χρησιμοποιηθούν ως εκτίμηση της αξιοπιστίας των ανατροφοδοτήσεων τους. Επειδή για να υπολογισθούν οι τιμές αυτές σε πραγματικό χρόνο απαιτείται μεγάλο υπολογιστικό κόστος, οι Xiong και Liu προτείνουν, κάθε μέλος να κρατά μια τοπική cache με τις τιμές εμπιστοσύνης όλων των μελών της κοινότητας τις οποίες έχουν υπολογίσει στο παρελθόν. Για όσα μέλη δεν έχουν υπολογισμένη τιμή, μπορούν να κρατούν μια εξ' ορισμού τιμή, ανάλογα με την προδιάθεση εμπιστοσύνης που έχουν (dispositional trust).

Τα παραπάνω οδηγούν στο συμπέρασμα ότι ο υπολογισμός της εμπιστοσύνης εμπεριέχει πολύ μεγάλο φόρτο επικοινωνίας για την συλλογή των δεδομένων, γεγονός που καθιστά το μοντέλο μη αποδοτικό.

## 2.4 Το μοντέλο των Mui - Mohtashemi - Halberstadt

Το 2002 οι L.Mui, M.Mohtashemi και A.Halberstadt παρουσίασαν ένα υπολογιστικό μοντέλο [29], που επιχειρεί να αναπαραστήσει την εμπιστοσύνη λαμβάνοντας υπόψη κάποια από τα κοινωνιολογικά χαρακτηριστικά της, με ιδιαίτερη έμφαση στην αμοιβαιότητα (reciprocity). Στήριξαν τη δουλειά τους κατά κύριο λόγο στη θεωρία πιθανοτήτων και τη στατιστική. Στόχος τους ήταν να δημιουργήσουν μια πρακτική και υλοποιήσιμη πρόταση που θα απέφευγε τα λάθη των ad hoc προσεγγίσεων των υπολοίπων γνωστών μοντέλων.

### 2.4.1 Παραδοχές- Χαρακτηριστικά

Το μοντέλο των Mui, Mohtashemi και Halberstadt ξεκινά από την απλή παρατήρηση ότι, στις επαναλαμβανόμενες αλληλεπιδράσεις, μέσα στις ανθρώπινες κοινωνίες, παίζουν κυρίαρχο ρόλο οι αμοιβαίες συμπεριφορές. Δηλαδή, τα μέλη μιας κοινωνίας συμπεριφέρονται σωστά, προσδοκώντας ότι θα τους ανταποδώσουν αυτή τη συμπεριφορά, ή αντίστοιχα, συμπεριφέρονται άσχημα σε μέλη που τους έχουν συμπεριφερθεί άσχημα. Οι συγγραφείς διατυπώνουν την άποψη ότι, η διερεύνηση του βαθμού

στον οποίο αναμένεται αμοιβαιότητα είναι ένα πρόβλημα εκτίμησης της εμπιστοσύνης. Κι αυτό γιατί βλέποντας την αμοιβαιότητα ως μια μεταβλητή μεταξύ δύο πρακτόρων, όσο υψηλότερη είναι η τιμή της τόσο μεγαλύτερη είναι η πιθανότητα ο ένας πράκτορας να ανταποκριθεί με αμοιβαίο τρόπο στις μελλοντικές πράξεις του άλλου. Έτσι, μέσα σε ένα περιβάλλον όπου τα άτομα, τακτικά, ακολουθούν νόρμες<sup>6</sup> αμοιβαιότητας, υπάρχει ένα κίνητρο για την απόκτηση καλής φήμης σχετικά με πράξεις αμοιβαιότητας.

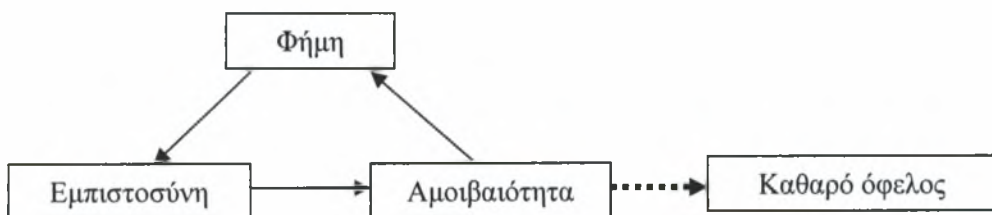
Με κεντρικό, λοιπόν, άξονα την αμοιβαιότητα, το μοντέλο μελετά τις τρεις βασικές παραμέτρους του προβλήματος της εμπιστοσύνης, την αμοιβαιότητα, τη φήμη και την εμπιστοσύνη. Οι ορισμοί που υιοθετεί είναι :

«**Αμοιβαιότητα** : Αμοιβαία ανταλλαγή πράξεων (όπως χάρη ή εκδίκηση).»

«**Φήμη** : Η αντίληψη που δημιουργεί ένας πράκτορας, μέσω των περασμένων ενεργειών του, για τις μελλοντικές του προθέσεις και νόρμες»

«**Εμπιστοσύνη**<sup>7</sup> : Η υποκειμενική προσδοκία που έχει ένας πράκτορας σχετικά με τη μελλοντική συμπεριφορά κάποιου άλλου, βασισμένη στο ιστορικό των συναντήσεών τους.»

Οι τρεις αυτές παράμετροι εξαρτώνται η μια από την άλλη, δημιουργώντας μια τριγωνική σχέση, όπως φαίνεται στο σχήμα 2-1.



Σχήμα 2-1 : Η σχέση φήμης, εμπιστοσύνης, αμοιβαιότητας

Δηλαδή, η αύξηση της φήμης ενός πράκτορα αυξάνει την εμπιστοσύνη που έχουν οι υπόλοιποι προς αυτόν. Η αύξηση της εμπιστοσύνης, με τη σειρά της, αυξάνει την πιθανότητα θετικής ανταπόκρισης των υπολοίπων στις ενέργειές του αλλά και του ίδιου προς τους άλλους. Αυτή, όμως, η αμοιβαία ανταπόκριση του πράκτορα στις πράξεις των υπολοίπων θα έχει ως αποτέλεσμα την αύξηση της φήμης του κ.οκ.

<sup>6</sup> Οι συγγραφείς υιοθετούν τον ορισμό της νόρμας από τον Ostrom [30] : “... ευρετικές μέθοδοι που υιοθετεί ένα άτομο, από ηθικής άποψης, υπό την έννοια ότι αυτό είναι το είδος των ενεργειών που επιθυμεί να ακολουθεί στη ζωή του.»

<sup>7</sup> Ο ορισμός της εμπιστοσύνης που υιοθετεί το μοντέλο βασίζεται στον ορισμό του Gambetta [2], με χρήση όμως του όρου «υποκειμενική προσδοκία» (subjective expectation) στη θέση της «υποκειμενικής πιθανότητας» (subjective probability).



Κατά παρόμοιο τρόπο, η μείωση οποιασδήποτε από τις τρεις παραμέτρους θα οδηγήσει τελικά σε μείωση και των υπολοίπων παραμέτρων.

Για λόγους απλότητας, το μοντέλο θεωρεί ότι η κοινωνία μέσα στην οποία αναπτύσσονται οι αλληλεπιδράσεις μεταξύ των πρακτόρων είναι στατική. Δηλαδή, κανένα καινούργιο μέλος δεν προστίθεται στην κοινωνία και κανένα μέλος δεν αποχωρεί από αυτή.

Τέλος, οι ενέργειες που μπορεί να κάνει ένας πράκτορας είναι μόνο δύο: «Συνεργασία» και «Άρνηση συνεργασίας».

## 2.4.2 Αναπαράσταση της εμπιστοσύνης

Η συμβολισμοί των τριών παραμέτρων που μελετά το μοντέλο καθώς και τα πεδία τιμών τους είναι :

- **Αμοιβαιότητα** :  $\gamma \in [0,1]$

Αναπαριστά το πλήθος των αμοιβαίων ενεργειών που συμβαίνουν σε μια κοινωνία, δηλ., των ενεργειών συνεργασίας, που απαντώνται με συνεργασία και των ενεργειών μη συνεργασίας, που απαντώνται με άρνηση συνεργασίας.

- **Φήμη** :  $\theta_{ij}(c) \in [0,1]$

όπου  $c$  : το πλαίσιο (context) των αλληλεπιδράσεων, που παίρνει τιμές από το σύνολο  $C$  των πλαισίων που ενδιαφέρουν την κοινωνία

Αναπαριστά την πιθανότητα ο πράκτορας  $a_j$  να ανταποδώσει τις ενέργειες του πράκτορα  $a_i$ . Όσο μεγαλύτερη είναι η τιμή της φήμης, τόσο μεγαλύτερη είναι η πρόθεση του  $a_j$  για ανταπόδοση των ενεργειών του  $a_i$ .

- **Εμπιστοσύνη** :  $\tau_{ij}(c) \in [0,1]$

Αναπαριστά την προσδοκία του  $a_i$  σε ότι αφορά την πρόθεση του  $a_j$  να ανταποδώσει τις ενέργειές του. Όσο υψηλότερη τιμή έχει η εμπιστοσύνη, τόσο υψηλότερη είναι αυτή η προσδοκία.

Εάν με  $D_{ij}(c)$  συμβολίσουμε το ιστορικό των συναντήσεων μεταξύ  $a_i$  και  $a_j$  που εντάσσονται στο πλαίσιο  $c$ , τότε η εμπιστοσύνη μπορεί να εκφραστεί ως :

$$\tau_{ij}(c) = E[\theta_{ij}(c) | D_{ij}(c)]$$

### 2.4.3 Αξιολόγηση της εμπιστοσύνης

Πριν ένας πράκτορας  $a$  αλληλεπιδράσει με ένα πράκτορα  $b$  θα πρέπει να εκτιμήσει την αξιοπιστία του  $b$ . Ο πράκτορας  $a$  θεωρείται ότι είναι πάντα συνεργάσιμος και ότι, μέσα από τη διαδικασία αξιολόγησης της αξιοπιστίας, προσπαθεί να εκτιμήσει την τάση του  $b$  να ανταποδώσει αυτή τη συνεργασιμότητα.

Έστω οι παρακάτω συμβολισμοί :

$\theta_{ab}$  : Η φήμη του  $b$  στα μάτια του  $a$

$e_{ab}(i)$  : Μια δυαδική τυχαία μεταβλητή, που αναπαριστά την  $i$ -οστή συνάντηση μεταξύ των  $a$  και  $b$ . Παίρνει την τιμή '1' εάν η ενέργεια του  $b$  είναι «Συνεργασία» και '0' εάν είναι «Άρνηση συνεργασίας».

$D_{ab}$  : Το ιστορικό των  $n$  προηγούμενων συναντήσεων των  $a$  και  $b$ . Προφανώς ισχύει ότι

$$D_{ab} = \{e_{ab}(1), e_{ab}(2), \dots, e_{ab}(n)\}.$$

$p$  : το πλήθος των συνεργασιών του  $b$  στις  $n$  προηγούμενες συναντήσεις του με τον  $a$

Προφανώς η φήμη  $\theta_{ab}$  του  $b$  θα είναι μια συνάρτηση των  $p$  και  $n$ . Μια απλή συνάρτηση αυτού του είδους θα μπορούσε να είναι η αναλογία των συνεργάσιμων ενεργειών του  $b$  στο σύνολο των  $n$  συναντήσεων του με τον  $a$ . Από τη στατιστική ισχύει ότι, μια τυχαία μεταβλητή αναλογίας (proportion random variable) μπορεί να μοντελοποιηθεί ως μια Β' κατανομή (Beta Distribution). Ισχύει λοιπόν ότι :

$$p(\hat{\theta}) = \text{Beta}(c_1, c_2)$$

όπου

$\hat{\theta}$  : ο εκτιμητής της συνάρτησης  $\theta$

$c_1, c_2$  : είναι παράμετροι που καθορίζονται από προηγούμενες υποθέσεις.

Αυτή η αναλογία των συνεργασιών κατά τις  $n$  πεπερασμένες συναντήσεις μπορεί να θεωρηθεί ένας απλός εκτιμητής για την  $\theta_{ab}$ . Δηλαδή

$$\hat{\theta}_{ab} = \frac{p}{n}$$

Αν υποθέσουμε ότι η πιθανότητα να συνεργαστεί ο  $b$  σε μια συνάντηση με τον  $a$  είναι ανεξάρτητη από τις άλλες συναντήσεις τους, τότε η πιθανότητα να έχουμε  $p$  συνεργασίες και  $n-p$  αρνήσεις συνεργασίας μπορεί να μοντελοποιηθεί ως :

$$L(D_{ab} | \hat{\theta}) = (\hat{\theta})^p (1 - \hat{\theta})^{n-p}$$

Η Β' κατανομή γίνεται τελικά ο πρότερος συζυγής (conjugate prior) αυτής της πιθανότητας [33].

Συνδυάζοντας τον πρότερο και την πιθανότητα, η επόμενη εκτίμηση για το  $\hat{\theta}$  γίνεται [32] :

$$p(\hat{\theta}_{ab} | D_{ab}) = \text{Beta}(c_1 + p, c_2 + n - p)$$

Οι στατιστικές ιδιότητες πρώτης τάξης της επόμενης εκτίμησης του  $\hat{\theta}$  είναι :

$$E[\hat{\theta}_{ab} | D_{ab}] = \frac{c_1 + p}{c_1 + c_2 + n}$$

$$\sigma_{\hat{\theta}_{ab}|D_{ab}}^2 = \frac{(c_1 + p)(c_2 + n - p)}{(c_1 + c_2 + n - 1)(c_1 + c_2 + n)^2}$$

Σε μια επόμενη συνάντηση των a και b, η εκτίμηση που μπορεί να κάνει ο a για την πιθανότητα που έχει ο b ισούται με :

$$\tau_{ab} = p(x_{ab}(n+1) = 1 | D_{ab}) = E[\hat{\theta}_{ab} | D_{ab}]$$

Δηλαδή, το μοντέλο προτείνει ότι η εμπιστοσύνη του a προς τον b είναι αυτή η υπό συνθήκη προσδοκία του  $\hat{\theta}_{ab}$  δοθέντος του  $D_{ab}$ .

Για την εκτίμηση  $\hat{\theta}_{ab}$  μπορούμε να έχουμε ένα όριο, βάσει του παρακάτω θεωρήματος Chernoff :

Έστω  $x_{ab}(1), x_{ab}(2), \dots, x_{ab}(m)$  μια ακολουθία m ανεξάρτητων δοκιμών Bernuli, κάθε μια με πιθανότητα επιτυχίας<sup>8</sup>  $E(x_{ab}) = \theta$ . Αν ορίσουμε ως εκτιμητή το

$$\hat{\theta} = \frac{x_{ab}(1) + x_{ab}(2) + \dots + x_{ab}(m)}{m}$$

το  $\hat{\theta}$  είναι μια τυχαία μεταβλητή που αναπαριστά την αναλογία της επιτυχίας και συνεπώς

$$E[\hat{\theta}] = \theta$$

Τότε για  $0 \leq \varepsilon \leq 1$  και  $0 \leq \delta \leq 1$ , θα ισχύει το όριο<sup>9</sup>

$$\Pr[|\hat{\theta} - \theta| \geq \varepsilon] \leq 2e^{-2m\varepsilon^2} \leq \delta$$

Το  $\varepsilon$ , ουσιαστικά, αντιστοιχεί στην απόκλιση του εκτιμητή από την πραγματική παράμετρο και υπ' αυτή την έννοια μπορεί να θεωρηθεί ως μια σταθερή παράμετρος λάθους (π.χ. με τιμή 0,05).

Από το θεώρημα Chernoff, το m αντιστοιχεί στο ελάχιστο πλήθος των συναντήσεων που είναι απαραίτητες για να επιτευχθεί το επιθυμητό επίπεδο λάθους εκτίμησης. Το ελάχιστο αυτό όριο μπορεί να υπολογιστεί ως :

<sup>8</sup> «Επιτυχία» στο μοντέλο των Mui-Mohteshemi-Halberstadt σημαίνει συνεργασία ή γενικότερα αμοιβαιότητα.

<sup>9</sup> Η απόδειξη είναι απευθείας εφαρμογή της αθροιστικής μορφής του ορίου Chernoff για της δοκιμές Bernulli

$$m \geq -\frac{1}{2\varepsilon^2} \ln\left(\frac{\delta}{2}\right)$$

### Βεβαιότητα για την εκτίμηση της εμπιστοσύνης

Ένα μέτρο της βεβαιότητας σχετικά με τον εκτιμητή  $\hat{\theta}$  θα μπορούσε να είναι το

$$\gamma_c = 1 - \delta$$

Όσο πλησιέστερα στο 1 είναι το  $\gamma_c$ , τόσο μεγαλύτερη βεβαιότητα έχουμε για την εκτίμηση της εμπιστοσύνης. Αυτό όμως απαιτεί και μεγαλύτερο  $m$ , ώστε να επιτευχθεί ένα συγκεκριμένο επίπεδο λάθους  $\varepsilon$ . Το  $\gamma_c$  μπορεί να επιλεγεί εξωγενώς, ως μια παράμετρος που καθορίζει το επίπεδο βεβαιότητας ενός πράκτορα για την αξιοπιστία των υπολογισμών της φήμης και της εμπιστοσύνης.

### Μέτρα της Αμοιβαιότητας

Στο μοντέλο των Mui-Mohteshemi-Halberstadt, η αμοιβαιότητα αντιπροσωπεύει ένα μέτρο των αμοιβαίων πράξεων μεταξύ των πρακτόρων. Συγκεκριμένα, περιγράφονται δύο διαφορετικά μέτρα για την αμοιβαιότητα :

- Η «Δυαδική Αμοιβαιότητα» (dyadic reciprocity), που ορίζεται πάνω στις συναντήσεις<sup>10</sup> δύο πρακτόρων

$$\text{Δυαδική Αμοιβαιότητα} = \frac{\text{Αμοιβαίες συνεργασίες} + \text{Αμοιβαίες απορρίψεις συνεργασίας}}{\text{Πλήθος συναντήσεων μεταξύ των δύο πρακτόρων}}$$

- Η «Κοινωνική Αμοιβαιότητα» (societal reciprocity), που ορίζεται πάνω στο σύνολο όλων των συναντήσεων ενός κοινωνικού δικτύου (social network)

$$\text{Κοινωνική Αμοιβαιότητα} = \frac{\text{Αμοιβαίες συνεργασίες} + \text{Αμοιβαίες απορρίψεις συνεργασίας}}{\text{Πλήθος όλων των συναντήσεων μέσα στο κοινωνικό δίκτυο}}$$

<sup>10</sup> Όπως ειπώθηκε και παραπάνω, όλες οι συναντήσεις θεωρούνται δυαδικές, δηλ. μεταξύ δύο μόνο πρακτόρων.

Συμβολίζοντας με  $\gamma_{ab}$  τη δυαδική αμοιβαιότητα μεταξύ δύο πρακτόρων a και b, εάν  $\gamma_{ab} < \gamma_c$ , τότε η υπολογισμένη φήμη και εμπιστοσύνη είναι κάτω από την εξωτερικά καθορισμένη κρίσιμη τιμή  $\gamma_c$  και συνεπώς δεν θεωρούνται αξιόπιστες.

### Πράκτορες άγνωστοι μεταξύ τους

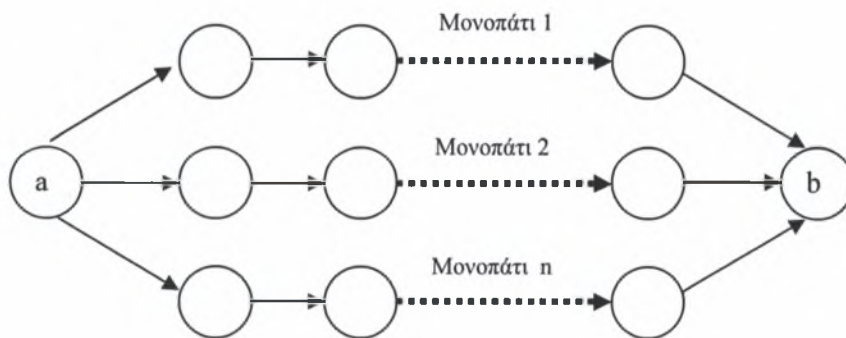
Εάν δύο πράκτορες a και b, που συναντώνται, είναι παντελώς άγνωστοι μεταξύ τους (δηλ. δεν είχαν καμιά προηγούμενη συνάντηση και επιπλέον δεν έχουν ούτε κοινούς γνωστούς), η εκτίμηση του καθενός για τη φήμη του άλλου θεωρείται ομοιόμορφα κατανεμημένη σ' όλη την περιοχή της φήμης. Δηλ.

$$p(\hat{\theta}) = \begin{cases} 1 & 0 < \hat{\theta} < 1 \\ 0 & \text{διαφορετικά} \end{cases}$$

Τέτοια ομοιόμορφη κατανομή δίνει ο Beta prior με τιμές  $c_1 = c_2 = 1$ .

### 2.4.4 Μηχανισμός διάδοσης της φήμης

Οι Mui-Mohteshemi-Halberstadt, προτείνουν συγκεκριμένο τρόπο για τον υπολογισμό της εμπιστοσύνης όταν υπάρχουν πολλά διαφορετικά μονοπάτια μέσω των οποίων φτάνει η φήμη ενός πράκτορα. Η πρότασή τους αναφέρεται μόνο σε παράλληλα μονοπάτια, δηλ. μονοπάτια των οποίων οι κόμβοι είναι εντελώς διαφορετικοί (σχήμα 2-2).



Σχήμα 2-2 : Σύνθεση της εμπιστοσύνης μέσα από παράλληλα μονοπάτια

Υποθέτουμε ότι ο πράκτορας a επιθυμεί να εκτιμήσει την αξιοπιστία του πράκτορα b. Αναζητώντας πληροφορίες για τον b, ο a βρίσκει n συστάσεις μέσα από n διαφορετικά μονοπάτια. Προφανώς, χρειάζεται να εκτιμηθεί η αξιοπιστία του κάθε μονοπατιού ώστε τελικά η πληροφορία της φήμης του b, που καταφθάνει από κάθε μονοπάτι, να συμμετάσχει στον υπολογισμό της εμπιστοσύνης του b με την ανάλογη βαρύτητα.

Η αξιοπιστία, όμως ενός μονοπατιού υπολογίζεται βάσει της αξιοπιστίας των συνδέσεών του. Ένα μέτρο για την αξιοπιστία-βαρύτητα μιας μεμονωμένης σύνδεσης μεταξύ δύο κόμβων  $i, j$  ενός μονοπατιού, που προτείνει το μοντέλο είναι:

$$w_{ij} = \begin{cases} \frac{m_{ij}}{m} & \text{εάν } m_{ij} < m \\ 1 & \text{διαφορετικά} \end{cases}$$

όπου

$m$  : το όριο των συναντήσεων που απαιτούνται (σύμφωνα με τα όσα ειπώθηκαν στην παράγραφο 1.2) για μια αξιόπιστη εκτίμηση της εμπιστοσύνης

$m_{ij}$  : το πλήθος το συναντήσεων μεταξύ  $i$  και  $j$

Η συνολική βαρύτητα  $w_i$  ενός μονοπατιού  $i$  μπορεί στη συνέχεια να υπολογισθεί συνδυάζοντας πολλαπλασιαστικά τη βαρύτητα των συνδέσεών του :

$$w_i = \prod_{j=1}^{l_i} w_{ij}$$

όπου

$l_i$  : το μήκος του μονοπατιού  $i$  (δηλ. το πλήθος των συνδέσεών του)

$w_{ij}$  : η βαρύτητα του  $j$ -οστού τμήματος του μονοπατιού  $i$

Χρησιμοποιώντας τις βαρύτητες των μονοπατιών, η συνολική φήμη  $r_{ab}$  του  $b$  που αντιλαμβάνεται ο  $a$ , όπως αυτή συντίθεται από τις επιμέρους φήμες που καταφθάνουν στον  $a$  μέσω των  $n$  διαφορετικών μονοπατιών, προκύπτει από τον τύπο :

$$r_{ab} = \sum_{i=1}^n r_{ab}(i) \overline{w}_i$$

όπου

$r_{ab}(i)$  : η φήμη του  $b$  που καταφτάνει στον  $a$  μέσω του μονοπατιού  $i$

$\overline{w}_i$  : το κανονικοποιημένο<sup>11</sup> βάρος του μονοπατιού  $i$

<sup>11</sup> Η κανονικοποίηση θα πρέπει να δίνει  $\sum_{i=1}^n \overline{w}_i = 1$ .

### 2.4.5 Δομές

Στο μοντέλο δεν προτείνονται συγκεκριμένες δομές δεδομένων.

### 2.4.6 Σχολιασμός - κριτική

Οι Mui-Mohteshemi-Halberstadt προσπάθησαν να προσεγγίσουν το πρόβλημα της εμπιστοσύνης μέσα από το χώρο της θεωρίας πιθανοτήτων. Πρότειναν ένα πιθανοτικό μηχανισμό εξαγωγής συμπερασμάτων χρησιμοποιώντας τις τρεις βασικές παραμέτρους του προβλήματος της εμπιστοσύνης : τη φήμη, την εμπιστοσύνη και την αμοιβαιότητα. Τα σημαντικότερα θετικά και αρνητικά στοιχεία το μοντέλου είναι :

#### Θετικά

- + Ο φορμαλισμός που προτείνεται είναι εύκολα υλοποιήσιμος
- + Χρησιμοποιεί τη φήμη και την αμοιβαιότητα

#### Αρνητικά

- Υιοθετεί την παραδοχή ότι υπάρχουν μόνο δύο εναλλακτικές για την έκβαση μιας αλληλεπίδρασης: "συνεργασία" και "μη συνεργασία"
- Κάνει την υπεραπλουστευμένη (και μη ρεαλιστική) παραδοχή ότι η πιθανότητα εμφάνισης μιας από τις δυνατές εκβάσεις μιας συγκεκριμένης αλληλεπίδρασης είναι ανεξάρτητη από τις υπόλοιπες συναντήσεις.
- Σε ότι αφορά τη διάδοση της φήμης λαμβάνει υπόψη και χειρίζεται μόνο παράλληλα μονοπάτια.
- Δεν αναφέρει με ποιο τρόπο συλλέγονται πληροφορίες για τη φήμη ενός πράκτορα, δηλ. πώς εντοπίζονται τα διάφορα μονοπάτια διάδοσης της φήμης.
- Η χρήση της θεωρίας πιθανοτήτων για την αναπαράσταση της εμπιστοσύνης έχει αμφισβητούμενη αποτελεσματικότητα κυρίως διότι απαιτεί τη διενέργεια επαναλαμβανόμενων "πειραμάτων", προσέγγιση που δεν μπορεί να εφαρμοστεί στον πραγματικό κόσμο.

## 2.5 Το μοντέλο των Abdul-Rahman & Hailes

Οι Abdul-Rahman και Hailes το 2000 ανέπτυξαν ένα μοντέλο εμπιστοσύνης [34] το οποίο βασιζόταν σε κοινωνιολογικά χαρακτηριστικά της εμπιστοσύνης από τον πραγματικό κόσμο. Ενσωμάτωσαν στο μοντέλο τους έννοιες όπως η φήμη (reputation) και οι συστάσεις (recommendations)

προκειμένου να δημιουργήσουν ένα μηχανισμό αξιολόγησης των υποκειμενικών στοιχείων που συλλέγονται κατά την αναζήτηση συστάσεων.

### 2.5.1 Παραδοχές- Χαρακτηριστικά

Το μοντέλο που πρότειναν οι Abdul-Rahman & Hailes αποτελεί ουσιαστικά ένα μηχανισμό αξιολόγησης των συστάσεων, βάσει των εμπειριών από προηγούμενες, παρόμοιες συστάσεις. Στο μηχανισμό αυτό, κάθε φορά που ένας πράκτορας A λαμβάνει μια σύσταση από έναν πράκτορα Γ σχετικά με την αξιοπιστία κάποιου πράκτορα Β, πριν αποφασίσει αν θα προχωρήσει σε συναλλαγή με τον Β ή όχι. Εάν ο Α επιλέξει να προχωρήσει σε συναλλαγή με τον Β, μετά το πέρας της υπολογίζει και καταγράφει την απόκλιση που διαπιστώνει μεταξύ της σύστασης που έλαβε από τον Γ και της προσωπικής του αξιολόγησης του Β όπως προκύπτει μέσα από τη συναλλαγή μαζί του. Η πληροφορία αυτή μπορεί να χρησιμοποιηθεί μελλοντικά από τον Α ως ένδειξη της αξιοπιστίας της πηγής της σύστασης, δηλαδή του πράκτορα Γ, για την αξιολόγηση άλλων συστάσεων από αυτόν.

Σύμφωνα με το μοντέλο δεν χρειάζεται να διατηρούνται σε κάποια Βάση Δεδομένων. προτάσεις της μορφής «ο πράκτορας Α εμπιστεύεται τον πράκτορα Β σχετικά με το πλαίσιο Γ». Η αξιοπιστίας κάθε πράκτορα προκύπτει από τη συγκεντρωτική αξιολόγηση των καταγεγραμμένων εμπειριών. Οι εμπειρίες που χειρίζεται το μοντέλο είναι εμπειρίες απευθείας εμπιστοσύνης, που προκύπτουν από τις απευθείας συναλλαγές μεταξύ των πρακτόρων, και εμπειρίες εμπιστοσύνης συστάσεων που προκύπτουν από την αξιολόγηση των συστάσεων που λαμβάνει ένας πράκτορας. Οι εμπειρίες αυτές καταγράφονται σε δύο ξεχωριστές δομές, το σύνολο Q και το σύνολο R αντίστοιχα.

Οι Abdul-Rahman & Hailes στο μοντέλο τους κάνουν τη σύμβαση ότι μια συναλλαγή, σε ότι αφορά την τελική της έκβαση, μπορεί να τοποθετηθεί σε ένα από τέσσερα διαφορετικά επίπεδα και να χαρακτηριστεί ως

'vg' (very good), 'g' (good), 'b' (bad) ή 'vb' (very bad)

Οι χαρακτηρισμοί αυτοί είναι διαβαθμισμένοι με τον 'vg' να είναι ο υψηλότερος και τον 'vb' να είναι ο χαμηλότερος και συνιστούν το σύνολο  $E=\{vg, g, b, vb\}$  των πιθανών τιμών έκβασης μιας συναλλαγής.

Το μοντέλο των Abdul-Rahman & Hailes εμφανίζει χαρακτηριστικά που αντικατοπτρίζουν τα κοινωνιολογικά χαρακτηριστικά και τις ιδιότητες της εμπιστοσύνης όπως αυτή είναι αντιληπτή στον πραγματικό κόσμο. Συγκεκριμένα το μοντέλο :

- Εξετάζει την εμπιστοσύνη σε σχέση με το γενικό πλαίσιο (context) μέσα στο οποίο λαμβάνουν χώρα οι συναλλαγές. Έτσι, ένας πράκτορας μπορεί να θεωρείται αξιόπιστος για την παροχή ενός συγκεκριμένου είδους υπηρεσιών (π.χ. εκτέλεση υπολογισμών), αλλά αναξιόπιστος για κάποιο άλλο είδος υπηρεσιών (π.χ. παροχή αποθηκευτικού χώρου).



- Λαμβάνει υπόψη και χρησιμοποιεί την συσσωρευμένη εμπειρία από προηγούμενες συναλλαγές. Για την εκτίμηση της αξιοπιστίας ενός πράκτορα, ενόψει μιας συνεργασίας-συναλλαγής, αναζητούνται (εάν υπάρχουν) αποτελέσματα από παλαιότερες συναλλαγές με αυτόν τον πράκτορα, του ίδιου ή παρόμοιου γενικού πλαισίου.
- Υποστηρίζει τόσο θετικές όσο και αρνητικές διαβαθμίσεις της εκτιμώμενης αξιοπιστίας, με τις θετικές να υποδηλώνουν την ύπαρξη αξιοπιστίας και τις αρνητικές την έλλειψη αξιοπιστίας.
- Οι πράκτορες μπορούν να ανταλλάσσουν μεταξύ τους, μέσα από την παροχή συστάσεων (recommendations), πληροφορίες που κατέχουν και αφορούν τη φήμη (ως προς την αξιοπιστία) άλλων πρακτόρων. Η ανταλλαγή αυτή συνιστά ουσιαστικά ένα μηχανισμό φήμης (reputation mechanism) που βοηθά τα μέλη της κοινότητας στη διαμόρφωση της εικόνας εμπιστοσύνης που έχουν το ένα για το άλλο.
- Θεωρεί ότι η εμπιστοσύνη δεν είναι μεταβατική. Έτσι κάθε σύσταση που λαμβάνεται, δεν χρησιμοποιείται αυτούσια, αλλά αξιολογείται με βάση την αξιοπιστία και τη φήμη της πηγής από την οποία προέρχεται.
- Αναγνωρίζει ότι η εμπιστοσύνη είναι υποκειμενική, δηλ. ότι διαφορετικοί παρατηρητές μπορεί να έχουν διαφορετικές εκτιμήσεις της αξιοπιστίας του ίδιου πράκτορα, ανάλογα με τις ιδιαίτερες πεποιθήσεις και τις απαιτήσεις του καθενός.
- Χειρίζεται την εμπιστοσύνη με δυναμικό τρόπο. Οι εμπειρίες και οι συστάσεις μεταβάλλουν διαρκώς (θετικά ή αρνητικά) την εικόνα εμπιστοσύνης που έχουν τα μέλη της κοινότητας μεταξύ τους.
- Παρέχει υποστήριξη μόνο για την απ' ευθείας εμπιστοσύνη μεταξύ πρακτόρων, την οποία και ονομάζει διαπροσωπική (interpersonal trust). Οι Abdul-Rahman & Hailes διακρίνουν άλλες δύο κατηγορίες εμπιστοσύνης, την απρόσωπη εμπιστοσύνη ή εμπιστοσύνη συστήματος (impersonal or system trust) που αφορά την εμπιστοσύνη που αναπτύσσεται προς το σύστημα μέσα στο οποίο διεκπεραιώνονται οι συναλλαγές (π.χ. νομισματικό σύστημα) και την εμπιστοσύνη προδιάθεσης ή βασική εμπιστοσύνη (dispositional or basic trust) που αφορά τη γενικότερη τάση για εμπιστοσύνη ή δυσπιστία που μπορεί να έχει κάποιος ανάλογα με την ιδιοσυγκρασία του. Το μοντέλο δεν ασχολείται με καμιά από τις δύο αυτές κατηγορίες εμπιστοσύνης.

## 2.5.2 Αναπαράσταση εμπιστοσύνης

### *Απευθείας εμπιστοσύνη (Direct trust)*

Η πεποίθηση ενός πράκτορα σχετικά με την αξιοπιστία κάποιου άλλου πράκτορα  $a$  συμβολίζεται στο μοντέλο με τη σχέση

$$t(a, c, td)$$

όπου

- a : ο πράκτορας του οποίου η αξιοπιστία εξετάζεται
- c : το πλαίσιο (context) μέσα στο οποίο εξετάζεται η αξιοπιστία
- td : το επίπεδο αξιοπιστίας

Οι τιμές που μπορεί να πάρει το επίπεδο αξιοπιστίας td και σημασία τους είναι :

- vt => very trustworthy (πολύ αξιόπιστος)
- t => trustworthy (αξιόπιστος)
- u => untrustworthy (αναξιόπιστος)
- vu => very untrustworthy (πολύ αναξιόπιστος)

Στο μοντέλο δεν υποδεικνύεται κάποιο αντίστοιχο πεδίο τιμών για το πλαίσιο c. Αυτό αφήνεται ανοικτό ώστε να προσαρμόζεται με τις εκάστοτε ανάγκες και να ορίζεται ανάλογα.

### ***Εμπιστοσύνη του εισηγητή (Recommender trust)***

Η πεποίθηση ενός πράκτορα x σχετικά με την αξιοπιστία κάποιου άλλου πράκτορα a, σε ότι αφορά την παροχή συστάσεων για κάποιο τυχαίο τρίτο πράκτορα y, συμβολίζεται στο μοντέλο με τη σχέση

$$rt(a, c, rtd)$$

όπου

- rtd : ακέραιος αριθμός στην περιοχή μεταξύ -3 και +3 που υποδεικνύει την απόκλιση ανάμεσα στην σύσταση του εισηγητή a και στην εικόνα της αξιοπιστίας του προτεινόμενου πράκτορα y όπως την αντιλαμβάνεται ο x.

Στο μοντέλο η παράμετρος αυτή ονομάζεται σημαντική απόσταση (semantic distance). Η αναγκαιότητα της ύπαρξής της οφείλεται αφενός στην υποκειμενικότητα της αξιολόγησης της αξιοπιστίας από τον ένα πράκτορα στον άλλο, και αφετέρου στην πιθανότητα παροχής ψευδών συστάσεων από κακόβουλους πράκτορες. Σε κάθε περίπτωση, η τιμή της υποδεικνύει το εύρος κατά το οποίο πρέπει να τροποποιηθεί η σύσταση ώστε να συμπέσει με την πραγματικότητα. Για παράδειγμα μια σύσταση 'vt' (Very Trustworthy) από ένα πράκτορα με rtd '-2' ουσιαστικά ισοδυναμεί με 'u' (Untrustworthy), δηλ. 2 επίπεδα χαμηλότερα.

### **2.5.3 Δομές Δεδομένων**

Το μοντέλο χρησιμοποιεί τις παρακάτω δομές :

#### ***Σύνολο Q : Εμπειρίες συναλλαγών***

Στο σύνολο αυτό, κάθε πράκτορας αποθηκεύει, για κάθε ένα πράκτορα a με το οποίο συναλλάσσεται και για κάθε πλαίσιο συναλλαγών c, τριάδες της μορφής (c, a, s), όπου

- c : το πλαίσιο (context) εκτέλεσης συναλλαγών
- a : ο πράκτορας με τον οποίο έγιναν οι συναλλαγές

$s$  : μια τετράδα αριθμών ( $s_{vg}, s_g, s_b, s_{vb}$ ) που αναπαριστά το πλήθος των συναλλαγών ανά επίπεδο έκβασης συναλλαγής. Π.χ. μια τετράδα (10, 3, 1, 8) σημαίνει ότι έλαβαν χώρα 10 συναλλαγές επιπέδου `very good`, 3 συναλλαγές επιπέδου `good`, 1 συναλλαγή επιπέδου `bad` και 8 συναλλαγές επιπέδου `very bad`.

### **Σύνολο $R$ : Εμπειρίες συστάσεων**

Για τις συστάσεις, κάθε πράκτορας χρειάζεται να κρατά ένα δείκτη που υποδηλώνει πόσο επιτυχημένη ήταν κάθε σύσταση. Ο δείκτης αυτός (στο μοντέλο ονομάζεται «σημασιολογική διαφορά» - semantic difference) παίρνει τιμές από το σύνολο  $G = \{-3, -2, -1, 0, 1, 2, 3\}$  και υποδεικνύει πόσο απέχει η συγκεκριμένη σύσταση από την πραγματικότητα. Μια τιμή `-2` δείχνει ότι το αποτέλεσμα της συναλλαγής ήταν 2 επίπεδα χαμηλότερα από ότι υπέδειξε η συγκεκριμένη σύσταση. Η σπουδαιότητα αυτού του δείκτη είναι ότι προσφέρει μια τιμή ρύθμισης (αύξησης ή μείωσης) της εκάστοτε σύστασης ώστε αυτή να γίνει περισσότερο αντικειμενική. Για παράδειγμα μια σύσταση `good` με σημασιολογική διαφορά 1 σημαίνει ότι ουσιαστικά είναι `very good`.

Για κάθε πράκτορα  $a$  και πλαίσιο  $c$  κρατιούνται τέσσερα σύνολα εμπειριών απόκλισης συστάσεων τα  $T_{vg}, T_g, T_{vb}, T_b$ . Κάθε σύνολο  $T_e$  περιέχει τιμές από το σύνολο  $G$  των δυνατών τιμών απόκλισης και τα στοιχεία του αντιστοιχούν στις αποκλίσεις που εμφανίστηκαν σε συστάσεις επιπέδου  $e \in E = \{vg, g, b, vb\}$ . Για παράδειγμα το σύνολο  $T_{vg} = \{0, 1, -1, 0\}$  απεικονίζει τις αποκλίσεις 4 συστάσεων επιπέδου very good.

Τελικά, στο σύνολο  $R$ , κάθε πράκτορας αποθηκεύει, για κάθε άλλο πράκτορα  $a$  που του παρέχει συστάσεις μέσα στο πλαίσιο  $c$ , τριάδες της μορφής  $(c, a, t)$  όπου

$c$  : το πλαίσιο (context) για το οποίο δίνονται συστάσεις

$a$  : ο πράκτορας που δίνει τις συστάσεις

$t$  : μια τετράδα συνόλων ( $T_{vg}, T_g, T_b, T_{vb}$ ) που αντιστοιχεί στις καταγεγραμμένες αποκλίσεις ανά επίπεδο σύστασης, για όλες τις συστάσεις που έδωσε ο πράκτορας  $a$  στο πλαίσιο  $c$ .

### **2.5.4 Υπολογισμοί**

Στο υπολογιστικό κομμάτι του μοντέλου περιλαμβάνονται οι παρακάτω υπολογισμοί :

#### **Υπολογισμός Απευθείας Εμπιστοσύνης**

Ο υπολογισμός του βαθμού εμπιστοσύνης ενός πράκτορα  $a$  προς ένα πράκτορα  $b$  σε σχέση με ένα πλαίσιο  $c$ , γίνεται με την αναζήτηση της σχέσης  $(c, b, s)$  μέσα στο σύνολο  $Q$  των εμπειριών απευθείας εμπιστοσύνης. Η θέση του μεγαλύτερου στοιχείου της τετράδας  $s$  υποδεικνύει και το βαθμό εμπιστοσύνης λαμβάνοντας υπόψη ότι  $s = (s_{vg}, s_g, s_b, s_{vb})$ . Τυπικά, έχουμε ότι :

$$\exists td \in E \forall se \in s, (se = \max(s)) \Rightarrow (td = e) \quad (1)$$

Εάν στο  $s$  δεν υπάρχει ένα μόνο μέγιστο στοιχείο, τότε έχουμε περίπτωση αβεβαιότητας. Οι διάφορες περιπτώσεις αβεβαιότητας που μπορεί να εμφανιστούν συνοψίζονται στον παρακάτω πίνακα (το σύμβολο  $?$  υποδηλώνει ένα ή κανένα άλλο στοιχείο) :

Μέγιστα στοιχεία	Βαθμός εμπιστοσύνης ( $td$ )	Ερμηνεία
'Vg', 'g', ?	$u^+$	Κυρίως 'good' και λίγο 'bad'
'vb', 'b', ?	$u^-$	Κυρίως 'bad' και λίγο 'good'
Οτιδήποτε άλλο	$u^0$	'bad' και 'good' σε όμοιο βαθμό

Πίνακας 2-1 : Περιπτώσεις αβεβαιότητας

### Υπολογισμός Εμπιστοσύνης Εισηγητή

Ο βαθμός εμπιστοσύνης ενός εισηγητή πράκτορα  $b$  υπολογίζεται με παρόμοιο τρόπο. Συγκεκριμένα εντοπίζεται η σχέση  $(c, a, t)$  μέσα στο σύνολο  $R$  των εμπειριών εισήγησης, όπου  $t=(T_{vg}, T_g, T_b, T_{vb})$ . Τα στοιχεία του  $t$  συγχωνεύονται στο σύνολο  $T^a = T_{vg} \cup T_g \cup T_b \cup T_{vb}$ . Το  $\text{mod}$  των απόλυτων τιμών των στοιχείων του  $T^a$  είναι ο ζητούμενος βαθμός εμπιστοσύνης του εισηγητή  $a$ . Τυπικά, έχουμε ότι :

$$rtd = \text{mod}(\{ \forall x \in T^a \mid |x| \}) \quad (2)$$

### Υπολογισμός Σημασιολογικής Διαφοράς

Κάθε πράκτορας εισηγητής χαρακτηρίζεται από τη σημασιολογική διαφορά των συστάσεών του. Ο υπολογισμός της γίνεται κατά τρόπο αντίστοιχο με αυτόν της εμπιστοσύνης του εισηγητή. Συγκεκριμένα, αρχικά εντοπίζεται η σχέση  $(c, a, t)$  μέσα στο σύνολο  $R$  των εμπειριών εισήγησης, όπου  $t=(T_{vg}, T_g, T_b, T_{vb})$ . Στη συνέχεια γίνεται υπολογισμός του  $\text{mod}$  των συνόλων  $T_e$ , όπου  $e \in E$ . Η τιμή που προκύπτει είναι η σημασιολογική διαφορά  $sd_e$  του επιπέδου  $e$ . Εάν κάποιο σύνολο  $T_e$  είναι multi-modal, τότε υπάρχει αβεβαιότητα ως προς τις συστάσεις επιπέδου  $e$  από το συγκεκριμένο πράκτορα. Στην περίπτωση αυτή γίνεται η σύμβαση ότι  $sd_e=0$ . Τυπικά έχουμε ότι :

$$\forall e \in E, sd_e = \text{mod}(T_e) \quad (3)$$

### Αξιολόγηση Συστάσεων

Για να μπορέσει να γίνει αξιόπιστη χρήση των συστάσεων που λαμβάνονται από κάποιον πράκτορα, θα πρέπει πρώτα αυτές να αξιολογηθούν, υπό την έννοια ότι θα πρέπει να ληφθεί υπόψη η απόκλιση που έχει εμφανιστεί σε άλλες παρόμοιες συστάσεις από τον ίδιο πράκτορα στο παρελθόν. Αυτό επιτυγχάνεται με τον υπολογισμό της σημασιολογικής διαφοράς του πράκτορα εισηγητή, πάντα στο πλαίσιο μέσα στο οποίο γίνεται η σύσταση.

Έτσι, για μια σύσταση  $\text{rec}(a, b, c, rd)$  ενός πράκτορα  $a$  για κάποιον πράκτορα  $b$ , μέσα στο πλαίσιο  $c$ , επιπέδου  $rd$ , γίνεται αρχικά, με βάση με την (3), ο υπολογισμός της σημασιολογικής διαφοράς του πράκτορα εισηγητή  $a$  για τα διάφορα επίπεδα  $e \in E$ . Στη συνέχεια γίνεται προσαρμογή του επιπέδου  $rd$  της σύστασης χρησιμοποιώντας τον τύπο :

$$rd^{\circ} = rd \oplus sd_{rd}$$

όπου

- $sd_{rd}$  : η σημασιολογική διαφορά για το επίπεδο rd.
- Ο τελεστής  $\oplus$  υποδηλώνει τη λειτουργία «αύξηση κατά», η οποία βέβαια στη περίπτωση αρνητικών τιμών του δεξιού operand ισοδυναμεί με μείωση. Για παράδειγμα εάν  $rd = g$  και  $sd_g = -1$  τότε  $rd^{\circ} = b$ .

### Ενημέρωση Εμπειριών

Μετά από κάθε συναλλαγή ενός πράκτορα a, με έναν άλλο πράκτορα b, μέσα σε ένα πλαίσιο c, η αντίστοιχη εγγραφή του b μέσα στο σύνολο Q ενημερώνεται. Αυτό πρακτικά σημαίνει ότι, ανάλογα με το επίπεδο e της έκβασης της συναλλαγής, αυξάνεται ο αντίστοιχος μετρητής ( $s_e = s_e + 1$ ).

Επιπρόσθετα, εάν για τη συναλλαγή υπήρχε και κάποια σύσταση, τότε ενημερώνεται και το σύνολο R των εμπειριών σύστασης. Εδώ η ενημέρωση αφορά την προσθήκη της διαφοράς που καταγράφηκε ανάμεσα στη αρχική σύσταση και το τελικό αποτέλεσμα, μέσα στο αντίστοιχο σύνολο  $T_e$  των αποκλίσεων των συστάσεων επιπέδου e. Τυπικά :

$$T_{rd} = T_{rd} \cup \{ (e \diamond rd) \}$$

όπου

Ο τελεστής  $\diamond$  υποδηλώνει τη διαφορά επιπέδων μεταξύ των δύο operands. Για παράδειγμα εάν έγινε μια σύσταση με επιπέδου  $rd = vg$  και στην πράξη εμφανίστηκε αποτέλεσμα επιπέδου  $e = g$ , τότε η διαφορά είναι  $e \diamond rd = -1$ .

### Σύνθεση Συστάσεων

Κατά τη διαδικασία συλλογής συστάσεων ένας πράκτορας μπορεί να βρει περισσότερες από μια συστάσεις, από διαφορετικούς πράκτορες, τις οποίες θα πρέπει με κάποιο τρόπο να συνδυάσει ώστε να βγάλει τελικά ένα βαθμό εμπιστοσύνης για τον υπό εξέταση πράκτορα. Η διαδικασία σύνθεσης που προτείνει το μοντέλο είναι η εξής :

- Αρχικά υπολογίζεται, με βάση τον (2), ο βαθμός εμπιστοσύνης  $rtd_i$  ( $i=1..n$ ) των n πρακτόρων εισηγητών που έδωσαν τις συστάσεις τους. Μόνο οι γνωστοί πράκτορες (αυτοί για τους οποίους υπάρχει εγγραφή στο σύνολο R των εμπειριών συστάσεων) λαμβάνονται υπόψη. Εάν υπάρχουν συστάσεις και από άγνωστους πράκτορες, αυτές αγνοούνται αφού δεν μπορεί να αξιολογηθεί η αξιοπιστία τους.
- Ανάλογα με το  $rtd_i$ , δίνεται σε κάθε πράκτορα εισηγητή μια βαρύτητα, με βάση τον παρακάτω πίνακα :

$rtd_i$	0	1	2	3	Άγνωστο
Βάρος	9	5	3	1	0

- Γίνεται μια ομαδοποίηση των συστάσεων ανά επίπεδο και για κάθε διαφορετικό επίπεδο αθροίζονται οι βαρύτητες των πρακτόρων που εισηγήθηκαν αυτές τις συστάσεις. Δηλ. :

$$\forall e \in E \quad \forall w_i \in L_e, \text{sum}_e = \sum_{i=1}^{|L_e|} w_i$$

Ο ζητούμενος βαθμός εμπιστοσύνης είναι το επίπεδο  $e \in E$  με το μεγαλύτερο άθροισμα  $\text{sum}_e$ . Εάν το μέγιστο δεν είναι μοναδικό, τότε υπάρχει περίπτωση αβεβαιότητας και ο βαθμός εμπιστοσύνης προκύπτει με βάση τον πίνακα 1.

### **Νέοι πράκτορες**

Σε ότι αφορά τους νεοεισαχθέντες πράκτορες, που δεν έχουν καθόλου εμπειρία και είναι ευάλωτοι σε παραπλάνηση και παραπληροφόρηση (αφού δεν γνωρίζουν ποιόν να εμπιστευτούν), το μοντέλο προτείνει, ότι θα πρέπει, για να μειωθούν οι κίνδυνοι και το ρίσκο, οι δημιουργοί τους να τους εφοδιάσουν, πριν τους ελευθερώσουν στην ηλεκτρονική κοινότητα, με κάποια αρχικοποιημένα σύνολα Q και R. Έτσι θα έχουν κάποια έμπιστη βάση (έστω και μικρή) για να στηριχτούν και εκεί πάνω να χτίσουν την εμπειρία τους.

### **2.5.5 Σχολιασμός - κριτική**

Οι Abdul-Rahman και Hailes προσπάθησαν να μοντελοποιήσουν την εμπιστοσύνη με ένα ρεαλιστικό τρόπο χρησιμοποιώντας χαρακτηριστικά του πραγματικού κόσμου. Το μοντέλο που πρότειναν είναι απλό, χωρίς διφορούμενα στοιχεία και εύκολα υλοποιήσιμο στον ψηφιακό κόσμο των πρακτόρων εμπιστοσύνης. Μια σύνοψη των θετικών και αρνητικών στοιχείων του μοντέλου τους είναι :

#### **Θετικά**

- + Αξιοποιεί την παροχή συστάσεων και εισάγει ένα αποδοτικό μηχανισμό διαχείρισης φήμης
- + Εισάγει ένα αντικειμενικό σύστημα αξιολόγησης των εισηγητών των συστάσεων
- + Εισάγει και χρησιμοποιεί την έννοια της εμπιστοσύνης εισηγητή
- + Λαμβάνει υπόψη τα χαρακτηριστικά του πλαισίου των συναλλαγών (context)

#### **Αρνητικά**

- Αναπαριστά την εμπιστοσύνη με προκαθορισμένα επίπεδα τιμών περιορίζοντας τις δυνατότητες έκφρασης οποιασδήποτε περίπτωσης
- Αντιμετωπίζει με επιείκεια τους κακόβουλες συμπεριφορές. Έτσι, ένας πράκτορας μπορεί να συμπεριφέρεται άσχημα στο 49% των συναλλαγών του και παρόλα αυτά να θεωρείται έμπιστος

- Δεν αξιοποιεί συνδυασμένα την εμπειρία και την φήμη. Οι υπολογισμοί που προτείνει για την αξιολόγηση της εμπιστοσύνης χρησιμοποιούν είτε τις προηγούμενες εμπειρίες είτε τη υπάρχουσα φήμη.

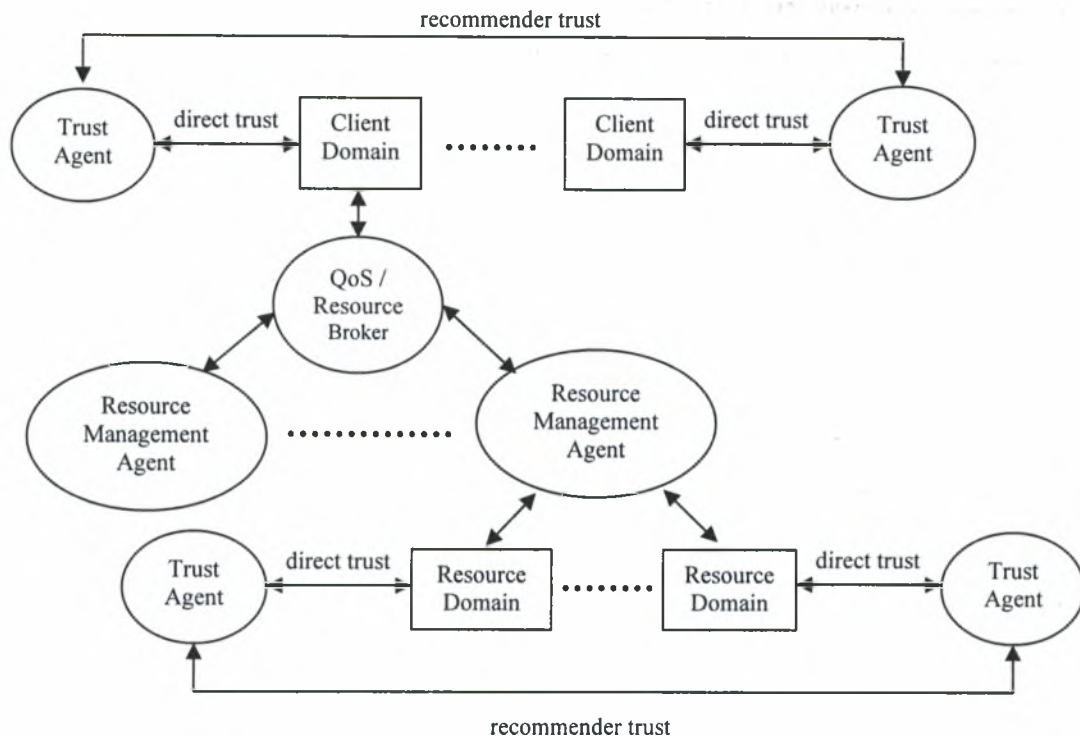
## 2.6 Το μοντέλο των *F. Azzerdin and M. Maheswaran*

Το 2002, σε συνέδριο της IEEE που έγινε στον Καναδά, οι F. Azzerdin και M. Maheswaran παρουσίασαν ένα μοντέλο εμπιστοσύνης προσανατολισμένο σε συστήματα Grid [35]. Έχοντας ως πεδίο εφαρμογής το κατανεμημένο περιβάλλον των Grids, μέσα στο οποίο αναπτύσσονται σχέσεις αλληλεπίδρασης κατά πρώτο λόγο μεταξύ κοινοτήτων και κατά δεύτερο λόγο μεταξύ οντοτήτων (χρηστών ή πόρων), προσπάθησαν να προσεγγίσουν την έννοια της εμπιστοσύνης από την πλευρά της συμπεριφοράς. Η προσέγγισή τους, συμπληρώνει και βελτιώνει την παραδοσιακή αντιμετώπιση της εμπιστοσύνης, η οποία στηρίζεται σχεδόν αποκλειστικά σε διαδικασίες αναγνώρισης της ταυτότητας και προσδιορισμού των δικαιωμάτων πρόσβασης (access rights). Έχοντας ως βάση τη δουλειά των Abdul-Rahman και Hailes [34], υιοθέτησαν αρκετές από τις σύγχρονες αντιλήψεις περί διαχείρισης εμπιστοσύνης και κατέληξαν σε μια πρόταση που διαθέτει τα απαιτούμενα χαρακτηριστικά για να ενσωματωθεί στην αρχιτεκτονική των συστημάτων Grid.

### 2.6.1 Παραδοχές- Χαρακτηριστικά

Το μοντέλο των Azzerdin & Maheswaran, ανήκει στην κατηγορία των βασισμένων στη φήμη (reputation based) μοντέλων. Ο μηχανισμός εκτίμησης της εμπιστοσύνης που διαθέτει, λαμβάνει υπόψη τόσο τις εμπειρίες από τις αλληλεπιδράσεις της οντότητας όσο και τη φήμη, δίνοντας σε κάθε συνιστώσα ξεχωριστή βαρύτητα.

Η προσέγγιση που ακολουθεί αντιμετωπίζει το χώρο ενός Grid ως ένα σύνολο περιοχών (Grid Domains – GD). Σε κάθε περιοχή οι οντότητες οργανώνονται και δραστηριοποιούνται μέσα σε δύο εικονικές περιοχές (virtual domains). Η μια περιοχή περιλαμβάνει τους πόρους (Resource Domain - RD) και η άλλη τους πελάτες (Client Domain - CD). Μέσα σε κάθε περιοχή GD υπάρχουν πράκτορες που διαθέτουν μηχανισμούς αφενός για να ενημερώνουν τις δομές αναπαράστασης της εμπιστοσύνης και αφετέρου για να επιτρέπουν σε άλλες οντότητες να ενταχθούν στην περιοχή κληρονομώντας τα χαρακτηριστικά της εμπιστοσύνης της. Μια απεικόνιση της προσέγγισης που ακολουθεί το μοντέλο, σε ότι αφορά τη δομή και τη λειτουργικότητα ενός Grid φαίνεται στο σχήμα 2-3.



Σχήμα 2-3 : Η εμπιστοσύνη στο Grid

Σχετικά με τον τρόπο χειρισμού της εμπιστοσύνης, το μοντέλο :

- Ορίζει την εμπιστοσύνη μέσα σε συγκεκριμένο πλαίσιο (context). Έτσι ένας πράκτορας A μπορεί να εμπιστευτεί έναν πράκτορα B για να χρησιμοποιήσει τους αποθηκευτικούς του πόρους, αλλά όχι για να του αναθέσει την εκτέλεση υπολογισμών.
- Ενημερώνει την εμπιστοσύνη με βάση τις εμπειρίες από τις απευθείας αλληλεπιδράσεις. Αλληλεπιδράσεις με θετική έκβαση αυξάνουν την εμπιστοσύνη, ενώ αρνητική έκβαση την μειώνουν.
- Εξαρτά την εμπιστοσύνη από το χρόνο, θεωρώντας ότι αυτή εξασθενεί με το πέρασμα του χρόνου. Έτσι ένας πράκτορας A μπορεί να εμπιστευόταν έναν πράκτορα B πριν από ένα χρόνο, αλλά σήμερα ίσως να μην τον εμπιστευτεί (ή τουλάχιστο όχι στο βαθμό που τον εμπιστευόταν), εκτός και εάν στο μεταξύ έλαβαν χώρα νέες αλληλεπιδράσεις με θετική έκβαση.

Επιπλέον, έχοντας ως πρώτη προτεραιότητα την δημιουργία ενός αποδοτικού και εφαρμόσιμου μηχανισμού, οι Azzerdin & Maheswaran προσπάθησαν να ελαχιστοποιήσουν την υπολογιστική επιβάρυνση του μοντέλου τους. Για το λόγο αυτό έκαναν κάποιες παραδοχές, που επιτρέπουν στο μοντέλο αφενός να έχει καλά χαρακτηριστικά επεκτασιμότητας (scalability) και αφετέρου να είναι υπολογιστικά εφικτό, παράμετρος ιδιαίτερα σημαντική για το απαιτητικό περιβάλλον των Grids. Συγκεκριμένα θεωρούν ότι :



- Όλοι οι πράκτορες (διαχειριστές πόρων ή πελάτες) κληρονομούν τα χαρακτηριστικά εμπιστοσύνης των περιοχών στις οποίες ανήκουν (RD ή CD).
- Η εμπιστοσύνη μεταβάλλεται με αργούς ρυθμούς. Για το λόγο αυτό η τιμή της ενημερώνεται μόνο εφόσον έχει λάβει χώρα ένας σημαντικός αριθμός αλληλεπιδράσεων.
- Τα διαφορετικά πλαίσια (contexts) εμπιστοσύνης περιορίζονται στους βασικότερους τύπους υπηρεσιών : εκτύπωση, αποθήκευση, υπολογισμοί.

## 2.6.2 Αναπαράσταση εμπιστοσύνης

Ο ορισμός της εμπιστοσύνης που υιοθετεί το μοντέλο είναι :

*«Εμπιστοσύνη είναι η ισχυρή πεποίθηση για την ικανότητα μιας οντότητας να ενεργήσει κατά τα αναμενόμενα, έτσι ώστε αυτή η ισχυρή πεποίθηση να μην είναι μια σταθερή τιμή, συσχετιζόμενη με την οντότητα, αλλά να υπόκειται στην συμπεριφορά της οντότητας και να ισχύει μέσα σε συγκεκριμένο πλαίσιο για δοθέν χρόνο.»*

Όπως αναφέρθηκε και προηγουμένως, η διαχείριση της εμπιστοσύνης γίνεται σε επίπεδο περιοχών. Έτσι, εάν  $D_i$  και  $D_j$  είναι δύο περιοχές Grid, η εμπιστοσύνη που αναπτύσσεται μεταξύ τους (από την περιοχή  $D_i$  προς την περιοχή  $D_j$ ), αναφορικά με ένα πλαίσιο  $c$ , κατά τη χρονική στιγμή  $t$  συμβολίζεται στο μοντέλο ως

$$\Gamma(D_i, D_j, t, c)$$

Η τιμή της κυμαίνεται μεταξύ έξι διακριτών επιπέδων, κλιμακούμενων από την ελάχιστη προς τη μέγιστη εμπιστοσύνη. Σε κάθε ένα επίπεδο έχει αντιστοιχηθεί ένα γράμμα (πίνακας 2-2)

<i>Επίπεδο Εμπιστοσύνης (Trust Level-TL)</i>	<i>Περιγραφή</i>
A	Πολύ χαμηλή εμπιστοσύνη
B	Χαμηλή εμπιστοσύνη
C	Μεσαία εμπιστοσύνη
D	Υψηλή εμπιστοσύνη
E	Πολύ Υψηλή εμπιστοσύνη
F	Εξαιρετικά υψηλή εμπιστοσύνη

**Πίνακας 2-2 : Επίπεδα εμπιστοσύνης**

Το επίπεδο F δεν αντιστοιχεί σε πραγματική εμπιστοσύνη (δεν επιτυγχάνεται ποτέ τόσο υψηλή εμπιστοσύνη), αλλά υπάρχει ως διακόπτης ασφάλειας και αποτροπής πραγματοποίησης οποιασδήποτε συναλλαγής. Περισσότερα θα αναφερθούν στην παράγραφο 2.6.4 .

Η εμπιστοσύνη συντίθεται από δύο συνιστώσες :

- i. Την άμεση εμπιστοσύνη (Direct Trust), που καθορίζεται από τις εμπειρίες προηγούμενων αλληλεπιδράσεων μεταξύ των περιοχών και η οποία στο μοντέλο συμβολίζεται ως :

$$\Theta(D_i, D_j, t, c)$$

- ii. Την εμπιστοσύνη φήμης (Reputation Trust), που αντικατοπτρίζει την εικόνα αξιοπιστίας που διαμορφώνεται μέσα από τη φήμη που δημιουργεί κάθε περιοχή και η οποία στο μοντέλο συμβολίζεται ως :

$$\Omega(D_j, t, c)$$

### 2.6.3 Δομές

Τα δεδομένα των συνιστωσών της εμπιστοσύνης κρατιούνται σε δύο πίνακες :

- τον πίνακα άμεσης εμπιστοσύνης (Direct Trust Table – DTT)
- τον πίνακα εμπιστοσύνης φήμης (Reputation Trust Table – RTT)

Στα περισσότερα πρακτικά συστήματα οι πράκτορες χρησιμοποιούν τις ίδιες πληροφορίες τόσο για να υπολογίσουν την άμεση εμπιστοσύνη όσο και για να δώσουν συστάσεις και συνεπώς οι πίνακες DTT και RTT είναι ίδιοι και κρατιούνται σε μια και μόνο δομή. Η δομή αυτή φαίνεται στο σχήμα 2-4 για την περιοχή  $D_k$ , όπου η τιμή  $TL_{kj}^{c_i}$  αντικατοπτρίζει το επίπεδο εμπιστοσύνης που έχει η περιοχή  $D_k$  προς την περιοχή  $D_j$  μέσα στο πλαίσιο  $c_i$

Πλαίσιο	Περιοχές		
	$D_1$	...	$D_n$
$c_1$	$TL_{k1}^{c_1}$	...	$TL_{kn}^{c_1}$
$c_2$	$TL_{k1}^{c_2}$	...	$TL_{kn}^{c_2}$
⋮	⋮	⋮	⋮
$c_m$	$TL_{km}^{c_m}$	...	$TL_{km}^{c_m}$

Σχήμα 2-4 : Δομή πινάκων DTT & RTT για την περιοχή  $D_k$ .

Δηλαδή για κάθε διαφορετικό πλαίσιο αποθηκεύονται η τιμές, μια για κάθε άλλη περιοχή του Grid με τους πράκτορες των οποίων υπάρχει (ή πρόκειται να υπάρξει) συνεργασία. Οι τιμές αυτές αντικατοπτρίζουν το επίπεδο εμπιστοσύνης που έχει η περιοχή  $D_k$  του πράκτορα που διατηρεί το πίνακα προς τις  $n$  άλλες περιοχές  $D_i$  ( $1 \leq i \leq n$ ).

Επιπλέον, διατηρείται και ο πίνακας R με τα recommender trust factors. Δυστυχώς το μοντέλο αναφέρει την ύπαρξη αυτού του πίνακα αλλά δεν περιγράφει τη δομή του.

## 2.6.4 Υπολογισμός της εμπιστοσύνης

Όταν ένας πράκτορας A μιας περιοχής  $D_i$  επιθυμεί να αλληλεπιδράσει με ένα πράκτορα B στην περιοχή  $D_j$  θα πρέπει να υπολογίσει την εμπιστοσύνη προς την περιοχή  $D_j$  ως συνάρτηση των δύο συνιστώσων, της άμεσης εμπιστοσύνης και της εμπιστοσύνης φήμης. Πιο συγκεκριμένα γίνεται ο υπολογισμός :

$$\Gamma(D_i, D_j, t, c) = \alpha \times \Theta(D_i, D_j, t, c) + \beta \times \Omega(D_j, t, c)$$

Οι παράγοντες  $\alpha$ ,  $\beta$  αποτελούν τη βαρύτητα της κάθε συνιστώσας και ρυθμίζονται ανάλογα με τις ιδιαίτερες συνθήκες εφαρμογής του μοντέλου και τις απαιτήσεις του κάθε συστήματος. Συνήθως, η εμπιστοσύνη που βασίζεται στις εμπειρίες, είναι πιο σημαντική από την εμπιστοσύνη που βασίζεται στη φήμη και συνεπώς, στα περισσότερα συστήματα, ισχύει ότι  $\alpha > \beta$ .

Η συνιστώσα της άμεσης εμπιστοσύνης υπολογίζεται ως το γινόμενο του επιπέδου εμπιστοσύνης (όπως αυτό είναι καταχωρημένο στον πίνακα DTT της άμεσης εμπιστοσύνης) επί μια συνάρτηση εξασθένησης (decay function), η χρήση της οποίας προσαρμόζει το επίπεδο εμπιστοσύνης στο παρόν. Δηλ.

$$\Theta(D_i, D_j, t, c) = DTT_{D_i}(D_j, c) \times Y(t - t_{ij}, c)$$

όπου

$t$  : η τρέχουσα χρονική στιγμή

$t_{ij}$  : η χρονική στιγμή κατά την οποία έλαβε χώρα η αλληλεπίδραση μεταξύ  $D_i$  και  $D_j$

Το μοντέλο δεν αναφέρει τίποτε για το τι υπολογίζει η συνάρτηση εξασθένησης, καθώς θεωρεί ότι η υλοποίησή της μπορεί να είναι ποικίλη, ανάλογα με τις ιδιαιτερότητες και τις απαιτήσεις της κάθε περιοχής εφαρμογής.

Η συνιστώσα της εμπιστοσύνης φήμης υπολογίζεται από τον τύπο :

$$\Omega(D_j, t, c) = \frac{\sum_{k=1}^n RTT_{D_k}(D_j, c) \times R(D_k, D_j) \times Y(t - t_{kj}, c)}{n}$$

όπου

$R(D_k, D_j)$  : ο παράγοντας εμπιστοσύνης εισηγητή (recommender trust factor) με τιμές στο διάστημα  $[0,1]$

Ο παραπάνω τύπος λέει ότι, ουσιαστικά, η εμπιστοσύνη φήμης είναι ο μέσος όρος της εικόνας εμπιστοσύνης που έχουν οι άλλες περιοχές για την υπό εξέταση περιοχή.

Ο παράγοντας R υποδεικνύει την αξιοπιστία του εισηγητή μιας σύστασης και χρησιμοποιείται για προστασία από προσπάθειες εξαπάτησης, σε περιπτώσεις συνεργαζόμενων κακόβουλων πρακτόρων

που δίνουν ψευδής συστάσεις ο ένας για τον άλλο. Μετά από κάθε αλληλεπίδραση (και εφόσον υπήρξαν συστάσεις για την περιοχή με την οποία έγινε η αλληλεπίδραση) αξιολογείται η εγκυρότητα των συστάσεων (μετά από σύγκριση με το αποτέλεσμα της αλληλεπίδρασης) και ενημερώνονται κατάλληλα τα στοιχεία του πίνακα R που αντιστοιχούν στις περιοχές που έδωσαν τις συστάσεις. Δυστυχώς, το μοντέλο, δεν περιγράφει κάποιον συγκεκριμένο τρόπο υπολογισμού του R, αφήνοντάς το ανοικτό για την υλοποίηση του μοντέλου.

Κάθε περιοχή ορίζει ένα απαιτούμενο επίπεδο εμπιστοσύνης (Requested Trust Level – RTL) για κάθε διαφορετικό πλαίσιο συναλλαγών. Αυτό το επίπεδο θα καθορίσει και ποιες από τις υποψήφιες συναλλαγές θα πραγματοποιηθούν. Μετά τον υπολογισμό της εμπιστοσύνης της περιοχής με την οποία διερευνάται η περίπτωση αλληλεπίδρασης, γίνεται σύγκριση αυτής της εμπιστοσύνης με το προκαθορισμένο RTL και εάν το καλύπτει η συναλλαγή προχωράει.

Το RTL, που παίρνει τιμές από την περιοχή τιμών της εμπιστοσύνης (πίνακας 2-2), μπορεί να λειτουργήσει και ως διακόπτης φραγής συναλλαγών εάν τεθεί στην υψηλότερη τιμή F που μπορεί να πάρει. Στην περίπτωση αυτή το RTL θα είναι απαγορευτικό, αφού δε θα υπάρχει περιοχή που να έχει τόσο υψηλή εμπιστοσύνη.

Μετά την ολοκλήρωση της συναλλαγής, και οι δύο πλευρές της αλληλεπίδρασης θα αξιολογήσουν την συναλλαγή ελέγχοντας το κατά πόσο τηρήθηκε ή παραβιάστηκε το RTL τους. Ως παραβίαση του RTL μπορεί να θεωρηθεί π.χ. η κατανάλωση περισσότερων πόρων από αυτούς που ζητήθηκαν ή η εκκίνηση μη προβλεπόμενων διεργασιών. Το ύψος της παραβίασης εξαρτάται από την κάθε περιοχή και από τις προδιαγραφές που έχει θέσει. Για παράδειγμα μια περιοχή με πολύ μεγάλους αποθηκευτικούς χώρους ίσως δεν θεωρήσει ιδιαίτερα σημαντική παραβίαση του RTL της την υπέρβαση της χρήσης των δίσκων της κατά 10 Gb. Μια άλλη περιοχή όμως, με περιορισμένους αποθηκευτικούς χώρους, μπορεί να εκλάβει αυτή τη συμπεριφορά ως εξαιρετικά κακή χρήση. Σε κάθε περίπτωση πάντως, η κάθε περιοχή θα αξιολογήσει την εμπειρία της συναλλαγής και θα καθορίσει το επίπεδο εμπιστοσύνης  $TL(t_{ij}, c)$  που της αναλογεί.

### 2.6.5 Ενημέρωση των δομών

Μετά από κάθε αλληλεπίδραση μεταξύ δύο περιοχών  $D_i$  και  $D_j$  και την επακόλουθη αξιολόγηση της άμεσης εμπιστοσύνης θα πρέπει να γίνει η ενημέρωση του πίνακα DTT. Ο τύπος που χρησιμοποιείται (από τη μεριά της περιοχής  $D_i$ ) είναι :

$$DTT_{D_i}(D_j, c) = (1-\delta) \times DTT_{D_i}(D_j, c) + \delta \times TL(t_{ij}, c)$$

όπου

$TL(t_{ij}, c)$  : η αξιολόγηση της άμεσης εμπιστοσύνης μετά από συναλλαγή μεταξύ  $D_i$  και  $D_j$  τη χρονική στιγμή  $t_{ij}$  από τη μεριά της περιοχής  $D_i$

Το  $\delta$  είναι ένας συντελεστής που ορίζεται σε κάθε περιοχή ξεχωριστά και παίρνει τιμές από το διάστημα  $[0,1]$ . Καθορίζει τη βαρύτητα επίδρασης του κάθε παράγοντα στη διαμόρφωση της νέας άμεσης εμπιστοσύνης. Έτσι εάν  $\delta > 0.5$ , σημαντικότερο ρόλο παίζει η εμπειρία της τρέχουσας συναλλαγής, ενώ για  $\delta < 0.5$  μετράει περισσότερο η εμπειρία των παλαιότερων συναλλαγών. Με παρόμοιο τρόπο γίνεται και η ενημέρωση του πίνακα RTT

### 2.6.6 Νέα μέλη & κληρονομικότητα

Στο δυναμικό περιβάλλον των Grids οντότητες εισέρχονται σε μια περιοχή και αποχωρούν απ' αυτή συνεχώς. Αυτό δημιουργεί την ανάγκη για αποδοτικές διαδικασίες αρχικοποίησης της κατάστασης των νέων μελών. Το μοντέλο των Azzerdin και Maheswaran στο σημείο αυτό εισάγει την έννοια της κληρονομικότητας. Κάθε νέα οντότητα που εγγράφεται σε μια περιοχή ως νέο μέλος, κληρονομεί την αξιολόγηση εμπιστοσύνης της περιοχής, όπως αυτή είναι καταγεγραμμένη στους πίνακες DTT και RTT.

Από την πλευρά του «έξω» κόσμου, τα νέα μέλη, προφανώς, δεν μπορούν να αντιμετωπίζονται με την ίδια εμπιστοσύνη που αντιμετωπίζονται και τα παλιά μέλη της περιοχής, των οποίων η συμπεριφορά έχει ήδη δοκιμασθεί και αξιολογηθεί. Για το λόγο αυτό το μοντέλο εισάγει τον παράγοντα βάρος μέλους (member weight). Το βάρος μέλους είναι ένας χαρακτηρισμός, που κυμαίνεται μεταξύ των τιμών «νέο» (new), «πρόσφατο» (recent) και «παλαιό» (old) και ο οποίος υποδεικνύει την παλαιότητα ενός μέλους μέσα στην περιοχή. Είναι στη δικαιοδοσία της κάθε περιοχής να αποφασίσει σε ποια κατηγορία εμπίπτει το κάθε μέλος της, καθώς και το πότε θα αλλάξει κατηγορία. Επίσης η κάθε περιοχή καθορίζει από μόνη της το πώς θα αντιμετωπίζει και θα εμπιστεύεται την κάθε κατηγορία βάρους μέλους κατά τις συναλλαγές της με άλλες περιοχές.

Νέα μέλη, όμως υπάρχουν όχι μόνο σε επίπεδο οντοτήτων εντός των ορίων μιας περιοχής, αλλά και σε επίπεδο περιοχών εντός των ευρύτερων ορίων της κοινότητας Grid. Έτσι μια περιοχή που εισέρχεται σε μια κοινότητα Grid αντιμετωπίζει το πρόβλημα του ποιον θα εμπιστευτεί, τουλάχιστο για το αρχικό διάστημα που θα ολοκληρώσει κάποιες συναλλαγές και θα έχει κάποια δεδομένα για να αξιολογήσει την εμπιστοσύνη των άλλων περιοχών. Για την περίπτωση αυτή, το μοντέλο προτείνει οι νέες περιοχές να θέτουν το RTL τους στην υψηλότερη τιμή F, ασφαλίζοντας έτσι τη χρήση των πόρων και των εφαρμογών της. Στη συνέχεια, καθώς θα συλλέγονται οι εμπειρίες από τις συναλλαγές θα μπορέσει να ρυθμίσει το RTL σε πιο χαλαρά επίπεδα αφού πλέον θα έχει διαμορφωθεί η εικόνα αξιοπιστίας των άλλων περιοχών και θα είναι σε θέση να κρίνει και να αποφασίσει για το ποιες αλληλεπιδράσεις θα επιτρέπει.

### 2.6.7 Σχολιασμός - κριτική

Οι Azzerdin και Maheswaran βασίστηκαν στην δουλειά των Abdul-Rahman και Hailes, την οποία βελτίωσαν στα εξής σημεία :

- i. Εισάγουν την έννοια της εξασθένησης, με το χρόνο, της επίδρασης των εμπειριών στη διαμόρφωση της εμπιστοσύνης
- ii. Υπολογίζουν την εμπιστοσύνη ως συνδυασμό των προσωπικών εμπειριών και της φήμης, με ρυθμιζόμενη βαρύτητα στην κάθε συνιστώσα.
- iii. Μελετούν τις σχέσεις εμπιστοσύνης μεταξύ κοινοτήτων και εισάγει την έννοια της κληρονομικότητας για τα νέα μέλη μιας κοινότητας.
- iv. Εστιάζουν το πεδίο εφαρμογής στην περιοχή των συστημάτων Grid που είναι ένας ραγδαία εξελισσόμενος χώρος, μέσα στον οποίο η ανάπτυξη σχέσεων εμπιστοσύνης είναι ζωτικής σημασίας.

Τα παραπάνω χαρακτηριστικά καθιστούν το μοντέλο τους ένα σημαντικό σημείο αναφοράς, ιδιαίτερα στο χώρο των συστημάτων Grid. Μια σύνοψη των θετικών και αρνητικών στοιχείων του μοντέλου είναι:

#### Θετικά

- + Εισάγει την έννοια της μείωσης της εμπιστοσύνης με το πέρασμα του χρόνου (Time Decay Function)
- + Είναι υλοποιήσιμο σε πραγματικό περιβάλλον grid
- + Λαμβάνει υπόψη τα χαρακτηριστικά των συναλλαγών (context)
- + Υπολογίζει την εμπιστοσύνη λαμβάνοντας υπόψη τόσο τις εμπειρίες από τις αλληλεπιδράσεις όσο και τη φήμη δίνοντας σε κάθε συνιστώσα ξεχωριστή βαρύτητα.

#### Αρνητικά

- Θεωρεί ότι η εμπιστοσύνη μεταβάλλεται με αργούς ρυθμούς και για το λόγο αυτό η ενημέρωση των δομών που χρησιμοποιεί γίνεται σε αραιά διαστήματα, αφού λάβει χώρα ένας σημαντικός αριθμός συναλλαγών
- Αφήνει πολλά ανοικτά θέματα (καθορισμός του member weight, ρύθμιση του RTL κλπ)

## **2.7 Το μοντέλο των Bin Yu & Munindar P. Singh**

Οι Bin Yu and Munindar Sign δημοσίευσαν το 2000 [36] ένα μοντέλο εμπιστοσύνης με κοινωνιολογικά χαρακτηριστικά, βασισμένο σε ένα μηχανισμό διαχείρισης φήμης (reputation management). Ο μηχανισμός αυτός επιτρέπει σε πράκτορες που απαρτίζουν ηλεκτρονικές κοινωνίες να αποφεύγουν τις αλληλεπιδράσεις με επικίνδυνες ή ανεπιθύμητες οντότητες. Η πρότασή τους κινείται στην κατεύθυνση της χαλαρής ασφάλειας (soft security) και έχει αποκεντρωτικά (decentralized) χαρακτηριστικά, καθώς υιοθετεί την αυτόνομη δράση των πρακτόρων χωρίς τη μεσολάβηση κάποιας τρίτης έμπιστης οντότητας.

### 2.7.1 Παραδοχές - Υπόβαθρο

Το μοντέλο επεκτείνει μια παλαιότερη πρόταση των συγγραφέων για τη δημιουργία κοινωνικών δικτύων για συλλογή πληροφοριών σε ηλεκτρονικές κοινωνίες [37, 38]. Στην αρχιτεκτονική αυτή, κάθε χρήστης σχετίζεται με ένα προσωπικό πράκτορα. Οι χρήστες υποβάλλουν αιτήματα στον πράκτορά τους και εκείνος αναλαμβάνει να τα εξυπηρετήσει, κάνοντας τις κατάλληλες επαφές και συναλλαγές με άλλους πράκτορες.

Όταν κάποιο αίτημα φθάσει σε ένα προσωπικό πράκτορα, αυτός αποφασίζει εάν αφορά τον χρήστη του. Στην περίπτωση που το αίτημα σχετίζεται με την εξειδίκευση του χρήστη, ο πράκτορας επιτρέπει στο αίτημα να φθάσει στο χρήστη του και να το δει. Αντίθετα, εάν ο πράκτορας κρίνει ότι ο χρήστης του δεν μπορεί να εξυπηρετήσει το αίτημα, μπορεί να απαντήσει με κάποιες αναφορές (referrals) προς άλλους πράκτορες που θεωρεί ότι είναι περισσότερο σχετικοί. Αυτό μπορεί να επαναληφθεί, μέχρι κάποιο ανώτατο όριο συνεχόμενων αναφορών, και τελικά να επιστραφεί πίσω στον πράκτορα που υπέβαλε το αρχικό αίτημα μια αλυσίδα αναφορών (referral chain).

Όταν ο αρχικός πράκτορας λάβει μια αλυσίδα αναφορών αποφασίζει εάν θα τη χρησιμοποιήσει ή αν θα την απορρίψει. Στην περίπτωση που λάβει μια κανονική απάντηση, θα τη χρησιμοποιήσει για να αξιολογήσει τον πράκτορα που την έδωσε. Μέσα από την αξιολόγηση αυτή, οι πράκτορες ενημερώνουν την εσωτερική τους αναπαράσταση της ηλεκτρονικής κοινωνίας μέσα στην οποία ζουν. Επιπλέον, με βάση τα αποτελέσματα της αξιολόγησης, επιλέγουν τους γείτονές τους, δηλ. τους πράκτορες με τους επιθυμούν να επικοινωνούν και να συνεργάζονται.

Μια αλυσίδα αναφορών από ένα πράκτορα  $A_0$  προς ένα πράκτορα  $A_n$  συμβολίζεται στο μοντέλο ως  $\langle A_0, A_1, \dots, A_n \rangle$ . Σε μια τέτοια αλυσίδα ο πράκτορας  $A_{i+1}$  είναι γείτονας του  $A_i$ . Οι αλυσίδες αναφορών που προκύπτουν κατά τη διαδικασία εξυπηρέτησης ενός αιτήματος, επιτρέπουν τη δημιουργία μιας εκτίμησης εμπιστοσύνης  $T_0(n)$  ενός πράκτορα  $A_0$  προς ένα (πιθανώς) άγνωστο πράκτορα  $A_n$  και συνιστούν ένα κατευθυνόμενο γράφο, που ονομάζεται δίκτυο εμπιστοσύνης (trust net).

Το μοντέλο έχει έντονες κοινωνιολογικές επιδράσεις και το κεντρικό του σημείο είναι ο μηχανισμός διαχείρισης φήμης που διαθέτει. Σύμφωνα με την προσέγγιση που ακολουθεί, ένας πράκτορας αξιολογητής αξιολογεί έναν άλλο πράκτορα και του αποδίδει μια εκτίμηση αξιοπιστίας με βάση :

- α) τις προσωπικές του παρατηρήσεις από τις μεταξύ τους αλληλεπιδράσεις και
- β) τις εκτιμήσεις των γειτόνων του αξιολογούμενου πράκτορα σε συνδυασμό με τη εκτίμηση της αξιοπιστίας των γειτόνων από τον πράκτορα αξιολογητή.

Για την επίδραση των αλληλεπιδράσεων μεταξύ πρακτόρων στην εκτίμηση της αξιοπιστίας γίνεται από το μοντέλο η παραδοχή ότι η έκβαση μιας αλληλεπίδρασης μπορεί να είναι είτε συνεργασία (cooperation) είτε απόρριψη συνεργασίας (defection). Στην πρώτη περίπτωση η αλληλεπίδραση θα έχει θετική επίδραση στην αξιολόγηση της εμπιστοσύνης του πράκτορα που αποδέχτηκε τη συνεργασία ενώ

στη δεύτερη η αξιοπιστία του πράκτορα θα υποστεί μείωση. Δεν γίνεται καθόλου λόγος για αξιολόγηση και διαβάθμιση του αποτελέσματος της συνεργασίας θεωρώντας ότι αυτό είναι πάντα μέγιστα θετικό.

Η εκτίμηση της αξιοπιστίας είναι μονοδιάστατη, δηλαδή υπολογίζεται μια μόνο τιμή που δεν εξαρτάται από την ειδικευση (expertise) του κάθε πράκτορα ή το πλαίσιο μέσα στο οποίο λαμβάνουν χώρα οι συναλλαγές.

## 2.7.2 Αναπαράσταση της εμπιστοσύνης

Οι Yu & Sign ακολούθησαν το φορμαλισμό του Marsh [1], σύμφωνα με τον οποίο η εμπιστοσύνη αναπαρίσταται με ένα πραγματικό αριθμό στην περιοχή  $[-1, 1)$  και μεταβάλλεται με το χρόνο. Έτσι, η εκτίμηση εμπιστοσύνης ενός πράκτορα  $j$  από ένα πράκτορα  $i$  κατά την χρονική στιγμή  $t$  συμβολίζεται ως

$$T_i(j)^t$$

όπου :  $-1 < T_i(j)^t < 1$  και  $T_i(j)^0 = 0$

Επίσης, υιοθετώντας πάλι την προσέγγιση του Marsh, το μοντέλο ορίζει ότι η απόφαση για το εάν ο πράκτορας  $i$  θα εμπιστευτεί τον πράκτορα  $j$  ώστε τελικά να προχωρήσει σε συνεργασία μαζί του, καθορίζεται από ένα άνω κατώφλι εμπιστοσύνης  $\omega_i$  και ένα κάτω κατώφλι  $\Omega_i$ , που κρατά ο  $i$ . Για τα  $\omega_i$  και  $\Omega_i$  ισχύουν τα εξής :

$$-1 < \omega_i < 1, \quad -1 < \Omega_i < 1 \text{ και } \omega_i > \Omega_i$$

$$\text{εάν } T_i(j)^t \geq \omega_i \quad \Rightarrow \quad \text{ο } i \text{ θα εμπιστευτεί τον } j \text{ και θα συνεργαστεί μαζί του}$$

$$\text{εάν } T_i(j)^t \leq \Omega_i \quad \Rightarrow \quad \text{ο } i \text{ θα απορρίψει τον } j$$

$$\text{εάν } \Omega_i < T_i(j)^t < \omega_i \quad \Rightarrow \quad \text{υπάρχει αβεβαιότητα και ο } i \text{ θα πρέπει να βρει επιπλέον πληροφορίες ή θα πρέπει να εκτιμήσει την εμπιστοσύνη με βάση άλλους παράγοντες.}$$

## 2.7.3 Υπολογισμοί

### Ενημέρωση Εμπιστοσύνης

Κάθε φορά που ένας πράκτορας  $i$  αλληλεπιδρά με ένα άλλο πράκτορα  $j$ , ενημερώνει την εκτίμησή του για την αξιοπιστία του  $j$ , ανάλογα με την έκβαση της αλληλεπίδρασης. Η συνεργασία θα δημιουργήσει μια θετική εμπειρία (στο μοντέλο αναφέρεται ως μαρτυρία - evidence) μεγέθους  $\alpha$ , ενώ η απόρριψη της συνεργασίας μια αρνητική εμπειρία μεγέθους  $\beta$ . Η εμπειρία αυτή χρησιμοποιείται από τον  $i$  για την ενημέρωση της εμπιστοσύνης του προς τον πράκτορα  $j$  σύμφωνα με τον παρακάτω πίνακα :



$T_i(j)^t$	Συνεργασία με τον $j$	Απόρριψη από τον $j$
$>0$	$T_i(j)^{t+1} = T_i(j)^t + \alpha (1 - T_i(j)^t)$	$T_i(j)^{t+1} = (T_i(j)^t + \beta) / (1 - \min\{ T_i(j)^t ,  \beta \})$
$<0$	$T_i(j)^{t+1} = (T_i(j)^t + \alpha) / (1 - \min\{ T_i(j)^t ,  \alpha \})$	$T_i(j)^{t+1} = T_i(j)^t + \beta (1 + T_i(j)^t)$
$=0$	$T_i(j)^{t+1} = \alpha$	$T_i(j)^{t+1} = \beta$

Το μοντέλο, προσπαθώντας να δημιουργήσει ένα κίνητρο συνεργασίας (ή γενικότερα «καλής» συμπεριφοράς), προτείνει τα  $\alpha$  και  $\beta$  να παίρνουν τιμές που να ικανοποιούν τη σχέση  $|\alpha| < |\beta|$ . Αυτό πρακτικά σημαίνει ότι η καλή φήμη θα χτίζεται δύσκολα, ενώ θα καταστρέφεται εύκολα, ακόμη και με λίγες «κακές» συμπεριφορές.

### Διάδοση της φήμης

Στο μοντέλο, κάθε πράκτορας διατηρεί ένα σύνολο γειτόνων, δηλ. πρακτόρων με τους οποίους έχει κοινά ενδιαφέροντα, τους οποίους εμπιστεύεται και με τους οποίους συνεργάζεται. Οι μαρτυρίες αυτών των γειτόνων χρησιμοποιούνται για τον υπολογισμό της φήμης άλλων, άγνωστων κυρίως, πρακτόρων.

Για το σκοπό αυτό ορίζεται ο τελεστής διάδοσης φήμης  $\otimes$  (για τη φήμη ενός πράκτορα  $y$ , όπως αυτή διαδίδεται από τον πράκτορα  $x$ ):

$$x \otimes y = \begin{cases} x \times y, & \text{εάν } (x \geq 0 \wedge y \geq 0) \\ -|x \times y|, & \text{διαφορετικά} \end{cases}$$

Ο παραπάνω ορισμός εξασφαλίζει ότι το επίπεδο εμπιστοσύνης που διαδίδεται μέσα από ένα «αρνητικό» σύνδεσμο, θα είναι πάντα αρνητικό. Αυτό πρακτικά σημαίνει ότι ακόμη και μια καλή φήμη, εάν μεταδοθεί από ένα αναξιόπιστο πράκτορα, θα αξιολογηθεί ως κακή.

Με χρήση του παραπάνω τελεστή διάδοσης φήμης ορίζεται η διάδοση της φήμης μέσα από μια αλυσίδα αναφορών  $\chi = \langle A_0, \dots, A_n \rangle$  από τον πράκτορα  $A_0$ , στον πράκτορα  $A_n$  κατά τη χρονική στιγμή  $t$ . Η αξιολόγηση της εμπιστοσύνης οποιουδήποτε πράκτορα  $k$  αυτής της αλυσίδας από τον αρχικό πράκτορα  $A_0$ , δίνεται από τη σχέση:

$$T_0^x(k)^t = T_0^x(1)^t \otimes T_1^x(2)^t \otimes \dots \otimes T_{k-1}^x(k)^t$$

όπου:  $0 \leq k \leq n$

Ο προτελευταίος πράκτορας  $n-1$  μιας αλυσίδας αναφορών έχει άμεση εμπειρία του υπό αξιολόγηση τελευταίου πράκτορα  $n$  και γι' αυτό ονομάζεται αυτόπτης μάρτυρας.

Η μαρτυρία (evidence) ενός πράκτορα  $k$  που δίνεται στον πράκτορα  $0$ , μέσα από μια αλυσίδα αναφορών  $\chi = \langle A_0, \dots, A_n \rangle$ , ορίζεται ως

$$E_0^x(k)^t = T_0^x(k)^t \cdot T_k^x(k+1)^t$$

Το μοντέλο προτείνει να λαμβάνονται υπόψη μόνο οι μαρτυρίες των επαρκώς αξιόπιστων πρακτόρων. Ως επαρκώς αξιόπιστος (ή απλά αξιόπιστος) θεωρείται ένας πράκτορας εάν τυγχάνει τουλάχιστον θετικής αξιολόγησης εμπιστοσύνης<sup>12</sup>.

Με την προϋπόθεση αυτή(των επαρκώς αξιόπιστων μαρτύρων), το μοντέλο δίνει τη δυνατότητα και στους λιγότερο αξιόπιστους (όχι όμως αναξιόπιστους) πράκτορες να συμμετέχουν στη διάδοση της φήμης.

Μέσα σε ένα δίκτυο εμπιστοσύνης, κατά την αξιολόγηση ενός πράκτορα  $k$  από ένα πράκτορα  $i$ , είναι πιθανό να προκύψουν δύο ή περισσότερες, διαφορετικές αλυσίδες αναφορών προς τον  $k$  που να έχουν την ίδια τελική μαρτυρία (του πράκτορα  $k-1$ ). Στην περίπτωση αυτή το μοντέλο προτείνει να ληφθεί υπόψη η αλυσίδα που δίνει στον  $k-1$  την υψηλότερη αξιολόγηση εμπιστοσύνης<sup>13</sup>.

### **Σύνθεση μαρτυριών από διαφορετικούς μάρτυρες**

Κατά την αξιολόγηση ενός πράκτορα  $n$ , μπορεί να εμφανιστούν πολλές διαφορετικές μαρτυρίες από διάφορους πράκτορες μάρτυρες, μέσα από διαφορετικές αλυσίδες αναφορών. Οι μαρτυρίες αυτές θα πρέπει να συνδυαστούν σε μια αξιολόγηση του πράκτορα  $n$ . Πριν δούμε όμως πως γίνεται αυτό, θα πρέπει να διερευνήσουμε λίγο περισσότερο την έννοια του συνόλου μαρτυριών.

Είναι σημαντικό, από ένα σύνολο μαρτυριών να μπορούμε να απομονώνουμε τα παρακάτω υποσύνολα :

- α) το υποσύνολο των διαφορετικών μαρτυριών (απομακρύνοντας τις διπλές μαρτυρίες)
- β) το υποσύνολο των υψηλότερων μαρτυριών μόνο των αξιόπιστων μαρτύρων<sup>14</sup>.

Συμβολίζοντας με  $e_w$  το μάρτυρα μιας μαρτυρίας  $e$ , έχουμε τους παρακάτω ορισμούς :

### **Ορισμός : Διακριτό Σύνολο Μαρτυριών (Distinct Set of Testimonies)**

Ένα σύνολο μαρτυριών  $E = \{e_1, \dots, e_L\}$  για κάποιον πράκτορα  $n$  είναι διακριτό (distinct) αν και μόνο αν οι αυτόπτες μάρτυρες όλων των μαρτυριών του  $E$  είναι διακριτοί. Δηλ. εάν  $|\{e_{1w}, \dots, e_{Lw}\}| = L$

### **Ορισμός : Μέγιστα Αξιόπιστο Διακριτό υποσύνολο συνόλου μαρτυριών (Maximally Reliable Distinct – MRD subset of a set of testimonies)**

Ένα υποσύνολο  $E'$  ενός συνόλου μαρτυριών  $E$  είναι Μέγιστα Αξιόπιστο Διακριτό (MRD) αν και μόνο αν :

<sup>12</sup> Ένας πράκτορας  $k$ , πάνω σε μια αλυσίδα αναφορών  $\chi$  από ένα πράκτορα  $i$ , θεωρείται αξιόπιστος, , αν και μόνο αν,  $T_i^x(k)^t > 0$ .

<sup>13</sup> Μεταξύ δύο αλυσίδων αναφορών  $\chi_1$  και  $\chi_2$  προς ένα πράκτορα  $k$  από ένα πράκτορα  $i$  λέμε ότι η  $\chi_1$  δίνει στον  $k$  υψηλότερη αξιολόγηση εμπιστοσύνης αν και μόνο αν  $T_i^{x_1}(k)^t > T_i^{x_2}(k)^t$

<sup>14</sup> Δηλ. λαμβάνονται υπόψη μόνο οι αξιόπιστοι μάρτυρες και απ' αυτούς μόνο η υψηλότερη μαρτυρία τους

- είναι διακριτό
- $E' \subseteq E$
- $\forall e : (e \in E \wedge T_i^{z_e}(e_w) > 0) \Rightarrow (\exists e' : e' \in E' \wedge e'_w = e_w \wedge T_i^{z_{e'}}(e'_w) \geq T_i^{z_e}(e_w))$

Με βάση τους παραπάνω ορισμούς, για μπορούμε να συνθέσουμε ένα πλήθος μαρτυριών για κάποιο πράκτορα  $A_n$ , που περιέχονται σε ένα σύνολο μαρτυριών  $E$  ακολουθούμε τα εξής βήματα :

1. Βρίσκεται το MRD υποσύνολο  $E'$  του συνόλου  $E$
2. Υπολογίζεται ο μέσος όρος  $\bar{e}$  των μαρτυριών του  $E'$   

$$\bar{e} = \frac{1}{L} \sum_{i=1}^{|E'|} e'_i$$
3. Υπολογίζεται η εκτίμηση εμπιστοσύνης του πράκτορα  $A_n$  από τον  $A_0$  με χρήση του παρακάτω πίνακα :

Εάν	Τότε
$T_0(n)^t > 0 \wedge \bar{e} > 0$	$T_0(n)^{t+1} = T_0(n)^t + \bar{e} (1 - T_0(n)^t)$
$T_0(n)^t > 0 \vee \bar{e} > 0$	$T_0(n)^{t+1} = T_0(n)^t + \bar{e} / (1 - \min\{ T_0(n)^t ,  \bar{e} \})$
$T_0(n)^t < 0 \wedge \bar{e} < 0$	$T_0(n)^{t+1} = T_0(n)^t + \bar{e} (1 + T_0(n)^t)$

### Κουτσομπολιό (gossip)

Οι Yu και Sign, προσπαθώντας να μεταφέρουν στο μοντέλο τους μηχανισμούς από τον πραγματικό κόσμο, προβλέπουν και μια κοινωνική συμπεριφορά των πρακτόρων ανάλογη με αυτή που οι ανθρώπινες κοινωνίες αποκαλούν κουτσομπολιό. Όταν ένας πράκτορας  $A$  συναντήσει ένα «κακό» ή γενικά μη συνεργάσιμο πράκτορα  $B$ , πέρα από το να μειώσει την εμπιστοσύνης του προς αυτόν, μπορεί να τον τιμωρήσει επιπλέον, ενημερώνοντας τους γείτονες για την κακή διαγωγή του  $B$ .

Ο μηχανισμός του κουτσομπολιού είναι διαφορετικός απ' αυτόν των αλυσίδων αναφορών, γιατί στο κουτσομπολιό ένας πράκτορας διαδίδει φήμη χωρίς αυτό να του έχει προηγούμενα ζητηθεί. Είναι καθαρά μια κοινωνική συμπεριφορά που αποσκοπεί στην ενίσχυση της ασφάλειας και της αξιοπιστίας της κοινωνίας στο σύνολό της. Επειδή όμως το κουτσομπολιό είναι μια οικειοθελής ενέργεια, η οποία στην περίπτωση κακόβουλων πρακτόρων μπορεί να δημιουργήσει πρόβλημα (με τη διάδοση ανυπόστατων φημών-συκοφαντιών), στο μοντέλο λαμβάνονται υπόψη μόνο εκτιμήσεις που προέρχονται από αξιόπιστους πράκτορες.

Έτσι, εάν ο πράκτορας  $i$  λάβει ένα μήνυμα  $T_k(n)$  από τον πράκτορα  $k$  για τον πράκτορα  $n$ , τότε εάν ο  $k$  είναι δεν είναι αξιόπιστος (δηλ.  $T_i(k) < 0$ ) το μήνυμα αγνοείται. Διαφορετικά ο  $i$  ενημερώνει την εικόνα του για τον  $n$  με βάση τον παρακάτω πίνακα :

Εάν	Τότε
$T_i(n)^t > 0 \wedge T_k(n) > 0$	$T_i(n)^{t+1} = T_i(n)^t + T_i(k)^t T_k(n) (1 - T_i(n)^t)$
$T_i(n)^t < 0 \wedge T_k(n) < 0$	$T_i(n)^{t+1} = T_i(n)^t + T_i(k)^t T_k(n) (1 + T_i(n)^t)$
$T_i(n)^t \times T_k(n) < 0$ (δηλ. με αντίθετα πρόσημα)	$T_i(n)^{t+1} = (T_i(n)^t + T_i(k)^t T_k(n)) / (1 - \min\{ T_i(n)^t ,  T_i(k)^t T_k(n) \})$

Σε όλους τους παραπάνω υπολογισμούς η εκτίμηση εμπιστοσύνης ποτέ δεν ξεφεύγει από την περιοχή τιμών (-1,1) όπως ορίζεται και στην παράγραφο 2.7.2

### 2.7.4 Απεικόνιση των ιδιοτήτων της εμπιστοσύνης

Ας δούμε πως απεικονίζονται από το μοντέλο οι σημαντικότερες ιδιότητες της έννοιας της εμπιστοσύνης.

#### Συμμετρικότητα (Symmetry)

Όπως είναι γνωστό η εμπιστοσύνη, γενικά, δεν είναι συμμετρική σχέση, αφού δύο πράκτορες μπορεί να εμπιστεύονται ο ένας τον άλλο σε διαφορετικό βαθμό. Παρόλα αυτά, μακροπρόθεσμα στο μοντέλο, μεταξύ αλληλεπιδρόμενων πρακτόρων, επέρχεται μια σχετική εξισορρόπηση της εικόνας εμπιστοσύνης του ενός προς τον άλλο. Αυτά συμβαίνει διότι εάν μεν και οι δύο πράκτορες είναι συνεργάσιμοι, βαθμιαία, θα αυξήσουν την εμπιστοσύνη τους προς τον άλλον, συγκλίνοντας προς το ανώτερο όριο του 1. Από την άλλη, εάν ο ένας από τους δύο πράκτορες αρνείται τη συνεργασία (λόγω της χαμηλής εκτίμησης εμπιστοσύνης που έχει προς τον άλλο), τότε ο άλλος πράκτορας θα μειώσει την εμπιστοσύνη του προς αυτόν, φθάνοντας στα δικά του επίπεδα. Γενικά λοιπόν, μεταξύ δύο πρακτόρων x, y θα έχουμε ότι :

$$T_x(y)^t \approx T_y(x)^t \text{ όταν } t \rightarrow \infty$$

#### Μεταβατικότητα (Transitivity)

Επίσης η μεταβατικότητα δεν ισχύει, γενικά, για την σχέση εμπιστοσύνη. Όμως, στο μοντέλο, μεταξύ ορθολογικών (rational) πρακτόρων θα ισχύει η μεταβατικότητα, δηλαδή ότι :

$$(T_x(y)^t > T_x(z)^t) \wedge (T_x(z)^t > T_x(w)^t) \Rightarrow (T_x(y)^t > T_x(w)^t)$$

#### Αυτοενίσχυση (Self-reinforcement)

Σύμφωνα με την οπτική με την οποία το μοντέλο αντιμετωπίζει την εμπιστοσύνη, υπάρχει μια σαφής τάση ενίσχυσης της ήδη υπάρχουσας διάθεσης μεταξύ των πρακτόρων. Κι αυτό γιατί οι πράκτορες που εμπιστεύονται ο ένας τον άλλο θα τείνουν να συνεργάζονται, αυξάνοντας την μεταξύ τους εμπιστοσύνη, ενώ αυτοί που δεν εκτιμούν ο ένας τον άλλο (χαμηλή εμπιστοσύνη) θα αποφεύγουν τη συνεργασία, επιδεινώνοντας περαιτέρω την ήδη μειωμένη εμπιστοσύνη. Αυτό βέβαια δεν αποτελεί γενική ιδιότητα

της εμπιστοσύνης, αλλά πηγάζει από την (ειδική) παραδοχή του μοντέλου ότι μια αλληλεπίδραση μεταξύ πρακτόρων θα οδηγήσει είτε σε συνεργασία είτε σε απόρριψη της συνεργασίας και ότι στην περίπτωση της συνεργασίας το αποτέλεσμα είναι πάντα θετικό (λιγότερο ή περισσότερο).

Τα παραπάνω μπορούν να εκφραστούν, χρησιμοποιώντας το φορμαλισμό του μοντέλου, ως εξής :

$$\text{Εάν } (T_x(y)^t > \omega_x) \wedge (T_y(x)^t > \omega_y) \Rightarrow (T_x(y)^{t+1} \geq T_x(y)^t) \wedge (T_y(x)^{t+1} \geq T_x(y)^t)$$

$$\text{Εάν } (T_x(y)^t < \Omega_x) \wedge (T_y(x)^t < \Omega_y) \Rightarrow (T_x(y)^{t+1} \leq T_x(y)^t) \wedge (T_y(x)^{t+1} \leq T_x(y)^t)$$

### **Διάδοση (Propagation)**

Διαισθητικά γνωρίζουμε ότι η εμπιστοσύνη που δείχνουμε σε κάποιον άγνωστο που μας τον έχει συστήσει κάποιος τρίτος, εξαρτάται κυρίως από την εμπιστοσύνη μας προς τον τρίτο. Και βέβαια, είναι φυσικό η εμπιστοσύνη που θα δείξουμε προς το άγνωστο πρόσωπο να μην υπερβαίνει ούτε την εμπιστοσύνη μας προς το τρίτο πρόσωπο αλλά ούτε και την εμπιστοσύνη του τρίτου πρόσωπου προς τον άγνωστο.

Η αρχή αυτή τηρείται και στο μοντέλο των Yu και Sign. Έτσι, για κάποιον ορθολογικό πράκτορα x, θα ισχύει ότι :

$$(T_x(z)^{t+1} \leq T_x(y)^t) \wedge (T_x(z)^{t+1} \leq T_y(z)^t)$$

### **2.7.5 Σχολιασμός - κριτική**

Οι Yu και Sign προσπαθούν με ένα όσο το δυνατόν πιο απλό τρόπο να αναπαραστήσουν την έννοια της εμπιστοσύνης στον κόσμο των ηλεκτρονικών κοινωνιών και των πρακτόρων. Δανείζονται γνωστούς μηχανισμούς των ανθρωπίνων κοινωνιών, όπως είναι η αναζήτηση συστάσεων και το κουτσομπολιό, για να δημιουργήσουν ένα πρακτικό υπολογιστικό μοντέλο που να είναι εύκολα υλοποιήσιμο. Σε πολλά σημεία ακολουθούν το φορμαλισμό και τις παραδοχές του Marsh, αντιμετωπίζοντας όμως αρκετά από τα κενά και τα προβλήματα που εμφάνιζε η αρχική πρότασή του. Γενικά, το μοντέλο τους αντιμετωπίζει, με απλουστεύσεις βέβαια, αλλά πολύ πρακτικά, το πρόβλημα της εμπιστοσύνης. Τα θετικά και αρνητικά στοιχεία του μοντέλου είναι :

#### **Θετικά**

- + Η εμπιστοσύνη αναπαρίσταται ως πραγματικός αριθμός στην περιοχή (-1,1) επιτρέποντας την έκφραση και της έλλειψης εμπιστοσύνης αλλά και της αδυναμίας υπολογισμού της
- + Η φήμη χτίζεται δύσκολα και χαλάει εύκολα
- + Χρησιμοποιεί το μηχανισμό του κουτσομπολιού

#### **Αρνητικά**

- Δεν λαμβάνεται υπόψη το πλαίσιο της αλληλεπίδρασης (context)

- Δεν υπάρχει πρόβλεψη για την αντιμετώπιση ψευδών φημών που διαδίδονται από κακόβουλους πράκτορες.
- Κατά τη διάδοση της φήμης, μπορεί μια καλή φήμη, εάν διαδοθεί από ένα κακό πράκτορα, να αξιολογηθεί ως κακή, αφού στο μοντέλο, λόγω του τελεστή διάδοσης φήμης, μέσα από αρνητικά μονοπάτια φθάνει πάντοτε αρνητική φήμη. Ενίοτε όμως, ακόμα και οι κακοί πράκτορες μπορεί να συμπεριφερθούν καλά.
- Στην περίπτωση πολλαπλών μονοπατιών αξιολόγησης προς ένα αυτόπτη μάρτυρα, το μοντέλο εμφανίζεται "αισιόδοξο" λαμβάνοντας υπόψη τη διαδρομή με την υψηλότερη αξιολόγηση εμπιστοσύνης προς τον μάρτυρα.
- Στο μηχανισμό του κουτσομπολιού λαμβάνονται υπόψη μόνο οι θετικές μαρτυρίες



# ΚΕΦΑΛΑΙΟ

## 3. Τα Χαρακτηριστικά των Μοντέλων Εμπιστοσύνης

Όπως φαίνεται από την παρουσίαση κάποιων αντιπροσωπευτικών μοντέλων εμπιστοσύνης, που έγινε στο προηγούμενο κεφάλαιο, υπάρχει ένα σύνολο βασικών χαρακτηριστικών που είναι κοινά σε όλα τα μοντέλα. Υπάρχει επίσης ένα σύνολο βασικών λειτουργιών ενός πράκτορα και μιας κοινωνίας πρακτόρων, πάνω στο οποίο βασίζονται είτε έμμεσα είτε άμεσα όλα τα μοντέλα εμπιστοσύνης. Στο κεφάλαιο αυτό καταγράφονται συστηματικά τα κοινά χαρακτηριστικά των μοντέλων εμπιστοσύνης και οι βασικές λειτουργίες ενός πράκτορα που δρα σε μια κοινωνία πρακτόρων, τα οποία, όπως θα δούμε στο επόμενο κεφάλαιο, αποτελούν τη βάση για τη σχεδίαση της πλατφόρμας προσομοίωσης.

### 3.1 Βασικά χαρακτηριστικά των μοντέλων εμπιστοσύνης

Τα βασικά χαρακτηριστικά που μοιράζονται όλα τα αντιπροσωπευτικά μοντέλα εμπιστοσύνης που μελετήθηκαν είναι :

- **Αναπαράσταση της εμπιστοσύνης**

Όλα τα μοντέλα υιοθετούν μια μορφή αναπαράστασης της εμπιστοσύνης. Η μορφή αυτή είναι σημαντική γιατί καθορίζει σε μεγάλο βαθμό την εκφραστικότητα του μοντέλου, τις απαιτήσεις του σχετικά με τον τύπο δεδομένων που διαχειρίζεται και τη λειτουργικότητά του σχετικά με το είδος των υπολογισμών που χρησιμοποιεί. Οι διαφορετικές μορφές αναπαραστάσεις και εντοπίστηκαν στα μοντέλα που μελετήσαμε είναι :

- θετικοί ακέραιοι αριθμοί
- θετικοί & αρνητικοί ακέραιοι αριθμοί
- θετικοί πραγματικοί αριθμοί
- θετικοί & αρνητικοί πραγματικοί αριθμοί
- λεκτικοί χαρακτηρισμοί που υποδεικνύουν διαφορετικά επίπεδα εμπιστοσύνης

Κάθε μια από τις παραπάνω αναπαραστάσεις έχει πλεονεκτήματα και μειονεκτήματα. Βέβαια, πρέπει να σημειωθεί ότι υπάρχει μια ξεκάθαρη κυριαρχία των αριθμητικών αναπαραστάσεων και μάλιστα με τη χρήση πραγματικών αριθμών. Το γεγονός αυτό, σε συνδυασμό με το ότι η εκδοχή των λεκτικών χαρακτηρισμών μπορεί να μεταφερθεί (με μια κατάλληλη αντιστοίχιση) στο χώρο των ακεραίων αριθμών, μας οδηγεί στο συμπέρασμα ότι η αναπαράσταση με χρήση πραγματικών αριθμών είναι γενικότερη και μπορεί να αποτελέσει τη βάση για όλα τα μοντέλα.

- **Αξιολόγηση της εμπιστοσύνης**

Κάθε μοντέλο εμπιστοσύνης προβλέπει ότι κάθε πράκτορας χρησιμοποιεί ένα μηχανισμό για την αξιολόγηση της εμπιστοσύνης που έχει προς κάποιον άλλο πράκτορα, προκειμένου να αποφασίσει αν θα αλληλεπιδράσει μαζί του ή όχι. Ο μηχανισμός αυτός συνήθως περιλαμβάνει :



- α) τον εντοπισμό και τη συλλογή κάποιων (απαραίτητων) δεδομένων
- β) την εκτέλεση ενός ή περισσότερων υπολογισμών πάνω στα συλλεχθέντα δεδομένα

Ο μηχανισμός αυτός μπορεί να χρησιμοποιηθεί είτε για να αξιολογηθούν πράκτορες εκ των προτέρων (πριν δηλαδή αποφασιστεί και εκτελεστεί μια αλληλεπίδραση μαζί τους) είτε για να αξιολογηθούν πράκτορες που παρέχουν συστάσεις, είτε για να αυτοαξιολογηθεί ένας πράκτορας για να εκτιμήσει την ικανότητά του σε κάποιο τομέα [29].

Εάν οι προβλεπόμενοι υπολογισμοί είναι περισσότεροι του ενός, τότε μπορεί να αφορούν είτε τις επιμέρους συνιστώσες της εμπιστοσύνης [1,28], είτε τα διάφορα είδη εμπιστοσύνης που χειρίζεται το μοντέλο [34, 35].

Η αξιολόγηση της εμπιστοσύνης, μαζί με την αναπαράστασή της, αποτελούν τα δύο σημαντικότερα χαρακτηριστικά ενός μοντέλου εμπιστοσύνης.

- **Καταγραφή δεδομένων**

Η εμπιστοσύνη είναι μια δυναμική παράμετρος των μοντέλων εμπιστοσύνης, που αποτελεί συνάρτηση δύο παραγόντων : των προσωπικών εμπειριών και της φήμης των πρακτόρων. Αυτό σημαίνει πως τα δεδομένα που χρησιμοποιούνται για την αξιολόγησή της θα πρέπει να ενημερώνονται διαρκώς. Για το λόγο αυτό, σε όλα τα μοντέλα , προβλέπεται :

- η ύπαρξη κατάλληλων δομών για τη φύλαξη των αναγκαίων για την αξιολόγηση της εμπιστοσύνης δεδομένων
- η καταγραφή των δεδομένων που αφορούν τις προσωπικές εμπειρίες των πρακτόρων όπως αυτές προκύπτουν από τις δραστηριότητες στις οποίες συμμετέχουν.
- η καταγραφή πληροφοριών που συλλέγονται από τον έξω κόσμο

Τα δεδομένα αυτά αποτελούν ένα είδος ιστορικού για τους πράκτορες που λειτουργεί ως πηγή γνώσης για τη λήψη αποφάσεων. Η χωρητικότητα των δομών στις οποίες καταγράφονται θεωρείται, στα περισσότερα τουλάχιστον μοντέλα, πεπερασμένη. Η παραδοχή αυτή αντικατοπτρίζει το χαρακτηριστικό του βάθους της μνήμης των πρακτόρων, και είναι απαραίτητη προκειμένου να είναι ένα μοντέλο υλοποιήσιμο.

- **Χρήση στοιχείων προσωπικότητας**

Στοιχεία από την προσωπικότητα των πρακτόρων μπορεί να χρησιμοποιηθούν για την αντιμετώπιση καταστάσεων έλλειψης γνώσης [1]. Χαρακτηριστικά όπως η προδιάθεση για εμπιστοσύνη (trusting disposition) ή η τάση για ανάληψη ρίσκου, μπορεί να χρησιμοποιηθούν για την διαμόρφωση μιας τιμής για την εμπιστοσύνη σε περιπτώσεις όπου τα υπάρχοντα δεδομένα δεν επαρκούν για την αξιολόγησή της (πρωτόγνωρες συνθήκες, άγνωστοι πράκτορες, αντιφατική πληροφορόρηση κλπ).

Επιπλέον, κάποιες από τις παραμέτρους που συμμετέχουν στην αξιολόγηση της εμπιστοσύνης, όπως είναι τα χαρακτηριστικά της εκάστοτε περίπτωσης συναλλαγής (κόστος, όφελος, επικινδυνότητα, ικανότητα της άλλης πλευράς κλπ) ή η αντιλαμβανόμενη ποιότητα των υπηρεσιών που προσφέρονται κατά τη διάρκεια μιας συναλλαγής, εμπεριέχουν την υποκειμενική κρίση του αξιολογητή πράκτορα. Η κρίση αυτή, προφανώς, επηρεάζεται και από τα στοιχεία της προσωπικότητάς του πράκτορα.

- **Υποστήριξη αποφάσεων**

Ως ζωντανές υπολογιστικές οντότητες, οι πράκτορες εμπιστοσύνης χρειάζεται κατά τη διάρκεια της δραστηριοποίησής τους να παίρνουν αποφάσεις και να κάνουν επιλογές. Παραδείγματα τέτοιων καταστάσεων είναι η επιλογή του καταλληλότερου πράκτορα για αλληλεπίδραση, από ένα σύνολο πρακτόρων που παρέχουν μια επιθυμητή υπηρεσία, η αποδοχή ή όχι μιας πρότασης για ανταλλαγή υπηρεσιών, η αποδοχή ή η απόρριψη κάποιας πληροφορίας που έρχεται από τον έξω κόσμο κλπ.

Σε όλα τα μοντέλα, προβλέπονται κατάλληλοι μηχανισμοί για την υποβοήθηση των πρακτόρων στην διαμόρφωση των επιλογών τους και τη λήψη σχετικών αποφάσεων. Οι μηχανισμοί αυτοί τις περισσότερες φορές [1, 29, 36] περιλαμβάνουν, μεταξύ άλλων, τη σύγκριση της τρέχουσας αξιολόγησης της εμπιστοσύνης με ένα προκαθορισμένο ή δυναμικά μεταβαλλόμενο κατώφλι (threshold) εμπιστοσύνης. Το κατώφλι εμπιστοσύνης αποτελεί μια πολύ σημαντική παράμετρο των μοντέλων εμπιστοσύνης καθώς ορίζει τη λεπτή διαχωριστική γραμμή μεταξύ των εναλλακτικών επιλογών της αποδοχής και της απόρριψης<sup>15</sup>. Ο καθορισμός του μπορεί να επηρεάσει δραματικά τη συμπεριφορά και την αποτελεσματικότητα ενός μοντέλου.

- **Διενέργεια αλληλεπιδράσεων**

Όπως έχουμε δει από το κεφάλαιο 1, το πεδίο εφαρμογής των μοντέλων εμπιστοσύνης είναι οι ψηφιακές κοινωνίες και ο κύριος στόχος τους η υποστήριξη των πρακτόρων σε θέματα διαφύλαξης συμφερόντων και προστασίας από κακόβουλες συμπεριφορές. Είναι λογικό λοιπόν σε κάθε μοντέλο να προβλέπονται διαδικασίες αναζήτησης, διενέργειας και αξιολόγησης αλληλεπιδράσεων μεταξύ των πρακτόρων. Για την ακρίβεια, η διενέργεια αλληλεπιδράσεων είναι πρωταρχικό μέλημα των πρακτόρων αφού μέσω αυτών και μόνο μπορούν να ικανοποιήσουν τις ανάγκες τους και να υλοποιήσουν τους προσωπικούς τους στόχους.

Από τεχνικής άποψης, στα περισσότερα μοντέλα δεν γίνεται αναλυτική περιγραφή ούτε του πλαισίου μέσα στο οποίο πρέπει να διεξάγονται οι αλληλεπιδράσεις, ούτε των κριτηρίων βάσει των οποίων πρέπει να γίνεται η αξιολόγησή τους. Έτσι, αναφέρεται μεν η αναγκαιότητα ύπαρξης αλληλεπιδράσεων και κάποιου τρόπου αξιολόγησής τους, όμως οι λεπτομέρειες αφήνονται για την

---

<sup>15</sup> Σε κάποιες περιπτώσεις αυτό το κατώφλι μπορεί να είναι διπλό, ορίζοντας έτσι και μια ενδιάμεση περιοχή που αντιστοιχεί στην αβεβαιότητα ή την ουδετερότητα [1,36].

υλοποίηση, ώστε να ληφθούν υπόψη τα ιδιαίτερα χαρακτηριστικά της εκάστοτε περιοχής εφαρμογής. Ωστόσο, πρέπει να τονισθεί ότι η αναπαράσταση της αξιολόγησης των αλληλεπιδράσεων δεν μπορεί να εντελώς ανεξάρτητη του μοντέλου εμπιστοσύνης. Κι αυτό γιατί σε όλα τα μοντέλα ο υπολογισμός της εμπιστοσύνης χρησιμοποιεί μεταξύ άλλων και τις αξιολογήσεις προηγούμενων αλληλεπιδράσεων και συνεπώς η αναπαράσταση της αξιολόγησης των αλληλεπιδράσεων θα πρέπει να είναι συμβατή με την αναπαράσταση της εμπιστοσύνης.

- **Χρήση κοινωνικών μηχανισμών**

Τα μοντέλα εμπιστοσύνης, στην προσπάθειά τους να απεικονίσουν και να διαχειριστούν με το μεγαλύτερο δυνατό ρεαλισμό την έννοια της εμπιστοσύνης έχουν υιοθετήσει αρκετούς μηχανισμούς που φαίνεται να λειτουργούν στις ανθρώπινες κοινωνίες. Οι σημαντικότεροι και πιο συχνά εμφανιζόμενοι κοινωνικοί μηχανισμοί είναι :

#### ***Φήμη***

Μέσω του μηχανισμού της φήμης καταγράφεται και αξιοποιείται η εικόνα της αξιοπιστίας των πρακτόρων όπως αυτή έχει διαμορφωθεί γενικά στην κοινωνία, μέσα από τη διάδοση και την ανάμιξη των επιμέρους εκτιμήσεων. Η διαχείριση της φήμης βοηθά τα μοντέλα που τη χρησιμοποιούν αφενός να αντιμετωπίσουν τον κίνδυνο που ελλοχεύει στις αλληλεπιδράσεις με άγνωστες οντότητες και αφετέρου να αναχαιτίσουν με αρκετά αποτελεσματικό τρόπο τις υποψήφιες κακόβουλες συμπεριφορές. Όπως θα δούμε και στο κεφάλαιο 5, η χρήση της φήμης μπορεί να βελτιώσει σημαντικά τις επιδόσεις ενός μοντέλου.

#### ***Συστάσεις***

Ο μηχανισμός των συστάσεων αφορά την αναζήτηση της γνώμης μιας τρίτης πηγής προκειμένου να αντισταθμιστεί η έλλειψη γνώσης κατά τη λήψη μιας απόφασης. Έχει αρκετά κοινά σημεία με το μηχανισμό της φήμης καθώς η γνώμη τρίτων σχετικά με ένα πράκτορα ουσιαστικά αποτελεί τη φήμη του πράκτορα. Η διαφοροποίηση των δύο μηχανισμών έγκειται στο ότι η πληροφορία που συλλέγεται μέσα από τις συστάσεις, εάν δεν υπάρχει ο μηχανισμός της φήμης, δεν καταγράφεται και συνεπώς επηρεάζει την αξιολόγηση της εμπιστοσύνης μόνο παροδικά. Έτσι, η χρήση συστάσεων μπορεί να ειπωθεί είτε ως ένας αυτόνομος μηχανισμός που βοηθά στην λήψη καλύτερων αποφάσεων, είτε ως συνιστώσα του μηχανισμού της φήμης.

#### ***Φιλίες***

Στις ανθρώπινες κοινωνίες η φιλία ανάμεσα σε δύο πρόσωπα υποδηλώνει την ύπαρξη αμοιβαίας, συνήθως υψηλής, εμπιστοσύνης. Η διαχείριση φιλιών στα μοντέλα εμπιστοσύνης έχει παρόμοια σημασιολογία και λειτουργικότητα. Μέσα από το μηχανισμό αυτό οι πράκτορες αξιοποιούν τις εμπειρίες τους για να διαμορφώσουν ένα σύνολο έμπιστων πρακτόρων στους οποίους μπορούν να ανατρέξουν για να αναζητήσουν υπηρεσίες ή συστάσεις. Επιπλέον όταν

τους ζητηθεί να προτείνουν ένα κατάλληλο πράκτορα για κάποια υπηρεσία η πρώτη τους επιλογή θα είναι μέσα από το σύνολο των φίλων.

### **Κουτσομπολιό**

Ο μηχανισμός του κουτσομπολιού αποτελεί μια ιδιαίτερη μορφή παροχής συστάσεων. Η διαφορά από τον μηχανισμό των συστάσεων είναι ότι με το κουτσομπολιό διαδίδονται προσωπικές γνώμες αλλά χωρίς αυτό να έχει προηγουμένως ζητηθεί. Ο μηχανισμός αυτός, στα μοντέλα εμπιστοσύνης δεν έχει απαραίτητα την αρνητική χροιά που του αποδίδεται στις ανθρώπινες κοινωνίες. Μπορεί βέβαια και εδώ να λειτουργήσει παραπλανητικά και ως μέσο διαστρέβλωσης της πραγματικότητας, στις περιπτώσεις που η πηγή είναι κακόβουλη. Συνήθως όμως χρησιμοποιείται για την επιτάχυνση της διάδοσης της κακής φήμης, με στόχο την προφύλαξη της κοινωνίας (ιδιαίτερα των φίλων, εάν υπάρχουν) από ανεπιθύμητες ή επικίνδυνες συμπεριφορές [36].

## **3.2 Η λειτουργικότητα των πρακτόρων εμπιστοσύνης**

Έχοντας καταγράψει τα σημαντικότερα χαρακτηριστικά των μοντέλων εμπιστοσύνης, μπορούμε να δούμε πως αυτά τα χαρακτηριστικά μετουσιώνονται σε λειτουργικότητα στο επίπεδο των πρακτόρων εμπιστοσύνης. Από την σκοπιά των πρακτόρων εμπιστοσύνης, η συμπεριφορά που πρέπει να επιδεικνύουν θα πρέπει να ικανοποιεί δύο προϋποθέσεις : α) να αποτελεί έκφραση της λογικής του μοντέλου εμπιστοσύνης και β) να υποστηρίζει τη βασική απαίτηση για ανάπτυξη κοινωνικότητας. Πρακτικά αυτό σημαίνει ότι, κάθε πράκτορας πρέπει να διαθέτει λειτουργίες και μηχανισμούς που να του επιτρέπουν να ζει και να δραστηριοποιείται σε κοινωνίες αποτελούμενες από άλλους ( παρόμοιους ή διαφορετικούς) πράκτορες κατά τρόπο ώστε, αφενός να ακολουθούνται οι γενικές αρχές του μοντέλου εμπιστοσύνης και, αφετέρου, να μην καταστρέφεται η ιδιότητα του μέλους των κοινωνιών αυτών.

### **3.2.1 Ένα υποθετικό σενάριο από τον πραγματικό κόσμο**

Για να μπορέσουμε να καταγράψουμε καλύτερα την λειτουργικότητα που πρέπει να διαθέτει ένας πράκτορας εμπιστοσύνης, ας μελετήσουμε ένα υποθετικό σενάριο που θα μπορούσε να διαδραματιστεί σε μια πραγματική ανθρώπινη κοινωνία.

*Το περιβάλλον :* Τρεις, άγνωστοι μεταξύ τους άνθρωποι, ο Αλέξης, η Βασιλική και ο Γιώργος μετακομίζουν (για λόγους που δεν μας ενδιαφέρουν) σε μια μεγαλούπολη. Κάθε ένας τους έχει μια ειδικευση-ειδικότητα. Ο Αλέξης είναι παθολόγος γιατρός, η Βασιλική είναι καθηγήτρια μουσικής (συγκεκριμένα κιθάρας) και ο Γιώργος είναι μηχανικός αυτοκινήτων. Δυστυχώς γι' αυτούς, στη μεγαλούπολη αυτή, δεν γνωρίζουν κανένα άλλο άνθρωπο και συνεπώς δεν έχουν φίλους που θα μπορούσαν να τους βοηθήσουν να ενταχθούν στην κοινωνία της.

*Συμβάν 1 :* Μια μέρα η Βασιλική εμφανίζει ένα εμπύρετο κρυολόγημα και χρειάζεται γιατρό. Μη γνωρίζοντας κανέναν στη μεγαλούπολη ανοίγει το χρυσό οδηγό και αρχίζει να ψάχνει για

παθολόγο. Από τα ονόματα που βρίσκει της τράβα την προσοχή (τυχαία) το όνομα του Αλέξη. Σημειώνει το τηλέφωνό του και αφού κλείνει ένα ραντεβού τον επισκέπτεται για να δει την περίπτωση της. Η φαρμακευτική αγωγή που της συστήνει τη βοηθά να ξεπεράσει γρήγορα το κρυολόγημά της. Έτσι η πρώτη εμπειρία της Βασιλικής από τον Αλέξη είναι θετική με αποτέλεσμα ο Αλέξης να θεωρείται πρόσωπο εμπιστοσύνης, σε ιατρικά θέματα, για τη Βασιλική.

*Συμβάν 2* : Μέσα από τη γνωριμία τους ο Αλέξης μαθαίνει από τη Βασιλική ότι είναι καθηγήτρια κιθάρας, γεγονός που αναζωπυρώνει την παιδική του επιθυμία να μάθει το συγκεκριμένο μουσικό όργανο. Μετά από μερικές μέρες, επικοινωνεί με τη Βασιλική και της ζητά να ξεκινήσουν μαθήματα. Ο Αλέξης κατά βάθος θέλει να μάθει να παίζει rock τραγούδια, τα οποία η Βασιλική (ίσως και λόγω προσωπικών προτιμήσεων) δεν τα διδάσκει σε ικανοποιητικό (για τον Αλέξη) βαθμό. Έτσι ο Αλέξης δεν μένει ιδιαίτερα ικανοποιημένος με αποτέλεσμα να χάσει τον αρχικό ενθουσιασμό του και μετά από μερικά εβδομάδες σταματά τα μαθήματα με το πρόσχημα της έλλειψης ελεύθερου χρόνου.

*Συμβάν 3* : Το αυτοκίνητο του Αλέξη, κάποια στιγμή, συμπληρώνει τον προβλεπόμενο αριθμό χιλιομέτρων και χρειάζεται service. Μη γνωρίζοντας κάποιο συνεργείο, επιλέγει να επισκεφτεί το συνεργείο της γειτονιάς του, το οποίο ανήκει στον Γιώργο. Το service εκτελείται με υποδειγματικό τρόπο και ο Αλέξης μένει απόλυτα ικανοποιημένος από την επιλογή του.

*Συμβάν 4* : Ο Αλέξης θέλει να αγοράσει ένα δεύτερο αυτοκίνητο για τη γυναίκα του. Έχοντας κάνει μια σχετική έρευνα αγοράς κατέληξε σε ένα συγκεκριμένο τύπο αυτοκινήτου, μιας συγκεκριμένης εταιρίας. Τον απασχολεί όμως το θέμα της εξυπηρέτησης από την αντιπροσωπεία, γεγονός που άπτεται σε μεγάλο βαθμό της επαγγελματικής αξιοπιστίας του ιδιοκτήτη της αντιπροσωπείας. Για το λόγο αυτό ζητά τη γνώμη του Γιώργου, μιας και αυτός αφενός είναι σχετικός με το χώρο και αφετέρου τυγχάνει της εμπιστοσύνης του (μετά από την θετική εμπειρία που είχε μαζί του). Ο Γιώργος, που γνωρίζει την αντιπροσωπεία, τον προειδοποιεί ότι ο συγκεκριμένος επαγγελματίας δεν έχει καλή φήμη στην αγορά. Τελικά ο Αλέξης αποφασίζει να μην αγοράσει αυτοκίνητο από τη συγκεκριμένη αντιπροσωπεία.

*Συμβάν 5* : Μια μέρα, το αυτοκίνητο της Βασιλικής παρουσιάζει πρόβλημα (δεν ξεκινά). Επειδή δεν γνωρίζει κανένα συνεργείο που να το εμπιστεύεται, απευθύνεται στον Αλέξη, καθώς είναι ένα από τα λίγα πρόσωπα που γνωρίζει και εκτιμά. Ο Αλέξης, όντας πολύ ευχαριστημένος από την εμπειρία του με το Γιώργο και το συνεργείο του, τον προτείνει στην Βασιλική ανεπιφύλακτα. Η Βασιλική καλεί το Γιώργο για να ελέγξει και να επισκευάσει το αυτοκίνητό της. Όμως η εξυπηρέτηση που λαμβάνει δεν ανταποκρίνεται στις προσδοκίες που τις είχαν δημιουργήσει οι συστάσεις του Αλέξη. Ο Γιώργος επισκευάζει μεν το αυτοκίνητό της, αλλά μετά από δέκα ολόκληρες ημέρες και με μια υπερβολικά μεγάλη (όπως κρίνει η Βασιλική) αμοιβή.

*Συμβάν 6* : Το καλοκαίρι η Βασιλική πάει διακοπές, μέσω του ταξιδιωτικού πρακτορείου "X-Travels", σε κάποιο ξενοδοχείο που όμως αποδεικνύεται απαράδεκτο. Επιπλέον κατά την έκφραση των παραπόνων της στο πρακτορείο διαπληκτίζεται έντονα μαζί τους καθώς δεν αποδέχονται την κρίση της για το ξενοδοχείο. Η Βασιλική έντονα δυσαρεστημένη από τις υπηρεσίες και τη στάση του πρακτορείου τηλεφωνεί στον Αλέξη (με τον οποίο έχει πλέον μια φιλική σχέση) και του εξιστορεί την περιπέτειά της προτρέποντάς τον να αποφύγει το συγκεκριμένο πρακτορείο. Ο Αλέξης με τη σειρά του, στην επόμενη επίσκεψή του για service στο Γιώργο, και κατά τη διάρκεια συζήτησης που κάνουν περί διακοπών, του μεταφέρει την ιστορία της Βασιλικής, ενημερώνοντάς τον για την άσχημη εξυπηρέτηση του συγκεκριμένου πρακτορείου.

*Συμβάν 7* : Μια μέρα ο Γιώργος, επιστρέφοντας στο σπίτι του από τη δουλειά, δέχεται ένα διαφημιστικό τηλεφώνημα από το πρακτορείο "X-Travels ". Το πρακτορείο του προσφέρει μια εβδομάδα φθηνών διακοπών σε πολυτελές ξενοδοχείο του εξωτερικού. Ο Γιώργος, έχοντας υπόψη του την κακή φήμη του συγκεκριμένου πρακτορείου, είναι πολύ επιφυλακτικός. Τελικά, αρνείται την προσφορά θεωρώντας ότι εμπεριέχει σημαντικό ποσοστό ρίσκου καθώς το κόστος των προσφερόμενων διακοπών, αν και συγκριτικά μικρό, είναι σημαντικό για τα οικονομικά του.

### 3.2.2 Οι βασικές συνιστώσες κοινωνικής συμπεριφοράς

Μέσα από τα συμβάντα του προηγούμενου σεναρίου είναι εύκολο να διακρίνει κανείς κάποιες χαρακτηριστικές συμπεριφορές που είναι κοινές σε όλες τις ανθρώπινες κοινωνίες. Είναι οι συμπεριφορές που αναπτύσσουν τα μέλη των κοινωνιών προκειμένου αφενός να ικανοποιήσουν τους στόχους τους (όπως αυτοί ορίζονται από τις προσωπικές τους ανάγκες) και αφετέρου να παραμείνουν μέλη της κοινωνίας<sup>16</sup>. Ας καταγράψουμε τις πιο χαρακτηριστικές απ' αυτές ώστε στη συνέχεια να προσπαθήσουμε να τις αντιστοιχίσουμε στον ψηφιακό κόσμο των πρακτόρων εμπιστοσύνης.

- Αναζήτηση υπηρεσιών

Τα μέλη έχουν ανάγκες που προσπαθούν να ικανοποιήσουν. Τις περισσότερες φορές οι ανάγκες αυτές ικανοποιούνται μέσα από υπηρεσίες που τους παρέχουν άλλα μέλη της κοινωνίας βάσει της εξειδίκευσης και των ικανοτήτων που διαθέτουν (ο Αλέξης προσφέρει ιατρικές υπηρεσίες, η Βασιλική υπηρεσίες εκπαίδευσης και ο Γιώργος υπηρεσίες μηχανικού αυτοκινήτων). Χρειάζεται λοιπόν να υπάρχει ένας μηχανισμός αναζήτησης των μελών που παρέχουν τις επιθυμητές υπηρεσίες. Ο μηχανισμός αυτός συνήθως εμπεριέχει και ένα δεύτερο (αλλά όχι δευτερεύοντα) μηχανισμό αξιολόγησης και κατάταξης των διαθέσιμων παροχών της επιθυμητής υπηρεσίας. Αυτός ο μηχανισμός κάνει δυνατή την επιλογή του καλύτερου (ανά περίπτωση) παροχέα.

---

<sup>16</sup> Τα δύο αυτά κριτήρια μπορεί να είναι ανταγωνιστικά στην περίπτωση των εγωκεντρικών ατόμων. Τα άτομα αυτά συνήθως επιδεικνύουν αντικοινωνικές συμπεριφορές που αποσκοπούν αποκλειστικά στο να μεγιστοποιήσουν το προσωπικό τους όφελος, αδιαφορώντας για τις επιπτώσεις στο κοινωνικό σύνολο. Με τον τρόπο αυτό όμως σταδιακά οδηγούνται σε κοινωνική απομόνωση.

- Διενέργεια αλληλεπιδράσεων (για παροχή / λήψη υπηρεσιών)

Για την εξυπηρέτηση των αναγκών τους τα μέλη θα πρέπει να μπορούν να αλληλεπιδρούν μεταξύ τους προκειμένου να προσφέρουν ή να λαμβάνουν τις υπηρεσίες που χρειάζονται. Η αλληλεπίδραση αυτή συνήθως περιλαμβάνει κάποιας μορφής επικοινωνία με ταυτόχρονη μεταφορά δεδομένων, αγαθών ή υπηρεσιών (ο Αλέξης εξετάζει τη Βασιλική και της δίνει την κατάλληλη θεραπεία, ο Γιώργος επισκευάζει το αυτοκίνητο της Βασιλικής κλπ) .

- Αξιολόγηση αλληλεπιδράσεων

Μετά από μια αλληλεπίδραση, τα εμπλεκόμενα μέρη με κάποιο τρόπο αξιολογούν το αποτέλεσμα της αλληλεπίδρασης. Η αξιολόγηση αυτή μπορεί να αφορά στην ποιότητα της παρασχεθείσας υπηρεσίας, την ηθική ή τη νομιμότητα της διαδικασίας, την ποιότητα των προσφερόμενων αγαθών κλπ. Μέσα από την αξιολόγηση αυτή είναι εφικτή η επανεκτίμηση της αξιοπιστίας της άλλης πλευράς έτσι ώστε, αφενός ο μηχανισμός της αξιολόγησης των διαθέσιμων επιλογών συνεχώς να βελτιώνεται, και αφετέρου, οι όποιες κακές επιλογές να αποφεύγονται στο μέλλον (η Βασιλική είναι ικανοποιημένη από τις ιατρικές υπηρεσίες του Αλέξη, αλλά δυσαρεστημένη από τις υπηρεσίες του Γιώργου). Δυστυχώς, η αξιολόγηση μιας αλληλεπίδρασης (τουλάχιστον σε ό,τι αφορά την ακριβή τιμή της) είναι τις περισσότερες φορές υποκειμενική καθώς εξαρτάται σε σημαντικό βαθμό από προσωπικά χαρακτηριστικά του αξιολογητή (κοινωνικές αντιλήψεις, στοιχεία χαρακτήρα, ψυχολογική κατάσταση κλπ). Ωστόσο μια γενική κατηγοριοποίηση μιας αλληλεπίδρασης ως "καλή" ή "κακή" είναι εφικτή και μπορεί να βασιστεί σε κάποια κοινά αποδεκτά κριτήρια (π.χ. εάν μετά την επισκευή το αυτοκίνητο εξακολουθεί να μην δουλεύει σωστά ή εάν μετά την θεραπευτική αγωγή τα συμπτώματα της ασθένειας δεν υποχωρούν).

- Αναζήτηση συστάσεων

Στην προσπάθειά τους να εντοπίσουν μια υπηρεσία που χρειάζονται, τα μέλη της κοινωνίας ίσως χρειαστεί να αναζητήσουν συστάσεις. Αυτό μπορεί να συμβεί όταν δεν γνωρίζουν ποιος διαθέτει τη συγκεκριμένη υπηρεσία ή όταν δεν μπορούν να εκτιμήσουν την αξιοπιστία του ατόμου που προσφέρει την υπηρεσία (ο Αλέξης δεν γνωρίζει τίποτε για την αντιπροσωπεία της εταιρεία αυτοκινήτων που τον ενδιαφέρει). Στην περίπτωση αυτή θα χρειαστεί να στραφούν είτε προς κάποια γνωστά και έμπιστα από προηγούμενες αλληλεπιδράσεις άτομα (ο Αλέξης ζητά την επαγγελματική γνώμη του Γιώργου), ή προς κάποια γενική πηγή συστάσεων που παρέχει υπηρεσίες καταλόγου (η Βασιλική ψάχνει στο Χρυσό Οδηγό για παθολόγο).

Συστάσεις μπορεί επίσης να χρειαστεί να αναζητήσουν και όταν κάποιος άγνωστος τους ζητά να αλληλεπιδράσουν προκειμένου να του προσφέρουν κάποια δική τους υπηρεσία. Εάν η αλληλεπίδραση αυτή εμπεριέχει ρίσκο (π.χ. εάν υπάρχει κίνδυνος να μην πληρωθούν για την υπηρεσία που θα προσφέρουν) τότε πριν προχωρήσουν θα προσπαθήσουν να συλλέξουν πληροφορίες για τη φήμη του άγνωστου ατόμου ώστε να μειώσουν το υφιστάμενο ρίσκο.

- Παροχή συστάσεων

Εφόσον χρειάζεται να αναζητούν συστάσεις, τα μέλη μιας κοινωνίας θα πρέπει να είναι σε θέση και να τις παρέχουν όταν τους ζητηθούν (ο Γιώργος δίνει στον Αλέξη την γνώμη του για την αντιπροσωπεία αυτοκινήτων). Η ικανότητα παροχής συστάσεων προϋποθέτει την ύπαρξη δομών και μηχανισμών αποθήκευσης των προσωπικών τους εμπειριών αλλά και των συστάσεων που συλλέγουν από τα άλλα μέλη τις κοινωνίας.

- Διάδοση προσωπικών απόψεων και φήμης (κουτσομπολιό)

Η παροχή συστάσεων συνήθως πραγματοποιείται μετά από κάποιο αίτημα. Υπάρχουν όμως και άτομα μέσα στην κοινωνία που διαδίδουν την προσωπική τους γνώμη ή αναμεταδίδουν τη γνώμη άλλων μελών, σε άλλα μέλη της κοινωνίας (συνήθως στους γνωστούς τους), χωρίς αυτό να τους έχει προηγουμένα ζητηθεί (ο Αλέξης μεταφέρει στο Γιώργο την άσχημη εμπειρία της Βασιλικής από το πρακτορείο X-Travel). Αυτή η συμπεριφορά έχει γίνει ευρέως γνωστή με τον όρο "κουτσομπολιό" και παρά το ότι της έχει αποδοθεί άσχημη χροιά, εάν χρησιμοποιηθεί στην κατεύθυνση της προειδοποίησης της κοινωνίας, μπορεί να αποδειχθεί εξαιρετικά χρήσιμος μηχανισμός, καθώς επιταχύνει την εξάπλωση της κακής φήμης κάποιου επικίνδυνου για την κοινωνία μέλους, γεγονός που κατά κανόνα μπορεί να προλάβει ανεπιθύμητες καταστάσεις.

Οι παραπάνω συμπεριφορές είναι ενδεικτικές και σε καμία περίπτωση δεν καλύπτουν το σύνολο των συμπεριφορών που αναπτύσσονται σε μια πραγματική κοινωνία. Μπορούν όμως να αποτελέσουν έναν οδηγό για τον προσδιορισμό των απαραίτητων μηχανισμών που πρέπει να διαθέτει ένας πράκτορας εμπιστοσύνης προκειμένου να εμφανίζει δραστηριότητα ανάλογη με αυτή των μελών μιας ανθρώπινης κοινωνίας.

### **3.2.3 Οι βασικοί μηχανισμοί συμπεριφοράς ενός πράκτορα εμπιστοσύνης**

Οι πράκτορες εμπιστοσύνης, σύμφωνα με όσα έχουν ειπωθεί στο κεφάλαιο 1, αποτελούν ψηφιακές υλοποιήσεις συγκεκριμένων μοντέλων κοινωνικής συμπεριφοράς. Θα πρέπει λοιπόν να υποστηρίζουν τουλάχιστον τις βασικές συνιστώσες κοινωνικής συμπεριφοράς που καταγράφηκαν στην προηγούμενη παράγραφο. Οι συνιστώσες αυτές είναι απαραίτητες στους πράκτορες προκειμένου να μπορούν να δημιουργούν κοινωνίες και να δραστηριοποιούνται μέσα σ' αυτές. Στη συνέχεια της παραγράφου καταγράφονται, με πιο τεχνικούς όρους, οι μηχανισμοί που πρέπει να ενσωματώνει η συμπεριφορά ενός πράκτορα εμπιστοσύνης ώστε αυτός να υποστηρίζει τις βασικές συνιστώσες κοινωνικής συμπεριφοράς.

- Αναζήτηση αλληλεπίδρασης

Με το μηχανισμό αυτό υλοποιείται η αναζήτηση μιας υπηρεσίας. Για την ακρίβεια υλοποιείται η αναζήτηση ενός πράκτορα που παρέχει μια συγκεκριμένη υπηρεσία. Η υπηρεσία συνήθως



χρειάζεται για την κάλυψη κάποιας ανάγκης του πράκτορα. Η αναζήτηση αλληλεπίδρασης περιλαμβάνει τις παρακάτω ενέργειες :

- Επιλογή μιας από τις ανάγκες του πράκτορα. Η διαδικασία της επιλογής, προφανώς θα λαμβάνει υπόψη τις προτεραιότητες και τους σχεδιαστικούς στόχους του πράκτορα
- Εντοπισμός των πρακτόρων που ειδικεύονται στο αντικείμενο της επιλεγμένης ανάγκης
- Επιλογή ενός απ' τους υποψήφιους για αλληλεπίδραση πράκτορες
- Πρόσκληση σε αλληλεπίδραση του επιλεγμένου πράκτορα και πιθανή διαπραγμάτευση μαζί του για τους όρους και τις συνθήκες εκτέλεσης της αλληλεπίδρασης

Εφόσον η διαπραγμάτευση είναι επιτυχής και η πρόσκληση γίνει αποδεκτή, θα ακολουθήσει ο μηχανισμός εκτέλεσης της αλληλεπίδρασης.

- ***Επιλογή ενός πράκτορα από ένα σύνολο πρακτόρων***

Ο μηχανισμός αυτός χρησιμοποιείται από το μηχανισμό αναζήτησης αλληλεπίδρασης προκειμένου να επιλεγεί ο καταλληλότερος από τους υποψήφιους για αλληλεπίδραση πράκτορες. Συνήθως έχει ως βασικό έργο την κατάταξη των υποψήφιων πρακτόρων σε μια διατεταγμένη σειρά. Ο κορυφαίος αυτής της σειράς επιστρέφεται ως ο επιλεγμένος πράκτορας.

Η κατάταξη των πρακτόρων γίνεται είτε βάσει κάποιων χαρακτηριστικών που διαθέτουν οι πράκτορες (π.χ. φήμη), είτε κάνοντας χρήση κάποιας προηγούμενης αξιολόγησής τους (π.χ. εμπιστοσύνη), είτε καλώντας το μηχανισμό αξιολόγησης που θα δώσει μια τιμή βασισμένη στα πιο πρόσφατα δεδομένα.

Στα περισσότερα μοντέλα η κατάταξη καθορίζεται από την τελευταία αξιολόγηση της εμπιστοσύνης των πρακτόρων [34,28,40,42].

- ***Αξιολόγηση πρακτόρων***

Ο μηχανισμός αυτός συνήθως καλείται μετά από μια αλληλεπίδραση ή μέσα από το μηχανισμό επιλογής και αποδίδει σε ένα πράκτορα μια τιμή βάσει κάποιων κριτηρίων αξιολόγησης. Η τιμή αυτή στη συνέχεια μπορεί να χρησιμοποιηθεί για τη σύγκριση του αξιολογούμενου πράκτορα με άλλους πράκτορες. Ο υπολογισμός της τιμής συνήθως γίνεται με κριτήρια που εμπεριέχονται στα σχεδιαστικά χαρακτηριστικά του πράκτορα αξιολογητή. Μεταξύ των κριτηρίων αυτών θα πρέπει (εφόσον μιλάμε για πράκτορες εμπιστοσύνης) να συμπεριλαμβάνεται πάντοτε και η τρέχουσα αξιολόγηση της εμπιστοσύνης του αξιολογούμενου πράκτορα.

Σε πολλά μοντέλα ο μηχανισμός αξιολόγησης συμπίπτει με τη διαδικασία αξιολόγησης της εμπιστοσύνης (αναφ.) οπότε μιλάμε για αξιολόγηση αξιοπιστίας.

- ***Εκτέλεση αλληλεπίδρασης***

Με το μηχανισμό αυτό, οι πράκτορες αλληλεπιδρούν μεταξύ τους ανταλλάσσοντας υπηρεσίες ή αγαθά. Η διαδικασία που ακολουθείται κατά την αλληλεπίδραση εξαρτάται σε μεγάλο βαθμό από

την θεματική περιοχή και το αντικείμενο των πρακτόρων που αλληλεπιδρούν. Συνήθως περιλαμβάνει μια αμφίδρομη επικοινωνία κατά την οποία αποστέλλονται όσα συμφωνήθηκαν κατά την φάση της διαπραγμάτευσης.

- **Χειρισμός πρόσκλησης σε αλληλεπίδραση**

Όταν ένας πράκτορας δεχτεί μια πρόσκληση σε αλληλεπίδραση θα πρέπει να αποφασίσει εάν θα την αποδεχτεί προχωρώντας στη διενέργεια της αλληλεπίδρασης ή αν θα την απορρίψει. Ο μηχανισμός αυτός μπορεί να περιλαμβάνει δύο φάσεις :

- i. την αξιολόγηση της επικινδυνότητας της αλληλεπίδρασης
- ii. την διαπραγμάτευση των όρων εκτέλεσης της αλληλεπίδρασης

Η αξιολόγηση της επικινδυνότητας περιλαμβάνει την αξιολόγηση της αξιοπιστίας του πράκτορα που στέλνει την πρόσκληση σε σχέση με τα ιδιαίτερα χαρακτηριστικά της επικείμενης αλληλεπίδρασης (πιθανοί κίνδυνοι, κόστος, χρησιμότητα-όφελος κλπ). Προφανώς η απόφαση αποδοχής της πρόσκλησης θα είναι ευθέως ανάλογη της αξιοπιστίας του πράκτορα και των "ακίνδυνων" χαρακτηριστικών της αλληλεπίδρασης (χρησιμότητα-όφελος) και αντιστρόφως ανάλογη των "επικίνδυνων" χαρακτηριστικών (κόστος, κίνδυνοι).

Η φάση της διαπραγμάτευσης είναι προαιρετική. Συνήθως καλείται μόνο εφόσον η απόφαση αποδοχής ή απόρριψης της αλληλεπίδρασης είναι οριακή και αποσκοπεί στην βελτίωση κάποιων χαρακτηριστικών της αλληλεπίδρασης ώστε να ελαχιστοποιηθεί η επικινδυνότητα και να μεγιστοποιηθεί το όφελος για τον πράκτορα που δέχεται την πρόσκληση.

Στα περισσότερα μοντέλα η απόφαση αποδοχής ή απόρριψης καθορίζεται από την απλή σύγκριση της εμπιστοσύνης του πράκτορα που στέλνει την πρόσκληση με ένα προκαθορισμένο κατώφλι (threshold) αξιοπιστίας. Το κατώφλι αυτό μπορεί να μεταβάλλεται ρυθμίζοντας την ισορροπία ανάμεσα στην ασφάλεια και την ελευθερία δράσης που προσφέρει το εκάστοτε μοντέλο.

Στην περίπτωση αποδοχής της πρότασης ακολουθεί ο μηχανισμός εκτέλεσης της αλληλεπίδρασης.

- **Αναζήτηση συστάσεων**

Ο μηχανισμός της αναζήτησης συστάσεων υποστηρίζει την αναγκαιότητα συλλογής στοιχείων για θέματα στα οποία υπάρχει ελλιπής πληροφόρηση. Τόσο ο μηχανισμός επιλογής ενός πράκτορα όσο και ο μηχανισμός χειρισμού μιας πρόσκλησης σε αλληλεπίδραση μπορεί να μην έχουν τα απαραίτητα δεδομένα για να λειτουργήσουν. Στην περίπτωση αυτή τα δεδομένα θα αναζητηθούν υπό τη μορφή συστάσεων (δηλ. προσωπικής εκτίμησης) από τα άλλα μέλη της κοινωνίας.

Η αναζήτηση συστάσεων περιλαμβάνει τρεις φάσεις :

- i. την επιλογή των πρακτόρων από τους οποίους θα ζητηθούν οι συστάσεις
- ii. την αποστολή στους επιλεγμένους πράκτορες κατάλληλων μηνυμάτων με τα ζητούμενα στοιχεία
- iii. την αναμονή των απαντήσεων και τη συλλογή των παρεχόμενων συστάσεων

Στη φάση της επιλογής υπάρχουν δύο εναλλακτικές :

α) είτε επιλέγονται όλοι οι σχετικοί πράκτορες (δηλ. όσοι σχετίζονται με τις ζητούμενες πληροφορίες)

β) είτε επιλέγονται μόνο οι πιο έμπιστοι από τους σχετικούς πράκτορες, με την εμπιστευτικότητα να προσδιορίζεται κυρίως μέσα από την αξιολόγηση της εμπιστοσύνης

Κάποια μοντέλα προτείνουν την καταγραφή των έμπιστων πρακτόρων (τους αναφέρουν ως γείτονες<sup>17</sup> [1,34,41]) ώστε να είναι εύκολος ο εντοπισμός και η αναφορά σ' αυτούς.

Κατά τη φάση της αποστολής των μηνυμάτων θα πρέπει να δομηθεί ένα σχετικό query σε κάποιο συντακτικό που να είναι κατανοητό από τους παραλήπτες πράκτορες. Οι συνηθέστερες πληροφορίες που μπορεί να ζητηθούν είναι : ποιοι είναι οι N καλύτεροι πράκτορες σε κάποιο αντικείμενο, τι φήμη έχει ένας πράκτορας σε κάποιο αντικείμενο, ποιες είναι οι N τελευταίες αξιολογήσεις των αλληλεπιδράσεων ενός πράκτορα κλπ.

Η φάση της συλλογής των συστάσεων, αν και φαινομενικά απλή, εμπεριέχει ένα σημαντικό βαθμό δυσκολίας και πολυπλοκότητας. Το πρόβλημα έγκειται στο ότι δεν απαντούν πάντοτε όλοι οι ερωτούμενοι πράκτορες, οπότε θα πρέπει να αποφασιστεί το πότε θα σταματήσει η διαδικασία της αναμονής των απαντήσεων. Το πρόβλημα περιπλέκει ακόμη περισσότερο το γεγονός ότι κάποιες απαντήσεις μπορεί στην πορεία να χαθούν ή να καθυστερήσουν λόγω τεχνικών προβλημάτων στην επικοινωνία. Αυτό καθιστά πολύ δύσκολο τον χαρακτηρισμό της συμπεριφοράς των ερωτηθέντων πρακτόρων (συνεργάσιμοι, αδιάφοροι).

- ***Χειρισμός εισερχόμενης αίτησης για παροχή συστάσεων***

Είναι ο μηχανισμός που επιτρέπει σε ένα πράκτορα να χειριστεί μια αίτηση για παροχή συστάσεων που θα δεχτεί. Ο μηχανισμός αυτός θα πρέπει κατ' αρχήν να μπορεί να απαντά σε ερωτήματα όπως :

- Εάν θα απαντηθεί ή θα αγνοηθεί η εισερχόμενη αίτηση. Αυτό μπορεί να εξαρτάται από διάφορους παράγοντες όπως για παράδειγμα η εμπειρία από προηγούμενες αλληλεπιδράσεις με τον αιτούντα πράκτορα, η καταγεγραμμένη φήμη του, η προσμονή ανταποδοτικού οφέλους κλπ
- Εάν μπορούν να δοθούν οι ζητούμενες συστάσεις. Μπορεί για παράδειγμα οι γνώσεις ή η ειδίκευση του πράκτορα να μην περιλαμβάνουν το ζητούμενο αντικείμενο.
- Εάν θα προωθηθεί η αίτηση σε άλλους (γνωστούς και σχετικούς) πράκτορες προκειμένου η απάντηση που θα δοθεί να είναι κατά το δυνατόν πληρέστερη.

Στην περίπτωση που επιλεγεί η εξυπηρέτηση της αίτησης, το επόμενο βήμα θα είναι η ενεργοποίηση του μηχανισμού της παροχής συστάσεων.

- ***Παροχή συστάσεων***

Ο μηχανισμός της παροχής συστάσεων έχει ως έργο :

- α) τη συγκέντρωση των απαραίτητων στοιχείων και τη δημιουργία των ζητούμενων συστάσεων σε μια μορφή κατανοητή από τον αιτούντα πράκτορα.
- β) την αποστολή των συστάσεων στον αιτούντα πράκτορα μέσα σε ένα κατάλληλο μήνυμα.

- **Ενημέρωση εσωτερικών δομών**

Ο κάθε πράκτορας διαθέτει κάποιες εσωτερικές δομές στις οποίες κρατά τα δεδομένα που χρειάζεται για να μπορέσει να λειτουργήσει. Οι δομές αυτές θα πρέπει να μπορούν να ενημερώνονται με τα δεδομένα της δραστηριότητας του πράκτορα. Το έργο αυτό υλοποιεί ο μηχανισμός της ενημέρωσης των εσωτερικών δομών.

Ο συγκεκριμένος μηχανισμός καλείται μετά από κάθε συμβάν που παράγει νέα δεδομένα για τον πράκτορα. Τέτοια συμβάντα είναι οι αλληλεπιδράσεις και η λήψη πληροφοριών (συνήθως συστάσεων) από τους άλλους πράκτορες. Συνήθως περιλαμβάνει τις εξής ενημερώσεις :

- Ενημέρωση της δομής για το ιστορικό των αλληλεπιδράσεων
- Ενημέρωση των δομών για την εμπιστοσύνη και τη φήμη
- Ενημέρωση των μεταβλητών στοιχείων της προσωπικότητας του πράκτορα

- **Διάδοση πληροφοριών:**

Ο μηχανισμός αυτός επιτρέπει σε ένα πράκτορα να αποστέλλει πληροφορίες σε πράκτορες της επιλογής του, χωρίς αυτό να του έχει προηγουμένα ζητηθεί. Χρησιμοποιείται κυρίως για την κοινοποίηση στα υπόλοιπα μέλη της κοινωνίας κάποια ανεπιθύμητης συμπεριφοράς που έγινε γνωστή μέσα από κάποια προσωπική εμπειρία (αλληλεπίδραση) ή από εξωτερική πληροφόρηση (μέσα από τη χρήση του συγκεκριμένου μηχανισμού).

Ο μηχανισμός λειτουργεί παρόμοια με το μηχανισμό παροχής συστάσεων και περιλαμβάνει τη δημιουργία και αποστολή κατάλληλου μηνύματος με περιεχόμενο τη διαδιδόμενη πληροφορία. Η αυτόκλητη διάδοση πληροφοριών αποτελεί για αρκετά μοντέλα [41,39,43] ένα σημαντικό μηχανισμό προστασίας των μελών μιας κοινωνίας πρακτόρων από ανεπιθύμητες ή κακόβουλες συμπεριφορές.

---

<sup>17</sup> Ο όρος "γείτονες" δεν σχετίζεται με την θέση των πρακτόρων στην τοπολογία του δικτύου άλλα με την κοινωνική διάσταση της συχνότερης επικοινωνίας και οικειότητας που έχουν οι γείτονες.



# ΚΕΦΑΛΑΙΟ

## 4. Ο προσομοιωτής

Η καρδιά της πλατφόρμας αξιολόγησης που προτείνουμε είναι ο προσομοιωτής. Πρόκειται για μια αυτόνομη εφαρμογή που παρέχει την απαραίτητη λειτουργικότητα για την αναπαράσταση, την προσομοίωση και τη αξιολόγηση μοντέλων εμπιστοσύνης. Στο κεφάλαιο αυτό θα δούμε τις βασικές απαιτήσεις, την αρχιτεκτονική, τα δομικά συστατικά και τη βασική λειτουργικότητα του προσομοιωτή. Η παρουσίαση θα είναι όσο γίνεται απλούστερη, περιλαμβάνοντας μόνο τις εντελώς απαραίτητες τεχνικές λεπτομέρειες, ώστε να μπορέσει και ο απλός αναγνώστης, που δεν έχει εξειδικευμένες τεχνικές γνώσεις, να διαμορφώσει μια εικόνα για τις δυνατότητες του περιβάλλοντος προσομοίωσης της πλατφόρμας.

### 4.1 Βασικές Απαιτήσεις

Ο κύριος στόχος του προσομοιωτή είναι να προσφέρει ένα περιβάλλον για τη μελέτη και την αξιολόγηση μοντέλων εμπιστοσύνης. Για την επίτευξη αυτού του στόχου, ο προσομοιωτής θα πρέπει να επιτρέπει :

- την αναπαράσταση και προσομοίωση οποιουδήποτε γνωστού μοντέλου εμπιστοσύνης, με πρόβλεψη και για μελλοντικά μοντέλα
- την προσομοίωση ακόμη και ημιτελών μοντέλων (δηλ. μοντέλων που δεν περιγράφουν λεπτομερώς το χειρισμό όλων των καταστάσεων που μπορεί να εμφανιστούν)
- την ρύθμιση των συνθηκών προσομοίωσης και των παραμέτρων του προσομοιούμενου μοντέλου
- την παρακολούθηση, σε πραγματικό χρόνο, κατά τη διάρκεια της προσομοίωσης, των χαρακτηριστικών τόσο της προσομοιούμενης κοινωνίας όσο και του προσομοιούμενου μοντέλου
- την καταγραφή της εξέλιξης της προσομοίωσης σε μια μορφή που να μπορεί να μελετηθεί, μετά το πέρας της προσομοίωσης, με χρήση κατάλληλων εργαλείων επεξεργασίας δεδομένων

Σε ότι αφορά τις συνθήκες προσομοίωσης, όπως έχουμε δει στο κεφάλαιο 1, το πεδίο εφαρμογής των μοντέλων εμπιστοσύνης είναι γενικά οι κοινωνίες. Συνεπώς, το περιβάλλον προσομοίωσης θα πρέπει να αποτελείται από μια κοινωνία στην οποία τα μέλη της θα είναι πράκτορες οι οποίοι θα έρχονται σε επαφή και θα αλληλεπιδρούν μεταξύ τους προκειμένου να επιτύχουν τους προσωπικούς τους στόχους. Το αντικείμενο των αλληλεπιδράσεων δεν έχει ιδιαίτερη σημασία. Το σημαντικό είναι να υπάρχει κάποιου είδους αλληλεξάρτηση μεταξύ των μελών της κοινωνίας, η οποία θα εκδηλώνεται μέσω των αλληλεπιδράσεων. Το προσομοιούμενο μοντέλο θα ενσωματώνεται στη λογική των πρακτόρων-μελών και θα καθοδηγεί τη συμπεριφορά τους στην επιλογή των αλληλεπιδράσεων.

## 4.2 Σχεδιαστικές επιλογές

### 4.2.1 Η πλατφόρμα

Για την πλατφόρμα ανάπτυξης του προσομοιωτή υπήρχαν τρεις επιλογές :

- 1) Να δημιουργηθεί από την αρχή ένα περιβάλλον προσομοίωσης, εξειδικευμένο για κοινωνίες πρακτόρων εμπιστοσύνης
- 2) Να χρησιμοποιηθεί κάποιο από τα υπάρχοντα περιβάλλοντα προσομοίωσης [44,46,47,48,49] ως βάση, και να επεκταθεί αυτό με την κατάλληλη λειτουργικότητα για την προσομοίωση πρακτόρων εμπιστοσύνης.
- 3) Να χρησιμοποιηθεί μια υπάρχουσα πλατφόρμα ανάπτυξης πολύ-πρακτορικών εφαρμογών ως βάση, και να επεκταθεί αυτή με την κατάλληλη λειτουργικότητα για προσομοίωση.

Οι δύο πρώτες επιλογές θα απαιτούσαν την ανάπτυξη διαδικασιών προσομοίωσης και την ανάπτυξη εξ' αρχής των βασικών λειτουργιών των πρακτόρων (νοημοσύνη, επικοινωνία κλπ), τα οποία δεν εμπίπτουν άμεσα στο αντικείμενο της μελέτης μας. Η τρίτη επιλογή, προτιμήθηκε, καθώς καθιστά δυνατή την ταχύτερη ανάπτυξη ενός ρεαλιστικού περιβάλλοντος προσομοίωσης, αξιοποιώντας τις υπάρχουσες βασικές λειτουργίες για την επικοινωνία μεταξύ πρακτόρων και επιτρέποντάς μας να εστιάσουμε στο θέμα που μας ενδιαφέρει, δηλαδή στην αξιολόγηση διαφόρων μοντέλων εμπιστοσύνης και την αποτίμησή τους με την ανάπτυξη κατάλληλων μετρικών. Επιπρόσθετα, η επιλογή μιας πραγματικής πλατφόρμας ανάπτυξης πολύ-πρακτορικών εφαρμογών προσφέρει περισσότερες πιθανότητες συμβατότητας του περιβάλλοντος προσομοίωσης με τα υπάρχοντα αλλά και τα μελλοντικά μοντέλα εμπιστοσύνης.

Για όλους τους παραπάνω λόγους, ακολουθήθηκε η τρίτη επιλογή. Η πλατφόρμα που επιλέχθηκε είναι το JADE.

### 4.2.2 Το JADE

Το JADE (Java Agent Development Framework) είναι ένα middle-ware που μπορεί να λειτουργήσει ως περιβάλλον ανάπτυξης και εκτέλεσης κατανεμημένων πολύ-πρακτορικών (multi-agent) εφαρμογών. Υποστηρίζει peer-to-peer αρχιτεκτονικές στις οποίες οι πόροι (resources) και ο έλεγχος (control) μπορούν να είναι πλήρως κατανεμημένα είτε πάνω σε υπολογιστές σταθερού δικτύου (fixed network) είτε πάνω σε κινούμενα τερματικά (mobile terminals). Το περιβάλλον εκτέλεσης που υποστηρίζει μπορεί να εξελίσσεται δυναμικά, επιτρέποντας την εμφάνιση ή την εξαφάνιση κόμβων, ανάλογα με τις ανάγκες και τις απαιτήσεις της εκάστοτε εφαρμογής. Η επικοινωνία μεταξύ των κόμβων είναι απόλυτα συμμετρική, με τον κάθε κόμβο να μπορεί να δράσει είτε εναρκτήριο είτε ανταποκρινόμενος σε μια σχετική πρόσκληση. Τα βασικότερα πλεονεκτήματα που διαθέτει είναι :

- **Ανάπτυξη σε Java**

Το JADE έχει αναπτυχθεί και προγραμματίζεται εξολοκλήρου σε Java., μια γλώσσα εξαιρετικά αναπτυσσόμενη, ιδιαίτερα ευρείας αποδοχής και για την οποία υπάρχει πλήθος έτοιμων εφαρμογών-εργαλείων κατηγορίας ανοικτού κώδικα (open source).

- **Διαλειτουργικότητα (interoperability)**

Το JADE είναι συμβατό με το πρότυπο FIPA<sup>18</sup>. Αυτό πρακτικά σημαίνει ότι στις εφαρμογές που βασίζονται στο JADE οι υποστηριζόμενοι πράκτορες μπορούν να αλληλεπιδρούν με οποιουδήποτε άλλους πράκτορες ακολουθούν το ίδιο πρότυπο.

- **Ομοιομορφία και μεταφερσιμότητα**

Το API του JADE χαρακτηρίζεται από ομοιογένεια και είναι ανεξάρτητο από τη υφιστάμενη δικτυακή υποδομή και την έκδοση της Java που χρησιμοποιείται. Οι εφαρμογές για JADE μπορούν πολύ εύκολα να μεταφερθούν από μια υπολογιστική πλατφόρμα σε μια άλλη.

- **Ευκολία χρήσης**

Η τεχνική πολυπλοκότητα του περιβάλλοντος του JADE καλύπτεται πίσω από το απλό και φιλικό περιβάλλον διεπαφής (interface) του API της πλατφόρμας. Επιπρόσθετα ο προγραμματιστής δεν είναι υποχρεωμένος να γνωρίζει και να χρησιμοποιεί όλες τις τεχνικές δυνατότητες της πλατφόρμας. Για τα χαρακτηριστικά που δεν χρησιμοποιούνται δεν απαιτείται καμιά ειδική ρύθμιση .

- **Σημαντικές πηγές πληροφόρησης & βοήθειας**

Για το JADE μπορεί κανείς να βρει εύκολα τεχνικές πληροφορίες, τόσο στο σχετικό ιστοχώρο που έχει δημιουργηθεί για το ερευνητικό πρόγραμμα του JADE, όσο και γενικότερα στο διαδίκτυο. Οι βιβλιοθήκες της πλατφόρμας συνοδεύονται από αναλυτικά εγχειρίδια χρήσης, tutorials και on-line βοήθεια, στοιχεία που βοηθούν σημαντικά το έργο του προγραμματιστή. Επιπρόσθετα, ολόκληρος ο κώδικας του JADE είναι διαθέσιμος κάτω από την πολιτική ανοικτού κώδικα Lesser GNU Public Licence<sup>19</sup>. Το πλέον σημαντικό είναι ότι υπάρχει μια διαρκώς αυξανόμενη κοινότητα χρηστών και προγραμματιστών του JADE, στην οποία μπορεί κανείς να προσφύγει για οποιαδήποτε βοήθεια.

Το JADE έχει σχεδιαστεί έτσι ώστε να απλοποιεί τη διαχείριση της επικοινωνίας και τη διακίνηση των μηνυμάτων μεταξύ των πρακτόρων. Αυτό το χαρακτηριστικό σε συνδυασμό και με την ευκολία εκμάθησης και χρήσης του API του, επιτάχυναν σε πολύ μεγάλο βαθμό το ρυθμό ανάπτυξης του προσομοιωτή.

---

<sup>18</sup> <http://www.fipa.org>

<sup>19</sup> <http://www.opensource.org/licences/lgpl-license.php>



Ως σχεδιαστική επιλογή, η πλατφόρμα του JADE παρέχει την απαραίτητη υποδομή για να δημιουργηθεί ένα περιβάλλον προσομοίωσης μοντέλων εμπιστοσύνης που να είναι πολύ κοντά στις πραγματικές συνθήκες εφαρμογής των μοντέλων. Αυτό σημαίνει ότι, μέσα από την προσομοίωση, ο μελετητής μπορεί να εξάγει ακριβέστερα και πιο αντικειμενικά συμπεράσματα για την συμπεριφορά και την αποτελεσματικότητα ενός μοντέλου εμπιστοσύνης.

### 4.2.3 Τα κυριότερα χαρακτηριστικά του JADE

Η πλατφόρμα του JADE περιλαμβάνει :

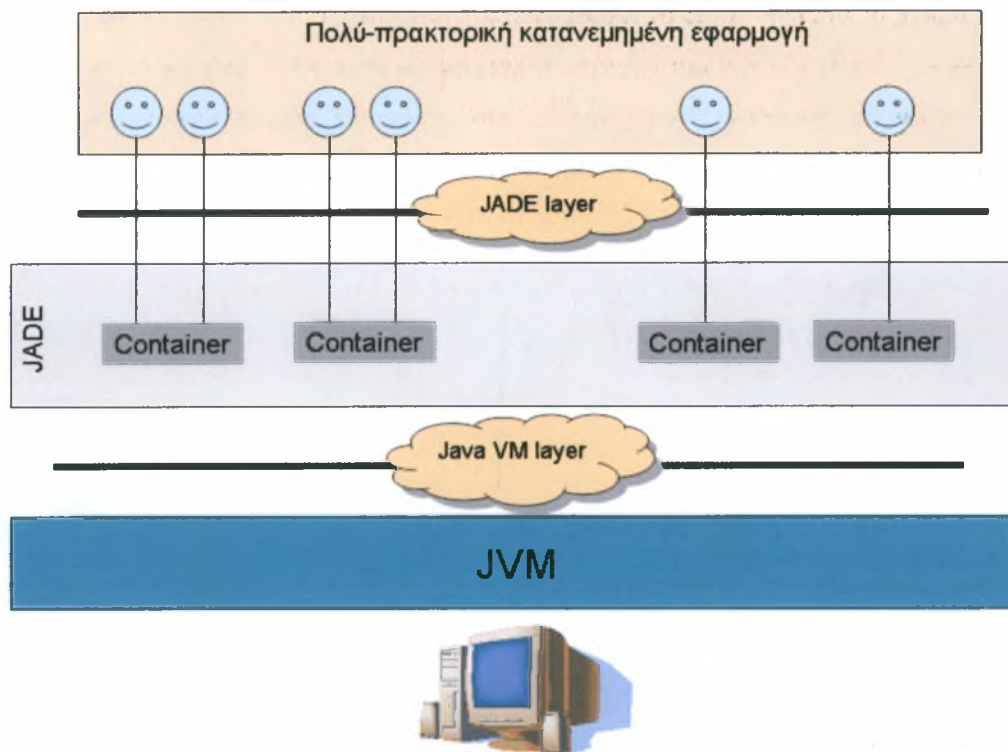
- i. Ένα σύνολο βιβλιοθηκών (Java κλάσεις) για την ανάπτυξη των πρακτόρων μιας εφαρμογής (application agents)
- ii. Το run-time περιβάλλον που παρέχει τις βασικές υπηρεσίες που χρειάζονται οι πράκτορες εφαρμογής για να λειτουργήσουν.
- iii. Μια σουίτα γραφικών εργαλείων που επιτρέπουν τη διαχείριση και την παρακολούθηση της δραστηριότητας των εκτελούμενων πρακτόρων

Μια εμφάνιση (instance) του run-time του JADE ονομάζεται container, επειδή μπορεί να περιέχει πολλούς πράκτορες. Το JADE επιτρέπει την ύπαρξη πολλών containers είτε πάνω στον ίδιο σταθμό είτε σε διαφορετικούς. Για να μπορέσουν να ξεκινήσουν οι πράκτορες θα πρέπει να είναι ενεργό τουλάχιστο ένα container. Το σύνολο όλων των ενεργών containers ονομάζεται πλατφόρμα (platform). Η πλατφόρμα ουσιαστικά είναι ένα αδιαφανές επίπεδο που κρύβει από τους πράκτορες (και από τον προγραμματιστή) την πολυπλοκότητα και την ποικιλομορφία των υποκείμενων επιπέδων (λειτουργικό σύστημα, hardware, δίκτυο, JVM). Μια απεικόνιση της αρχιτεκτονικής του JADE φαίνεται στο σχήμα 4-1.

Από λειτουργική άποψη, το JADE παρέχει τις όλες βασικές υπηρεσίες που είναι απαραίτητες για την εκτέλεση μιας κατανεμημένης πολύ-πρακτορικής εφαρμογής. Οι υπηρεσίες αυτές επιτρέπουν στους πράκτορες της εφαρμογής να δραστηριοποιούνται, εκτελώντας τις εργασίες που τους έχουν ανατεθεί, εντοπίζοντας με δυναμικό τρόπο άλλους πράκτορες και επικοινωνώντας μαζί τους σύμφωνα με το peer-to-peer μοντέλο.

Από την οπτική γωνία της εφαρμογής, κάθε πράκτορας διαθέτει τα παρακάτω λειτουργικά χαρακτηριστικά :

- έχει ένα μοναδικό όνομα που τον προσδιορίζει και τον διακρίνει από τους άλλους πράκτορες
- μπορεί να παρέχει υπηρεσίες τις οποίες έχει τη δυνατότητα να δηλώσει ώστε να γνωστοποιηθούν σε κάθε ενδιαφερόμενο
- μπορεί να αναζητήσει πράκτορες που παρέχουν συγκεκριμένες υπηρεσίες
- μπορεί να επικοινωνήσει με άλλους πράκτορες και να αλληλεπιδράσει μαζί τους
- μπορεί να ελέγξει τον κύκλο ζωής του



Σχήμα 4-1: Η αρχιτεκτονική του JADE

Η επικοινωνία των πρακτόρων γίνεται με την ανταλλαγή ασύγχρονων μηνυμάτων τα οποία διατυπώνονται στη γλώσσα επικοινωνίας πρακτόρων που καθορίζεται από το πρότυπο FIPA (FIPA ACL). Το JADE μπορεί να υποστηρίξει πολύπλοκες συζητήσεις μεταξύ των πρακτόρων, ενώ ταυτόχρονα παρέχει σκελετούς τυπικών μοτίβων αλληλεπίδρασης που επιτρέπουν την εκτέλεση εξειδικευμένων εργασιών όπως διαπραγματεύσεις, δημοπρασίες, ανάθεση αρμοδιοτήτων κλπ. Μέσα απ' αυτούς τους σκελετούς (που είναι υλοποιημένοι ως Java abstract κλάσεις) ο προγραμματιστής απαλλάσσεται από το φόρτο για το χειρισμό τεχνικών ζητημάτων που δεν αφορούν άμεσα την εφαρμογή που τον ενδιαφέρει

Για την αύξηση της επεκτασιμότητας και την υποστήριξη εφαρμογών σε περιβάλλον με περιορισμένους πόρους το JADE υλοποιεί ένα πράκτορα με ένα μόνο νήμα Java (Java thread). Επιπρόσθετα, μέσα στο νήμα του πράκτορα, παρέχεται η δυνατότητα εκτέλεσης πολλών παράλληλων εργασιών, με τις πιο σύνθετες εργασίες να μπορούν να δομηθούν από απλούστερες εργασίες ως μηχανές πεπερασμένων καταστάσεων (Finite States Machines).

Μια σημαντική υπηρεσία που υποστηρίζει το JADE είναι αυτή του «χρυσού οδηγού» (yellow pages service) μέσω της οποίας οι πράκτορες μπορούν να καταχωρήσουν τις υπηρεσίες που παρέχουν ή να αναζητήσουν αυτές που τους ενδιαφέρουν. Η υπηρεσία αυτή ονομάζεται Directory Facilitator και μπορεί να κατανεμηθεί ακόμη και πάνω σε διαφορετικούς σταθμούς.

Τέλος, αξίζει να σημειωθεί ότι το JADE έχει ήδη δοκιμαστεί με επιτυχία σε αρκετά περιβάλλοντα όπως Bluetooth, GPRS, W-LAN και Internet. Η επί μακρόν χρησιμοποίησή του σε πολύ μεγάλο αριθμό απαιτητικών εφαρμογών όπως βιομηχανικούς αυτοματισμούς, έλεγχο κυκλοφορίας και διαχείριση δικτύων, το καθιστά μια αποδεδειγμένα αξιόπιστη πλατφόρμα για την ανάπτυξη πολύ-πρακτορικών εφαρμογών.

Για περισσότερες τεχνικές πληροφορίες σχετικά με το JADE ο αναγνώστης μπορεί να ανατρέξει στον επίσημο ιστοχώρο της πλατφόρμας [44].

#### 4.2.4 Η γλώσσα ανάπτυξης

Η ανάπτυξη του προσομοιωτή έγινε στη γλώσσα Java. Η επιλογή αυτή ουσιαστικά υπαγορεύτηκε από την επιλογή της πλατφόρμας JADE, αφού ολόκληρο το API αλλά και ο πυρήνας του JADE είναι γραμμένα σε Java. Τα πολλά και σημαντικά πλεονεκτήματα της Java έναντι άλλων γλωσσών προγραμματισμού αποτελούν ένα επιπλέον επιχείρημα για την επιλογή του JADE. Ενδεικτικά, για λόγους πληρότητας, αναφέρουμε τα σημαντικότερα πλεονεκτήματα της Java, τα οποία με τον ένα ή τον άλλο τρόπο είχαν θετική συνεισφορά στην παρούσα εργασία :

- Η Java είναι μια σύγχρονη αντικειμενοστραφής γλώσσα, αρκετά απλούστερη από άλλες γλώσσες ίδιου επιπέδου. Ενσωματωμένα χαρακτηριστικά όπως ο garbage collector, η μη χρήση pointers και η υποστήριξη της πολλαπλής κληρονομικότητας με απλούστερες δομές (interface) βοηθούν σημαντικά στη μείωση της πολυπλοκότητας του παραγόμενου κώδικα.
- Με τους μηχανισμούς και τις δομές που διαθέτει αποτρέπει αρκετά από τα συνηθισμένα προγραμματιστικά λάθη διαμορφώνοντας κατάλληλες συνθήκες για την παραγωγή πιο αξιόπιστου κώδικα.
- Επιτυγχάνει καλύτερη διαχείριση νημάτων (threads) επιτρέποντας την εύκολη ανάπτυξη αξιόπιστων πολύ-νηματικών (multi-threading) εφαρμογών.
- Είναι δυναμική, καθώς έχει σχεδιαστεί για να προσαρμόζεται σε μεταβαλλόμενα περιβάλλοντα (φόρτωση κλάσεων κατά το χρόνο εκτέλεσης), επιτρέποντας με τον τρόπο αυτό τη δημιουργία ευέλικτων και αυτοδύναμων εφαρμογών.
- Ενσωματώνει ένα σημαντικό αριθμό υποστηρικτικών εργαλείων, ενώ παράλληλα υπάρχει διαθέσιμο ένα πλήθος βιβλιοθηκών ανοικτού κώδικα που μπορούν να χρησιμοποιηθούν από τον προγραμματιστή, μειώνοντας σημαντικά το χρόνο ανάπτυξης μιας εφαρμογής.
- Το μεγαλύτερο πλεονέκτημα της Java είναι η εξαιρετική μεταφορσιμότητα που προσφέρει, καθώς υποστηρίζει ένα πολύ μεγάλο αριθμό λειτουργικών συστημάτων και hardware. Είναι χαρακτηριστικό το σλόγκαν που την ακολουθεί: "Program once, Run anywhere".

Τα πολλά και σημαντικά πλεονεκτήματα της Java της έχουν χαρίσει τεράστια δημοτικότητα και αποδοχή. Η συνεχής εξέλιξή της αποτελεί εχέγγυο για τη μελλοντική υποστήριξη νέων τεχνολογιών, γεγονός που ως ένα βαθμό εξασφαλίζει και τη δυνατότητα μελλοντικής συντήρησης και αναβάθμισης των εφαρμογών.

#### 4.2.5 Η λειτουργικότητα των πρακτόρων στο JADE

Πριν αναλύσουμε τον τρόπο με τον οποίο μοντελοποιείται στον προσομοιωτή η λειτουργικότητα των πρακτόρων εμπιστοσύνης, είναι απαραίτητο να δούμε προηγουμένως εν συντομία τον τρόπο με τον οποίο το JADE υλοποιεί γενικά την λειτουργικότητα ενός πράκτορα.

Στο JADE, η πραγματική εργασία που εκτελεί ένας πράκτορας (αυτή που αφορά το αντικείμενο της εφαρμογής) περιλαμβάνεται μέσα σε ενότητες δράσης που ονομάζονται "συμπεριφορές" (behaviours). Η συμπεριφορά αντιπροσωπεύει μια εργασία που μπορεί να εκτελέσει ένας πράκτορας και υλοποιείται ως ένα αντικείμενο μιας κλάσης που επεκτείνει τη βασική κλάση των συμπεριφορών `java.core.behaviours.Behaviour`. Προκειμένου ένας πράκτορας να εκτελέσει την εργασία που περιγράφει μια κλάση συμπεριφοράς, αρκεί να προστεθεί στον πράκτορα ένα αντικείμενο αυτής της κλάσης. Αυτό γίνεται μέσω της μεθόδου `addBehaviour()` που διαθέτει η κλάση `Agent` των πρακτόρων. Συμπεριφορές μπορούν να προστεθούν σ' ένα πράκτορα οποιαδήποτε στιγμή, είτε κατά την εκκίνηση του πράκτορα είτε μέσα από οποιαδήποτε άλλη συμπεριφορά. Κάθε συμπεριφορά που προστίθεται με την `addBehaviour()` τοποθετείται σε μια εσωτερική (μη ορατή στον προγραμματιστή) λίστα εργασιών (task list) και περιμένει εκεί μέχρι να έρθει η σειρά της για εκτέλεση. Η εκτέλεση των συμπεριφορών γίνεται σειριακά με ένα non-pre-emptive τρόπο που υπαγορεύεται από το γεγονός ότι κάθε πράκτορας υλοποιείται από ένα και μόνο νήμα Java. Αυτό πρακτικά σημαίνει ότι για την εκτέλεση σύνθετων συμπεριφορών χρειάζεται κάποια ειδική διαχείριση της λίστας εργασιών. Ευτυχώς για τον προγραμματιστή, το JADE προσφέρει για το σκοπό αυτό αρκετά έτοιμα πρότυπα ειδικών συμπεριφορών, υλοποιημένα ως Java κλάσεις που επεκτείνουν τη βασική κλάση `Behaviour`. Δύο τέτοια πρότυπα που χρειάστηκαν στην περίπτωση του προσομοιωτή είναι η παράλληλη συμπεριφορά (`ParallelBehaviour`) και η συμπεριφορά πεπερασμένων καταστάσεων (`FSMBehaviour`).

Το πρότυπο `ParallelBehaviour` υλοποιεί σύνθετες συμπεριφορές που περιλαμβάνουν πολλές επιμέρους εργασίες που πρέπει να εκτελεστούν παράλληλα. Κάθε μια από τις επιμέρους εργασίες υλοποιείται ως μια ξεχωριστή συμπεριφορά.

Το πρότυπο `FSMBehaviour` υλοποιεί σύνθετες συμπεριφορές των οποίων η εργασία αναλύεται σε επιμέρους εργασίες η αλληλουχία εκτέλεσης των οποίων περιγράφεται από μια μηχανή πεπερασμένων καταστάσεων (finite state machine). Το συγκεκριμένο πρότυπο παρέχει αδιαφανώς όλη τη διαχείριση

που απαιτείται για την εναλλαγή των συνιστωσών υπο-εργασιών βάσει των συνθηκών που διαμορφώνονται μετά την εκτέλεση της κάθε συνιστώσας.

Η λειτουργικότητα ενός πράκτορα εμπιστοσύνης, όπως έχει αναλυθεί στο κεφάλαιο 3, περιλαμβάνει κάποιες κύριες λειτουργίες που υποστηρίζουν την εκδήλωση της κοινωνικότητας ενός πράκτορα (αναζήτηση αλληλεπιδράσεων, χειρισμός προσκλήσεων αλληλεπίδρασης, παροχή συστάσεων). Αυτές οι λειτουργίες, όπως θα εξηγηθεί σε επόμενη παράγραφο, είναι σύνθετες διαδικασίες που πρέπει να εκτελούνται παράλληλα προκειμένου να αποφευχθούν καταστάσεις αμοιβαίου αποκλεισμού (deadlock) που οδηγούν στην εμφάνιση ακούσιας αντικοινωνικής συμπεριφοράς (μη απόκριση σε προσκλήσεις). Για το λόγο αυτό, οι κύριες λειτουργίες ενσωματώθηκαν ως συνιστώσες σε μια σύνθετη υπερ-συμπεριφορά, η οποία υλοποιήθηκε κάνοντας χρήση του προτύπου *ParallelBehaviour*.

Κάθε μια από τις κύριες λειτουργίες αναλύεται σε ένα σύνολο επιμέρους εργασιών οι οποίες όμως δυστυχώς δεν ακολουθούν ένα προκαθορισμένο μοτίβο εκτέλεσης. Σε αρκετά σημεία της δραστηριοποίησής του ο πράκτορας χρειάζεται να παίρνει αποφάσεις και ανάλογα με τις επικρατούσες συνθήκες να κάνει επιλογές που διαφοροποιούν κάθε φορά τη σειρά εκτέλεσης των συνιστωσών εργασιών. Οι συνθήκες εκτέλεσης μιας συνιστώσας επηρεάζουν το αποτέλεσμα της εκτέλεσής της, το οποίο με τη σειρά του επηρεάζει την πορεία εκτέλεσης της περιβάλλουσας λειτουργίας. Για το λόγο αυτό οι κύριες λειτουργίες μοντελοποιήθηκαν ως μηχανές πεπερασμένων καταστάσεων και υλοποιήθηκαν κάνοντας χρήση του προτύπου συμπεριφοράς *FSMBehaviour*.

### 4.3 Η Αρχιτεκτονική του προσομοιωτή

Ο προσομοιωτής (TrustSIM) αποτελείται από δύο υποσυστήματα, το Configuration Center και το Simulation Center. Τα υποσυστήματα αυτά υλοποιούν τις δύο κύριες συνιστώσες της λειτουργικότητάς του που είναι, αφενός, η ρύθμιση των παραμέτρων προσομοίωσης και, αφετέρου, η διαχείριση της προσομοίωσης. Σχηματικά η αρχιτεκτονική του TrustSIM μπορεί να συνοψισθεί στο διάγραμμα που εμφανίζεται στο Σχήμα 4-2.

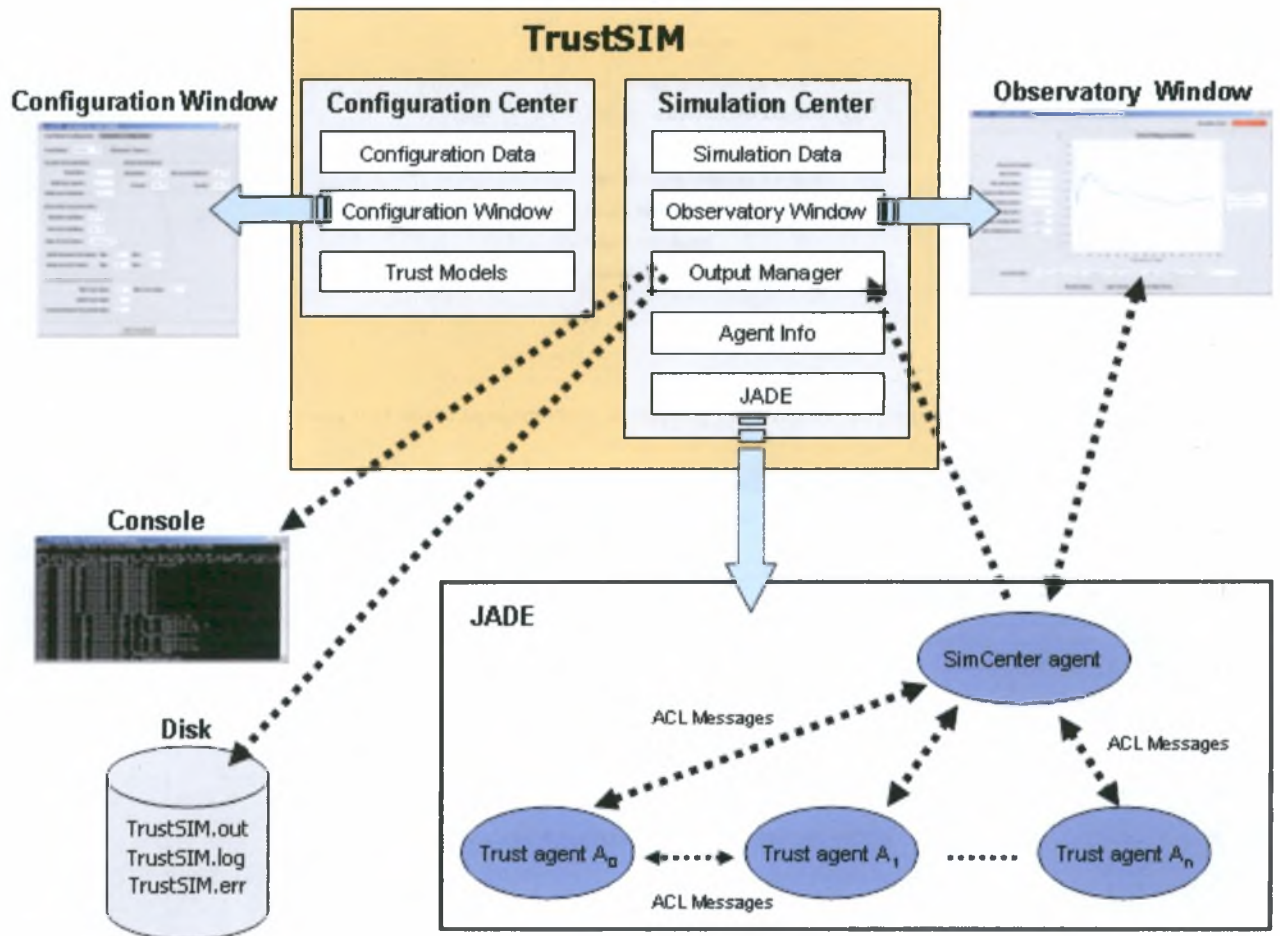
Στη συνέχεια ακολουθεί μια σύντομη περιγραφή των υποσυστημάτων του TrustSIM και των βασικών λειτουργιών τους.

- **Configuration Data**

Είναι η δομή δεδομένων που κρατά τις τιμές των παραμέτρων προσομοίωσης. Εδώ αποθηκεύονται όλες οι τιμές των παραμέτρων που ρυθμίζουν το πώς ακριβώς θα εκτελεστεί η προσομοίωση. Τα δεδομένα αυτά χρησιμοποιούνται, αφενός, από το υποσύστημα του Simulation Center για την εκκίνηση και την διαχείριση της προσομοίωσης, και, αφετέρου, από τους πράκτορες, προκειμένου να ακολουθήσουν την προβλεπόμενη από το υπό εξέταση μοντέλο συμπεριφορά. Αναλυτικά, οι παράμετροι προσομοίωσης που περιέχονται στη δομή Configuration Data παρουσιάζονται στο παράρτημα Α.

• **Configuration window**

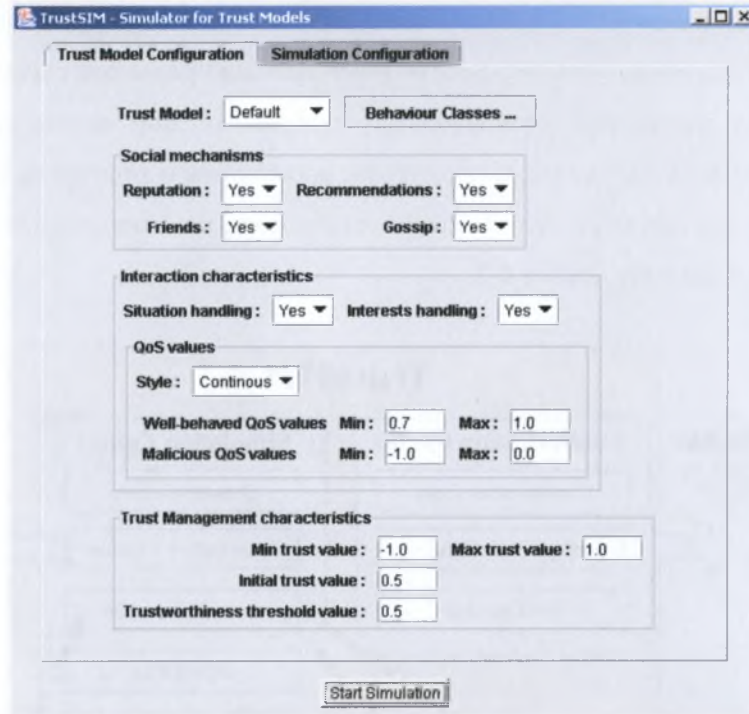
Είναι το παράθυρο διεπαφής χρήστη (user interface) μέσω του οποίου ο χρήστης εισάγει τις τιμές των παραμέτρων προσομοίωσης. Το παράθυρο αυτό παραμένει ανοικτό και κατά τη διάρκεια της εκτέλεσης της προσομοίωσης προκειμένου ο χρήστης να μπορεί να ανατρέχει στις επιλογές που έχει κάνει. Δύο ενδεικτικά στιγμιότυπα του Configuration window φαίνονται στην εικόνα 4-1 και στην εικόνα 4-2.



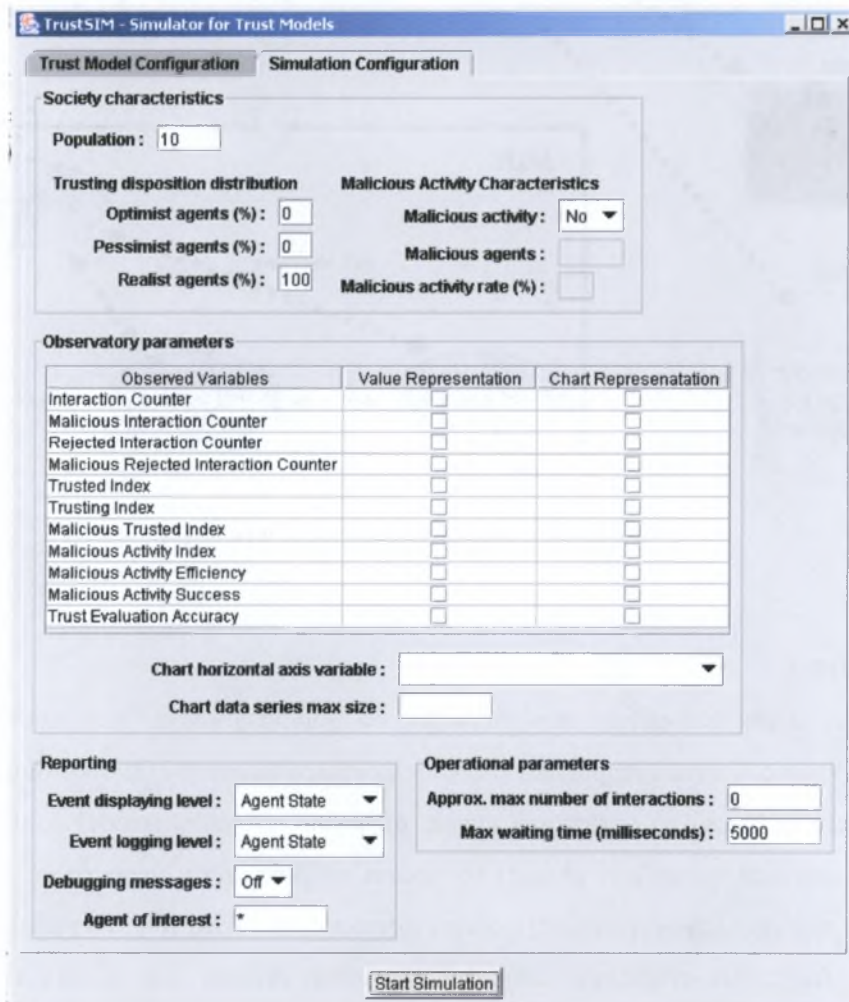
Σχήμα 4-2 : Η αρχιτεκτονική του TrustSIM

• **Trust Models**

Είναι η δομή δεδομένων που περιέχει τα χαρακτηριστικά των διαθέσιμων μοντέλων εμπιστοσύνης. Τα χαρακτηριστικά του κάθε μοντέλου φορτώνονται στη δομή κατά το ξεκίνημα του TrustSIM, από το αντίστοιχο αρχείο ρύθμισης μοντέλου (model configuration file). Τα αρχεία ρύθμισης μοντέλων (\*.mcf) θα πρέπει να βρίσκονται μέσα στον υποκατάλογο *Trust Models* του καταλόγου εργασίας (working directory) του TrustSIM. Το συντακτικό των αρχείων αυτών (όπως υποστηρίζεται από την τρέχουσα έκδοση του TrustSIM) αναλύεται στο παράρτημα Β.



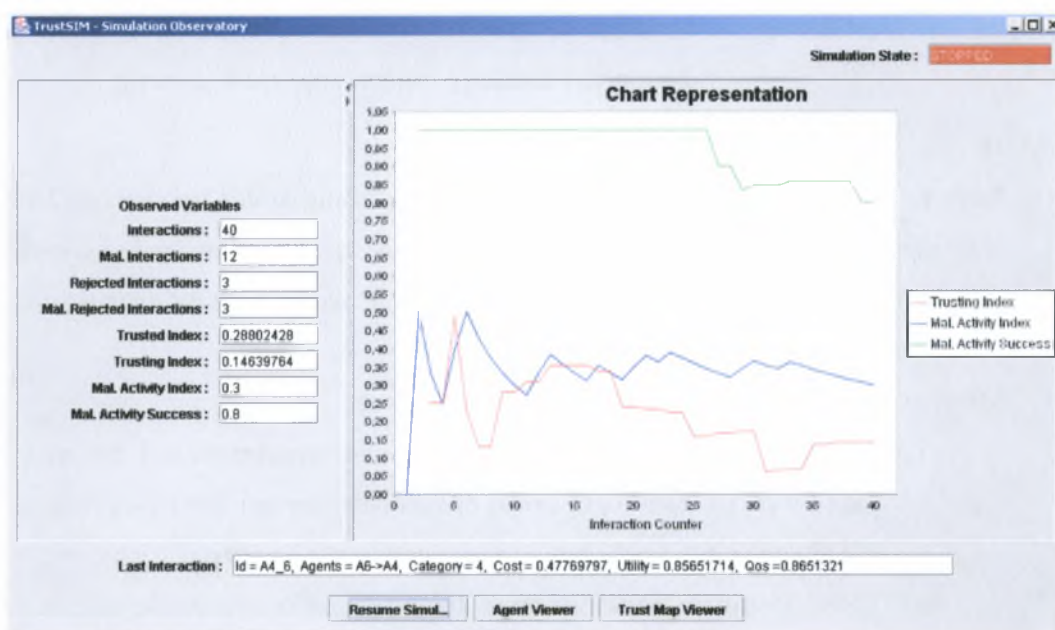
Εικόνα 4-1 : Configuration window (οι ρυθμίσεις των παραμέτρων του μοντέλου εμπιστοσύνης)



Εικόνα 4-2 : Configuration window (οι ρυθμίσεις του περιβάλλοντος προσομοίωσης)

- **Observatory window**

Είναι το παράθυρο - παρατηρητήριο της προσομοίωσης. Στο παράθυρο αυτό εμφανίζονται οι τιμές των μεταβλητών παρακολούθησης είτε ως απλά αριθμητικά δεδομένα είτε και ως γράφημα. Η απεικόνιση των μεταβλητών παρακολούθησης, στην τρέχουσα έκδοση του TrustSIM, γίνεται με δυναμικό τρόπο. Δηλ. οι τιμές εμφανίζονται σε πραγματικό χρόνο όπως ακριβώς διαμορφώνονται από τα στοιχεία που αποστέλλουν οι πράκτορες. Σε μια επόμενη έκδοση, το παρατηρητήριο θα μπορούσε να υλοποιηθεί και ως μια ανεξάρτητη εφαρμογή που θα διαβάζει τα δεδομένα από το αρχείο εξόδου (\*.out) που δημιουργεί η προσομοίωση στο δίσκο. Μια τέτοια υλοποίηση έχει δύο βασικά πλεονεκτήματα : α) επιτρέπει στο χρήστη να μελετήσει μια προσομοίωση off-line, δηλ. μετά την ολοκλήρωσή της και β) δίνει τη δυνατότητα σε περισσότερους από ένα χρήστες (ακόμη και απομακρυσμένους) να παρακολουθούν την εξέλιξη μιας συγκεκριμένης προσομοίωσης, ξεκινώντας από οποιοδήποτε χρονική στιγμή επιθυμούν. Ένα χαρακτηριστικό στιγμιότυπο του Observatory window φαίνεται στην εικόνα 4-3



Εικόνα 4-3 : Το Observatory window

- **Output Manager**

Είναι το υποσύστημα που υλοποιεί τις διαδικασίες παροχής στοιχείων της προσομοίωσης προς τον έξω κόσμο μέσω των περιφερειακών συσκευών. Στην τρέχουσα έκδοση του TrustSIM, υποστηρίζονται δύο περιφερειακές συσκευές, η κονσόλα και ο δίσκος. Στην κονσόλα εμφανίζονται στοιχεία που αφορούν την πορεία της προσομοίωσης με στόχο την on-line ενημέρωση του χρήστη. Τέτοια στοιχεία είναι για παράδειγμα τα συμβάντα που λαμβάνουν χώρα κατά τη διάρκεια της προσομοίωσης (π.χ. ότι έγινε πρόταση για αλληλεπίδραση από τον



πράκτορα X στον πράκτορα Y ή ότι ο πράκτορας Y αξιολόγησε τον πράκτορα X κλπ). Στο δίσκο καταγράφονται οι ίδιες πληροφορίες που εμφανίζονται και στην κονσόλα, οργανωμένες όμως σε ξεχωριστά αρχεία, ανάλογα με το είδος τους. Επιπλέον στο δίσκο καταγράφονται και οι τιμές των μεταβλητών παρακολούθησης που απεικονίζονται στο παρατηρητήριο. Συνοπτικά, τα στοιχεία που παρέχει ο output manager ανά περιφερειακή συσκευή είναι :

Στοιχείο	Console	Disk
Τιμές μεταβλητών παρακολούθησης		✓
Συμβάντα προσομοίωσης	✓	✓
Μηνύματα προειδοποίησης & λάθους	✓	✓

Τα στοιχεία που αποστέλλονται στο δίσκο καταγράφονται σε τρία ξεχωριστά αρχεία, τα οποία έχουν ίδιο κυρίως όνομα και διαφορετική κατάληξη. Τα αρχεία αυτά είναι :

TrustSIM.out : Περιέχει τις τιμές των μεταβλητών παρακολούθησης

TrustSIM.log : Περιέχει τα συμβάντα (events) της προσομοίωσης

TrustSIM.err : Περιέχει τα λάθη (errors) και τις προειδοποιήσεις (warnings) που προέκυψαν κατά την διάρκεια της προσομοίωσης

#### • JADE

Είναι το τμήμα του simulation center που κρατά τα δεδομένα που αφορούν το JADE, δηλ. την πλατφόρμα πάνω στην οποία γεννιούνται και ζουν οι πράκτορες της προσομοίωσης. Το τμήμα αυτό περιέχει τις παραμέτρους λειτουργίας του JADE και παρέχει τις απαραίτητες διαδικασίες για τη φόρτωση και τον τερματισμό του.

Μέσα στο JADE δημιουργούνται δύο ειδών πράκτορες :

- i. Οι πράκτορες εμπιστοσύνης (trust agents) που απαρτίζουν την προσομοιούμενη κοινωνία. Οι πράκτορες αυτοί δραστηριοποιούνται αυτόνομα, ακολουθώντας τις διαδικασίες που προβλέπει το προσομοιούμενο μοντέλο.
- ii. Ένας κεντρικός πράκτορας που λειτουργεί ως ο ενδιάμεσος κόμβος επικοινωνίας ανάμεσα στους πράκτορες εμπιστοσύνης και στο παρατηρητήριο.

Ο κεντρικός πράκτορας είναι λειτουργικά συνδεδεμένος με το παρατηρητήριο<sup>20</sup>, τροφοδοτώντας το με στοιχεία από την προσομοίωση. Ο ρόλος του είναι διπλός :

- συλλέγει τα στοιχεία που στέλνουν οι πράκτορες εμπιστοσύνης σχετικά με την πορεία της προσομοίωσης και τα προωθεί στο παρατηρητήριο και τον Output Manager
- ενημερώνει τους πράκτορες εμπιστοσύνης για θέματα τους αφορούν (π.χ. για το θάνατο ή τη γέννηση κάποιου πράκτορα).

<sup>20</sup> Μιλώντας με τεχνικούς όρους του JADE, ο κεντρικός πράκτορας είναι ένας GUI agent με το GUI του να είναι το Observatory Window.

Οι πράκτορες επικοινωνούν μεταξύ τους και με τον κεντρικό πράκτορα ανταλλάσσοντας μηνύματα τύπου ACL. Τα είδη των μηνυμάτων που χρησιμοποιούνται από τους πράκτορες παρουσιάζονται στο παράρτημα Γ. Σε ότι αφορά τη μορφή των μηνυμάτων έγινε προσπάθεια να τηρηθούν οι προδιαγραφές της FIPA ώστε να είναι όσο το δυνατόν ευκολότερη είτε η προσθήκη συμπεριφορών από υπάρχουσες βιβλιοθήκες συμπεριφορών πρακτόρων είτε η ενσωμάτωση του περιβάλλοντος του TrustSIM μέσα σε κάποιο πραγματικό multi agent περιβάλλον με σκοπό την παρακολούθησή και τη μελέτη των λειτουργικών χαρακτηριστικών του.

- **Agent info**

Είναι μια ομάδα δεδομένων που χρειάζεται ο simulation center για το χειρισμό των πρακτόρων που απαρτίζουν την κοινωνία της προσομοίωσης. Τα δεδομένα αυτά είναι :

- i. Οι ελεγκτές (controllers) των πρακτόρων. Οι ελεγκτές προέρχονται από το περιβάλλον του JADE και χρησιμοποιούνται για τον έλεγχο της κατάστασης λειτουργίας των πρακτόρων (διαδικασίες suspend, resume, kill).
- ii. Ο χάρτης των κακόβουλων πρακτόρων. Η πληροφορία για το ποιοι πράκτορες είναι κακόβουλοι χρειάζεται κατά τον υπολογισμό των μετρικών που αξιολογούν χαρακτηριστικά των κακόβουλων συμπεριφορών.
- iii. Ο χάρτης των "ζωντανών" πρακτόρων. Κάποιοι πράκτορες μπορεί για κάποιο λόγο να τερματίσουν τη λειτουργία τους ή μπορεί να χρειαστεί να τερματιστούν πριν την ολοκλήρωση της προσομοίωσης. Χρειάζεται λοιπόν ο simulation center να γνωρίζει σε κάθε στιγμή ποιοι πράκτορες είναι "ζωντανοί", ώστε να μην γίνει κάποια αναφορά (επικοινωνία ή αλλαγή κατάστασης) σε ανύπαρκτο πράκτορα.

#### **4.4 Οι βασικές λειτουργίες του TrustSIM**

Ξεκινώντας, ο TrustSIM προσπαθεί να βρει και να φορτώσει τα διαθέσιμα μοντέλα εμπιστοσύνης. Για το σκοπό, ελέγχει το φάκελο *Trust Models* ψάχνοντας για αρχεία ρύθμισης μοντέλου (αρχεία με κατάληξη \*.mcf). Τα αρχεία που θα εντοπίσει θα τα διαβάσει ενημερώνοντας τη δομή Trust Models με τα χαρακτηριστικά των διαθέσιμων για προσομοίωση μοντέλων εμπιστοσύνης. Αυτός είναι και ο μοναδικός τρόπος μέσω του οποίου μπορεί να εισαχθεί ένα καινούργιο μοντέλο μέσα στον TrustSIM. Δηλ. για κάθε νέο μοντέλο εμπιστοσύνης που ο χρήστης επιθυμεί να μελετήσει, θα πρέπει να δημιουργήσει ένα κατάλληλο \*.mcf αρχείο με τα χαρακτηριστικά του μοντέλου. Το αρχείο αυτό θα πρέπει να το προσθέσει στον φάκελο *Trust Models* ώστε να διαβαστεί κατά το επόμενο startup του TrustSIM..

Η διαδικασία εντοπισμού των διαθέσιμων μοντέλων, εκτός από το διάβασμα των αρχείων ρύθμισης μοντέλου του φακέλου *Trust Models*, παράγει και ένα σύνολο ρυθμίσεων που ορίζουν ένα

ψευδομοντέλο. Το μοντέλο αυτό, που χαρακτηρίζεται ως Default Model, περιλαμβάνει τις εξ' ορισμού τιμές για όλες τις παραμέτρους προσομοίωσης του TrustSIM. Η ύπαρξη του default μοντέλου είναι απαραίτητη προκειμένου να υπάρχει ένα υπόβαθρο αρχικών ρυθμίσεων πάνω στο οποίο να τοποθετούνται οι ρυθμίσεις του μοντέλου που πρόκειται να προσομοιωθεί. Με τον τρόπο αυτό ένα μοντέλο μπορεί να μην περιλαμβάνει όλες τις δυνατές ρυθμίσεις που χρειάζεται ο TrustSIM αλλά ένα υποσύνολο αυτών. Αυτό είναι ιδιαίτερα χρήσιμο αφού, όπως αναφέρθηκε και στο κεφάλαιο της επισκόπησης της βιβλιογραφίας, σχεδόν στο σύνολό τους τα μοντέλα εμπιστοσύνης εστιάζουν το ενδιαφέρον τους σε συγκεκριμένες περιοχές του προβλήματος της διαχείρισης εμπιστοσύνης, προτείνοντας μηχανισμούς και διαδικασίες μόνο γι' αυτές τις περιοχές (αναφ.). Για τις υπόλοιπες πλευρές του προβλήματος, που ένα μοντέλο δεν περιγράφει τι πρέπει να γίνει, χρησιμοποιούνται οι ρυθμίσεις και οι μηχανισμοί του default μοντέλου.

Το επόμενο βήμα στην λειτουργία του TrustSIM είναι η εμφάνιση του Configuration Window που επιτρέπει στο χρήστη να καθορίσει τις επιλογές του σχετικά με την επικείμενη προσομοίωση. Ο χρήστης, μέσω του Configuration Window, μπορεί να ρυθμίσει όλες τις παραμέτρους που περιέχονται στη δομή Configuration Data. (βλ. παράρτημα Α). Το περιβάλλον του Configuration Window είναι σχεδιασμένο έτσι ώστε να προσφέρει στο χρήστη τη μέγιστη δυνατή ελευθερία και να μπορεί να επιλεγεί οποιοσδήποτε δυνατός συνδυασμός ρυθμίσεων. Η ρυθμίσεις αυτές που θα καθορίσουν τα χαρακτηριστικά και τις γενικές κατευθύνσεις της προσομοίωσης, δεν μπορούν να αλλάξουν κατά τη διάρκεια εκτέλεσης της προσομοίωσης. Μόνο μετά τον τερματισμό της επιτρέπεται να αλλάξουν από το χρήστη και στην περίπτωση αυτή θα ισχύσουν στην επόμενη προσομοίωση.

Μετά την ρύθμιση των παραμέτρων ακολουθεί η εκτέλεση της προσομοίωσης. Η προσομοίωση μπορεί να χωριστεί σε τρεις φάσεις : την εκκίνηση (startup), την εκτέλεση (execution) και τον τερματισμό (shut down).

Κατά τη φάση της εκκίνησης :

- Δημιουργείται το Simulation Center που θα διαχειριστεί την προσομοίωση
- Φορτώνεται η πλατφόρμα του JADE (δημιουργείται το main container πάνω στο οποίο θα δημιουργηθούν και θα τρέξουν οι πράκτορες της προσομοίωσης)
- Δημιουργείται ο κεντρικός πράκτορας που θα λειτουργήσει ως κέντρο επικοινωνίας
- Δημιουργούνται οι πράκτορες εμπιστοσύνης της προσομοιούμενης κοινωνίας

Κατά τη φάση της εκτέλεσης οι πράκτορες εμπιστοσύνης δραστηριοποιούνται ακολουθώντας τους μηχανισμούς του προσομοιούμενου μοντέλου και σύμφωνα με τις ρυθμίσεις που έχουν καθοριστεί από το χρήστη. Στη φάση αυτή, μετά από κάθε ενέργεια που μπορεί να έχει επιφέρει αλλαγή σε κάποια μεταβλητή παρακολούθησης, οι πράκτορες αποστέλλουν τα δεδομένα που σχετίζονται με την ενέργεια στον κεντρικό πράκτορα, ώστε εκείνος με τη σειρά του να ενημερώσει τα simulation data και το παρατηρητήριο. Το ίδιο συμβαίνει και για τα συμβάντα που λαμβάνουν χώρα (π.χ. προσκλήσεις για

αλληλεπίδραση, αιτήσεις για συστάσεις κλπ) και τα οποία ο χρήστης έχει ορίσει (μέσω του configuration window) ότι τον ενδιαφέρουν.

Όταν ο χρήστης επιλέξει να τερματίσει την προσομοίωση ή εναλλακτικά όταν πραγματοποιηθεί ο προκαθορισμένος από το χρήστη αριθμός αλληλεπιδράσεων, ξεκινά η φάση του τερματισμού. Κατά τη φάση αυτή, διαδοχικά :

- Τερματίζονται όσοι πράκτορες εμπιστοσύνης εξακολουθούν να είναι ενεργοί
- Τερματίζεται ο κεντρικός πράκτορας
- Τερματίζεται η πλατφόρμα του JADE
- Τερματίζεται το simulation center
- Ενεργοποιείται το παράθυρο ρυθμίσεων (Configuration window) και ξεκινά ένας νέος γύρος καθορισμού ρυθμίσεων για την επόμενη προσομοίωση.

#### 4.5 Το βασικό μοντέλο του TrustSIM

Όπως φαίνεται από την ανάλυση της αρχιτεκτονικής του TrustSIM , κατά την διάρκεια μιας προσομοίωσης αναπτύσσονται δύο είδη δραστηριότητας : η δραστηριότητα του κεντρικού πράκτορα, ο οποίος συντονίζει την προσομοίωση, και η δραστηριότητα των πρακτόρων εμπιστοσύνης της προσομοιούμενης κοινωνίας. Η δεύτερη προσδιορίζεται από το συγκεκριμένο μοντέλο εμπιστοσύνης που προσομοιώνεται. Όπως έχει αναφερθεί και κατά τη συζήτηση των προδιαγραφών και της αρχιτεκτονικής, ο TrustSIM έχει ενσωματωμένο ένα βασικό μοντέλο εμπιστοσύνης (default model). Το μοντέλο αυτό είναι προεπιλεγμένο ώστε να λειτουργεί ως υπόβαθρο για οποιοδήποτε άλλο μοντέλο επιλεγεί από το χρήστη για να προσομοιωθεί. Ουσιαστικά, το βασικό μοντέλο προσφέρει μια βασική λειτουργικότητα, που μπορεί να χρησιμοποιηθεί για την κάλυψη των κενών λειτουργικότητας που πιθανόν υπάρχουν σε ένα μοντέλο εμπιστοσύνης. Με τον τρόπο αυτό είναι εφικτή η προσομοίωση ακόμη και ημιτελών μοντέλων ή μοντέλων που απλά εστιάζουν την προσοχή τους σε κάποια συγκεκριμένη περιοχή του προβλήματος της διαχείρισης εμπιστοσύνης.

Όπως θα δούμε στη συνέχεια, το βασικό μοντέλο αποτελείται από ένα σύνολο ανεξάρτητων συμπεριφορών<sup>21</sup> που εκτελούνται από τους πράκτορες της προσομοιούμενης κοινωνίας. Η λειτουργικότητα αυτή θα μπορούσε ενδεχομένως να αποτελέσει και ένα οδηγό για κάθε μελλοντικό σχεδιαστή ενός μοντέλου εμπιστοσύνης, αφού περιλαμβάνει όλα τα δομικά στοιχεία που είναι απαραίτητα για τη σύνθεση μιας ολοκληρωμένης κοινωνικής δραστηριότητας σε κοινωνίες πρακτόρων εμπιστοσύνης. Σημειώνουμε ότι για καθεμιά από τις προτεινόμενες συμπεριφορές υπάρχουν πολλοί διαφορετικοί τρόποι υλοποίησης. Η υλοποίηση που υιοθετήθηκε για το βασικό μοντέλο προέκυψε συνθέτοντας διάφορες επιλογές που εμφανίζονται στα μοντέλα εμπιστοσύνης που μελετήθηκαν. Δεν

<sup>21</sup> Ο όρος συμπεριφορά χρησιμοποιείται σε συμφωνία με τον αντίστοιχο όρο που αναφέρεται στην πλατφόρμα JADE

δαπανήθηκε όμως κάποια ιδιαίτερη προσπάθεια για τη βελτιστοποίησή της, καθώς ο ρόλος του βασικού μοντέλου είναι απλά για να προσφέρει μια εφεδρική λειτουργικότητα στην περίπτωση που αυτή απουσιάζει από το προσομοιούμενο μοντέλο.

#### 4.5.1 Συμβάσεις

Για την υλοποίηση της βασικής λειτουργικότητας των πρακτόρων έγιναν οι παρακάτω συμβάσεις :

- Ένας πράκτορας μπορεί να είναι είτε καλοπροαίρετος είτε κακόβουλος.
- Κάθε πράκτορας έχει κάποια ενδιαφέροντα τα οποία αντικατοπτρίζουν τις ανάγκες του.
- Κάθε πράκτορας μπορεί να έχει κάποια ειδικευση η οποία αντιστοιχεί σε κάποιο από τα ενδιαφέροντά του<sup>22</sup>. Τα ενδιαφέροντα βρίσκονται σε αντιστοιχία με τις ειδικεύσεις, δηλ. για κάθε ενδιαφέρον υπάρχει αντίστοιχη ειδικευση.
- Οι πράκτορες προσφέρουν υπηρεσίες στους άλλους πράκτορες βάσει της ειδικεισής τους. Οι πράκτορες που δεν διαθέτουν ειδικευση δεν μπορούν προσφέρουν υπηρεσίες γι' αυτό και κανείς πράκτορας δεν τους ζητά αλληλεπίδραση. Έτσι αρκούνται απλά στο να ζητούν υπηρεσίες από τους άλλους.
- Οι πράκτορες αρχικά δεν γνωρίζουν τίποτε για την κοινωνία στην οποία ανήκουν. Οι αρχικές αναζητήσεις αλληλεπίδρασης βασίζονται αποκλειστικά σε κάποια υπηρεσία ευρετηρίου που παρέχεται από την κοινωνία. Η γνώση τους για τους άλλους πράκτορες της κοινωνίας "κτίζεται" μέσα από τις αλληλεπιδράσεις τους.
- Κατά τη φάση αναζήτησης αλληλεπίδρασης, ο πράκτορας επιλέγει ένα από τα ενδιαφέροντά του και αναζητά τον καταλληλότερο ειδικευμένο πράκτορα για να του παρέχει την αντίστοιχη υπηρεσία. Για τις ανάγκες της προσομοίωσής μας αυτή η επιλογή γίνεται τυχαία. Σε μια πολυπρακτορική εφαρμογή για συγκεκριμένο τομέα, κάθε πράκτορας θα αναζητά παροχή συγκεκριμένων υπηρεσιών, ανάλογα με τους τρέχοντες στόχους του.
- Κάθε αλληλεπίδραση έχει κάποια χαρακτηριστικά τα οποία ορίζουν την κατάσταση (situation) της αλληλεπίδρασης. Τα χαρακτηριστικά αυτά είναι : α) το είδος της προσφερόμενης υπηρεσίας β) το κόστος παροχής της υπηρεσίας και γ) η ωφελιμότητα από την απόκτηση της υπηρεσίας.
- Σε κάθε αλληλεπίδραση που λαμβάνει χώρα, οι συναλλασσόμενοι πράκτορες κάνουν μια ανταλλαγή υπηρεσιών (ή αγαθών). Μετά την ολοκλήρωση της αλληλεπίδρασης η κάθε πλευρά προβαίνει σε κάποιου είδους αξιολόγηση της ποιότητας εξυπηρέτησης (quality of service). Οι δύο αξιολογήσεις είναι γενικά διαφορετικές και χαρακτηρίζουν την αλληλεπίδραση από την κάθε πλευρά ξεχωριστά.

<sup>22</sup> Αν και δεν ισχύει πάντα, είναι λογικό να υποθέτουμε ότι ένας πράκτορας δεν μπορεί να είναι ειδικός σε ένα αντικείμενο που δεν τον ενδιαφέρει

- Οι κακόβουλοι πράκτορες, εκτός από κακές υπηρεσίες, μπορούν να δίνουν και ψευδής συστάσεις.
- Οι πράκτορες δραστηριοποιούνται ομοιόμορφα. Αυτό σημαίνει ότι δεν υπάρχουν κάποιοι πράκτορες που να είναι περισσότερο ενεργοί από τους άλλους.
- Οι πράκτορες έχουν περιορισμένη μνήμη. Μπορούν δηλ. να συγκρατήσουν μόνο ένα πεπερασμένο αριθμό στοιχείων (αξιολογήσεων, δεδομένων αλληλεπιδράσεων κλπ). Έτσι, όταν γεμίσει μια αποθηκευτική τους δομή, για κάθε επιπλέον στοιχείο που καταγράφεται σ' αυτή διαγράφεται κάποιο παλαιότερο.
- Η εμπιστοσύνη και η φήμη αναπαρίστανται ως πραγματικοί αριθμοί στην περιοχή  $[-1,1)$  Η μορφή αυτή αφενός είναι η πιο συνηθισμένη στα μοντέλα εμπιστοσύνης και αφετέρου είναι η πιο εκφραστική αφού επιτρέπει μια πολύ μεγάλη διακύμανση της εμπιστοσύνης.
- Η αξιολόγηση της ποιότητας εξυπηρέτησης (QoS) αναπαρίσταται ως πραγματικός αριθμός στην περιοχή  $[-1,1]$ .
- Η βασική λειτουργικότητα χαρακτηρίζεται από μια θετική φιλοσοφία θεωρώντας ότι οι περισσότερες αλληλεπιδράσεις θα είναι "καλές", δηλαδή δεν θα είναι επιβλαβείς για τους πράκτορες. Έτσι, όπως θα δούμε στη συνέχεια, κάθε πράκτορας διαδίδει αυτόκλητα πληροφορίες (κουτσομπολεύει) μόνο σε σχέση με τις κακές εμπειρίες που έχει από τις αλληλεπιδράσεις του με άλλους πράκτορες.

#### 4.5.2 Οι συμπεριφορές του βασικού μοντέλου

Οι συμπεριφορές που απαρτίζουν το βασικό μοντέλο του TrustSIM ουσιαστικά αποτελούν υλοποιήσεις των βασικών μηχανισμών συμπεριφοράς που καταγράφηκαν στην ενότητα 3.2.3 του 3<sup>ου</sup> κεφαλαίου. Μέσω των μηχανισμών αυτών ένας πράκτορας εμπιστοσύνης μπορεί να δραστηριοποιείται και να αλληλεπιδρά με άλλους πράκτορες συνάπτοντας μαζί τους σχέσεις, σε μια προσπάθεια να εκπληρώσει τους προσωπικούς του στόχους και να ικανοποιήσει τις ανάγκες του.

Για την υλοποίηση των συμπεριφορών του βασικού μοντέλου υιοθετήθηκε μια αρθρωτή (modular) αρχιτεκτονική σύμφωνα με την οποία οι σύνθετες συμπεριφορές δομούνται από άλλες απλούστερες, ακολουθώντας μια λογική συναρμολόγησης διακριτών και αυτόνομων δομικών στοιχείων. Για το σκοπό αυτό, κάθε συμπεριφορά (απλή ή σύνθετη) έχει σχεδιαστεί ως ένα μαύρο κουτί που δέχεται μια είσοδο, επιτελεί κάποια λειτουργία και παράγει μια έξοδο. Μ' αυτή την top-down προσέγγιση υποστηρίζονται δύο από τις βασικές σχεδιαστικές προδιαγραφές του TrustSIM : α) η εύκολη ενσωμάτωση νέων μοντέλων και β) η δυνατότητα εξερεύνησης χαρακτηριστικών που δεν προβλέπονται στον αρχικό σχεδιασμό ενός μοντέλου. Κι αυτό γιατί μέσα από την παραπάνω προσέγγιση, ένα μοντέλο εμπιστοσύνης στο περιβάλλον του TrustSIM δεν είναι παρά ένα σύνολο διακριτών υλοποιήσεων των βασικών συμπεριφορών που μπορούν εύκολα να αντικατασταθούν από άλλες υλοποιήσεις, δημιουργώντας ενδιαφέροντες και ίσως ανεξερεύνητους συνδυασμούς.

Οι σύνθετες συμπεριφορές έχουν υλοποιηθεί ως μηχανές πεπερασμένων καταστάσεων (finite state machine), όπου κάθε κατάσταση καταστάσεις είναι άλλη, απλή ή σύνθετη συμπεριφορά. Κάθε σύνθετη συμπεριφορά περικλείει τις συνιστώσες καταστάσεις-συμπεριφορές της ανάμεσα σε μια αρχική και μια τελική κατάσταση. Μέσα στις καταστάσεις αυτές συνήθως λαμβάνουν χώρα αρχικοποιήσεις μεταβλητών, έλεγχοι των δεδομένων εισόδου, μορφοποιήσεις δεδομένων εξόδου και γενικά λειτουργίες που παρέχουν τη διασύνδεση των συμπεριφορών<sup>23</sup>. Στη συνέχεια του κεφαλαίου γίνεται μια σύντομη παρουσίαση των συμπεριφορών του βασικού μοντέλου.

## Η βασική συμπεριφορά

Στο υψηλότερο επίπεδο σύνθεσης συναντάμε τη βασική συμπεριφορά. Η βασική συμπεριφορά περιλαμβάνει όλη τη λειτουργικότητα ενός πράκτορα εμπιστοσύνης, την οποία αναπτύσσει μέσα από πέντε παράλληλες δράσεις. Οι δράσεις αυτές είναι :

- η αναζήτηση αλληλεπίδρασης
- ο χειρισμός αίτησης για αλληλεπίδραση
- ο χειρισμός αίτησης για παροχή πληροφοριών
- ο χειρισμός διαδιδόμενων πληροφοριών
- ο χειρισμός των μηνυμάτων του κεντρικού πράκτορα

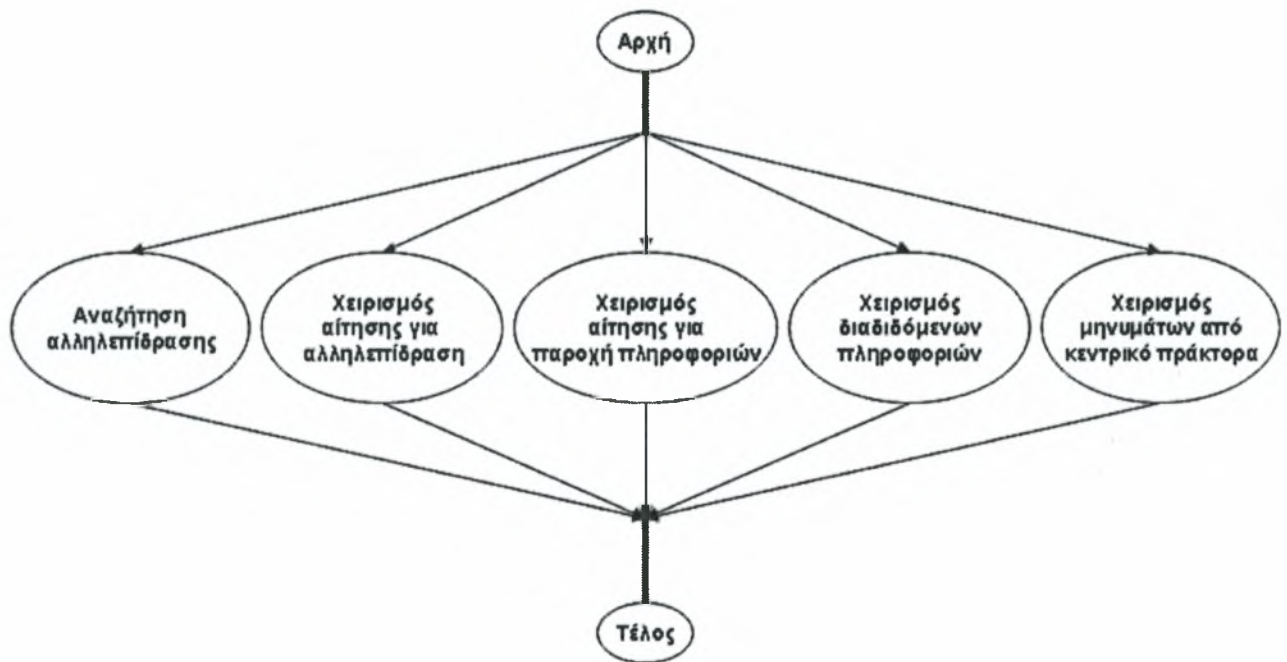
Κάθε μια από τις δράσεις αυτές υλοποιείται από μια ξεχωριστή και ανεξάρτητη συνιστώσα συμπεριφορά. Η αναζήτηση αλληλεπίδρασης αποβλέπει στην ικανοποίηση των προσωπικών αναγκών και στόχων του πράκτορα, ο χειρισμός των διαδιδόμενων πληροφοριών και των αιτήσεων για αλληλεπίδραση και παροχή πληροφοριών στην ανάπτυξη της κοινωνικότητας του πράκτορα, ενώ ο χειρισμός των μηνυμάτων του κεντρικού πράκτορα αφορά τη συγκεκριμένη υλοποίηση και υποστηρίζει τις διαδικασίες διαχείρισης της προσομοίωσης. Όλες οι συνιστώσες συμπεριφορές είναι αέναες, δηλαδή επαναλαμβάνονται διαρκώς, καθ' όλη τη διάρκεια της ζωής του πράκτορα.

Η παράλληλη εκτέλεση των συνιστωσών συμπεριφορών είναι επιβεβλημένη προκειμένου να αποφευχθούν καταστάσεις αμοιβαίου αποκλεισμού (deadlock) ή λανθασμένης απόρριψης συνεργασίας λόγω κακού συγχρονισμού. Τέτοιες καταστάσεις μπορεί να εμφανιστούν όταν, για παράδειγμα, ενώ ο πράκτορας Α εκτελεί τη συμπεριφορά αναζήτησης αλληλεπίδρασης και περιμένει απάντηση από τον πράκτορα Β (στον οποίο έχει κάνει αίτηση αλληλεπίδρασης), ο Β δεν απαντά γιατί περιμένει ανάλογη απάντηση από τον Γ (στον οποίο έχει κάνει παρόμοια αίτηση), ο οποίος όμως έχει επίσης κάνει αίτηση στην Α και περιμένει απάντηση. Παρόμοια, εάν ένας πράκτορας δεχτεί ένα αίτημα (για αλληλεπίδραση ή παροχή πληροφοριών) ενώ βρίσκεται εν μέσω

<sup>23</sup> Αυτές οι καταστάσεις δεν αντιστοιχούν σε κάποιες γενικές δράσεις της συνολικής δραστηριότητας των πρακτόρων, αλλά αποτελούν αναγκαία δομικά στοιχεία της συγκεκριμένης υλοποίησης. Γι' αυτό και θα

μιας αλληλεπίδρασης, θα πρέπει να μπορεί να το χειριστεί άμεσα. Διαφορετικά, θα στείλει την απάντησή του μετά το πέρας της αλληλεπίδρασης και μπορεί ο άλλος πράκτορας να έχει πάψει να περιμένει, οπότε η απάντησή του θα αγνοηθεί<sup>24</sup>. Από τα παραπάνω, γίνεται φανερό ότι ένας πράκτορας είναι απαραίτητο να εκτελεί κάθε συμπεριφορά που εμπεριέχει αναμονή παράλληλα, ώστε να είναι σε θέση, κατά τη διάρκεια της αναμονής, να επικοινωνεί και να αποκρίνεται σε κλήσεις από το εξωτερικό περιβάλλον. Αυτή είναι μια γενική αρχή που τηρήθηκε με συνέπεια κατά την υλοποίηση όλων των συμπεριφορών του βασικού μοντέλου.

Το διάγραμμα καταστάσεων της βασικής συμπεριφοράς φαίνεται στο σχήμα 4-3. Η σειρά με την οποία παρουσιάζονται οι συνιστώσες συμπεριφορές είναι τυχαία και δεν υποδηλώνει προτεραιότητα ή σπουδαιότητα.



Σχήμα 4-3 : Η βασική συμπεριφορά

### Αναζήτηση αλληλεπίδρασης

Η αναζήτηση αλληλεπίδρασης είναι μια σύνθετη συμπεριφορά που έχει ως στόχο την εκτέλεση μιας αλληλεπίδρασης για την κάλυψη κάποιας ανάγκης του πράκτορα. Η λειτουργία της περιλαμβάνει τις παρακάτω δράσεις :

- Δημιουργία των χαρακτηριστικών της περίπτωσης αλληλεπίδρασης.

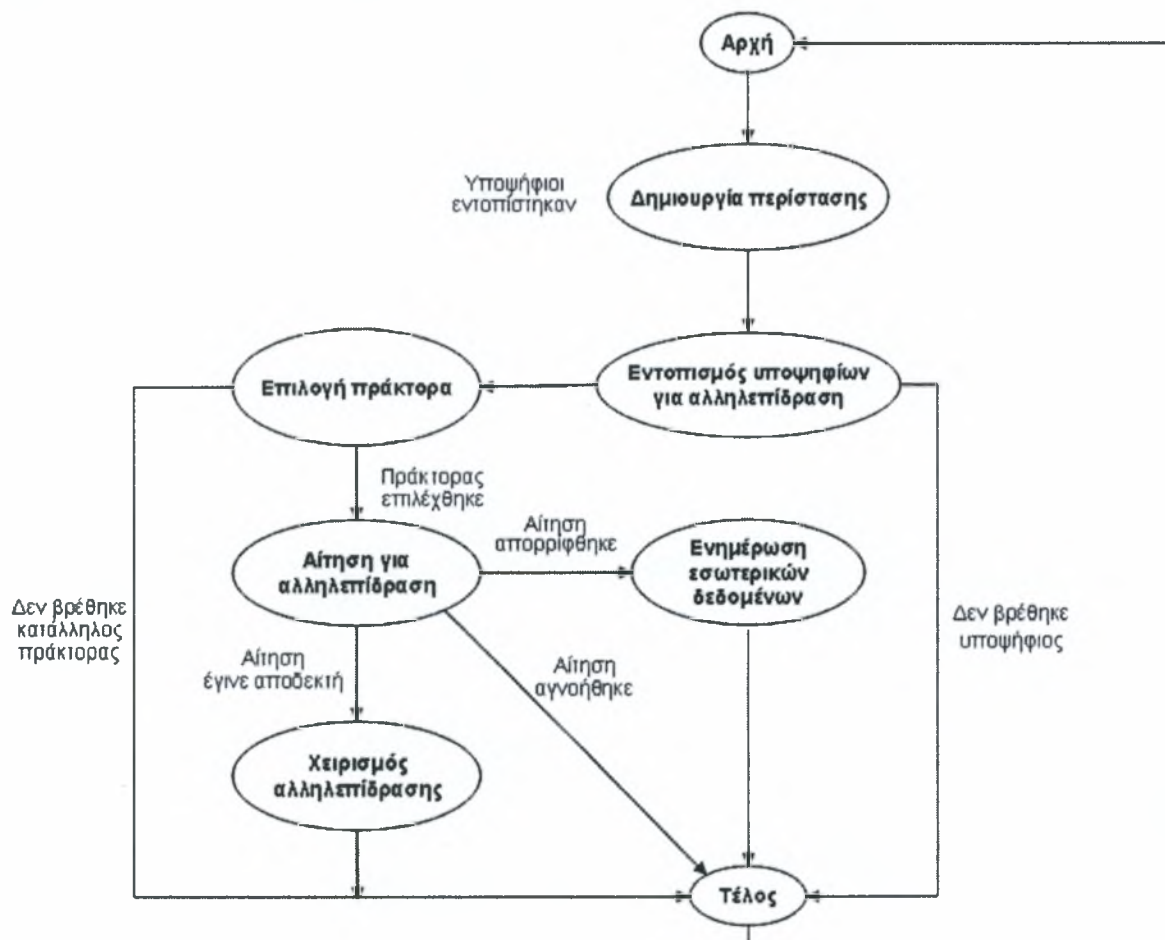
συζητηθούν μόνο σε ότι αφορά τη λειτουργικότητα που προσφέρουν στις συμπεριφορές στις οποίες περιλαμβάνονται.

<sup>24</sup> Το πρόβλημα της αναμονής μιας απάντησης συζητείται αναλυτικά κατά την παρουσίαση της συμπεριφοράς εκτέλεσης αλληλεπίδρασης (παρ. 0 )



- Εντοπισμό των υποψήφιων (κατάλληλων) πρακτόρων για την συγκεκριμένη αλληλεπίδραση
- Επιλογή ενός εκ των υποψηφίων πρακτόρων
- Υποβολή αίτησης για αλληλεπίδραση στον επιλεγμένο πράκτορα
- Χειρισμό της αλληλεπίδρασης (στην περίπτωση που η αίτηση γίνει αποδεκτή).

Οι παραπάνω δράσεις υλοποιούνται από ξεχωριστές ανεξάρτητες συμπεριφορές οι οποίες εκτελούνται σε μια αλληλουχία που υποδεικνύεται από το διάγραμμα καταστάσεων που εμφανίζεται στο σχήμα 4-4. Όπως φαίνεται και από το διάγραμμα, η αναζήτηση δεν είναι πάντα επιτυχημένη. Η επιθυμητή αλληλεπίδραση μπορεί να μην πραγματοποιηθεί είτε λόγω έλλειψης κατάλληλων πρακτόρων είτε λόγω απόρριψης της σχετικής αίτησης από τον επιλεγθέντα πράκτορα. Στην περίπτωση που η αίτηση απορριφθεί, ο πράκτορας ενημερώνει τα εσωτερικά του δεδομένα, κυρίως σε ό,τι αφορά την άποψή του για τον πράκτορα που προέβη στην απόρριψη. Η περίπτωση της αγνόησης της αίτησης, αν και μοιάζει παρόμοια με την απόρριψη, είναι κάπως διαφορετική καθώς μπορεί να οφείλεται σε προβλήματα επικοινωνίας και όχι σε αντικοινωνική συμπεριφορά. Γι' αυτό και αντιμετωπίζεται καλοπροαίρετα, χωρίς να επισύρει την ενημέρωση των εσωτερικών δεδομένων όπως συμβαίνει με την απόρριψη.



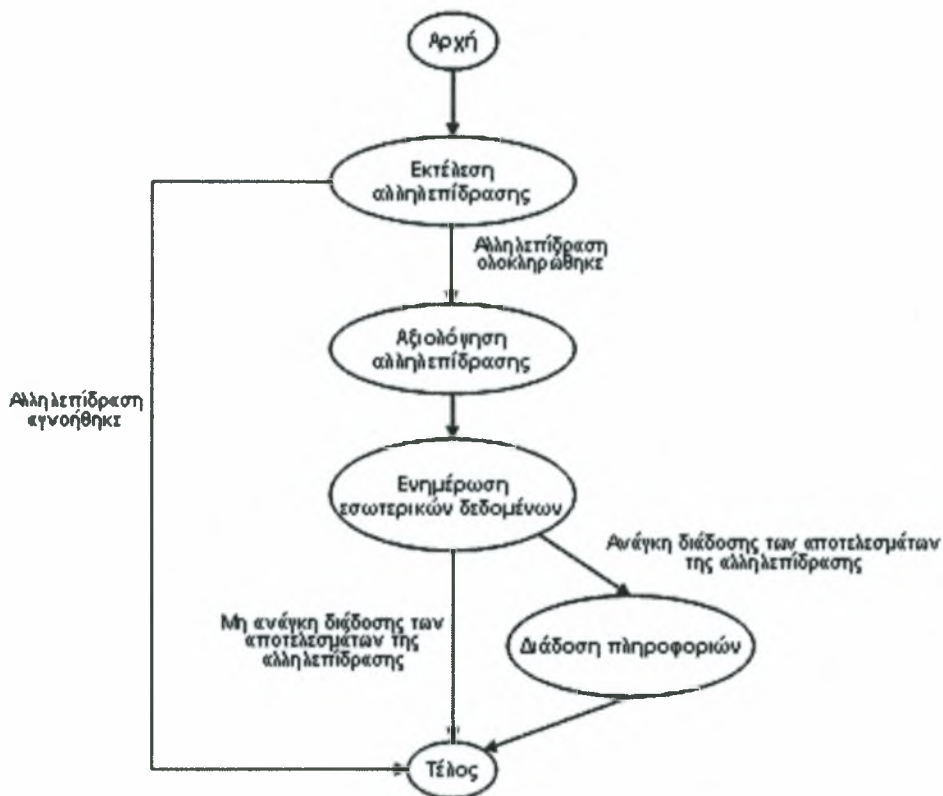
Σχήμα 4-4 : Η συμπεριφορά αναζήτησης αλληλεπίδρασης

## Χειρισμός αλληλεπίδρασης

Ο χειρισμός αλληλεπίδρασης είναι μια σύνθετη συμπεριφορά που αφορά τις ενέργειες που πρέπει να γίνουν προκειμένου να εκτελεστεί μια αλληλεπίδραση. Περιλαμβάνει τις εξής δράσεις :

- Εκτέλεση της αλληλεπίδρασης
- Αξιολόγηση της αλληλεπίδρασης
- Ενημέρωση των εσωτερικών δεδομένων
- Διάδοση πληροφοριών

Το διάγραμμα καταστάσεων της συμπεριφοράς φαίνεται στο σχήμα 4-5.



Σχήμα 4-5 : Η συμπεριφορά χειρισμού αλληλεπίδρασης

Η συμπεριφορά προϋποθέτει ότι έχει ήδη γίνει η συνεννόηση-διαπραγμάτευση των πρακτόρων και έχει συμφωνηθεί η διεξαγωγή της αλληλεπίδρασης. Ως είσοδο η συμπεριφορά χρειάζεται το όνομα του πράκτορα με τον οποίο θα γίνει η αλληλεπίδραση καθώς και τα στοιχεία της περίπτωσης (situation). Με τα στοιχεία αυτά θα τροφοδοτηθεί η πρώτη κατά στη σειρά εκτέλεσης συνιστώσα συμπεριφορά που είναι εκτέλεση της αλληλεπίδρασης. Μετά την εκτέλεση ακολουθεί η αξιολόγηση της αλληλεπίδρασης, το αποτέλεσμα της οποίας δίνεται ως είσοδος στην ενημέρωση των εσωτερικών δομών προκειμένου, μεταξύ άλλων, να επανεκτιμηθούν η εμπιστοσύνη προς τον αλληλεπιδρώντα πράκτορα και ίσως κάποια στοιχεία της προσωπικότητας. Μεταξύ των δράσεων προβλέπεται και διάδοση πληροφοριών η οποία όμως εκτελείται μόνο εφόσον η αλληλεπίδραση

έδωσε χαμηλή αξιολόγηση. Στόχος είναι η ενημέρωση της κοινωνίας για την κακή (και πιθανόν επιβλαβή) συμπεριφορά του αλληλεπιδρώντος πράκτορα. Επειδή το βασικό μοντέλο χαρακτηρίζεται από θετική και καλοπροαίρετη φιλοσοφία, αντιμετωπίζει τις αλληλεπιδράσεις με υψηλή αξιολόγηση ως φυσιολογική περίπτωση και έτσι δεν προβλέπει ενημέρωση της κοινωνίας αντίστοιχη με αυτή που γίνεται για τις επιβλαβείς αλληλεπιδράσεις. Η μέριμνα του μοντέλου εστιάζεται κυρίως στην αποθάρρυνση των ανεπιθύμητων συμπεριφορών και την απομόνωση των κακόβουλων πρακτόρων.

### **Εκτέλεση αλληλεπίδρασης**

Η εκτέλεση μιας αλληλεπίδρασης περιλαμβάνει, στη γενική περίπτωση, την ανταλλαγή μεταξύ των πρακτόρων των συμφωνηθέντων υπηρεσιών ή αγαθών. Η φύση και το πραγματικό περιεχόμενο των ανταλλασσόμενων ειδών δεν έχει ιδιαίτερη σημασία, τουλάχιστο για τις ανάγκες της προσομοίωσης. Το σημαντικό σε μια αλληλεπίδραση (και πάλι από τη σκοπιά της προσομοίωσης) είναι η αξιολόγησή της ως "θετική" ή "αρνητική", "καλή" ή "κακή", "ωφέλιμη" ή "επιβλαβή" κλπ. Για το λόγο αυτό, στο βασικό μοντέλο η συμπεριφορά της εκτέλεσης αλληλεπίδρασης περιλαμβάνει μια και μόνο δράση : την ανταλλαγή μιας τιμής που αντιπροσωπεύει το quality of service της κάθε πλευράς.

Η εκτέλεση αλληλεπίδρασης υλοποιείται ως μια απλή συμπεριφορά που δέχεται ως είσοδο το όνομα του αλληλεπιδρώντος πράκτορα, τα στοιχεία της περίπτωσης που χαρακτηρίζει την συγκεκριμένη αλληλεπίδραση και τον τύπο της διαγωγής (καλή ή κακής) που θα υιοθετηθεί. Ο πράκτορας, ανάλογα με τον πως έχει αποφασίσει να φερθεί, δημιουργεί μια τιμή για το QoS της δικής του υπηρεσίας και την στέλνει στην άλλη πλευρά. Στη συνέχεια περιμένει το QoS της άλλης πλευράς. Η σειρά αποστολής των QoS δεν έχει σημασία. Αυτό που έχει σημασία είναι ο πράκτορας πρώτα να στέλνει το δικό του QoS και στη συνέχεια να αναμένει το QoS της άλλης πλευράς. Κι αυτό για μην υπάρχει περίπτωση να αλλάξει η προαποφασισμένη διαγωγή εξαιτίας της γνώσης της διαγωγής της άλλης πλευράς.

Ένα σημείο που αξίζει να προσεχθεί είναι το πώς αντιμετωπίζεται η φάση της αναμονής. Εδώ προκύπτουν δύο ερωτήματα :

- α) πόσο θα περιμένει ο πράκτορας για να φθάσει το QoS της άλλης πλευράς;
- β) εάν τελικά εκπνεύσει η περίοδος αναμονής και δεν παραληφθεί το QoS πως θα ερμηνευτεί το γεγονός;

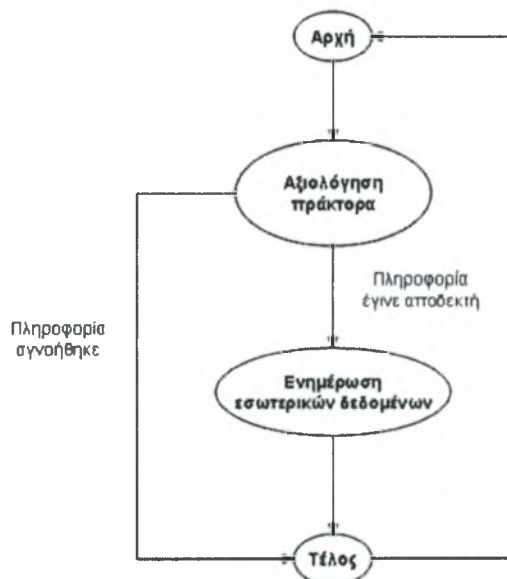
Για το πρώτο ερώτημα δεν υπάρχει ξεκάθαρη απάντηση. Από τη μια, η περίοδος αναμονής δεν μπορεί να είναι απεριόριστη καθώς το QoS ίσως και να μη φτάσει ποτέ, είτε γιατί η άλλη πλευρά συνειδητά δεν το απέστειλε (υποδηλώνοντας με τον τρόπο αυτό τη μη τήρηση των συμφωνημένων υποχρεώσεων), είτε γιατί προέκυψε πρόβλημα επικοινωνίας με αποτέλεσμα το σχετικό μήνυμα να χαθεί. Από την άλλη, ο χρόνος που θα χρειαστεί για να φτάσει το μήνυμα του QoS δεν είναι

συγκεκριμένος ή προβλέψιμος, καθώς εξαρτάται από διάφορους παράγοντες όπως π.χ. ο φόρτος επικοινωνίας. Το ενδεχόμενο της μη έγκαιρης άφιξης ή της απώλειας του μηνύματος με το QoS καθιστά και το δεύτερο ερώτημα δύσκολο να απαντηθεί. Διότι ο πράκτορας δεν μπορεί να ξέρει τι ακριβώς έχει συμβεί (αφερέγγυα συμπεριφορά ή πρόβλημα επικοινωνίας) προκειμένου να χειριστεί κατάλληλα την περίπτωση.

Το βασικό μοντέλο αντιμετωπίζει τα προβλήματα αυτά με ένα σχετικά απλό αλλά ξεκάθαρο τρόπο. Ο χρόνος αναμονής δεν είναι συγκεκριμένος αλλά καθορίζεται από το χρήστη της προσομοίωσης και είναι διαθέσιμος στους πράκτορες ως μια παράμετρος λειτουργίας. Στην περίπτωση που δεν παραληφθεί QoS από την άλλη πλευρά, μέσα στον καθορισμένο χρόνο, ακολουθείται η αρχή ότι "κάθε πράκτορας είναι καλοπροαίρετος μέχρι αποδείξεως του εναντίου" και η αλληλεπίδραση θεωρείται ως μη γενόμενη χωρίς όμως να υπάρχουν κυρώσεις.

### Χειρισμός διαδιδόμενων πληροφοριών

Οι διαδιδόμενες πληροφορίες είναι εισερχόμενες πληροφορίες που φτάνουν στον πράκτορα με πρωτοβουλία του αποστολέα, δηλαδή χωρίς αυτό να το έχει ζητήσει ο πράκτορας-παραλήπτης. Αντιστοιχούν στον κοινωνικό μηχανισμό του κουτσομπολιού και συνήθως μεταφέρουν στοιχεία σχετικά με την αξιοπιστία ή την πρόσφατη συμπεριφορά κάποιου άλλου πράκτορα. Τα στοιχεία αυτά μπορεί είτε να αποτελούν προσωπικές εκτιμήσεις του αποστολέα, είτε να προέρχονται από μια τρίτη πηγή και απλά να προωθούνται από τον αποστολέα. Το διάγραμμα καταστάσεων της φαίνεται στο σχήμα 4-6.



Σχήμα 4-6 : Η συμπεριφορά χειρισμού διαδιδόμενων πληροφοριών

Η διαδικασία χειρισμού των διαδιδόμενων πληροφοριών υλοποιείται ως μια σύνθετη συμπεριφορά που περιλαμβάνει δύο συνιστώσες. Αρχικά αξιολογείται ο αποστολέας για να διαπιστωθεί αφενός

το επίπεδο αξιοπιστίας του και αφετέρου η σχετικότητα του με το αντικείμενο των πληροφοριών που μεταδίδει. Εάν η αξιολόγηση δώσει ικανοποιητικό αποτέλεσμα, οι πληροφορίες χρησιμοποιούνται για την ενημέρωση των εσωτερικών δομών (καταγραφή των παρεχόμενων στοιχείων και επανεκτίμηση των δεδομένων εμπιστοσύνης).

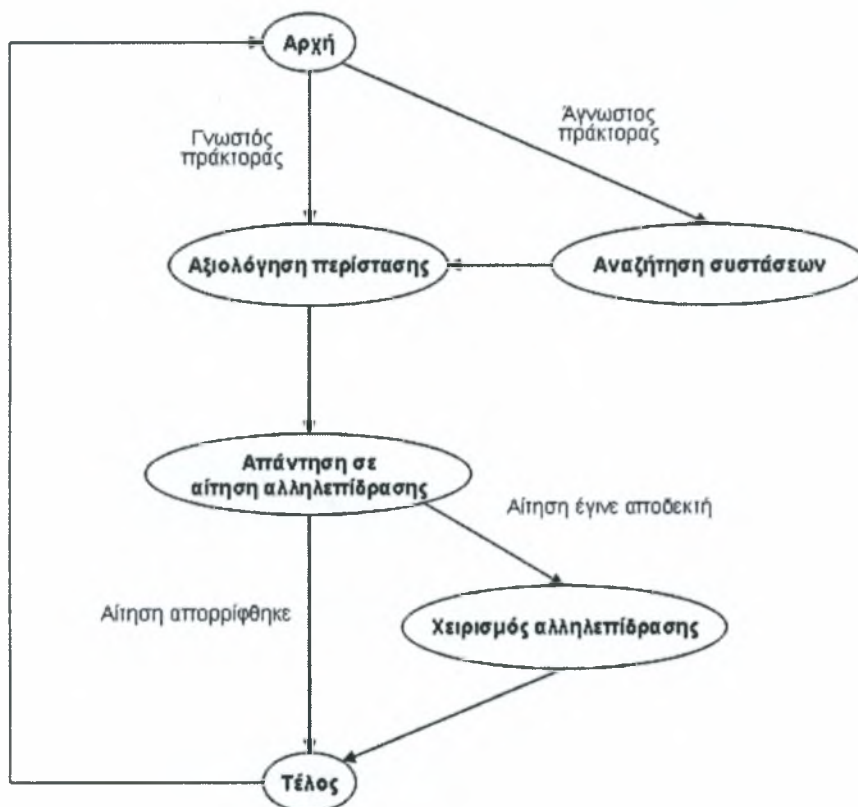
Η συμπεριφορά χειρισμού διαδιδόμενων πληροφοριών χρειάζεται ως είσοδο το όνομα του αποστολέα πράκτορα και την εγγραφή με τις διαδιδόμενες πληροφορίες.

### Χειρισμός αίτησης για αλληλεπίδραση

Ο χειρισμός αίτησης για αλληλεπίδραση περιλαμβάνει τις ενέργειες που απαιτούνται για τον χειρισμό των αιτήσεων που καταφθάνουν στον πράκτορα με ζητούμενο την εκτέλεση αλληλεπίδρασης. Πρόκειται για σύνθετη συμπεριφορά, συνιστώσα της βασικής συμπεριφοράς, που εκτελείται παράλληλα με τις υπόλοιπες συνιστώσες. Περιλαμβάνει τέσσερις βασικές δράσεις :

- Παρακολούθηση της εισόδου για άφιξη αίτησης αλληλεπίδρασης
- Αξιολόγηση του αιτούντα πράκτορα
- Λήψη απόφασης αποδοχής ή απόρριψης και αποστολή ανάλογης απάντησης
- Χειρισμό της αλληλεπίδρασης

Το διάγραμμα καταστάσεων της συμπεριφοράς φαίνεται στο σχήμα 4-7.

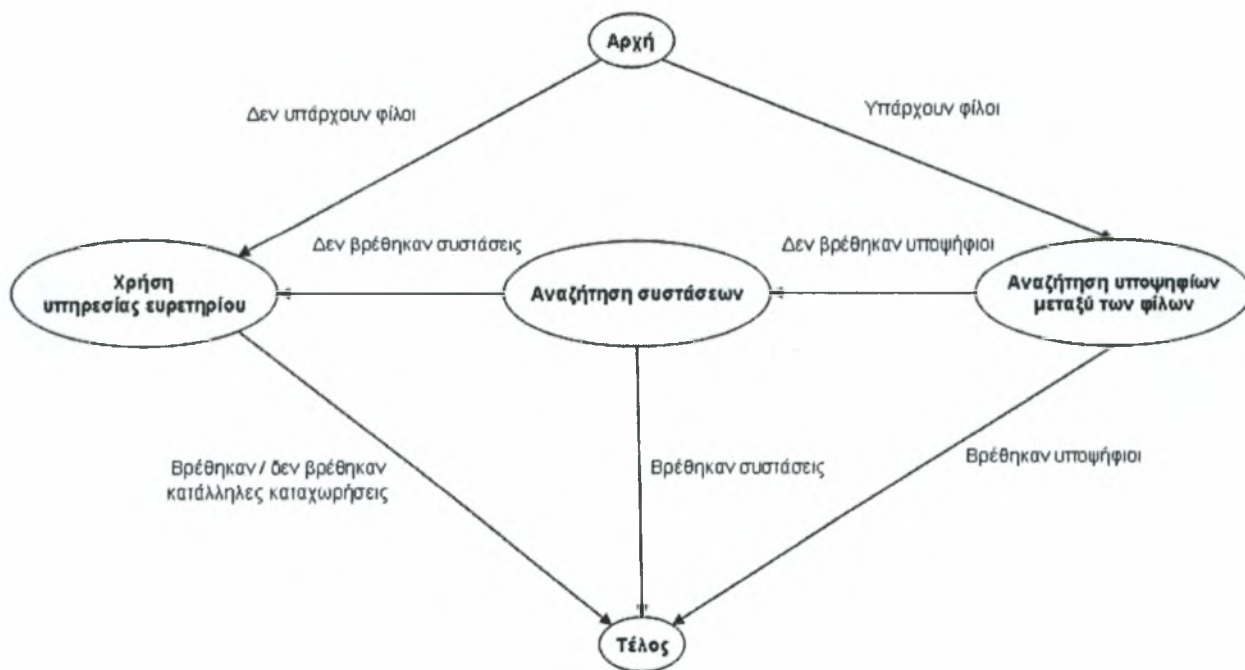


Σχήμα 4-7 : Η συμπεριφορά χειρισμού αίτησης για αλληλεπίδραση

Στην αρχική κατάσταση η συμπεριφορά ελέγχει διαρκώς την ουρά των μηνυμάτων για την άφιξη μηνυμάτων με περιεχόμενο αίτηση αλληλεπίδρασης. Μόλις εμφανιστεί ένα τέτοιο μήνυμα ελέγχεται εάν ο αποστολέας είναι γνωστός, δηλ. εάν υπάρχουν στοιχεία γι' αυτόν από προηγούμενες αλληλεπιδράσεις ή από εξωτερική πληροφόρηση (φήμη). Στην περίπτωση που ο αιτών πράκτορας είναι άγνωστος αναζητούνται συστάσεις προκειμένου να υπάρχουν κάποια στοιχεία για να γίνει αξιολόγηση του πράκτορα. Στη συνέχεια ακολουθεί η αξιολόγηση του αιτούντα πράκτορα με χρήση όσων στοιχείων υπάρχουν. Εάν η αναζήτηση συστάσεων δεν είχε αποτέλεσμα, αναγκαστικά η αξιολόγηση θα βασιστεί στην προδιάθεση του πράκτορα που λαμβάνει την αίτηση για αλληλεπίδραση. Το αποτέλεσμα της αξιολόγησης χρησιμοποιείται για την λήψη απόφασης σχετικά με την αποδοχή ή την απόρριψη της αίτησης. Σε κάθε περίπτωση, την απόφαση συνοδεύει η αποστολή ανάλογου ενημερωτικού μηνύματος προς τον αιτούντα πράκτορα. Τέλος, εάν επιλεγεί η αποδοχή της αίτησης, καλείται η συμπεριφορά χειρισμού της αλληλεπίδρασης με αντικείμενο την εκτέλεση της αιτούμενης αλληλεπίδρασης.

### Αναζήτηση υποψηφίων για αλληλεπίδραση

Η αναζήτηση υποψηφίων για αλληλεπίδραση είναι μια σύνθετη συμπεριφορά που αποτελεί συνιστώσα συμπεριφορά της αναζήτησης αλληλεπίδρασης. Ως είσοδο χρειάζεται μόνο τα στοιχεία της περίπτωσης, προκειμένου να εντοπίσει τους πράκτορες που είναι κατάλληλοι για την επιζητούμενη αλληλεπίδραση. Κατάλληλοι θεωρούνται οι πράκτορες που είναι ειδικοί στο αντικείμενο της περίπτωσης. Το διάγραμμα καταστάσεων της συμπεριφοράς αναζήτησης υποψηφίων για αλληλεπίδραση φαίνεται στο σχήμα 4-8.



Σχήμα 4-8 : Η συμπεριφορά αναζήτησης υποψηφίων για αλληλεπίδραση

Αρχικά, ελέγχεται εάν ο πράκτορας έχει φίλους κι αν έχει, ελέγχεται ποιοι απ' αυτούς είναι κατάλληλοι. Εάν κανείς φίλος δεν είναι κατάλληλος, γίνεται προσπάθεια να αναζητηθούν συστάσεις. Αν και η αναζήτηση συστάσεων αποβεί άκαρπη, τότε γίνεται κλήση της υπηρεσίας ευρετηρίου για τον εντοπισμό των πρακτόρων που έχουν καταχωρηθεί ως ειδικοί του ζητούμενου αντικειμένου. Η υπηρεσία ευρετηρίου θα χρησιμοποιηθεί και στην περίπτωση που ο πράκτορας δεν έχει φίλους, οπότε και δεν μπορούν να αναζητηθούν ούτε συστάσεις<sup>25</sup>. Η συμπεριφορά ολοκληρώνεται είτε έχοντας εντοπίσει κατάλληλους πράκτορες που είναι και υποψήφιοι για την επιζητούμενη αλληλεπίδραση, είτε χωρίς να έχει εντοπιστεί κάποιος κατάλληλος. Στην τελευταία περίπτωση, η περιβάλλουσα αναζήτηση αλληλεπίδρασης θα αποτύχει.

### **Επιλογή πράκτορα για αλληλεπίδραση**

Είναι μια σύνθετη συμπεριφορά που δέχεται ως είσοδο μια λίστα υποψηφίων για αλληλεπίδραση πρακτόρων και τα στοιχεία της περίπτωσης της αλληλεπίδρασης. Από τη λίστα των υποψηφίων επιλέγεται και επιστρέφεται ο καταλληλότερος πράκτορας ως προτεινόμενος για να του ζητηθεί η διενέργεια αλληλεπίδρασης. Η συμπεριφορά περιλαμβάνει δύο δράσεις :

- Την ατομική αξιολόγηση του κάθε υποψήφιου
- Την κατάταξη των υποψηφίων βάσει της ατομικής τους αξιολόγησης

Η συμπεριφορά ενεργοποιεί τη συμπεριφορά ελέγχου και κατάταξης των υποψηφίων η οποία διαπερνά τη λίστα των υποψηφίων και για κάθε υποψήφιο που βρίσκει καλεί τη συμπεριφορά αξιολόγησης. Η αξιολόγηση, λαμβάνοντας υπόψη τα χαρακτηριστικά της επικείμενης αλληλεπίδρασης και τα συσσωρευμένα στοιχεία του πράκτορα (προηγούμενες αλληλεπιδράσεις, ειδικότητα, φήμη κλπ), επιστρέφει μια τιμή, η οποία στη συνέχεια χρησιμοποιείται για την κατάξη του υποψηφίου. Ο πρώτος της κατάταξης θα δοθεί τελικά στην έξοδο ως ο επιλεγμένος και προτεινόμενος πράκτορας για την αλληλεπίδραση. Εδώ θα πρέπει να τονιστεί ότι η αξιολόγηση μπορεί να δώσει και αρνητική τιμή υποδηλώνοντας ότι ο συγκεκριμένος πράκτορας είναι ακατάλληλος. Στην περίπτωση που όλοι οι υποψήφιοι δώσουν αρνητική αξιολόγηση, η συνιστώσα συμπεριφορά της κατάταξης των υποψηφίων μπορεί και μην είναι σε θέση να επιλέξει κάποιο πράκτορα. Γι ' αυτό και είναι σημαντικό οι συμπεριφορές που θα καλέσουν τη συγκεκριμένη συμπεριφορά να μην θεωρήσουν ως δεδομένο ότι θα επιλεγεί οπωσδήποτε κάποιος πράκτορας για αλληλεπίδραση. Το διάγραμμα καταστάσεων της συμπεριφοράς φαίνεται στο σχήμα 4-9.

<sup>25</sup> Δείτε περισσότερα στην συμπεριφορά αναζήτησης συστάσεων



Σχήμα 4-9 : Η συμπεριφορά επιλογής πράκτορα για αλληλεπίδραση

### Αναζήτηση συστάσεων

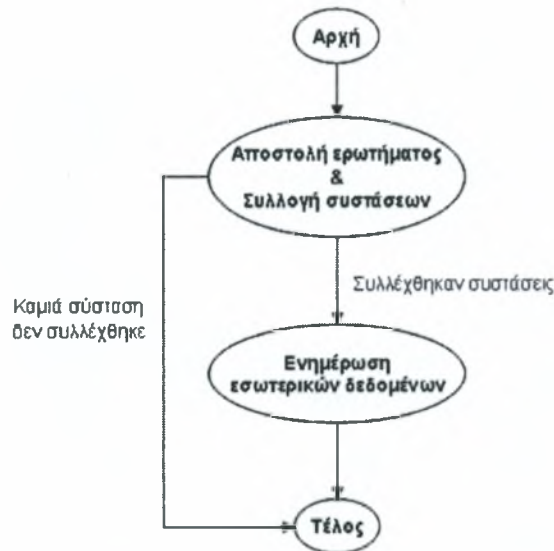
Αρκετά συχνά ένας πράκτορας χρειάζεται να ζητήσει τη βοήθεια άλλων πρακτόρων προκειμένου να πάρει μια απόφαση ή να κάνει κάποια επιλογή. Αυτό συμβαίνει όταν πρέπει να εκτιμήσει την αξιοπιστία ή τη φερεγγυότητα κάποιου πράκτορα για τον οποίο δεν έχει καθόλου πληροφορίες. Στις περιπτώσεις αυτές ζητά από άλλους πράκτορες τις πληροφορίες που χρειάζεται. Οι πληροφορίες αυτές παρέχονται υπό τη μορφή συστάσεων και μπορεί να αφορούν την προσωπική γνώμη ή την εμπειρία του ερωτώμενου πράκτορα σχετικά με τον υπό αξιολόγηση πράκτορα.

Η συμπεριφορά της αναζήτησης συστάσεων εκτελεί το έργο της εύρεσης των απαιτούμενων πληροφοριών μέσα από τη επίκληση των κατάλληλων πρακτόρων. Υλοποιείται ως σύνθετη συμπεριφορά με δύο συνιστώσες : α) την αποστολή του σχετικού ερωτήματος και τη συλλογή των απαντήσεων-συστάσεων και β) την καταγραφή των πληροφοριών που συλλέχθηκαν στις εσωτερικές δομές του πράκτορα. Η καταγραφή των πληροφοριών στις εσωτερικές δομές εξασφαλίζει ότι αυτές θα είναι διαθέσιμες σε όλες τις συμπεριφορές του πράκτορα από το σημείο εκείνο και έπειτα.

Στο βασικό μοντέλο η συλλογή συστάσεων χρησιμοποιείται κυρίως σε δύο περιπτώσεις. Όταν αξιολογείται ένας πράκτορας (είτε γενικά είτε αναφορικά με την ικανότητά του σε συγκεκριμένο αντικείμενο) και όταν αναζητούνται οι καταλληλότεροι πράκτορες πάνω σε συγκεκριμένο αντικείμενο. Η διάκριση των περιπτώσεων γίνεται με βάση την παρεχόμενη είσοδο που μπορεί να είναι το όνομα ενός πράκτορα, τα στοιχεία μιας περίπτωσης ή και τα δύο μαζί. Εάν στην είσοδο δίνεται το όνομα ενός πράκτορα, αναζητούνται συστάσεις για το συγκεκριμένο πράκτορα και μόνο. Εάν δεν δίνεται κάποιο όνομα τότε αναζητούνται συστάσεις για το ποιοι είναι οι καταλληλότεροι πράκτορες σε κάποιο αντικείμενο. Εάν επιπλέον στην είσοδο υπάρχουν και τα στοιχεία κάποιας



περίστασης, αυτό υποδηλώνει ότι οι ζητούμενες συστάσεις θα αφορούν συγκεκριμένο αντικείμενο. Το διάγραμμα καταστάσεων της συμπεριφοράς φαίνεται στο σχήμα 4-10.



Σχήμα 4-10 : Η συμπεριφορά αναζήτησης συστάσεων

### Ενημέρωση εσωτερικών δεδομένων

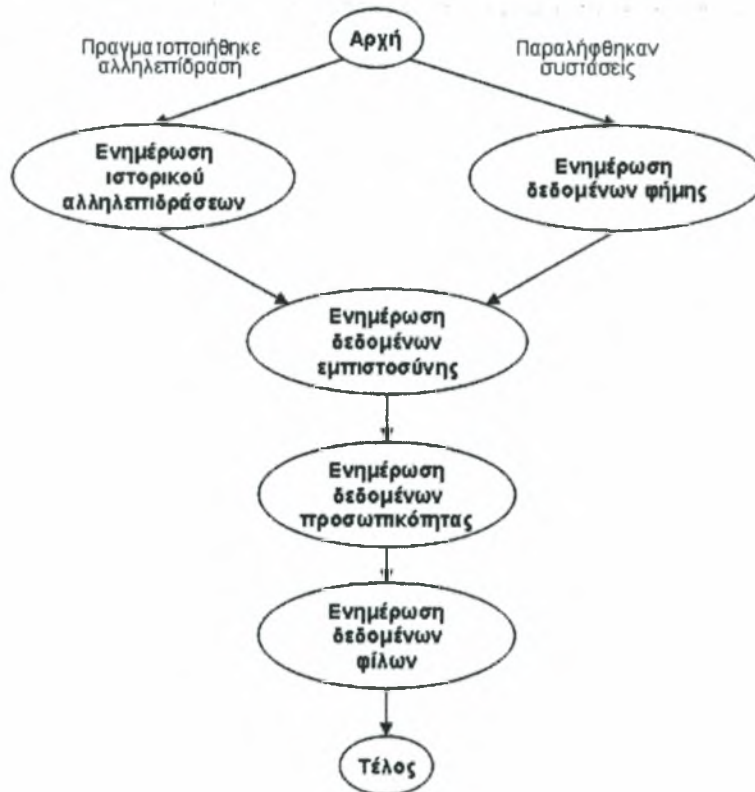
Η ενημέρωση των εσωτερικών δομών είναι μια πολύ βασική συμπεριφορά που καλείται κάθε φορά που προκύπτουν νέα δεδομένα τα οποία επηρεάζουν την λειτουργία των υπολοίπων συμπεριφορών. Πρόκειται για σύνθετη συμπεριφορά που περιλαμβάνει μια ξεχωριστή συνιστώσα ενημέρωσης για κάθε μια από τις κύριες δομές δεδομένων που διατηρεί ο πράκτορας και οι οποίες είναι :

- το ιστορικό των αλληλεπιδράσεων
- ο πίνακας φήμης
- ο πίνακας εμπιστοσύνης
- τα χαρακτηριστικά προσωπικότητας του πράκτορα
- η λίστα των φίλων

Η συμπεριφορά ενεργοποιείται από δύο είδη συμβάντων :

- α) την διενέργεια αλληλεπίδρασης
- β) την παραλαβή συστάσεων

Και στις δύο περιπτώσεις χρειάζεται να ενημερωθεί η πλειονότητα των εσωτερικών δομών. Η μόνη διαφοροποίηση αφορά τις δομές του ιστορικού των αλληλεπιδράσεων και του πίνακα της φήμης. Εάν η συμπεριφορά έχει ενεργοποιηθεί μετά από αλληλεπίδραση ενημερώνεται το ιστορικό, ενώ εάν έχει προηγηθεί άφιξη συστάσεων ενημερώνεται ο πίνακας φήμης. Το διάγραμμα καταστάσεων της συμπεριφοράς φαίνεται στο σχήμα 4-11.



Σχήμα 4-11 : Η συμπεριφορά ενημέρωσης εσωτερικών δεδομένων

Η ύπαρξη των συνιστωσών δεν υπαγορεύει υποχρεωτικά και την μεταβολή των περιεχομένων των δομών. Απλά υποδεικνύει ότι πρέπει να γίνεται έλεγχος για το εάν χρειάζεται να επανεκτιμηθούν κάποια δεδομένα. Για παράδειγμα εάν πραγματοποιηθεί θετική αλληλεπίδραση με ένα ήδη έμπιστο πράκτορα, τα στοιχεία της αλληλεπίδρασης θα καταγραφούν στη δομή του ιστορικού αλλά δεν θα χρειαστεί να αλλάξει κάτι στον πίνακα εμπιστοσύνης, την προσωπικότητα ή τη λίστα των φίλων.

Η είσοδος της συμπεριφοράς μπορεί να είναι είτε μια εγγραφή με τα στοιχεία μιας αλληλεπίδρασης είτε μια λίστα συστάσεων.

### Χειρισμός αίτησης για παροχή πληροφοριών

Είναι μια απλή συμπεριφορά που ενεργοποιείται κάθε φορά που καταφθάνει ένα μήνυμα με περιεχόμενο μια αίτηση για παροχή πληροφοριών. Στο βασικό μοντέλο οι ζητούμενες πληροφορίες είναι πάντα συστάσεις (βλ. συμπεριφορά Αναζήτηση συστάσεων σχήμα 4-10), ωστόσο επειδή σε κάποια μοντέλα υπάρχει η ανάγκη αναζήτησης και άλλων στοιχείων (Αναφορ. μοντέλο PeerTrust) η συμπεριφορά έχει υλοποιηθεί κατά τρόπο που να μπορεί να χειρίζεται δύο διαφορετικά είδη αιτήσεων: α) για παροχή συστάσεων και β) για παροχή εξειδικευμένων πληροφοριών. Για καθένα από τα δύο είδη η συμπεριφορά διαθέτει μια μέθοδο χειριστή (handler). Για τις μεν συστάσεις ο σχετικός χειριστής (handler) έχει υλοποιηθεί και η λειτουργία του περιγράφεται παρακάτω, για τις εξειδικευμένες πληροφορίες όμως ο σχετικός χειριστής είναι κενός (εμφανίζει απλά ένα warning)

και χρειάζεται να δοθεί ως επικαλυπτόμενη μέθοδος από το εκάστοτε μοντέλο που χρησιμοποιεί αυτή τη δυνατότητα.

Η μέθοδος-χειριστής της παροχής συστάσεων αρχικά προσπαθεί να αναγνωρίσει το είδος των ζητούμενων συστάσεων. Αυτό το πετυχαίνει ελέγχοντας τα στοιχεία που περιέχει η αίτηση που δίνεται ως είσοδος στη συμπεριφορά. Τα στοιχεία αυτά είναι είτε το όνομα κάποιου πράκτορα είτε ένα αντικείμενο είτε και τα δύο. Η απουσία συγκεκριμένου ονόματος πράκτορα υποδηλώνει αίτηση για συστάσεις σχετικά με τους καταλληλότερους πράκτορες σε κάποιο αντικείμενο. Η απουσία συγκεκριμένου αντικειμένου υποδηλώνει ότι ζητείται μια γενική σύσταση σχετικά με κάποιο πράκτορα. Στη συνέχεια και ανάλογα με το είδος της ζητούμενης σύστασης η συμπεριφορά ελέγχει τις κατάλληλες εσωτερικές δομές (πίνακας εμπιστοσύνης, πίνακας φήμης, λίστα φίλων κλπ) προκειμένου να βρει τα δεδομένα που απαιτούνται για την δημιουργία της σύστασης. Η αναζήτηση των απαραίτητων δεδομένων ακολουθεί τις εξής γενικές αρχές :

- Οι καταλληλότεροι πράκτορες αναζητούνται μόνο μεταξύ των φίλων. Αυτό γίνεται διότι η λίστα των φίλων, εξ' ορισμού, περιλαμβάνει τους πιο αξιόπιστους γνωστούς πράκτορες. Εάν δεν βρεθεί κατάλληλος πράκτορας εκεί, τότε δεν υπάρχει κανείς για να προταθεί.
- Τα δεδομένα που λαμβάνονται υπόψη και χρησιμοποιούνται κατά σειρά σπουδαιότητας είναι : η τιμή της εμπιστοσύνης στο ζητούμενο αντικείμενο, η γενική τιμή της εμπιστοσύνης, η φήμη στο ζητούμενο αντικείμενο, η γενική φήμη

Η έξοδος της συμπεριφοράς είναι το αντικείμενο της σύστασης που δημιουργήθηκε από τα στοιχεία που συλλέχθηκαν.

### **Αξιολόγηση περίπτωσης**

Η αξιολόγηση περίπτωσης είναι μια από τις πλέον απαραίτητες λειτουργίες ενός πράκτορα εμπιστοσύνης. Κατά τη διάρκεια της δραστηριοποίησής του χρειάζεται συνεχώς να παίρνει αποφάσεις που σχετίζονται με τον τρόπο που χειρίζεται τις αλληλεπιδράσεις του (αποδοχή ή απόρριψη προτάσεων, επιλογή του καταλληλότερου υποψήφιου κλπ). Οι αποφάσεις αυτές κατά κύριο λόγο στηρίζονται σε κάποια διαδικασία αξιολόγησης των συνθηκών κάτω από τις οποίες διενεργείται η εκάστοτε αλληλεπίδραση. Αυτό είναι ο ρόλος της συγκεκριμένης συμπεριφοράς.

Η Αξιολόγηση περίπτωσης χρησιμοποιεί δεδομένα από τις εσωτερικές δομές του πράκτορα και σε συνδυασμό με τα στοιχεία της περίπτωσης αλληλεπίδρασης προσπαθεί να δώσει μια ένδειξη για το κατά πόσο πρέπει να εκτελεστεί η αλληλεπίδραση. Είναι λοιπόν σημαντικό πριν την ενεργοποίηση αυτής της συμπεριφοράς να προηγηθεί η εκτέλεση της συμπεριφοράς ενημέρωσης των εσωτερικών δομών, ώστε εάν υπάρχουν νέα στοιχεία να χρησιμοποιηθούν κι αυτά.

Η συμπεριφορά διαβάζει από την είσοδο το όνομα του υποψήφιου πράκτορα και τα στοιχεία της περίπτωσης και εκτελεί τρεις υπολογισμούς :

α) Πρώτα υπολογίζεται το λεγόμενο κατώφλι περίστασης (situation threshold) από τον τύπο

$$\text{Κατώφλι περίστασης} = \text{Κόστος} * (1 - \text{Όφελος})$$

Το αποτέλεσμα του υπολογισμού είναι μια πραγματική τιμή στην περιοχή  $[0,1]$  η οποία αντιπροσωπεύει ένα συνδυασμό της επικινδυνότητας και της σημασίας εκτέλεσης της σχετικής αλληλεπίδρασης. Όσο μεγαλύτερη είναι αυτή η τιμή τόσο πιο επικίνδυνη και λιγότερο σημαντική είναι η αλληλεπίδραση και το αντίθετο.

β) Στη συνέχεια υπολογίζεται η αξιοπιστία του υποψήφιου πράκτορα. Για το σκοπό αυτό ενεργοποιείται η συμπεριφορά *Αξιολόγηση πράκτορα* της οποίας η έξοδος χρησιμοποιείται

γ) Στο τέλος υπολογίζεται η διαφορά

$$\text{Κατώφλι περίστασης} - \text{Αξιοπιστία πράκτορα}$$

Η τιμή της διαφοράς υποδεικνύει το κατά πόσο πρέπει να προχωρήσει η αξιολογούμενη αλληλεπίδραση. Θετική τιμή σημαίνει ότι η αλληλεπίδραση μπορεί να προχωρήσει και μάλιστα όσο μεγαλύτερη είναι αυτή η τιμή τόσο πιο ασφαλής ή σημαντική είναι για τον πράκτορα η αλληλεπίδραση. Αρνητική τιμή σημαίνει ότι η αλληλεπίδραση είναι επικίνδυνη και πρέπει να αποφευχθεί.

Η τιμή της διαφοράς που προκύπτει από τον τελευταίο υπολογισμό αποτελεί και την έξοδο της συμπεριφοράς.

### **Αξιολόγηση πράκτορα**

Η Αξιολόγηση πράκτορα είναι μια απλή συμπεριφορά που χρησιμοποιείται κάθε φορά που χρειάζεται μια ένδειξη της αξιοπιστίας ενός πράκτορα. Αυτό συμβαίνει όταν αξιολογείται η περίσταση μιας αλληλεπίδρασης (κατά την παραλαβή σχετικής αίτησης ή κατά την αναζήτηση αλληλεπίδρασης) ή όταν παραλαμβάνονται διαδιδόμενες πληροφορίες.

Η συμπεριφορά δέχεται στην είσοδο το όνομα του αξιολογούμενου πράκτορα και προαιρετικά τα στοιχεία περίστασης αλληλεπίδρασης (εφόσον η αξιολόγηση αφορά κάποια αλληλεπίδραση). Αντλώντας στοιχεία από τις εσωτερικές δομές υπολογίζει και επιστρέφει μια τιμή που υποδεικνύει την αξιοπιστία του αξιολογούμενου πράκτορα.

Στην τρέχουσα έκδοση του βασικού μοντέλου η αξιοπιστία ενός πράκτορα αντιπροσωπεύεται από την τρέχουσα τιμή της εμπιστοσύνης ή της φήμης του. Συγκεκριμένα χρησιμοποιείται το πρώτο κατά σειρά διαθέσιμο στοιχείο από τα παρακάτω :

- η τιμή εμπιστοσύνης πάνω στο αντικείμενο της περίστασης (εφόσον δίνεται η περίσταση)
- η γενική τιμή εμπιστοσύνης
- η φήμη πάνω στο αντικείμενο της περίστασης (εφόσον δίνεται η περίσταση)
- η γενική φήμη

Εάν ο αξιολογούμενος πράκτορας είναι άγνωστος κανένα από τα παραπάνω στοιχεία δεν θα είναι διαθέσιμο. Στην περίπτωση αυτή λαμβάνεται υπόψη η προδιάθεση εμπιστοσύνης (trusting disposition) και χρησιμοποιείται μια διαφορετική τιμή για κάθε τύπο διάθεσης (Αισιόδοξος = 1, Απαισιόδοξος = 0, Ρεαλιστής = 0.5).

### **Έλεγχος και κατάταξη υποψηφίων**

Πρόκειται για εξειδικευμένη συμπεριφορά που η ύπαρξή της υπαγορεύεται από το συγκεκριμένο σχεδιασμό του βασικού μοντέλου. Ενεργοποιείται μέσα στη συμπεριφορά της επιλογής πράκτορα για αλληλεπίδραση η οποία με τη σειρά της αποτελεί συνιστώσα της συμπεριφοράς αναζήτησης αλληλεπίδρασης. Ο ρόλος της είναι καθαρά τεχνικός και περιλαμβάνει δύο δράσεις :

- i. τη σάρωση της λίστας των υποψηφίων για αλληλεπίδραση πρακτόρων και τη διαδοχική ενεργοποίηση της συμπεριφοράς αξιολόγησης περίπτωσης για καθένα απ' αυτούς
- ii. την διάταξη της λίστας σύμφωνα με τις παραγόμενες αξιολογήσεις.

Η συμπεριφορά δέχεται ως είσοδο τη λίστα των υποψηφίων πρακτόρων καθώς και τα στοιχεία περίπτωσης της αξιολογούμενης αλληλεπίδρασης και επιστρέφει τη λίστα των υποψηφίων διατεταγμένη.

### **Αποστολή ερωτήματος & Συλλογή συστάσεων**

Η συμπεριφορά αυτή ενεργοποιείται μέσα από την Αναζήτηση συστάσεων και περιλαμβάνει δύο φάσεις

- α) τη δημιουργία και αποστολή κατάλληλου μηνύματος με τα στοιχεία της ζητούμενης σύστασης
- β) τη συλλογή των παρεχόμενων συστάσεων

Ως παραλήπτες του ερωτήματος ορίζονται οι φίλοι του πράκτορα. Εάν ο πράκτορας δεν διαθέτει φίλους, τότε παραλήπτες ορίζονται όλοι οι γνωστοί πράκτορες που θεωρούνται αξιόπιστοι, δηλ. η αξιολόγηση της γενικής εμπιστοσύνης είναι μεγαλύτερη ή ίση από το τρέχον κατώφλι αξιοπιστίας.

Η συστάσεις που συλλέγονται οργανώνονται σε μια λίστα που δίνεται στην έξοδο της συμπεριφοράς.

### **Αίτηση για αλληλεπίδραση**

Είναι μια απλή συμπεριφορά που περιλαμβάνει την δημιουργία και αποστολή κατάλληλου μηνύματος και την αναμονή της σχετικής απάντησης. Στο περιεχόμενο του μηνύματος τοποθετούνται τα στοιχεία της περίπτωσης της αιτούμενης αλληλεπίδρασης, προκειμένου ο άλλος πράκτορας να αποφανθεί εάν δέχεται ή απορρίπτει την αλληλεπίδραση. Ενεργοποιείται κατά την Αναζήτηση αλληλεπίδρασης . Χρειάζεται στην είσοδο το όνομα του υποψήφιου για αλληλεπίδραση

πράκτορα και τα στοιχεία της περίπτωσης αλληλεπίδρασης. Στην έξοδο παρέχει μια ένδειξη για την αποδοχή ή την απόρριψη της αίτησης.

### **Απάντηση σε αίτηση για αλληλεπίδραση**

Η συμπεριφορά αυτή ενεργοποιείται κατά το χειρισμό μιας αίτησης για αλληλεπίδραση. Δέχεται ως είσοδο το αποτέλεσμα της αξιολόγησης της περίπτωσης της αλληλεπίδρασης και επιλέγει εάν η αίτηση θα γίνει αποδεκτή ή εάν θα απορριφθεί. Το κριτήριο είναι το πρόσημο της αξιολόγησης της περίπτωσης. Θετική τιμή σημαίνει αποδοχή, αρνητική σημαίνει απόρριψη. Η συμπεριφορά ολοκληρώνεται με την αποστολή σχετικού μηνύματος απόρριψης ή αποδοχής προς τον αιτούντα πράκτορα.

### **Αξιολόγηση αλληλεπίδρασης**

Η Αξιολόγηση αλληλεπίδρασης είναι μια συμπεριφορά που καλείται μετά την ολοκλήρωση μιας αλληλεπίδρασης. Το αποτέλεσμά της θα χρησιμοποιηθεί για την ενημέρωση των εσωτερικών δομών του πράκτορα. Στην τρέχουσα έκδοση του βασικού μοντέλου η αξιολόγηση αλληλεπίδρασης περιλαμβάνει απλά την χρήση του QoS που στέλνει η άλλη πλευρά ως τεκμήριο της συμμετοχής της στην αλληλεπίδραση. Θετικές τιμές του QoS υποδεικνύουν θετική αλληλεπίδραση με τις υψηλές τιμές να είναι οι πιο επιθυμητές. Το αντίθετο ισχύει για τις αρνητικές τιμές του QoS.

Αυτή η υλοποίηση, αν και εξαιρετικά απλοϊκή, είναι επαρκής για τις ανάγκες της προσομοίωσης. Είναι λογικό όμως το κάθε μοντέλο εμπιστοσύνης να ακολουθεί τη δική του εξειδικευμένη προσέγγιση που θα ταιριάζει καλύτερα στις αρχιτεκτονικές του επιλογές.

### **Διάδοση πληροφοριών**

Η διάδοση πληροφοριών όπως έχουμε πει αντιστοιχεί στον κοινωνικό μηχανισμό του κουτσομπολιού. Ενεργοποιείται όταν ο πράκτορας κρίνει ότι πρέπει να ενημερώσει τα άλλα μέλη της κοινωνίας για κάποιο θέμα. Στο βασικό μοντέλο η ανάγκη αυτή προκύπτει μόνο όταν λάβει χώρα μια ανεπιθύμητη αλληλεπίδραση, δηλ. μια αλληλεπίδραση με αρνητικό QoS. Σε μια τέτοια περίπτωση, ο πράκτορας (που δέχεται το αρνητικό QoS) ελέγχει εάν ο δράστης (δηλ. ο πράκτορας που έστειλε το αρνητικό QoS) είναι πλέον αναξιόπιστος (μετά και την ενημέρωση της αξιοπιστίας βάσει της τελευταίας του συμπεριφοράς). Εάν βρεθεί αναξιόπιστος, δημιουργεί μια σύσταση με περιεχόμενο το όνομα του δράστη, το αντικείμενο της αρνητικής αλληλεπίδρασης και την προφανώς χαμηλή αξιολόγηση εμπιστοσύνης του. Τη σύσταση αυτή την αποστέλλει σε όλους τους φίλους του προκειμένου να τους προειδοποιήσει για την επιβλαβή δραστηριότητα του δράστη. Στην περίπτωση που το θύμα δεν έχει φίλους, είτε γιατί είναι νέο μέλος της κοινωνίας είτε γιατί ο κοινωνικός μηχανισμός της διαχείρισης φίλων είναι απενεργοποιημένος, η σύσταση αποστέλλεται σε όλους τους πράκτορες της κοινωνίας, πλην, φυσικά, του δράστη.

## Χρήση υπηρεσίας ευρετηρίου

Η υπηρεσία ευρετηρίου, γνωστή και ως υπηρεσία «χρυσού οδηγού» (yellow pages) είναι μια αρκετά συνηθισμένη υπηρεσία που υπάρχει σε όλες σχεδόν τις κοινωνίες. Χρησιμοποιείται για την εύρεση πληροφοριών κυρίως από τα νέα μέλη της κοινωνίας όταν ακόμη δεν έχουν συσσωρεύσει επαρκή εμπειρία. Στην τρέχουσα έκδοση του βασικού μοντέλου, οι πληροφορίες που αναζητούνται, αφορούν το ποιοι πράκτορες έχουν μια συγκεκριμένη ειδικότητα και κατά συνέπεια μπορούν να παρέχουν τη σχετική υπηρεσία.

Η Χρήση υπηρεσίας ευρετηρίου χρησιμοποιεί την υπηρεσία DF (Directory Facility) του JADE που είναι συμβατή με το πρότυπο FIPA. Για να είναι εφικτό αυτό, κάθε πράκτορας μετά την εκκίνησή του κάνει registration στην υπηρεσία DF δηλώνοντας την ειδικότητά του. Δεν δίνονται επιπλέον πληροφορίες προκειμένου να αξιοποιηθούν και άλλες δυνατότητες της υπηρεσίας DF (π.χ. οντολογίες) διότι αυτό θα αύξανε την πολυπλοκότητα της υλοποίησης χωρίς να υπάρχει κάποιο όφελος για την προσομοίωση.

Σε ότι αφορά τη λειτουργία της, η συμπεριφορά καλείται μόνο μέσα από την Αναζήτηση αλληλεπίδρασης και εφόσον δεν έχουν βρεθεί κάποιοι γνωστοί πράκτορες που να μπορούν να παρέχουν τη ζητούμενη υπηρεσία. Δέχεται ως είσοδο τα στοιχεία περιστασης της αναζητούμενης αλληλεπίδρασης, από τα οποία απομονώνει το αντικείμενο της σχετικής υπηρεσίας. Στη συνέχεια υποβάλλει στο DF ένα κατάλληλο ερώτημα ζητώντας πληροφορίες για το ποιοι πράκτορες έχουν την αντίστοιχη ειδικότητα. Το DF, αφού ελέγξει τα registrations που έχουν γίνει, στέλνει ως απάντηση μια λίστα με τα ονόματα των σχετικών πρακτόρων. Από τα ονόματα αυτά διαμορφώνεται, με τυχαίο τρόπο<sup>26</sup>, μια λίστα υποψηφίων για αλληλεπίδραση η οποία δίνεται στην έξοδο της συμπεριφοράς.

## Δημιουργία περίπτωσης

Η Δημιουργία περίπτωσης είναι μια απλή συμπεριφορά που έχει ως ρόλο τον καθορισμό των χαρακτηριστικών μιας επικείμενης αλληλεπίδρασης. Είναι η πρώτη συνιστώσα συμπεριφορά της Αναζήτησης αλληλεπίδρασης. Στο βασικό μοντέλο η περίπτωση προσδιορίζεται από τρία στοιχεία :

- i. Το αντικείμενο της αλληλεπίδρασης  
Αντιπροσωπεύει το είδος της αναζητούμενης υπηρεσίας.
- ii. Το κόστος της αλληλεπίδρασης  
Αντιπροσωπεύει, για τον μεν αναζητούντα την αλληλεπίδραση, τη σπουδαιότητα του ανταλλάγματος που χρειάζεται να δοθεί στον παροχέα της υπηρεσίας, για το δε παροχέα, το έργο που θα χρειαστεί να δαπανήσει για την παροχή της υπηρεσίας.

<sup>26</sup> Η τυχαία διάταξη εξασφαλίζει ότι σε διαδοχικές κλήσεις της υπηρεσίας DF δεν θα εμφανίζεται πάντα πρώτος ο ίδιος πράκτορας, με κίνδυνο να επιλέγεται συνεχώς ο ίδιος.

iii. Το όφελος που θα προκύψει από την αλληλεπίδραση

Αντιπροσωπεύει την βελτίωση που θα επέλθει στην κατάσταση του πράκτορα μετά την λήψη της υπηρεσίας.

Το αντικείμενο της περίπτωσης προκύπτει με τυχαία επιλογή κάποιου από τα ενδιαφέροντα του πράκτορα. Το κόστος και το όφελος είναι πραγματικοί αριθμοί που επιλέγονται επίσης τυχαία μέσα στην περιοχή  $[0,1]$ , με το 0 να αντιστοιχεί σε ανύπαρκτο κόστος ή όφελος και το 1 στο μέγιστο δυνατό.

Το κόστος θα ήταν ορθότερο να καθορίζεται αρχικά από τον παροχέα της υπηρεσίας και η τελική του τιμή να διαμορφώνεται μετά από διαπραγμάτευση των εμπλεκόμενων πρακτόρων. Η συγκεκριμένη επιλογή για καθορισμό του κόστους εξ' αρχής, από τον πράκτορα που αναζητά την αλληλεπίδραση, είναι μια σύμβαση που δεν επηρεάζει επί της ουσίας τις συνθήκες διενέργειας της αλληλεπίδρασης, ενώ επιπλέον απλοποιεί σημαντικά την υλοποίηση. Ο αναγνώστης μπορεί να θεωρήσει απλά ότι αυτό το κόστος αντιπροσωπεύει το ανώτερο όριο που είναι μπορεί να διαθέσει ο πράκτορας για την απόκτηση της υπηρεσίας και ότι ο παροχέας της υπηρεσίας, εφόσον αποδεχτεί την αίτηση για αλληλεπίδραση, συμφωνεί με την τιμή αυτή.

### **Αναζήτηση υποψηφίων μεταξύ των φίλων**

Είναι μια απλή συμπεριφορά που έχει ως ρόλο τον εντοπισμό των φίλων πρακτόρων που είναι κατάλληλοι για μια συγκεκριμένη αλληλεπίδραση. Δέχεται ως είσοδο τα στοιχεία περίπτωσης της αλληλεπίδρασης και διατρέπει τη λίστα των φίλων ελέγχοντας ποιοι απ' αυτούς έχουν ως ειδικότητα το αντικείμενο της αλληλεπίδρασης. Δεν διενεργεί άλλο έλεγχο καθώς εξ' ορισμού οι φίλοι είναι έμπιστοι πράκτορες, με την αξιοπιστία τους να έχει ελεγχθεί από τις κατάλληλες συμπεριφορές. Σε μια πληρέστερη εκδοχή, η συμπεριφορά θα μπορούσε να ελέγχει και την εξειδικευμένη αξιοπιστία του κάθε πράκτορα πάνω στο συγκεκριμένο αντικείμενο της αλληλεπίδρασης. Η επιλογή αυτή δεν υιοθετήθηκε γιατί θα μείωνε σημαντικά το πλήθος των διενεργούμενων αλληλεπιδράσεων αφού δεν θα ήταν εύκολο να βρεθούν υποψήφιοι.

Η έξοδος της συμπεριφοράς είναι μια λίστα με τα ονόματα των φίλων πρακτόρων που είναι κατάλληλοι για την αλληλεπίδραση. Εάν δεν βρεθεί κατάλληλος φίλος η λίστα θα είναι κενή.

Αξίζει εδώ να σημειωθεί ότι η Αναζήτηση υποψηφίων μεταξύ των φίλων είναι μια μάλλον εξειδικευμένη συμπεριφορά που η ύπαρξή της υπαγορεύεται από το συγκεκριμένο σχεδιασμό του βασικού μοντέλου. Είναι πιθανό να μην χρειάζεται σε άλλα μοντέλα εμπιστοσύνης.

### **Χειρισμός μηνυμάτων από κεντρικό πράκτορα**

Πρόκειται για μια τεχνική συμπεριφορά που σχετίζεται αποκλειστικά με τη λειτουργία της προσομοίωσης και όχι την κοινωνική συμπεριφορά των πρακτόρων. Αν και δεν αποτελεί κομμάτι



του προσομοιούμενου μοντέλου εμπιστοσύνης η λειτουργικότητά της αναλύεται απλά για λόγους πληρότητας της παρουσίασης της βασικής συμπεριφοράς των πρακτόρων.

Η συμπεριφορά παρακολουθεί διαρκώς την ουρά των εισερχόμενων μηνυμάτων για μηνύματα με αποστολέα τον κεντρικό πράκτορα. Μόλις εντοπιστεί ένα τέτοιο μήνυμα ελέγχεται το περιεχόμενό του. Στην τρέχουσα έκδοση του TrustSIM, τα μηνύματα από τον κεντρικό πράκτορα μπορεί να περιέχουν είτε απλές πληροφορίες, χρήσιμες στους πράκτορες, είτε οδηγίες για την εκτέλεση κάποιων ενεργειών. Οι πληροφορίες αφορούν την ενημέρωση του πράκτορα για τη γέννηση ή τον τερματισμό κάποιου άλλου πράκτορα. Οι οδηγίες αφορούν τον τερματισμό και την επανεκκίνηση των αλληλεπιδράσεων (δηλ. ενεργοποίηση / απενεργοποίηση της συμπεριφοράς αναζήτησης αλληλεπίδρασης) και την εκκίνηση ή τον τερματισμό της διαδικασίας συνεχούς αποστολής στοιχείων για τα εσωτερικά δεδομένα του πράκτορα.

Τα μηνύματα από τον κεντρικό πράκτορα εξυπηρετούνται παράλληλα με τις υπόλοιπες δραστηριότητες του πράκτορα. Αυτό συμβαίνει ώστε σε περίπτωση που ζητηθεί η αλλαγή συμπεριφοράς αυτό να μην επηρεάσει τις υπό εξέλιξη δραστηριότητες.

### 4.5.3 Σύνοψη

Κάθε μια από τις συμπεριφορές του βασικού μοντέλου αποτελεί ένα αυτόνομο και διακριτό δομικό στοιχείο, το οποίο μπορεί να αντικατασταθεί από οποιαδήποτε υλοποίηση τηρεί τους κανόνες διασύνδεσης των συμπεριφορών (δηλ. τη μορφή και πλήθος των δεδομένων εισόδου - εξόδου). Ένα μοντέλο εμπιστοσύνης όμως δεν χρειάζεται να παρέχει υλοποιήσεις για όλες τις συμπεριφορές. Εάν η παρεχόμενη υλοποίηση μιας σύνθετης συμπεριφοράς είναι πλήρης χωρίς να κάνει χρήση συνιστωσών συμπεριφορών, οι υλοποιήσεις των συνιστωσών συμπεριφορών που περιγράφονται από το βασικό μοντέλο μπορούν να παραληφθούν. Για παράδειγμα, κάποιο μοντέλο μπορεί να υλοποιεί τη συμπεριφορά αναζήτησης αλληλεπίδρασης κατά τρόπο που να μην χρειάζεται καμιά από τις συνιστώσες που προτείνει το βασικό μοντέλο (δηλ. δημιουργία περιστασης, εντοπισμό υποψηφίων κλπ). Αυτές οι συνιστώσες, δεν είναι υποχρεωτικό να υλοποιηθούν, εφόσον βέβαια δεν χρειάζονται και σε κάποια άλλη συμπεριφορά.

Η υλοποίηση των βασικών συμπεριφορών από ένα μοντέλο δεν υπόκειται σε κανένα περιορισμό παρά μόνο στο ότι θα πρέπει η κάθε συμπεριφορά να χρησιμοποιεί μόνο τις εισόδους που παρέχονται από τις προηγούμενες και να παρέχει την προβλεπόμενη έξοδο ώστε να μην δημιουργείται λειτουργικό πρόβλημα στις επόμενες. Αυτό το χαρακτηριστικό παρέχει εξαιρετικά μεγάλη ευελιξία επιτρέποντας την εύκολη κωδικοποίηση των περισσότερων μοντέλων εμπιστοσύνης.

## 4.6 Οι κυριότερες δομές δεδομένων

Ένας πράκτορας εμπιστοσύνης (όπως κάθε πράκτορας) διαθέτει κατάλληλες δομές δεδομένων μέσα στις οποίες διατηρεί τα δεδομένα που του χρειάζονται προκειμένου να υλοποιεί την προβλεπόμενη λειτουργία του. Οι δομές αυτές, ανάλογα με τη χρήση τους, μπορούν να χωριστούν σε δύο κατηγορίες :

- α) στις δομές γενικής χρήσης που υποστηρίζουν τις λειτουργικές ανάγκες του πράκτορα που δημιουργούνται από τη δραστηριοποίησή του μέσα στο πολυπρακτορικό περιβάλλον
- β) στις δομές ειδικής χρήσης που χρησιμοποιούνται για την υποστήριξη των μηχανισμών διαχείρισης εμπιστοσύνης που διαθέτει ένας πράκτορας εμπιστοσύνης

Οι δομές της δεύτερης κατηγορίας, αν και παίζουν ένα πολύ σημαντικό ρόλο στην αρχιτεκτονική και τη λογική ενός μοντέλου εμπιστοσύνης, τις περισσότερες φορές δεν τυγχάνουν της απαιτούμενης ανάλυσης από το συγγραφέα του μοντέλου. Ωστόσο, σε οποιαδήποτε προσπάθεια υλοποίησης ενός μοντέλου, το πρώτο πρόβλημα που θα χρειαστεί να αντιμετωπιστεί είναι ο σχεδιασμός και η επιλογή των κατάλληλων δομών δεδομένων.

Στην παράγραφο αυτή παρουσιάζουμε τις σημαντικότερες δομές δεδομένων που υιοθετήθηκαν για τους πράκτορες εμπιστοσύνης του περιβάλλοντος προσομοίωσης. Οι δομές αυτές θα μπορούσαν ίσως να αποτελέσουν σημείο αναφοράς για τις ελάχιστες απαραίτητες δομές δεδομένων που απαιτούνται για την υλοποίηση ενός μοντέλου εμπιστοσύνης.

### Δομή αξιολόγηση εμπιστοσύνης

Στη δομή αυτή διατηρούνται δεδομένα για την τελευταία αξιολόγηση της εμπιστοσύνης των γνωστών του πράκτορα. Η δομή αυτή θα πρέπει να ενημερώνεται κάθε φορά που προκύπτουν νέα στοιχεία, όπως συμβαίνει για παράδειγμα κατά τη διεξαγωγή αλληλεπιδράσεων, συλλογή συστάσεων ή την παραλαβή πληροφοριών φήμης (μέσω κουτσομπολιού) τα οποία επηρεάζουν την αξιολόγηση.

Η δομή αυτή υλοποιήθηκε ως μια λίστα αντιστοιχίσεων (mappings) μεταξύ ονομάτων πρακτόρων και εγγραφών αξιολόγησης εμπιστοσύνης. Η δομή περιέχει μια αντιστοίχιση για κάθε πράκτορα του οποίου η εμπιστοσύνη αξιολογήθηκε έστω και μια φορά. Μια εγγραφή αξιολόγησης εμπιστοσύνης ουσιαστικά είναι μια λίστα από τριάδες της μορφής

*(Ειδικότητα, Αξιολόγηση εμπιστοσύνης, Χρονική στιγμή)*

όπου

*Ειδικότητα* : η ειδικότητα για την οποία έγινε η αξιολόγηση εμπιστοσύνης

*Αξιολόγηση εμπιστοσύνης* : η τιμή της αξιολόγησης εμπιστοσύνης

*Χρονική στιγμή*: η χρονική στιγμή κατά την οποία έγινε η αξιολόγηση εμπιστοσύνης

Οι τριάδες της λίστας αντιστοιχούν στις επιμέρους αξιολογήσεις εμπιστοσύνης για τις διάφορες ειδικότητες που μπορεί να έχει ο αξιολογούμενος πράκτορας. Ανάμεσα στα στοιχεία της λίστας περιλαμβάνεται και μια τριάδα για τη γενική αξιολόγηση εμπιστοσύνης όπως αυτή θα καθορίζεται από

το υλοποιούμενο μοντέλο (π.χ. ως ο μέσος όρος των επιμέρους αξιολογήσεων). Εάν ένα μοντέλο δεν χειρίζεται την εμπιστοσύνη ξεχωριστά για κάθε ειδικότητα, η λίστα περιέχει μια μόνο τριάδα, αυτή της γενικής αξιολόγησης.

Η πληροφορία της χρονικής στιγμής τελευταίας ενημέρωσης συμπεριλαμβάνεται σε κάθε τριάδα αξιολόγησης διότι χρειάζεται στα μοντέλα εμπιστοσύνης που λαμβάνουν υπόψη την εξασθένηση της ισχύος μιας αξιολόγησης με το χρόνο [15,35].

Τεχνικά, η δομή υλοποιήθηκε ως αντικείμενο της κλάσης `jade.core.behaviours.DataStore`, η οποία παρέχεται από το API του JADE και είναι επέκταση της κλάσης `java.util.HashMap`. Η συγκεκριμένη αυτή κλάση επιλέχθηκε, έναντι άλλων εναλλακτικών (π.χ. πίνακας ή λίστα), γιατί προσφέρει πολύ εύκολη συντήρηση (προσθήκη, διαγραφή στοιχείων) και αναζήτηση περιεχομένων (κυρίως πρακτόρων).

Τέλος, αξίζει να σημειωθεί ότι η δομή της αξιολόγησης εμπιστοσύνης μπορεί να λειτουργήσει και ως ευρετήριο γνωστών πρακτόρων: κάθε γνωστός πράκτορας θα διαθέτει μια εγγραφή στα περιεχόμενα της δομής, καθώς θα του έχει γίνει τουλάχιστον μια αξιολόγηση εμπιστοσύνης.

Ένα παράδειγμα περιεχομένων αυτής της δομής (σύμφωνα με τις συμβάσεις που ακολουθήθηκαν στην συγκεκριμένη υλοποίηση) για ένα τυχαίο πράκτορα A? φαίνεται παρακάτω :

Δομή αξιολόγηση εμπιστοσύνης :

A12 →	<table border="1"> <tr> <td>("General", 0.94, 12/4/06 01:23:44)</td> </tr> <tr> <td>("8", 0.90, 22/4/06 10:12:05)</td> </tr> </table>	("General", 0.94, 12/4/06 01:23:44)	("8", 0.90, 22/4/06 10:12:05)	
("General", 0.94, 12/4/06 01:23:44)				
("8", 0.90, 22/4/06 10:12:05)				
A38 →	<table border="1"> <tr> <td>("General", 0.865, 25/3/06 20:01:17)</td> </tr> <tr> <td>("4", 0.81, 20/2/06 12:20:09)</td> </tr> <tr> <td>("1", 0.92, 25/3/06 20:01:17)</td> </tr> </table>	("General", 0.865, 25/3/06 20:01:17)	("4", 0.81, 20/2/06 12:20:09)	("1", 0.92, 25/3/06 20:01:17)
("General", 0.865, 25/3/06 20:01:17)				
("4", 0.81, 20/2/06 12:20:09)				
("1", 0.92, 25/3/06 20:01:17)				
⋮	⋮			
A55 →	<table border="1"> <tr> <td>("General", 0.88, 1/5/06 14:18:27)</td> </tr> <tr> <td>("2", 0.88, 1/5/06 14:18:27)</td> </tr> </table>	("General", 0.88, 1/5/06 14:18:27)	("2", 0.88, 1/5/06 14:18:27)	
("General", 0.88, 1/5/06 14:18:27)				
("2", 0.88, 1/5/06 14:18:27)				

Στο παράδειγμα αυτό, ο τυχαίος πράκτορας A? διατηρεί την εξής πληροφορία:

- Εμπιστεύεται γενικά τον πράκτορα A12 κατά 94% (τιμή γενικής εμπιστοσύνης = 0.94). Αυτή η εικόνα της γενική εμπιστοσύνης διαμορφώθηκε (τελευταία ενημέρωση) στις 12/4/06 και ώρα 1:23:44.
- Εμπιστεύεται τον πράκτορα A12 ως ειδικό του αντικειμένου "8" κατά 90%. Η τελευταία ενημέρωση της ειδικής εμπιστοσύνης έγινε στις 22/4/06 και ώρα 10:12:05
- Εμπιστεύεται γενικά τον πράκτορα A38 κατά 86,5% (τιμή γενικής εμπιστοσύνης = 0.865). Η τελευταία ενημέρωση της γενικής εμπιστοσύνης έγινε στις 25/3/06 και ώρα 20:01:17.
- .....

## Δομή φήμης

Κρατά δεδομένα που σχετίζονται με τη φήμη άλλων πρακτόρων. Η δομή αυτή έχει παρόμοια οργάνωση με τη δομή της αξιολόγησης εμπιστοσύνης: τα περιεχόμενά της είναι, και πάλι, αντιστοιχίσεις μεταξύ ονομάτων πρακτόρων και εγγραφών φήμης. Η δομή αυτή ενημερώνεται με στοιχεία που προέρχονται από συστάσεις που συλλέγονται, είτε κατόπιν σχετικής αίτησης, είτε μέσω της διαδικασίας του κουτσομπολιού.

Οι εγγραφές φήμης είναι τετράδες της μορφής

( *Ειδικότητα, Φήμη, Αποστολέας, Χρονική στιγμή* )

όπου

*Ειδικότητα* : η ειδικότητα την οποία αφορά η φήμη

*Φήμη* : η τιμή της φήμης

*Αποστολέας* : το όνομα του πράκτορα που έστειλε την σχετική πληροφορία φήμης

*Χρονική στιγμή*: η στιγμή κατά την οποία παραλήφθηκε η πληροφορία φήμης

Όπως και στη δομή αξιολόγησης εμπιστοσύνης, η πληροφορία της χρονικής στιγμής δίνει (στα μοντέλα που έχουν σχετική πρόβλεψη) τη δυνατότητα κανονικοποίησης της τιμής της φήμης βάσει της παλαιότητάς της (όσο πιο παλιά, τόσο λιγότερο σημαντική θα πρέπει να είναι).

Τεχνικά, κι αυτή η δομή υλοποιήθηκε ως αντικείμενο της κλάσης `jade.core.behaviours.DataStore`.

Ένα παράδειγμα περιεχομένων αυτής της δομής (σύμφωνα με τις συμβάσεις που ακολουθήθηκαν στην συγκεκριμένη υλοποίηση) για ένα τυχαίο πράκτορα A? φαίνεται παρακάτω :

Δομή φήμης :

A2 →	("General", 0.94, A8, 9/5/06 11:25:40)
	("3", 0.55, A5, 12/5/06 10:12:05)
A41 →	("1", 0.72, A59, 1/3/06 10:11:27)
	("4", 0.99, A12, 20/3/06 15:12:09)
	("1", 0.89, A44, 30//3/06 20:01:17)
	("General", 0.67, A59, 30/3/06 10:11:27)
⋮	⋮
A100 →	("General", 0.98, A41, 4/5/06 16:28:27)
	("7", 0.88, A33, 10/5/06 11:10:07)
	("General", 0.72, A1, 16/5/06 04:23:11)

Στο παράδειγμα αυτό, ο τυχαίος πράκτορας A? διατηρεί την εξής πληροφορία:

- Έχει δεχτεί από τον πράκτορα A8 μια προσωπική εκτίμηση (σύσταση) για την γενική αξιοπιστία του πράκτορα A12 που λέει ότι ο A12 είναι γενικά αξιόπιστος κατά 94% (τιμή σύστασης εμπιστοσύνης = 0.94). Η σύσταση παραλήφθηκε στις 9/5/06 και ώρα 11:25:40.

- Έχει δεχτεί από τον πράκτορα A5 μια προσωπική εκτίμηση (σύσταση) για την αξιοπιστία του πράκτορα A12 ως ειδικού στο αντικείμενο "3", που λέει ότι ο A12 είναι αξιόπιστος στο αντικείμενο "3" κατά 55% (τιμή σύστασης εμπιστοσύνης = 0.55). Η σύσταση παραλήφθηκε στις 12/5/06 και ώρα 10:12:05.

- ....

Γενικά οι δομές της φήμης και της αξιολόγησης εμπιστοσύνης δεν είναι υποχρεωτικό να περιέχουν τα ίδια ονόματα πρακτόρων. Όμως, επειδή, όπως και στα περισσότερα μοντέλα, η αξιολόγηση της εμπιστοσύνης είναι και συνάρτηση της τρέχουσας φήμης, το σύνολο των ονομάτων πρακτόρων της δομή της φήμης είναι υποσύνολο του αντίστοιχου της δομή αξιολόγησης της εμπιστοσύνης.

### Ιστορικό αλληλεπιδράσεων

Στη δομή αυτή διατηρούνται στοιχεία για όλες τις αλληλεπιδράσεις που διεξάγει ο συγκεκριμένος πράκτορας. Αυτό που ενδιαφέρει κυρίως ένα πράκτορα εμπιστοσύνης (από την οπτική γωνία της διαχείρισης εμπιστοσύνης) είναι η αξιολόγηση της κάθε αλληλεπίδρασης. Κι αυτό γιατί σ' όλα τα μοντέλα εμπιστοσύνης, η αξιολόγηση των αλληλεπιδράσεων είναι ο κυριότερος παράγοντας διαμόρφωσης της εμπιστοσύνης.

Περιεχόμενο της δομής είναι και πάλι αντιστοιχίσεις μεταξύ ονομάτων πρακτόρων και του ιστορικού των αλληλεπιδράσεών τους. Το ιστορικό των αλληλεπιδράσεων κάθε πράκτορα υλοποιείται ως μια λίστα εγγραφών αλληλεπίδρασης. Μια εγγραφή αλληλεπίδρασης περιλαμβάνει τα παρακάτω στοιχεία :

- Κωδικός* : μοναδικό αλφαριθμητικό που ξεχωρίζει την αλληλεπίδραση από οποιαδήποτε άλλη
- Περίσταση* : τα χαρακτηριστικά περίστασης της αλληλεπίδρασης (αντικείμενο, κόστος, όφελος)
- Χρονική στιγμή* : η χρονική στιγμή κατά την οποία ολοκληρώθηκε η αλληλεπίδραση
- Αξιολόγηση* : η αξιολόγηση της αλληλεπίδρασης

Η αξιολόγηση της αλληλεπίδρασης συνήθως είναι μια τιμή που αντιπροσωπεύει το επίπεδο ικανοποίησης από το αποτέλεσμα της αλληλεπίδρασης (δηλ. την ποιότητα της προσφερθείσας υπηρεσίας ή του συμφωνηθέντος ανταλλάγματος). Η αναπαράστασή της ποικίλει ανάλογα με το μοντέλο εμπιστοσύνης.

Τεχνικά, και η δομή του ιστορικού των αλληλεπιδράσεων υλοποιήθηκε ως αντικείμενο της κλάσης `jade.core.behaviours.DataStore` .

Παρακάτω δίνεται ένα παράδειγμα περιεχομένων της δομής για τον πράκτορα A1 :

- Δομή ιστορικού αλληλεπιδράσεων :

A13 →	("A13_5", ["1", 0.94, 1.00], 9/5/06 11:25:40, 0,97)
	("A1_89", ["4", 0.50, 0.60], 11/5/06 15:20:34, 0,68)
A89 →	("A1_5", ["4", 0.73, 0.80], 10/5/06 11:25:40, 0,97)
	("A89_10", ["1", 0.15, 0.98], 13/5/06 21:32:42, 0,70)
⋮	⋮
A50 →	("A50_15", ["1", 0.44, 1.00], 10/5/06 01:15:22, -0,54)

Στο παράδειγμα αυτό, ο πράκτορας A1 διατηρεί την εξής πληροφορία:

- Έχει αλληλεπιδράσει με τον πράκτορα A13 στις 9/5/06 και ώρα 11:25:40. Η αλληλεπίδραση είχε κωδικό "A13\_5" (δηλ. ήταν η 5<sup>η</sup> αλληλεπίδραση που ξεκίνησε με πρωτοβουλία του A13), αφορούσε το αντικείμενο "1", είχε κόστος 0.94 και όφελος 1. Η ποιότητα της αλληλεπίδρασης έφτασε το 97% (τιμή QoS = 0.97).
- Έχει αλληλεπιδράσει με τον πράκτορα A13 στις 11/5/06 και ώρα 15:20:34. Η αλληλεπίδραση είχε κωδικό "A1\_89" (δηλ. ήταν η 89<sup>η</sup> αλληλεπίδραση που ξεκίνησε με πρωτοβουλία του A1), αφορούσε το αντικείμενο "4", είχε κόστος 0.50 και όφελος 0.60. Η ποιότητα της αλληλεπίδρασης έφτασε το 68% (τιμή QoS = 0.68).
- ...



# ΚΕΦΑΛΑΙΟ

## 5. Μετρικές Αξιολόγησης Μοντέλων Εμπιστοσύνης

Παρά το γεγονός ότι έχουν προταθεί πολλά διαφορετικά μοντέλα εμπιστοσύνης, δεν έχει γίνει, έως σήμερα, συστηματική αξιολόγησή τους και συγκριτική μελέτη τους. Μάλιστα, δεν υπάρχει, έως σήμερα, ενιαία πλατφόρμα που θα μπορούσε να χρησιμοποιηθεί για τη μελέτη και αξιολόγηση όλων των μοντέλων, ούτε κοινά αποδεκτά κριτήρια που θα μπορούσαν να χρησιμοποιηθούν για την αποτίμησή τους. Πολύ συχνά, τα προτεινόμενα μοντέλα εμπιστοσύνης δεν συνοδεύονται από πειραματική αξιολόγηση της αποτελεσματικότητάς τους, ενώ στις περιπτώσεις που διενεργείται κάποια αξιολόγηση οι μετρικές που χρησιμοποιούνται είναι εξειδικευμένες, και πάνω στην περιοχή εφαρμογής του εκάστοτε μοντέλου, χωρίς να έχουν κοινά στοιχεία με αξιολογήσεις άλλων μοντέλων. Ωστόσο, είναι εξαιρετικά σημαντικό να καθοριστεί ένα σύνολο κοινών κριτηρίων και μετρικών γιατί έτσι θα διευκολυνόταν η συγκριτική αξιολόγηση των μοντέλων.

Η παρούσα εργασία επιχειρεί μια πρώτη συστηματική προσέγγιση στο πρόβλημα της αξιολόγησης των μοντέλων εμπιστοσύνης προτείνοντας μια σειρά κατάλληλων μετρικών. Οι μετρικές αυτές, που παρουσιάζονται σ' αυτό το κεφάλαιο, είναι γενικές και αποσκοπούν, κυρίως, στην εκτίμηση της αποτελεσματικότητας με την οποία ένα μοντέλο υλοποιεί τους γενικούς του στόχους. Οι στόχοι αυτοί, όπως είδαμε και στο 1ο κεφάλαιο αφορούν :

- την παροχή αξιόπιστων μηχανισμών αναγνώρισης και εντοπισμού κακόβουλων μελών
- την αποτροπή και την πρόληψη ανεπιθύμητων συμπεριφορών
- την ενθάρρυνση και προώθηση των επιθυμητών συμπεριφορών
- τη μη παρακώλυση των προβλεπόμενων δραστηριοτήτων της κοινωνίας (π.χ. συνεργασίες)
- τη διαμόρφωση κλίματος εμπιστοσύνης μεταξύ των μελών της κοινωνίας

Όλες οι μετρικές έχουν σχεδιαστεί έτσι ώστε να είναι όσο το δυνατόν πιο αντικειμενικές, χωρίς να εξαρτώνται από το συγκεκριμένο τρόπο με τον οποίο το κάθε μοντέλο προσπαθεί να επιτύχει τους βασικούς του στόχους. Για το λόγο αυτό, στον υπολογισμό τους, χρησιμοποιούν μόνο γενικές μεταβλητές, οι οποίες συναντώνται σε όλα τα μοντέλα εμπιστοσύνης.

### 5.1 *Trusting index*

Είναι ένας δείκτης που αναπαριστά το βαθμό κατά τον οποίο, κατά μέσο όρο, τα μέλη της κοινωνίας εμπιστεύονται τους γύρω τους. Υπολογίζεται ως ο μέσος όρος της μέσης εμπιστοσύνης που έχουν τα μέλη της κοινωνίας προς τα άλλα μέλη. Η μαθηματική του έκφραση είναι :

$$\text{Trusting index} = \frac{\sum_{i=1}^N \left( \frac{\sum_{j=1}^N T_{ij}}{N} \right)}{N}$$



όπου

$N$  : το πλήθος των μελών της κοινωνίας

$T_{ij}$  : η εμπιστοσύνη που έχει το μέλος  $i$  προς το μέλος  $j$

Από κοινωνιολογικής άποψης, η μετρική αυτή αντικατοπτρίζει το βαθμό επιφυλακτικότητας (ή ανασφάλειας) που διακατέχει τα μέλη μιας κοινωνίας. Όσο μικρότερη είναι η τιμή της, τόσο πιο επιφυλακτική είναι στο σύνολό της η κοινωνία. Προφανώς, το επιθυμητό για ένα μοντέλο είναι να επιτυγχάνει όσο γίνεται μεγαλύτερη τιμή στη μετρική αυτή, αφού αυτό συνεπάγεται υψηλές τιμές εμπιστοσύνης γεγονός που δημιουργεί ευνοϊκές συνθήκες για την διενέργεια αλληλεπιδράσεων και την ανάπτυξη σχέσεων συνεργασίας μεταξύ των μελών της κοινωνίας.

Θα πρέπει εδώ να τονισθεί ότι, επειδή ο υπολογισμός της μετρικής είναι ένας μέσος όρος επιμέρους τιμών εμπιστοσύνης, ο τύπος δεδομένων της μετρικής ποικίλει από μοντέλο σε μοντέλο και εξαρτάται άμεσα από την αναπαράσταση εμπιστοσύνης που υιοθετείται.

## 5.2 Trusted index

Η μετρική αυτή μπορεί να ειπωθεί ως συμμετρική της *trusting index*, υπό την έννοια ότι αξιολογεί το κλίμα εμπιστοσύνης που επικρατεί στην κοινωνία αλλά από την πλευρά των εμπιστευόμενων. Συγκεκριμένα, αναπαριστά το βαθμό κατά τον οποίο, κατά μέσο όρο, τα μέλη της κοινωνίας απολαμβάνουν εμπιστοσύνη από τους γύρω τους. Υπολογίζεται από τη μαθηματική έκφραση :

$$\text{Trusted index} = \frac{\sum_{j=1}^N \left( \frac{\sum_{i=1}^N T_{ij}}{N} \right)}{N}$$

όπου

$N$  : το πλήθος των μελών της κοινωνίας

$T_{ij}$  : η εμπιστοσύνη που έχει το μέλος  $i$  προς το μέλος  $j$

Κοινωνιολογικά, η μετρική αυτή παρέχει μια ένδειξη για το επίπεδο αξιοπιστίας ή φερεγγυότητας που χαρακτηρίζει συνολικά την κοινωνία. Όσο μεγαλύτερη είναι η τιμή της τόσο πιο αξιόπιστα θεωρούνται τα μέλη της αφού κατά μέσο όρο απολαμβάνουν υψηλή εμπιστοσύνη, ενώ αντίθετα μικρή τιμή υποδηλώνει ότι η κοινωνία απαρτίζεται από μέλη χαμηλής αξιοπιστίας. Και σ' αυτή τη μετρική το επιθυμητό είναι η μεγιστοποίηση της τιμής της.

Γενικά, στα περισσότερα μοντέλα, οι δείκτες *trusting index* και *trusted index* κινούνται σε παραπλήσια επίπεδα. Μάλιστα είναι χαρακτηριστικό ότι όσο πιο γεμάτος είναι ο πίνακας

εμπιστοσύνης<sup>27</sup>, τόσο πιο κοντά διαμορφώνονται οι δύο δείκτες, με την απόλυτη εξίσωσή τους να συμβαίνει όταν ο πίνακας εμπιστοσύνης είναι πλήρης. Ένα αρκετά ενδιαφέρον θέμα μελλοντικής δουλειάς θα ήταν να διερευνηθεί ποια μπορεί να είναι η μέγιστη δυνατή απόκλιση των δύο δεικτών, υπό ποιες προϋποθέσεις και τι μπορεί αυτό να σημαίνει για την κοινωνία.

### 5.3 Trustworthiness evaluation accuracy

Με την μετρική αυτή αξιολογείται η εγκυρότητα με την οποία ένα μοντέλο εκτιμά την αξιοπιστία των πρακτόρων. Είναι από τα σημαντικότερα κριτήρια αξιολόγησης της αποτελεσματικότητας ενός μοντέλου, δεδομένου ότι αφορά το μηχανισμό εκτίμησης της εμπιστοσύνης που είναι το κεντρικό χαρακτηριστικό κάθε μοντέλου εμπιστοσύνης. Η ορθότητα του μηχανισμού αυτού καθορίζει σε πολύ μεγάλο βαθμό και την ικανότητα του μοντέλου να εντοπίζει τα ανεπιθύμητα ή κακόβουλα μέλη της κοινωνίας. Ο υπολογισμός της μετρικής είναι :

$$\text{Trustworthiness evaluation accuracy} = \frac{E_c}{E}$$

όπου

$E$  : το πλήθος των μελών των οποίων η αξιοπιστία έχει αξιολογηθεί

$E_c$  : το πλήθος των σωστών αξιολογήσεων αξιοπιστίας

Μια αξιολόγηση της αξιοπιστίας ενός πράκτορα  $j$  από ένα πράκτορα  $i$  θεωρείται σωστή εάν για την τιμή εμπιστοσύνης  $T_{ij}$  που υπολογίζει ο πράκτορας  $i$  ισχύει ότι :

$T_{ij} \geq T_{\text{threshold}}$  και ο  $A_j$  είναι καλοπροαίρετος

ή

$T_{ij} < T_{\text{threshold}}$  και ο  $A_j$  είναι κακόβουλος

όπου

$T_{\text{threshold}}$  : το κατώφλι εμπιστοσύνης πάνω από το οποίο ένας πράκτορας θεωρείται, σύμφωνα με το κάθε μοντέλο, έμπιστος

Στον τύπο της μετρικής, το πλήθος των αξιολογήσεων προκύπτει αθροίζοντας μόνο τους πράκτορες για τους οποίους έχει γίνει μια τουλάχιστο εκτίμηση της εμπιστοσύνης τους, ενώ το πλήθος των σωστών αξιολογήσεων προκύπτει λαμβάνοντας υπόψη την τελευταία αξιολόγηση εμπιστοσύνης κάθε πράκτορα.

<sup>27</sup> Γεμάτος είναι ο πίνακας εμπιστοσύνης όταν  $\exists T_{ij} \forall i, j$ . Αυτό πρακτικά σημαίνει ότι κάθε μέλος της κοινωνίας έχει μια εκτίμηση για την εμπιστοσύνη κάθε άλλου μέλους.

Ουσιαστικά η μετρική υπολογίζει το ποσοστό των σωστών εκτιμήσεων αξιοπιστίας των μελών επί του συνόλου των μελών τα οποία έχουν αξιολογηθεί. Προφανώς όσο υψηλότερο είναι αυτό το ποσοστό τόσο πιο αξιόπιστο και αποτελεσματικό είναι το μοντέλο.

Πρέπει εδώ να τονίσουμε ότι το κατώφλι εμπιστοσύνης που χρησιμοποιείται στη μετρική είναι το ίδιο που χρησιμοποιείται από τους πράκτορες για τον καθορισμό των διαφόρων επιλογών τους (αποδοχή ή απόρριψη αλληλεπιδράσεων, συστάσεων κλπ). Είναι προφανές λοιπόν ότι η μετρική εξαρτάται σε πολύ μεγάλο βαθμό από την τιμή αυτού του κατωφλίου. Λανθασμένα επιλεγμένο κατώφλι εμπιστοσύνης (π.χ. πολύ υψηλή ή πολύ χαμηλή τιμή) μπορεί να οδηγήσει σε πολλές λάθος εκτιμήσεις με ανάλογο αντίκτυπο στην τιμή της μετρικής.

## 5.4 Malicious activity index

Ένα μοντέλο εμπιστοσύνης δεν μπορεί να θεωρείται επιτυχημένο εάν στην κοινωνία όπου εφαρμόζεται αναπτύσσεται μεγάλος αριθμός κακόβουλων συμπεριφορών. Η συγκεκριμένη μετρική αξιολογεί αυτήν ακριβώς την πλευρά των μοντέλων εμπιστοσύνης, δηλ. την ικανότητά τους να περιορίζουν τις κακόβουλες συμπεριφορές. Ο υπολογισμός της μετρικής περιγράφεται από τον τύπο :

$$\text{Malicious activity index} = \frac{I_m}{I}$$

όπου

$I_m$  : το πλήθος των κακόβουλων ή ανεπιθύμητων αλληλεπιδράσεων που έχουν λάβει χώρα

$I$  : το σύνολο των αλληλεπιδράσεων που έχουν διεξαχθεί στην κοινωνία

Από μαθηματική άποψη, η μετρική υπολογίζει το ποσοστό των κακόβουλων αλληλεπιδράσεων επί του συνόλου των αλληλεπιδράσεων που λαμβάνουν χώρα στην κοινωνία. Κοινωνιολογικά, αντικατοπτρίζει την έκταση των ανεπιθύμητων συμπεριφορών που εμφανίζονται σε μια κοινωνία και κατ' επέκταση μια ισχυρή ένδειξη για το πόσο υγιής είναι η κοινωνία. Η περιοχή τιμών της μετρικής είναι το διάστημα  $[0,1]$  με το αριστερό άκρο να είναι το επιθυμητό.

Στο σημείο αυτό θα πρέπει να σημειώσουμε ότι, το πλήθος των κακόβουλων αλληλεπιδράσεων που εμφανίζονται σε μια κοινωνία εξαρτάται άμεσα τόσο από την αναλογία κακόβουλων – καλοπροαίρετων μελών, όσο και από την συχνότητα με την οποία τα κακόβουλα μέλη επιδεικνύουν κακόβουλη συμπεριφορά. Η βελτίωση της συγκεκριμένης μετρικής ώστε να ανεξαρτητοποιηθεί από τις συγκεκριμένες παραμέτρους αποτελεί αντικείμενο μελλοντικής δουλειάς και συζητείται στο κεφάλαιο 6.

### 5.5 Malicious activity success

Στόχος και αυτής της μετρικής είναι η εκτίμηση της ικανότητας ενός μοντέλου να αντιμετωπίζει τις κακόβουλες συμπεριφορές, αξιολογώντας την αποτελεσματικότητα με την οποία μπλοκάρονται οι προσπάθειες για κακόβουλες αλληλεπιδράσεις. Συγκεκριμένα, μετρά το ποσοστό των επιτυχημένων κακόβουλων προσπαθειών μέσα από τον υπολογισμό :

$$\text{Malicious activity success} = \frac{I_m}{A_m}$$

όπου

$I_m$  : το πλήθος των κακόβουλων αλληλεπιδράσεων που έλαβαν χώρα

$A_m$  : το πλήθος των κακόβουλων προσπαθειών που επιχειρήθηκαν

Συγκριτικά με την προηγούμενη μετρική, η Malicious activity success δεν επηρεάζεται από την αναλογία των κακόβουλων μελών στην κοινωνία, ή από τη συχνότητα με την οποία συμπεριφέρονται κακόβουλα. Κι αυτό γιατί, για δεδομένη αποτελεσματικότητα παρεμπόδισης κακόβουλων συμπεριφορών μόνο ένα συγκεκριμένο ποσοστό των επιχειρούμενων κακόβουλων προσπαθειών θα καταλήξουν σε κακόβουλες αλληλεπιδράσεις. Το πεδίο τιμών της μετρικής είναι το διάστημα  $[0,1]$  με επιθυμητό, όπως και στην προηγούμενη μετρική, το κάτω άκρο.

### 5.6 Malicious activity efficiency

Η απλή καταγραφή του πλήθους των αναχαιτισμένων κακόβουλων προσπαθειών δεν είναι πάντα αρκετή για να εκτιμηθεί η αποτελεσματικότητα ενός μοντέλου στην αντιμετώπιση των κακόβουλων συμπεριφορών. Αυτό ισχύει ιδιαίτερα σε εφαρμογές που οι αλληλεπιδράσεις διαθέτουν κάποιο χαρακτηριστικό που τους προσδίδει βαρύτητα. Στις περιπτώσεις αυτές θα πρέπει να προσμετρηθεί και η βαρύτητα της κάθε αλληλεπίδρασης προκειμένου να εκτιμηθεί η απόδοση ενός μοντέλου. Στην κατεύθυνση αυτή κινείται η συγκεκριμένη μετρική, της οποίας ο υπολογισμός δίνεται από τον τύπο :

$$\text{Malicious activity efficiency} = \frac{\sum_{i=1}^{I_m} W_{I_m}(i)}{\sum_{i=1}^{A_m} W_{A_m}(i)}$$

όπου

$W_{I_m}(i)$  : η βαρύτητα της  $i$ -οστής κακόβουλης αλληλεπίδρασης

$W_{A_m}(i)$  : η βαρύτητα της  $i$ -οστής προσπάθειας για κακόβουλη αλληλεπίδραση

$I_m$  : το πλήθος των κακόβουλων αλληλεπιδράσεων που διενεργήθηκαν

$A_m$  : το πλήθος των κακόβουλων προσπαθειών που επιχειρήθηκαν

Η μετρική αυτή, λαμβάνοντας υπόψη τη βαρύτητα των κακόβουλων προσπαθειών που επιχειρήθηκαν και των κακόβουλων αλληλεπιδράσεων που πραγματοποιήθηκαν, προσπαθεί να κάνει μια εκτίμηση για την έκταση της ζημιάς που έχει προκληθεί σε σχέση με τη μέγιστη ζημιά που θα μπορούσε να προκληθεί εάν ήταν επιτυχημένες όλες οι κακόβουλες προσπάθειες.

Προφανώς, η μετρική έχει νόημα μόνο στις εφαρμογές όπου υπάρχει ποιοτική διαφοροποίηση των αλληλεπιδράσεων μέσω χαρακτηριστικών όπως το κόστος, η σπουδαιότητα, η επικινδυνότητα κλπ. Στις εφαρμογές αυτές η μετρική μπορεί να αξιολογήσει από μια άλλη οπτική τις αμυντικές επιδόσεις ενός μοντέλου, καθώς κάποιες φορές μπορεί να είναι προτιμότερο να "ξεφεύγουν" από τους αμυντικούς μηχανισμούς πολλές αλλά ασήμαντες (μικρής βαρύτητας) κακόβουλες αλληλεπιδράσεις, παρά λίγες και σημαντικές.

Τεχνικά, και στη μετρική αυτή, λόγω της ποσοστιαίας φύσης του υπολογισμού της, το πεδίο τιμών είναι η περιοχή  $[0,1]$ , με τις τιμές κοντά στο 0 να είναι οι επιθυμητές.

## 5.7 Interaction prevention index

Εάν ο στόχος ενός μοντέλου εμπιστοσύνης ήταν απλά ο περιορισμός των κακόβουλων αλληλεπιδράσεων, τότε θα μπορούσε, πολύ απλά, να οριστεί ένα πολύ υψηλό κατώφλι εμπιστοσύνης, από το οποίο δεν θα περνούσε καμιά προσπάθεια αλληλεπίδρασης. Όμως το δόγμα "ασφαλείς αλληλεπιδράσεις είναι μόνο αυτές που δεν πραγματοποιούνται" δεν είναι αποδεκτό σε εφαρμογές του πραγματικού κόσμου. Από τη σκοπιά αυτή, η απόδοση ενός μοντέλου εμπιστοσύνης θα πρέπει να κρίνεται και από το βαθμό κατά τον οποίο παρεμποδίζει τις υγιείς δραστηριότητες μιας κοινωνίας.

Η συγκεκριμένη μετρική κινείται σ' αυτή την κατεύθυνση και δίνει ένα δείκτη για την επίδραση που έχει ένα μοντέλο στις καλοπροαίρετες αλληλεπιδράσεις. Συγκεκριμένα υπολογίζει το ποσοστό των προσπαθειών που έγιναν για καλοπροαίρετες αλληλεπιδράσεις και απορρίφθηκαν, πάνω στο σύνολο των προσπαθειών αυτού του είδους. Ο τύπος του υπολογισμού της είναι :

$$\text{Interaction prevention index} = \frac{A_g^R}{A_g}$$

όπου

$A_g^R$  : το πλήθος των προσπαθειών που έγιναν για καλοπροαίρετες αλληλεπιδράσεις και απορρίφθηκαν

$A_g$  : το σύνολο των προσπαθειών που έγιναν για καλοπροαίρετες αλληλεπιδράσεις

Η περιοχή τιμών της μετρικής είναι το διάστημα  $[0,1]$ . Προφανώς, ένα σωστά σχεδιασμένο μοντέλο θα πρέπει να εμποδίζει όσο γίνεται λιγότερες προσπάθειες για καλοπροαίρετες αλληλεπιδράσεις, οπότε οι μικρές τιμές είναι οι επιθυμητές για τη μετρική.

### 5.8 Interaction prevention efficiency

Η μετρική αυτή κινείται στην κατεύθυνση της Malicious activity efficiency, αλλά από τη σκοπιά των καλοπροαίρετων αλληλεπιδράσεων. Παρέχει ένα μέτρο της επίπτωσης που έχει ένα μοντέλο στην ομαλή λειτουργία μιας κοινωνίας, αλλά λαμβάνοντας υπόψη τη σπουδαιότητα των αλληλεπιδράσεων. Αυτή η οπτική μπορεί να αποκαλύψει κρυμμένες διαστάσεις στις επιδόσεις ενός μοντέλου. Για παράδειγμα, ένα μοντέλο μπορεί να φαίνεται ότι παρεμποδίζει σημαντικό αριθμό χρήσιμων αλληλεπιδράσεων, οι οποίες όμως δεν έχουν μεγάλη αξία, ενώ αντίθετα σε κάποιο άλλο μοντέλο οι χρήσιμες αλληλεπιδράσεις που παρεμποδίζονται μπορεί να είναι λίγες αλλά σημαντικές.

Από τεχνική άποψη, η μετρική υπολογίζει το ποσοστό της βαρύτητας των προσπαθειών που έγιναν για καλοπροαίρετες αλληλεπιδράσεις και οι οποίες απορρίφθηκαν, επί του συνόλου των προσπαθειών αυτού του είδους. Ο υπολογισμός της περιγράφεται από τον τύπο :

$$\text{Interaction prevention efficiency} = \frac{\sum_{i=1}^{A_g^R} W_{A_g^R}(i)}{\sum_{i=1}^{A_g} W_{A_g}(i)}$$

όπου

$W_{A_g^R}(i)$  : η βαρύτητα της  $i$ -οστής προσπάθειας για καλοπροαίρετη αλληλεπίδραση, που απορρίφθηκε

$W_{A_g}(i)$  : η βαρύτητα της  $i$ -οστής προσπάθειας για καλοπροαίρετη αλληλεπίδραση

$A_g^R$  : Το πλήθος των προσπαθειών για καλοπροαίρετες αλληλεπιδράσεις που απορρίφθηκαν

$A_g$  : Το σύνολο των προσπαθειών για καλοπροαίρετες αλληλεπιδράσεις

Προφανώς, ή μετρική (όπως και η Malicious activity efficiency) είναι χρήσιμη σε εφαρμογές που οι αλληλεπιδράσεις διαθέτουν κάποιο χαρακτηριστικό γνώρισμα που τους προσδίδει βαρύτητα (κόστος, ωφελιμότητα κλπ).

Τεχνικά, η βαρύτητα μιας αλληλεπίδρασης μπορεί να αντιστοιχεί σε ένα από τα χαρακτηριστικά της ή και σε συνδυασμό αυτών. Η εκλογή όμως του τύπου που θα υπολογίζει τη βαρύτητα της κάθε αλληλεπίδρασης δεν επηρεάζει το αποτέλεσμα της μετρικής. Τέλος, περιοχή τιμών της μετρικής είναι το διάστημα  $[0,1]$  με τις μικρές τιμές να είναι και εδώ οι επιθυμητές, αφού η συνολική βαρύτητα των απορριπτόμενων καλοπροαίρετων αλληλεπιδράσεων πρέπει να παραμένει όσο το δυνατό χαμηλότερα.



# ΚΕΦΑΛΑΙΟ

## 6. Αξιολόγηση Μοντέλων Εμπιστοσύνης με τον Προσομοιωτή

Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα από την εκτέλεση μιας σειράς προσομοιώσεων. Για λόγους που αφορούν κυρίως τους χρονικούς περιορισμούς στην εκπόνηση της παρούσας εργασίας, δεν ήταν εφικτό να κωδικοποιήσουμε και να μελετήσουμε όλα τα μοντέλα εμπιστοσύνης που παρουσιάστηκαν στο κεφάλαιο 2. Μας δόθηκε όμως η δυνατότητα να δοκιμάσουμε και να ελέγξουμε το σύνολο της διαδικασίας κωδικοποίησης ενός μοντέλου, της προσθήκης του στον προσομοιωτή και της προσομοίωσής του για το μοντέλο PeerTrust. Αυτό βέβαια δεν στέρησε από την εργασία τη δυνατότητα εξαγωγής χρήσιμων συμπερασμάτων, καθώς η ύπαρξη του μοντέλου PeerTrust, σε συνδυασμό με το βασικό μοντέλο που είναι ενσωματωμένο στον προσομοιωτή, επέτρεψε τη διενέργεια συγκριτικής αξιολόγησης των αποτελεσμάτων της προσομοίωσης. Η περαιτέρω κωδικοποίηση επιπλέον μοντέλων και η εισαγωγή τους στον προσομοιωτή για μελέτη και αξιολόγηση αποτελεί αντικείμενο μελλοντικής δουλειάς, που δεν προβλέπουμε ότι παρουσιάζει κάποια ιδιαίτερη πρόκληση από την πλευρά της υλοποίησης.

Όπως είδαμε στο κεφάλαιο 4, ο προσομοιωτής λαμβάνει υπόψη και χειρίζεται ένα πλήθος παραμέτρων που αφορούν αφενός τα λειτουργικά χαρακτηριστικά των προσομοιούμενων μοντέλων και αφετέρου τη δομή της προσομοιούμενης κοινωνίας. Οι τιμές των παραμέτρων αυτών ορίζουν τα σενάρια προσομοίωσης. Από τις δοκιμές που έγιναν, διαπιστώσαμε ότι, τα αποτελέσματα μιας προσομοίωσης, ακόμη και του ίδιου μοντέλου, εξαρτώνται σημαντικά από το σενάριό της. Μάλιστα, ακόμη και δύο παραπλήσια σενάρια, τα οποία διαφοροποιούνται σε μια μόνο παράμετρο, μπορεί να δώσουν τελείως διαφορετικά αποτελέσματα. Στην ενότητα αυτή, παρουσιάζουμε κάποια ενδεικτικά σενάρια, στα οποία χρησιμοποιήθηκαν συγκεκριμένες τιμές για κάποιες από τις παραμέτρους προσομοίωσης, και δοκιμάστηκε ένας αριθμός εναλλακτικών τιμές στις υπόλοιπες. Η επιλογή των σεναρίων έγινε με γνώμονα την επίτευξη δύο απλών στόχων :

- α) την επίδειξη των δυνατοτήτων που προσφέρει το περιβάλλον του προσομοιωτή για τη μελέτη και αξιολόγηση μοντέλων εμπιστοσύνης
- β) τον έλεγχο της χρησιμότητας των προτεινόμενων μετρικών αξιολόγησης

Οι τιμές των παραμέτρων προσομοίωσης που διατηρήθηκαν σταθερές σε όλα τα σενάρια φαίνονται στον παρακάτω πίνακα:



Παράμετρος	Τιμή
Μέγεθος κοινωνίας	50 πράκτορες
Σύνθεση κοινωνίας	100% ρεαλιστές
Τύπος τιμών QoS	Δυαδικές
Ελάχιστο QoS	-1
Μέγιστο QoS	1
Μέγιστο πλήθος αλληλεπιδράσεων	1000
Διαχείριση περίπτωσης	Ναι
Εξειδίκευση & Ενδιαφέροντα πρακτόρων	Όχι
Περιοχή τιμών εμπιστοσύνης	[-1,1]

Στη συνέχεια αυτού του κεφαλαίου παρουσιάζονται τα αποτελέσματα των προσομοιώσεων που διενεργήθηκαν για τα δύο μοντέλα (βασικό και PeerTrust) μαζί με τις παρατηρήσεις στις οποίες οδήγησαν.

## 6.1 Τα σενάρια των προσομοιώσεων

Για να δοκιμάσουμε τις δυνατότητες που προσφέρει η πλατφόρμα του προσομοιωτή για συγκριτική αξιολόγηση των μοντέλων εμπιστοσύνης προσομοιώσαμε τα δύο μοντέλα πάνω στα ίδια ακριβώς σενάρια μελετώντας την απόδοσή τους μέσα από τις ίδιες μετρικές. Τα σενάρια αυτά είναι :

- **Σενάριο 1** : Διαμόρφωση εμπιστοσύνης χωρίς κακόβουλα μέλη

### Περιγραφή

Στο σενάριο αυτό μελετήθηκε η επίδοση του βασικού μοντέλου στη διαμόρφωση κλίματος εμπιστοσύνης χωρίς τη χρήση κοινωνικών μηχανισμών, σε μια υγιή κοινωνία χωρίς κακόβουλα μέλη.

### Ρυθμίσεις

Παράμετρος	Τιμή
Κακόβουλες συμπεριφορές	Όχι
Κοινωνικοί μηχανισμοί	Όχι
Αρχική εμπιστοσύνη	0.5
Κατώφλι εμπιστοσύνης	0.5

### Μετρικές

Trusting index, Trusted index

- **Σενάριο 2** : Διαμόρφωση εμπιστοσύνης & ακρίβεια στον εντοπισμό των κακόβουλων μελών σε κοινωνία με κακόβουλα μέλη

#### Περιγραφή

Στο σενάριο αυτό μελετήθηκε η επίδοση του βασικού μοντέλου στη διαμόρφωση κλίματος εμπιστοσύνης σε μια κοινωνία με αρκετά κακόβουλα μέλη (50% του πληθυσμού) τα οποία δρουν κακόβουλα στο 50% των ενεργειών τους. Παράλληλα, αξιολογήθηκε και η αποτελεσματικότητα με την οποία το μοντέλο εντοπίζει τα κακόβουλα μέλη. Και στο σενάριο αυτό δεν ενεργοποιήθηκαν κοινωνικοί μηχανισμοί.

#### Ρυθμίσεις

Παράμετρος	Τιμή
Κακόβουλες συμπεριφορές	Ναι
Κακόβουλοι πράκτορες	25 (50%)
Ρυθμός κακόβουλων ενεργειών	50%
Κοινωνικοί μηχανισμοί	Όχι
Αρχική εμπιστοσύνη	0.5
Κατώφλι εμπιστοσύνης	0.5

#### Μετρικές

Trusting index, Trusted index, Trustworthiness evaluation accuracy

- **Σενάριο 3** : Διαμόρφωση εμπιστοσύνης & ακρίβεια στον εντοπισμό των κακόβουλων μελών, σε κοινωνία με κακόβουλα μέλη, με χρήση κοινωνικών μηχανισμών

#### Περιγραφή

Στο σενάριο αυτό επαναλάβαμε τις ρυθμίσεις του 2<sup>ου</sup> σεναρίου αλλά προσθέτοντας την ενεργοποίηση όλων των διαθέσιμων κοινωνικών μηχανισμών. Στόχος ήταν να δούμε την επίδραση των κοινωνικών μηχανισμών στη διαμόρφωση της εμπιστοσύνης στην κοινωνία καθώς και στον εντοπισμό των κακόβουλων μελών.

#### Ρυθμίσεις

Παράμετρος	Τιμή
Κακόβουλες συμπεριφορές	Ναι
Κακόβουλοι πράκτορες	25 (50%)
Ρυθμός κακόβουλων ενεργειών	50%
Κοινωνικοί μηχανισμοί	Φήμη, Συστάσεις,
Αρχική εμπιστοσύνη	0.5
Κατώφλι εμπιστοσύνης	0.5

**Μετρικές**

Trusting index, Trusted index, Trust evaluation accuracy

- **Σενάριο 4** : Αντιμετώπιση κακόβουλων συμπεριφορών χωρίς κοινωνικούς μηχανισμούς

**Περιγραφή**

Στο σενάριο αυτό μελετήσαμε την αμυντική ικανότητα του βασικού μοντέλου, δηλ. την ικανότητά του να αποτρέπει και να προλαμβάνει τις κακόβουλες συμπεριφορές. Η αξιολόγηση βασίστηκε στο πλήθος των κακόβουλων συμπεριφορών που "πέρασαν" (διενεργήθηκαν) καθώς και στο πλήθος των προσπαθειών για κακόβουλες συμπεριφορές που απορρίφθηκαν. Στα χαρακτηριστικά του σεναρίου, η κοινωνία ρυθμίστηκε να έχει τα μισά μέλη της κακόβουλα (50%), τα οποία συμπεριφερόταν κακόβουλα στις μισές τους ενέργειες. Στη συγκεκριμένη προσομοίωση μελετήσαμε την αμυντική ικανότητα του μοντέλου χωρίς την βοήθεια κοινωνικών μηχανισμών.

**Ρυθμίσεις**

Παράμετρος	Τιμή
Κακόβουλες συμπεριφορές	Ναι
Κακόβουλοι πράκτορες	25 (50%)
Ρυθμός κακόβουλων ενεργειών	50%
Κοινωνικοί μηχανισμοί	Όχι
Αρχική εμπιστοσύνη	0.5
Κατώφλι εμπιστοσύνης	0.5

**Μετρικές**

Malicious activity index, Malicious activity success, Malicious activity efficiency

- **Σενάριο 5** : Αντιμετώπιση κακόβουλων συμπεριφορών με χρήση κοινωνικών μηχανισμών

**Περιγραφή**

Το σενάριο αυτό αποτελεί παραλλαγή του προηγούμενου δηλαδή έχει τα ίδια χαρακτηριστικά με την επιπλέον προσθήκη της ενεργοποίησης των κοινωνικών μηχανισμών. Στόχος ήταν να δούμε την επίδραση των κοινωνικών μηχανισμών στην αμυντική ικανότητα του βασικού μοντέλου όπως αυτή καταγράφηκε στην προηγούμενη προσομοίωση. Για να είναι δυνατή η σύγκριση των δύο περιπτώσεων, χρησιμοποιήθηκαν οι ίδιες μεταβλητές παρακολούθησης και μετρικές.

**Ρυθμίσεις**

Παράμετρος	Τιμή
Κακόβουλες συμπεριφορές	Ναι
Κακόβουλοι πράκτορες	25 (50%)
Ρυθμός κακόβουλων ενεργειών	50%
Κοινωνικοί μηχανισμοί	Φήμη, Συστάσεις,
Αρχική εμπιστοσύνη	0.5
Κατώφλι εμπιστοσύνης	0.5

**Μετρικές**

Malicious activity index, Malicious activity success, Trust evaluation accuracy

- **Σενάριο 6** : Αντιμετώπιση έντονα κακόβουλων συμπεριφορών με χρήση κοινωνικών μηχανισμών

**Περιγραφή**

Το 6<sup>ο</sup> σενάριο βασίστηκε στο σενάριο 5. Η διαφορά ήταν ότι χρησιμοποιήθηκαν έντονα κακόβουλες συμπεριφορές. Στόχος ήταν να μελετήσαμε τη αμυντική ικανότητα του βασικού μοντέλου σε έντονα κακόβουλο περιβάλλον. Στα χαρακτηριστικά του σεναρίου, επιλέξαμε μια κοινωνία με πολλούς και έντονα κακόβουλους πράκτορες (70% επί του συνολικού πληθυσμού και κακόβουλοι σε όλες οι ενέργειές τους), ενώ παράλληλα ενεργοποιήθηκαν οι κοινωνικοί μηχανισμοί. Στα στοιχεία μελέτης παρακολουθήσαμε επιπλέον και τις παρενέργειες (side effects) του αμυντικού μηχανισμού του μοντέλου στα έντιμα μέλη της κοινωνίας. Για το σκοπό αυτό, στις μεταβλητές παρακολούθησης προστέθηκε και το πλήθος των προσπαθειών για καλοπροαίρετες αλληλεπιδράσεις που απορρίφθηκαν.

**Ρυθμίσεις**

Παράμετρος	Τιμή
Κακόβουλες συμπεριφορές	Ναι
Κακόβουλοι πράκτορες	35 (70%)
Ρυθμός κακόβουλων ενεργειών	100%
Κοινωνικοί μηχανισμοί	Φήμη, Συστάσεις,
Αρχική εμπιστοσύνη	0.5
Κατώφλι εμπιστοσύνης	0.5

**Μετρικές**

Malicious activity index, Malicious activity success, Malicious activity efficiency

- **Σενάριο 7** : Αντιμετώπιση έντονα κακόβουλων συμπεριφορών με χρήση κοινωνικών μηχανισμών και αυστηρά κριτήρια επιλογής αλληλεπιδράσεων

### Περιγραφή

Στο σενάριο αυτό επεκτείναμε το σενάριο 6 και μελετήσαμε την αμυντική ικανότητα του βασικού μοντέλου σε έντονα κακόβουλο περιβάλλον, αλλά αυτή τη φορά ρυθμίζοντας το κατώφλι εμπιστοσύνης σχετικά ψηλά, ώστε να είναι γίνει πιο αυστηρός ο αμυντικός μηχανισμός. Το ενδιαφέρον εστιάστηκε και πάλι όχι μόνο στην αμυντική επίδοση του μοντέλου αλλά και στις παρενέργειές της πάνω στη φυσιολογική δραστηριότητα της κοινωνίας. Επιπλέον, επειδή οι πράκτορες ξεκινούν χωρίς καμιά γνώση για την υπόλοιπη κοινωνία, για να δουλέψει το σενάριο, ήταν αναγκαίο να αυξήσουμε και την αρχική εμπιστοσύνη, διαφορετικά δεν θα μπορούσε να διενεργηθεί καμιά αλληλεπίδραση (λόγω του υψηλού κατωφλίου).

### Ρυθμίσεις

Παράμετρος	Τιμή
Κακόβουλες συμπεριφορές	Ναι
Κακόβουλοι πράκτορες	35 (70%)
Ρυθμός κακόβουλων ενεργειών	100%
Κοινωνικοί μηχανισμοί	Φήμη, Συστάσεις,
Αρχική εμπιστοσύνη	0.8
Κατώφλι εμπιστοσύνης	0.8

### Μετρικές

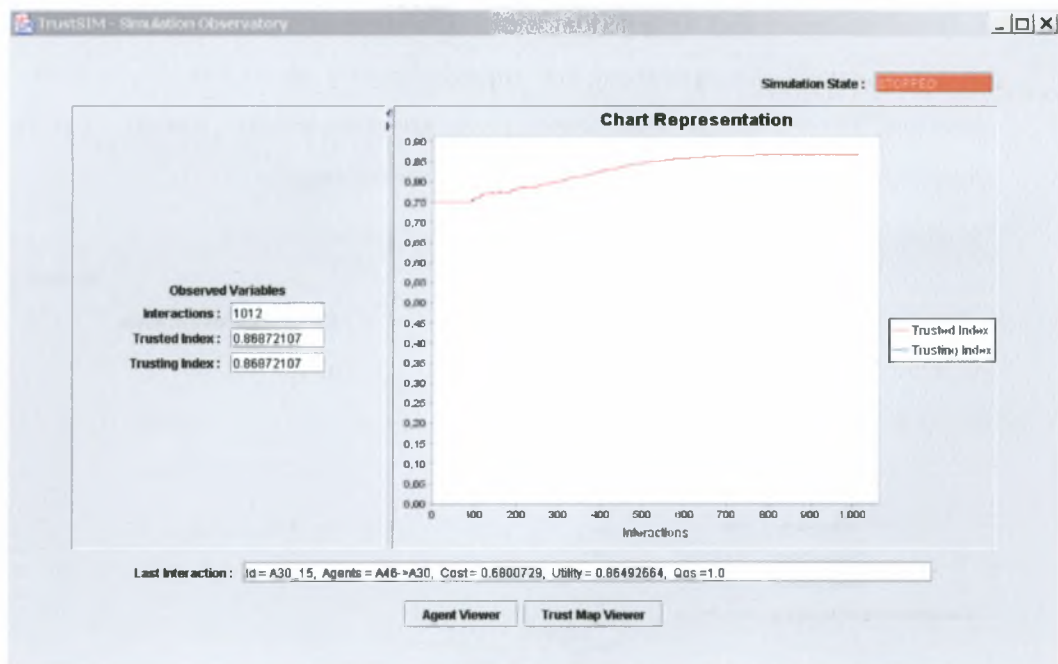
Malicious activity index, Malicious activity success, Malicious activity efficiency

## 6.2 Προσομοιώσεις του Βασικού Μοντέλου

- **Σενάριο 1** : Διαμόρφωση εμπιστοσύνης χωρίς κακόβουλα μέλη

### Παρατηρήσεις - Σχολιασμός

Η τελικές τιμές των επιλεγμένων μετρικών και η διακύμανση κατά τη διάρκεια της προσομοίωσης φαίνονται στην εικόνα 6-1 .



**Εικόνα 6-1 : Βασικό μοντέλο - Αποτελέσματα 1<sup>ου</sup> σεναρίου**

Όπως βλέπουμε, το μοντέλο βοηθά στην αποκατάσταση μιας αρκετά υψηλής μέσης εμπιστοσύνης στην κοινωνία, της τάξης του 0.86. Επίσης παρατηρούμε ότι, οι πράκτορες εμπιστεύονται τους άλλους στον ίδιο βαθμό που τους εμπιστεύονται και εκείνοι. Βέβαια, λόγω της απουσίας κακόβουλων συμπεριφορών, ίσως να ανέμενε κανείς ακόμη υψηλότερη μέση εμπιστοσύνη, κοντά στην μέγιστη δυνατή τιμή (1). Ο λόγος που δεν συμβαίνει αυτό είναι ότι οι πράκτορες από ένα σημείο και μετά διαμορφώνουν προτιμήσεις και επιλέγουν για τις αλληλεπιδράσεις τους αυτούς με την υψηλότερη εμπιστοσύνη. Το αποτέλεσμα είναι να συνδιαλέγονται συνεχώς με τους λίγους εκλεκτούς τους (που μπορεί να είναι και μόνο ένας) αδιαφορώντας για τους υπόλοιπους των οποίων η εμπιστοσύνη παραμένει χαμηλότερα (στην τιμή που διαμορφώθηκε μετά την πρώτη αλληλεπίδραση).

- **Σενάριο 2** : Διαμόρφωση εμπιστοσύνης & ακρίβεια στον εντοπισμό των κακόβουλων μελών σε κοινωνία με κακόβουλα μέλη

### Παρατηρήσεις - Σχολιασμός

Η προσομοίωση έδειξε ότι με το βασικό μοντέλο η μέση εμπιστοσύνη στην κοινωνία πέφτει σημαντικά όταν υπάρχουν κακόβουλες συμπεριφορές (εικόνα 6-2). Το αποτέλεσμα είναι λογικό αφού θα υπάρχουν αρκετές αρνητικές αξιολογήσεις εμπιστοσύνης με αρνητική τιμή που θα αντιστοιχούν στην έλλειψη εμπιστοσύνης. Ανεξάρτητα όμως από το επίπεδο στο οποίο σταθεροποιείται η μέση εμπιστοσύνη της κοινωνίας, παρατηρούμε ότι οι πράκτορες εξακολουθούν να επιδεικνύουν εμπιστοσύνη στον ίδιο βαθμό που απολαμβάνουν εμπιστοσύνη. Ως προς το μηχανισμό αξιολόγησης της αξιοπιστίας, το μοντέλο εμφανίζει αρκετά μεγάλη

ακρίβεια, που φτάνει το 87%. Μάλιστα, οι λανθασμένες εκτιμήσεις οφείλονται περισσότερο στο ότι οι κακόβουλοι πράκτορες δεν συμπεριφέρονται πάντα κακόβουλα (ένας κακόβουλος πράκτορας μπορεί να έχει συμπεριφερθεί σε κάποιους εντελώς έντιμα), και λιγότερο στην ανεπάρκεια του μηχανισμού αξιολόγησης της εμπιστοσύνης.



Εικόνα 6-2 : Βασικό μοντέλο - Αποτελέσματα 2<sup>ου</sup> σεναρίου

- **Σενάριο 3** : Διαμόρφωση εμπιστοσύνης & ακρίβεια στον εντοπισμό των κακόβουλων μελών, σε κοινωνία με κακόβουλα μέλη, με χρήση κοινωνικών μηχανισμών

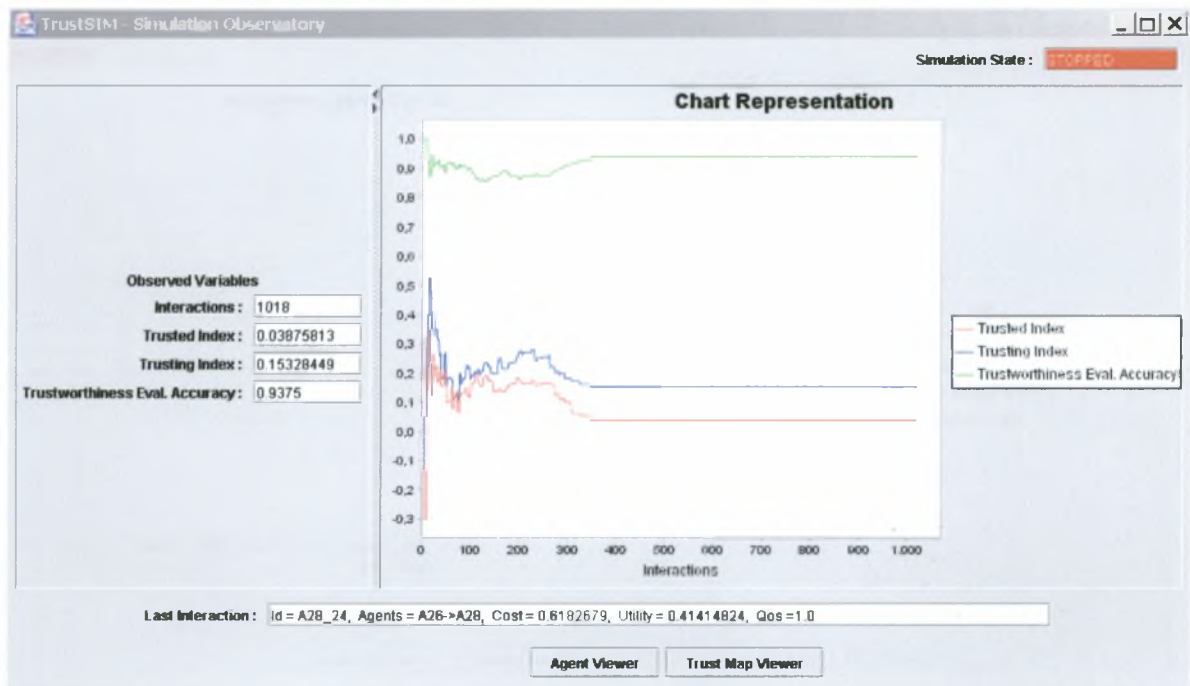
### Παρατηρήσεις - Σχολιασμός

Τα αποτελέσματα της προσομοίωσης του 3ου σεναρίου φαίνονται στην εικόνα 6-3. Όπως βλέπουμε η μέση εμπιστοσύνη της κοινωνίας δεν σημείωσε ιδιαίτερη πρόοδο. Υπήρξαν όμως δύο αλλαγές σε σχέση με το προηγούμενο σενάριο :

- Διαφοροποιήθηκε ο βαθμός κατά τον οποίο τα μέλη της κοινωνίας δείχνουν εμπιστοσύνη από το βαθμό κατά τον οποίο τους εμπιστεύονται.
- Η μέση εμπιστοσύνη σταθεροποιήθηκε πολύ νωρίτερα απ' ότι στο σενάριο 2

Επιπλέον παρατηρήσαμε ότι υπήρξε μια μικρή αλλά εμφανής αύξηση στην ακρίβεια εκτίμησης της αξιοπιστίας (της τάξης του 10%).

Οι αλλαγές αυτές οφείλονται κατά κύριο λόγο στο γεγονός ότι, με τους μηχανισμούς διαχείρισης φίλων και φήμης οι πράκτορες, αφενός ενημερώνονται γρηγορότερα και αφετέρου τείνουν να συναναστρέφονται μόνο τους έντιμους, με αποτέλεσμα οι κακόβουλοι πράκτορες να μην βρίσκουν θύματα.

Εικόνα 6-3 : Βασικό μοντέλο - Αποτελέσματα 3<sup>ου</sup> σεναρίου

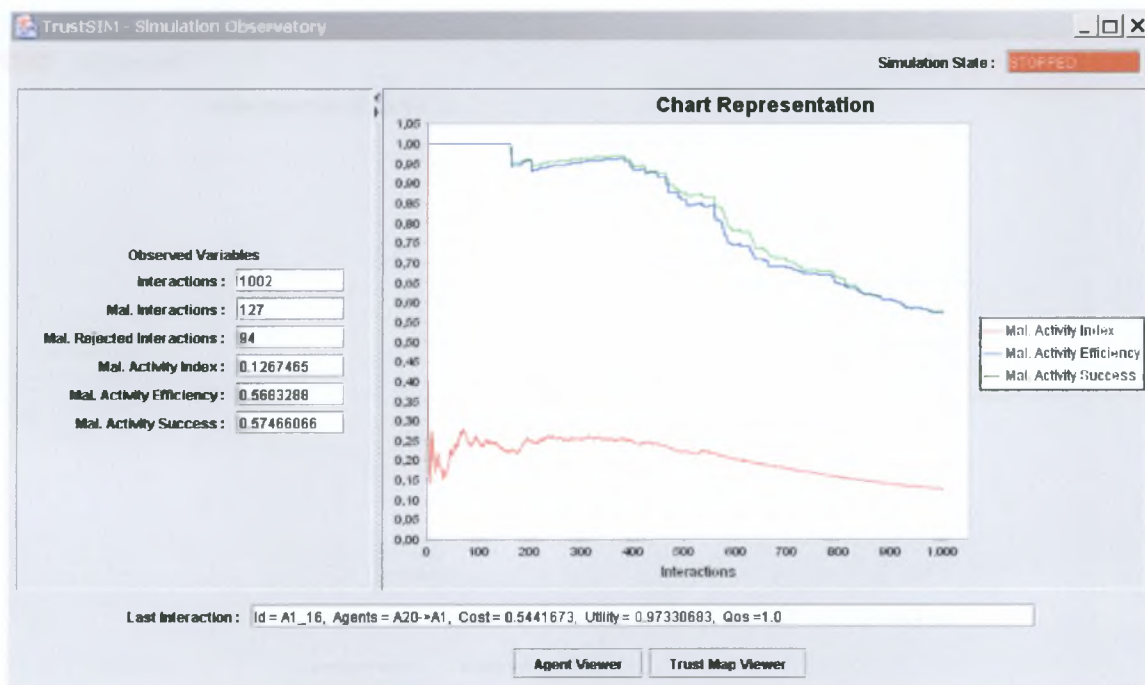
- **Σενάριο 4** : Αντιμετώπιση κακόβουλων συμπεριφορών χωρίς κοινωνικούς μηχανισμούς

#### Παρατηρήσεις - Σχολιασμός

Τα αποτελέσματα της προσομοίωσης φαίνονται στην εικόνα 6-4. Η ύπαρξη αρκετών κακόβουλων μελών στην κοινωνία είχε ως αποτέλεσμα να εκδηλωθούν αρκετές κακόβουλες συμπεριφορές, συνολικά 211. Απ' αυτές ήταν επιτυχημένες οι 127, ποσοστό αρκετά υψηλό όπως υποδεικνύει και ο δείκτης επιτυχίας κακόβουλης δραστηριότητας (*Malicious activity success*) που κυμάνθηκε πάνω από το 50%, στα ίδια επίπεδα με τον δείκτη απόδοσης της κακόβουλης δραστηριότητας (*Malicious activity efficiency*). Βέβαια, πρέπει να σημειώσουμε ότι η πορεία των δύο δεικτών ήταν σταθερά φθίνουσα, γεγονός αναμενόμενο αφού από ένα σημείο και μετά οι κακόβουλοι πράκτορες έγιναν αντιληπτοί και οι προτάσεις τους για αλληλεπιδράσεις άρχισαν να εισπράττουν αρνήσεις. Πάντως, λόγω της απουσίας κοινωνικών μηχανισμών, είναι λογικό να αναμένουμε ότι κάθε κακόβουλος πράκτορας θα κατάφερνε να εξαπατήσει τουλάχιστο μια φορά κάθε ένα από τους άλλου πράκτορες. Δηλ. μακροπρόθεσμα, θα εκδηλώνονταν τουλάχιστο  $25 \times 49 = 1225$  κακόβουλες αλληλεπιδράσεις.

Τέλος, ένα άλλο στοιχείο που αξίζει να προσέξουμε ότι η αμυντική λειτουργία του μοντέλου άρχισε να αποδίδει μετά τις πρώτες 400 περίπου αλληλεπιδράσεις, οπότε και άρχισε η φθίνουσα πορεία των δεικτών απόδοσης της κακόβουλης συμπεριφοράς.



Εικόνα 6-4 : Βασικό μοντέλο - Αποτελέσματα 4<sup>ου</sup> σεναρίου

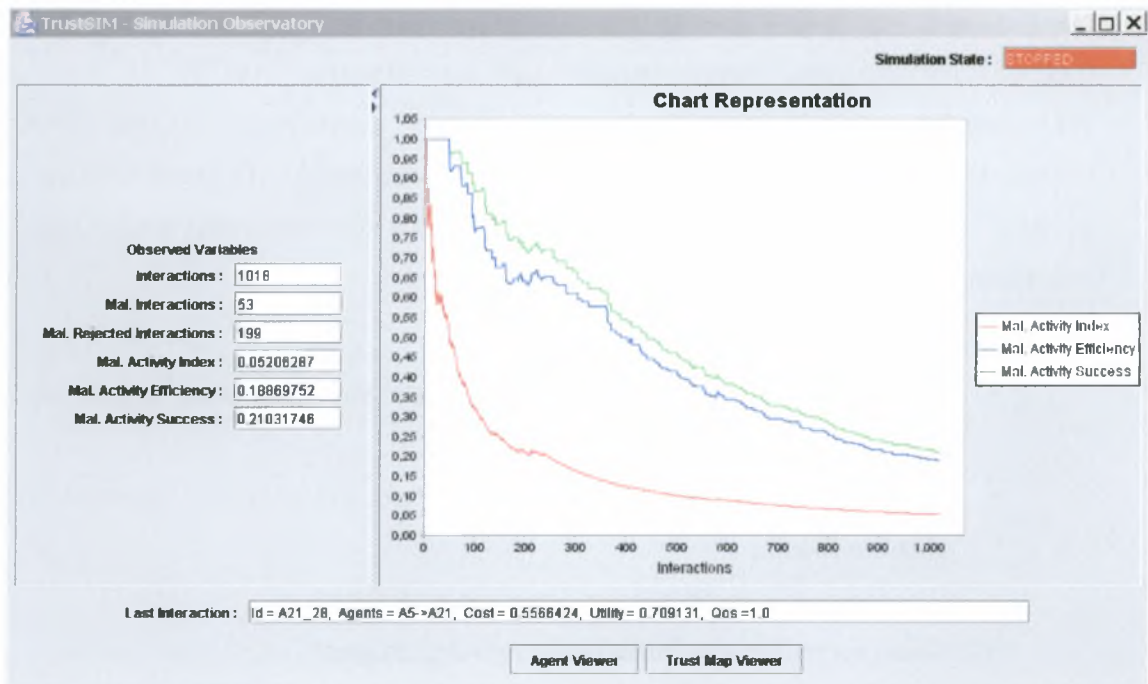
- **Σενάριο 5** : Αντιμετώπιση κακόβουλων συμπεριφορών με χρήση κοινωνικών μηχανισμών

#### Παρατηρήσεις - Σχολιασμός

Όπως φαίνεται και από το γράφημα με τη διακύμανση των δεικτών απόδοσης της κακόβουλης συμπεριφοράς (εικόνα 6-5), η αμυντική ικανότητα του μοντέλου βελτιώθηκε κατακόρυφα με τη χρήση των κοινωνικών μηχανισμών. Αυτό προκύπτει από τα εξής στοιχεία :

- Το ποσοστό των επιτυχημένων προσπαθειών για κακόβουλες αλληλεπιδράσεις έπεσε από 57% σε 21%.
- Η αμυντική λειτουργία του μοντέλου απέδωσε από τις πρώτες κιόλας αλληλεπιδράσεις
- Οι κακόβουλες συμπεριφορές είχαν σημαντικά μειωμένη απόδοση (*Malicious activity efficiency* = 0.18).

Η βελτίωση αυτή οφείλεται κυρίως στην ύπαρξη των συστάσεων και του κουτσομπολιού. Μέσω των μηχανισμών αυτών οι κακόβουλες συμπεριφορές έγιναν γρηγορότερα αντιληπτές με αποτέλεσμα κάθε κακόβουλος πράκτορας να μπορέσει να διαπράξει κατά μέσο όρο μόνο δύο εξαπατήσεις, σε αντίθεση με την προηγούμενη προσομοίωση (χωρίς τους κοινωνικούς μηχανισμούς) όπου διέπραξε πέντε.

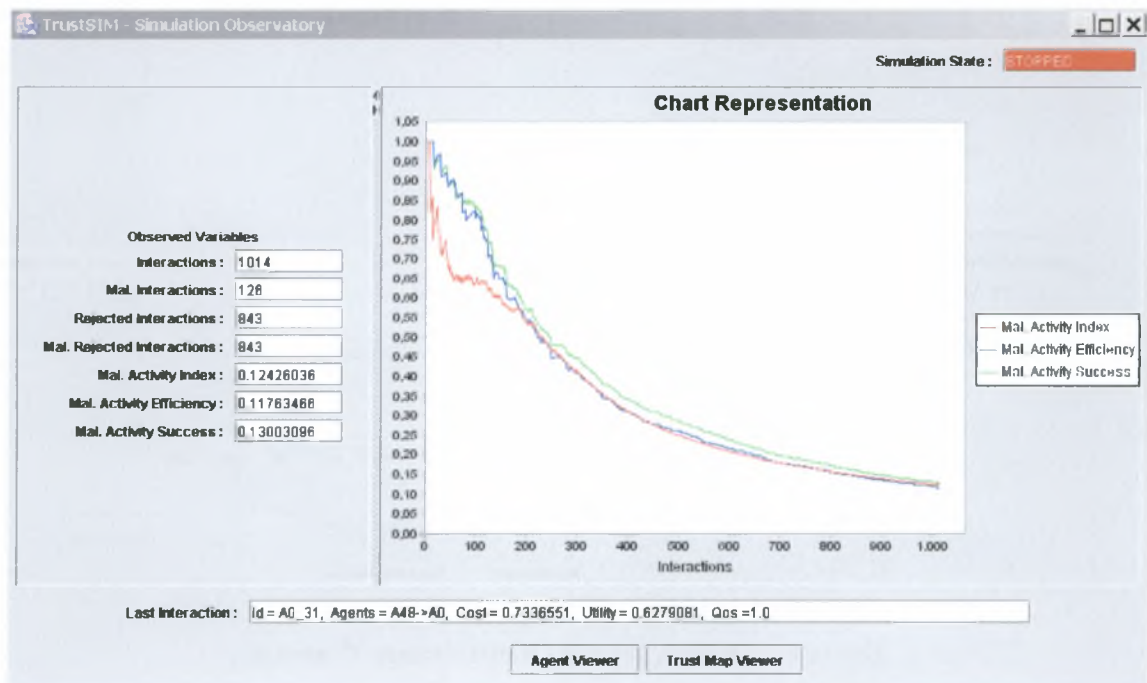


Εικόνα 6-5 : Βασικό μοντέλο - Αποτελέσματα 5<sup>ου</sup> σεναρίου

- **Σενάριο 6** : Αντιμέτωπιση έντονα κακόβουλων συμπεριφορών με χρήση κοινωνικών μηχανισμών

**Παρατηρήσεις - Σχολιασμός**

Τα αποτελέσματα της προσομοίωσης του 6<sup>ου</sup> σεναρίου φαίνονται στην εικόνα 6-6.



Εικόνα 6-6 : Βασικό μοντέλο - Αποτελέσματα 6<sup>ου</sup> σεναρίου

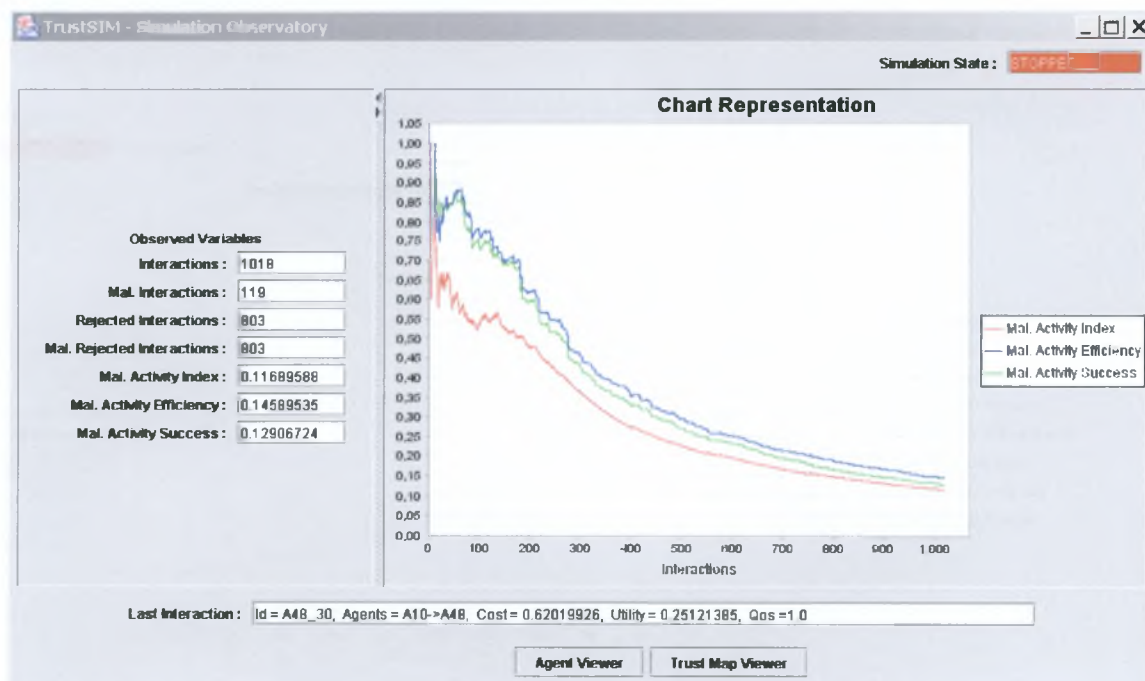
Όπως φαίνεται, το βασικό μοντέλο εξακολουθεί να είναι το ίδιο αποτελεσματικό και στις αντίξοες συνθήκες ενός έντονα κακόβουλου περιβάλλοντος. Αν και οι κακόβουλες αλληλεπιδράσεις αυξήθηκαν (λογικό, αφού υπήρχαν περισσότερα και πιο δραστήρια κακόβουλα μέλη), το ποσοστό επιτυχίας τους ήταν πολύ μικρό. Το σημαντικότερο όμως στοιχείο είναι ότι η αμυντική λειτουργία του μοντέλου δεν είχε καμιά παρενέργεια στις καλοπροαίρετες δραστηριότητες της κοινωνίας.

Τα κυριότερα ευρήματα της προσομοίωσης είναι :

- Το ποσοστό των επιτυχημένων προσπαθειών για κακόβουλες αλληλεπιδράσεις ήταν μόλις 13%
  - Η αμυντική λειτουργία του μοντέλου απέδωσε και πάλι από τις πρώτες κιόλας αλληλεπιδράσεις
  - Η απόδοση των κακόβουλων συμπεριφορών μειώθηκε.
  - Καμιά προσπάθεια για καλοπροαίρετη αλληλεπίδραση δεν παρεμποδίστηκε
- **Σενάριο 7** : Αντιμέτωπιση έντονα κακόβουλων συμπεριφορών με χρήση κοινωνικών μηχανισμών και αυστηρά κριτήρια επιλογής αλληλεπιδράσεων

### Παρατηρήσεις - Σχολιασμός

Τα αποτελέσματα της προσομοίωσης όπως απεικονίζονται στην εικόνα 6-7.



Εικόνα 6-7 : Βασικό μοντέλο - Αποτελέσματα 7<sup>ου</sup> σεναρίου

Όπως γίνεται αντιληπτό ότι μοντέλο δεν διαφοροποίησε σημαντικά την επίδοσή του. Όλοι οι δείκτες και οι μεταβλητές παρακολούθησης που καταγράφηκαν, έδωσαν παραπλήσια τιμές με αυτές του 6<sup>ου</sup> σεναρίου. Οι μικρές αποκλίσεις που εμφανίστηκαν είναι μέσα σε φυσιολογικά όρια και δικαιολογούνται από τη αυτόνομη και τυχαία δραστηριοποίηση των πρακτόρων. Το ενδιαφέρον στοιχείο της προσομοίωσης είναι το γεγονός ότι το αυξημένο κατώφλι δεν προκάλεσε την απόρριψη καμιάς προσπάθειας για καλοπροαίρετη αλληλεπίδραση. Αυτό οφείλεται στη υψηλή αρχική εμπιστοσύνη. Η εμπιστοσύνη των έντιμων πρακτόρων, σύμφωνα με το μηχανισμό ενημέρωσης της εμπιστοσύνης του μοντέλου, αυξάνει συνεχώς και έτσι, από τη στιγμή που ξεκινάει πάνω από το κατώφλι εμπιστοσύνης, παραμένει υψηλότερα με αποτέλεσμα οι έντιμοι πράκτορες να ικανοποιούν πάντα το κριτήριο επιλογής. Η παρατήρηση αυτή μας οδηγεί στο συμπέρασμα ότι, στο βασικό μοντέλο, προσπάθειες για καλοπροαίρετες αλληλεπιδράσεις μπορούν να απορριφθούν μόνο εάν προέρχονται από κακόβουλους πράκτορες.

### 6.3 Προσομοιώσεις του μοντέλου PeerTrust

Κατά τις προσομοιώσεις του μοντέλου PeerTrust, πέραν όλων των άλλων παρατηρήσεων που έγιναν και οι οποίες παρουσιάζονται στο σχολιασμό των σεναρίων, διαπιστώσαμε και μια σημαντική διαφορά (σε σχέση με το βασικό μοντέλο) στην χρονική διάρκεια των προσομοιώσεων. Το στοιχείο αυτό επιβεβαίωσε την θεωρητική εκτίμηση που είχε γίνει κατά τη μελέτη του μοντέλου στο κεφάλαιο 2 και η οποία προέβλεπε ότι το μοντέλο είναι βαρύ υπολογιστικά, εξαιτίας της συλλογής των εμπειριών από όλα τα μέλη της κοινωνίας κατά την αξιολόγηση της εμπιστοσύνης ενός μέλους.

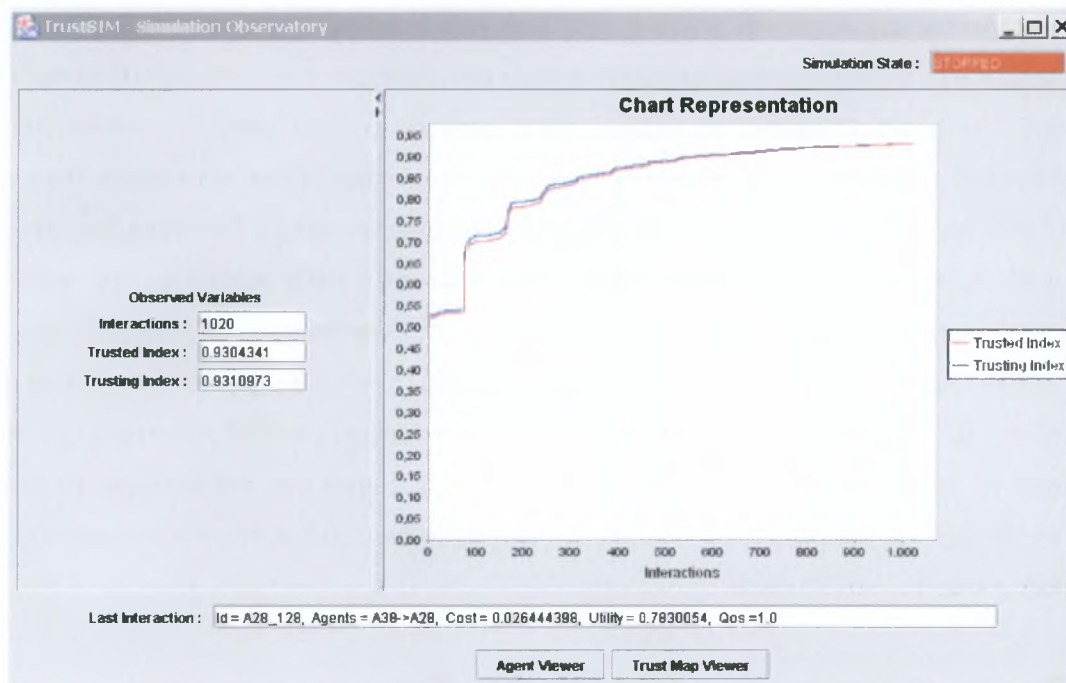
- **Σενάριο 1:** Διαμόρφωση εμπιστοσύνης χωρίς κακόβουλα μέλη

#### Παρατηρήσεις - Σχολιασμός

Τα αποτελέσματα της προσομοίωσης του 1<sup>ου</sup> σεναρίου για το μοντέλο PeerTrust φαίνονται στην εικόνα 6-8. Σε κοινωνίες χωρίς κακόβουλα μέλη, το μοντέλο PeerTrust φαίνεται να επιτυγχάνει ελαφρώς καλύτερα επίπεδα μέσης εμπιστοσύνης στην κοινωνία απ' ότι το βασικό μοντέλο. Η τελική τιμή των δύο δεικτών εμπιστοσύνης στο PeerTrust κυμαίνεται πάνω από 0.93. Επιπλέον υπάρχουν και άλλες δύο μικρές διαφοροποιήσεις σε σχέση με το βασικό μοντέλο :

- Οι δείκτες εμπιστοσύνης ξεκινούν χαμηλότερα, από την τιμή 0.5 που είναι η τιμή της αρχικής εμπιστοσύνης
- Η μέγιστη τιμή των δεικτών εμπιστοσύνης δεν φαίνεται να έχει επιτευχθεί μετά το πέρας των 1000 αλληλεπιδράσεων.

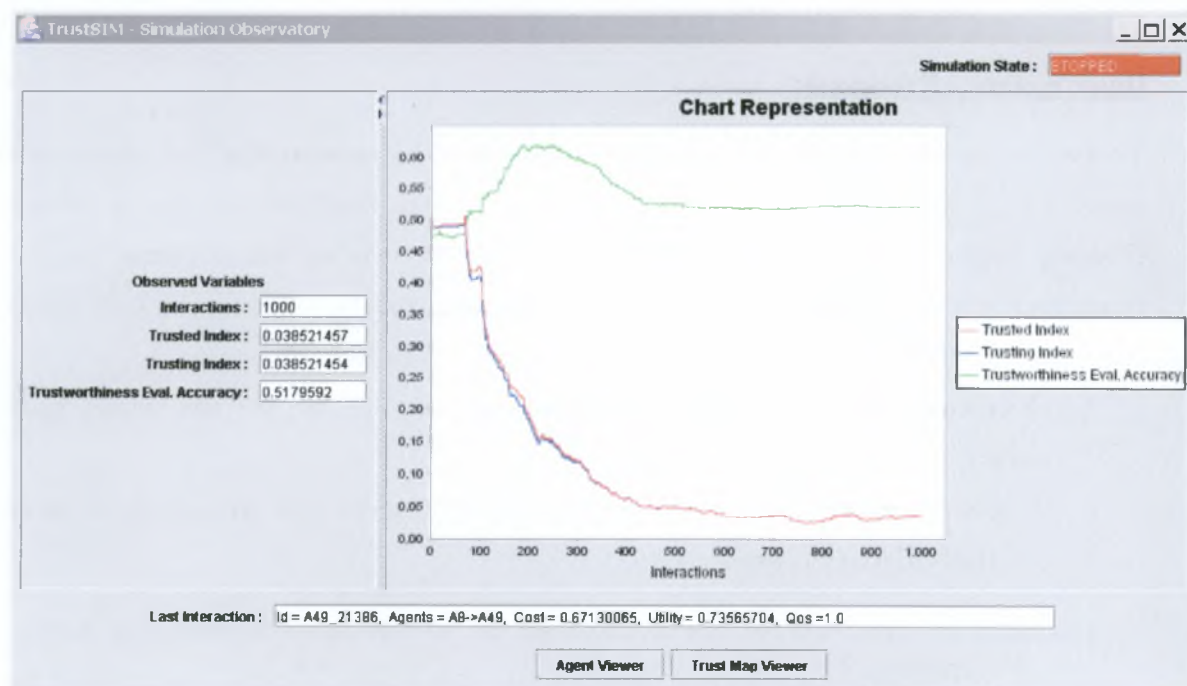
Τα παραπάνω στοιχεία οδηγούν στο συμπέρασμα ότι το μοντέλο PeerTrust είναι κάπως πιο επιφυλακτικό στην αρχή, αλλά μακροπρόθεσμα επιτυγχάνει καλύτερο κλίμα εμπιστοσύνης σε μια κοινωνία χωρίς κακόβουλα μέλη.

Εικόνα 6-8 : PeerTrust - Αποτελέσματα 1<sup>ου</sup> σεναρίου

- **Σενάριο 2** : Διαμόρφωση εμπιστοσύνης & ακρίβεια στον εντοπισμό των κακόβουλων μελών σε κοινωνία με κακόβουλα μέλη

### Παρατηρήσεις - Σχολιασμός

Η προσομοίωση του 2<sup>ου</sup> σεναρίου έδειξε το μοντέλο PeerTrust σε κοινωνίες με κακόβουλα μέλη αντιμετωπίζει αρκετά προβλήματα (εικόνα 6-9).

Εικόνα 6-9 : PeerTrust - Αποτελέσματα 2<sup>ου</sup> σεναρίου

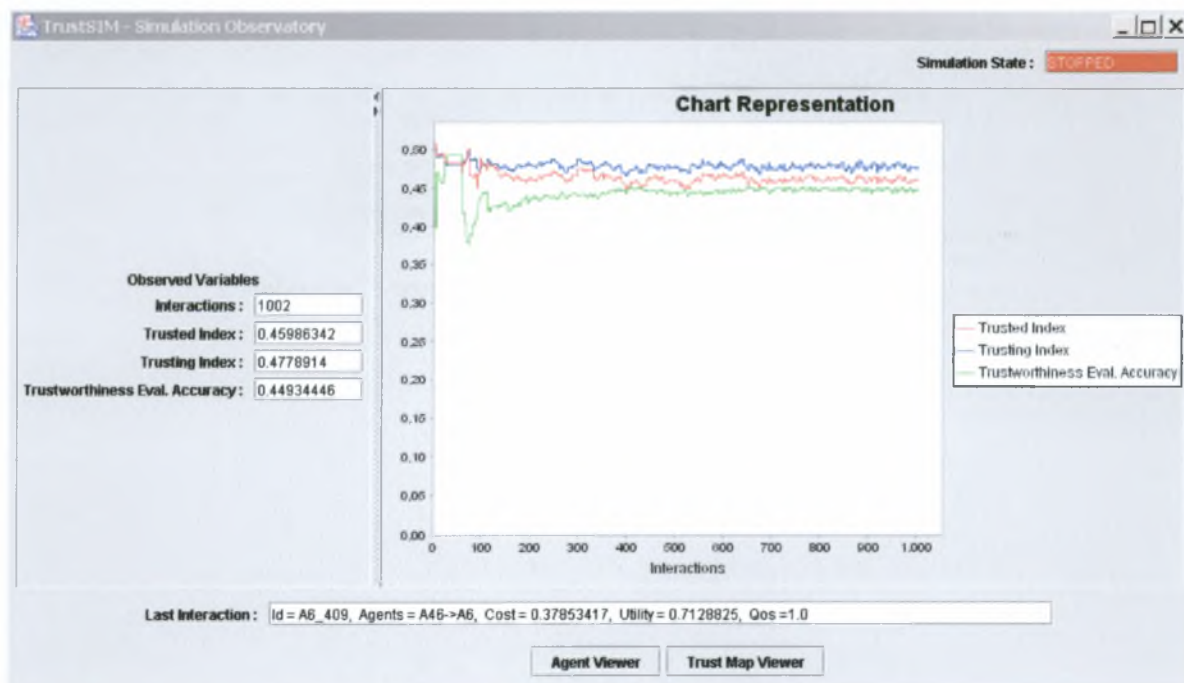
Όπως βλέπουμε, το κλίμα εμπιστοσύνης στην κοινωνία δεν καθόλου καλό, με τους δείκτες εμπιστοσύνης να κυμαίνονται κοντά στο 0, ενώ και η ακρίβεια του μηχανισμού αξιολόγησης της αξιοπιστίας έχει επιτυχία μόλις 51%. Μάλιστα η εμπιστοσύνη φαίνεται να ξεκινά την πτωτική της πορεία αρκετά σύντομα (μετά την 90<sup>η</sup> αλληλεπίδραση) και να παίρνει πολύ γρήγορα τις χαμηλές τιμές στις οποίες τελικά κυμάνθηκε. Σε σχέση με το βασικό μοντέλο, το μοντέλο PeerTrust φαίνεται να έχει ελαφρά χειρότερη επίδοση στη διαμόρφωση κλίματος εμπιστοσύνης στην κοινωνία, αλλά σημαντικά χειρότερη στην ακρίβεια με την οποία εντοπίζει τα κακόβουλα μέλη.

- **Σενάριο 3** : Διαμόρφωση εμπιστοσύνης & ακρίβεια στον εντοπισμό των κακόβουλων μελών, σε κοινωνία με κακόβουλα μέλη, με χρήση κοινωνικών μηχανισμών

### Παρατηρήσεις - Σχολιασμός

Με την ενεργοποίηση των κοινωνικών μηχανισμών, το μοντέλο PeerTrust παρουσίασε μερική μόνο βελτίωση. Όπως φαίνεται και από τα αποτελέσματα της προσομοίωσης στην εικόνα 6-10, η μέση εμπιστοσύνη της κοινωνίας κυμάνθηκε κοντά στην αρχική εμπιστοσύνη, επίδοση σημαντικά υψηλότερη από την τιμή που διαμορφώθηκε στο προηγούμενο σενάριο, χωρίς τη χρήση κοινωνικών μηχανισμών. Όμως η ικανότητα εντοπισμού των κακόβουλων μελών μειώθηκε πέφτοντας στο 45% από το 51% του προηγούμενου σεναρίου.

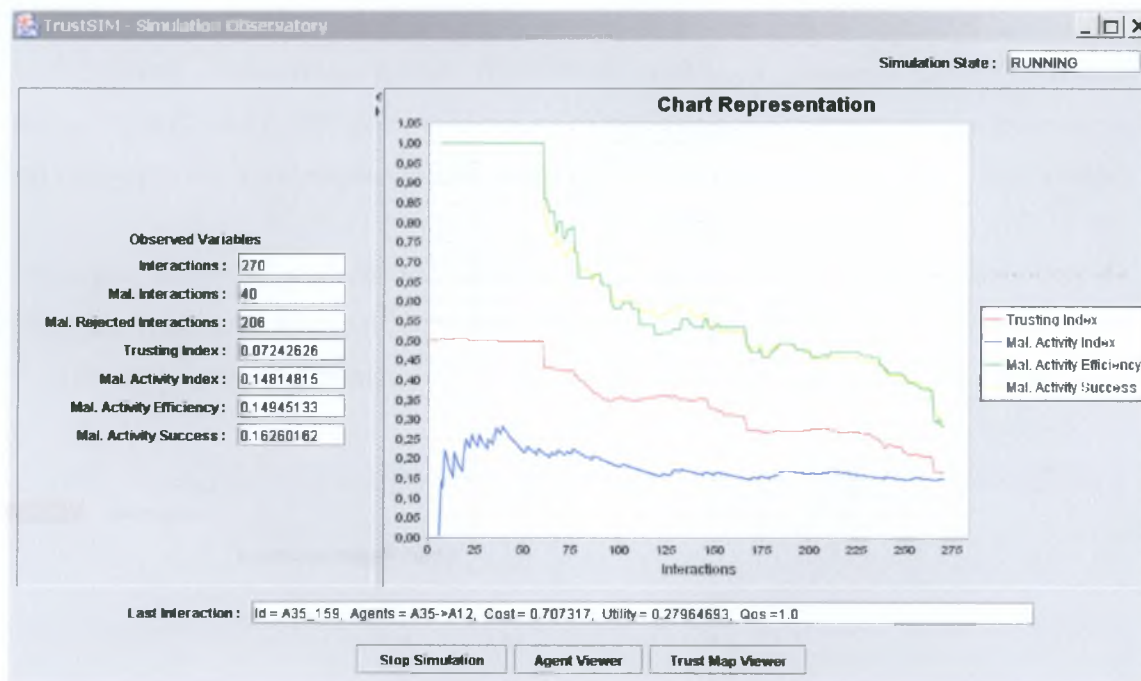
Οι επιδόσεις αυτές, συγκρινόμενες με τις αντίστοιχες του βασικού μοντέλου δείχνουν ότι το PeerTrust υπερέχει μεν στη διαμόρφωση κλίματος εμπιστοσύνης αλλά μειονεκτεί στον εντοπισμό των κακόβουλων μελών, γεγονός που περιμένουμε να έχει επίπτωση και στην αμυντική ικανότητα του μοντέλου απέναντι στις κακόβουλες συμπεριφορές.

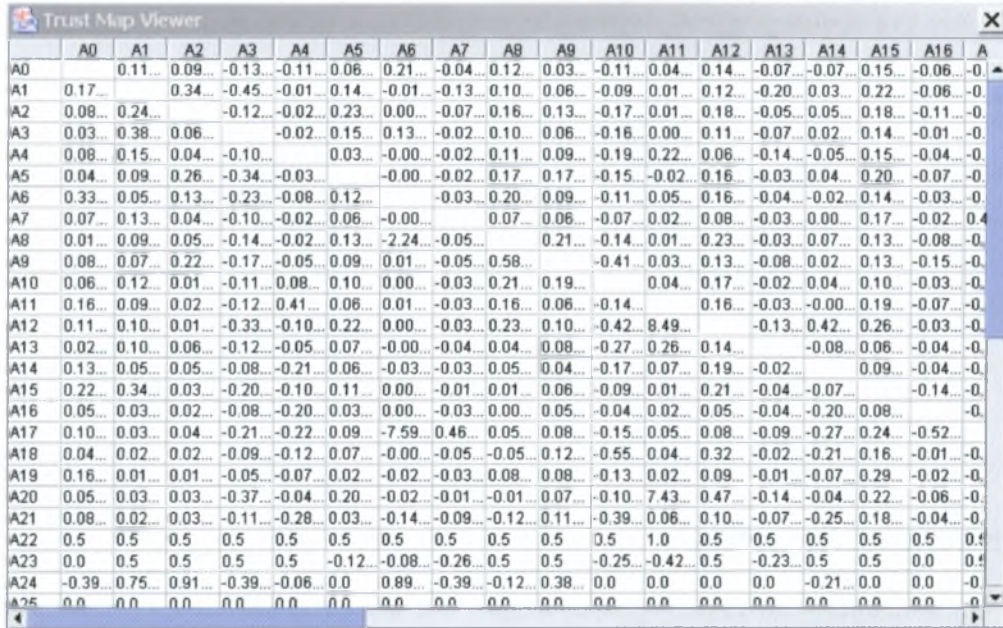


Εικόνα 6-10 : PeerTrust - Αποτελέσματα 3<sup>ου</sup> σεναρίου

**Σενάριο 4** : Αντιμέτωπιση κακόβουλων συμπεριφορών χωρίς κοινωνικούς μηχανισμούς**Παρατηρήσεις - Σχολιασμός**

Τα αποτελέσματα της προσομοίωσης του 4<sup>ου</sup> σεναρίου για το μοντέλο PeerTrust φαίνονται στην εικόνα 6-11. Στο σενάριο αυτό, εμφανίστηκε μια περίεργη κατάσταση. Στην πρώτη προσπάθεια εκτέλεσης της προσομοίωσης, η προσομοίωση φάνηκε να σταματά περίπου στην 260<sup>η</sup> αλληλεπίδραση. Το ίδιο συνέβη και σε επόμενες προσπάθειες (με μικρές αποκλίσεις στο πλήθος των αλληλεπιδράσεων) ανεξάρτητα από το χρόνο που δόθηκε σε κάθε μια εκτέλεση. Το μυστήριο λύθηκε ρίχνοντας μια ματιά στον χάρτη εμπιστοσύνης της κοινωνίας (εικόνα 6-12). Όλες οι αξιολογήσεις της εμπιστοσύνης ήταν κάτω από το κατώφλι εμπιστοσύνης (και μάλιστα στην συντριπτική πλειοψηφία σημαντικά χαμηλότερα) με αποτέλεσμα να μην μπορεί να περάσει καμιά προσπάθεια για αλληλεπίδραση είτε καλοπροαίρετη είτε κακόβουλη. Το γεγονός επιβεβαιώνει και ο δείκτης *Trusting index*, που διαμορφώθηκε σε μια πολύ χαμηλή τιμή, κοντά στο 0.

Εικόνα 6-11 : PeerTrust - Αποτελέσματα 4<sup>ου</sup> σεναρίου

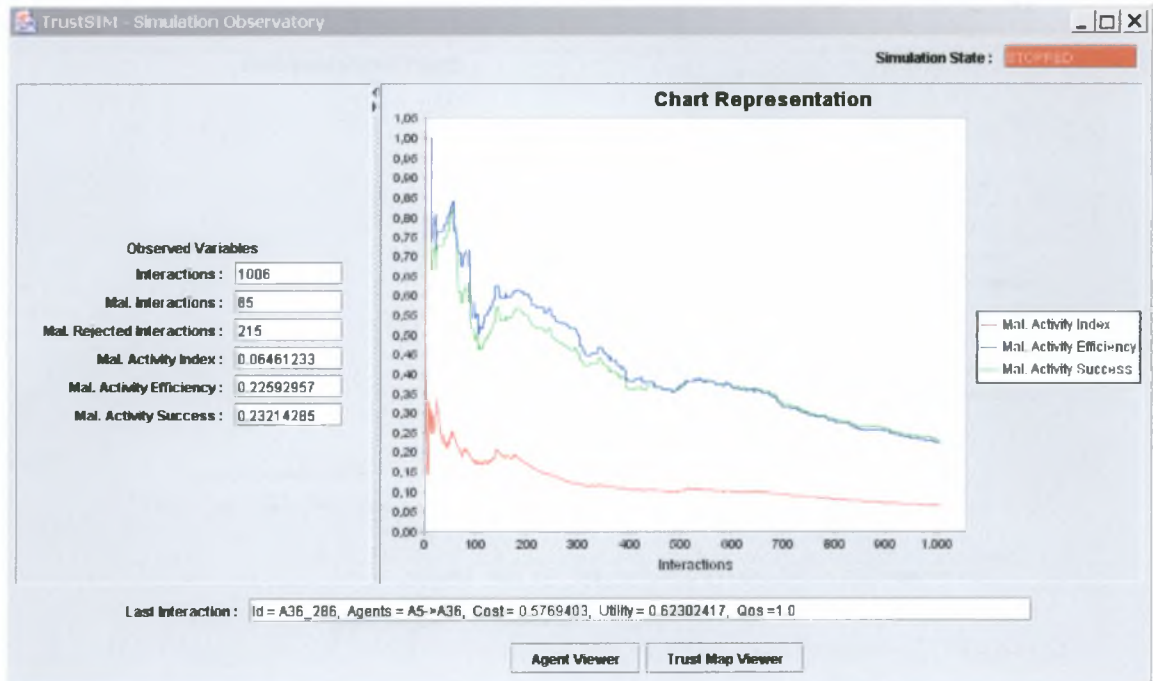


Εικόνα 6-12 : PeerTrust - Τμήμα του χάρτη εμπιστοσύνης του 4<sup>ου</sup> σεναρίου

- **Σενάριο 5** : Αντιμετώπιση κακόβουλων συμπεριφορών με χρήση κοινωνικών μηχανισμών

**Παρατηρήσεις - Σχολιασμός**

Τα αποτελέσματα της προσομοίωσης του 5<sup>ου</sup> σεναρίου για το μοντέλο PeerTrust φαίνονται στην εικόνα 6-13.



Εικόνα 6-13 : PeerTrust - Αποτελέσματα 5<sup>ου</sup> σεναρίου



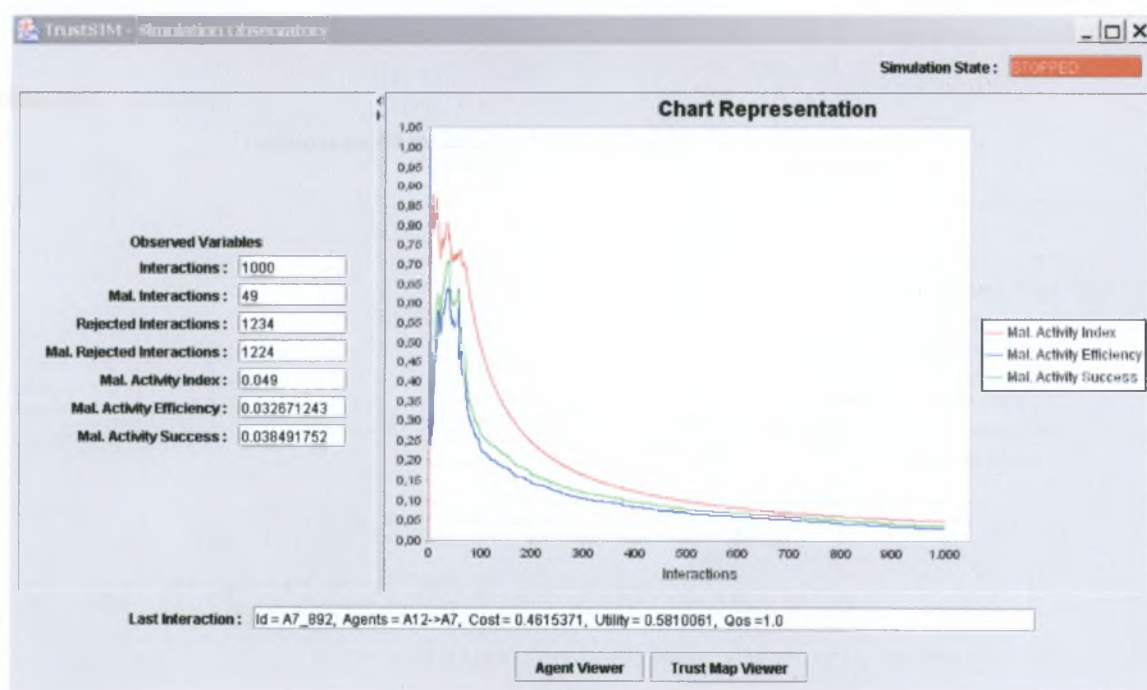
Δυστυχώς στο σενάριο αυτό δεν υπάρχει η δυνατότητα σύγκρισης με το προηγούμενο καθώς το 4<sup>ο</sup> σενάριο δεν ολοκληρώθηκε (πραγματοποιήθηκαν μόνο 270 αλληλεπιδράσεις). Μπορούμε όμως να χρησιμοποιήσουμε ως μέτρο σύγκρισης τις αντίστοιχες επιδόσεις του βασικού μοντέλου. Σε σχέση λοιπόν με το βασικό μοντέλο, το PeerTrust εμφανίζει παρόμοια ικανότητα άμυνας. Αν και εκ πρώτης όψεως φαίνεται να διενεργήθηκαν περισσότερες κακόβουλες αλληλεπιδράσεις, μια προσεκτικότερη παρατήρηση δείχνει ότι οι προσπάθειες που επιχειρήθηκαν ήταν επίσης περισσότερες. Αυτό αποδεικνύεται και από το ποσοστό επιτυχίας των κακόβουλων προσπαθειών (*Malicious activity success*) το οποίο κυμαίνεται σε παρόμοια επίπεδα και στις δύο περιπτώσεις.

- **Σενάριο 6** : Αντιμετώπιση έντονα κακόβουλων συμπεριφορών με χρήση κοινωνικών μηχανισμών

### Παρατηρήσεις - Σχολιασμός

Στο σενάριο αυτό, στο οποίο προσομοιώθηκε ένα έντονα κακόβουλο περιβάλλον, το μοντέλο PeerTrust έδειξε να τα πηγαίνει εξαιρετικά. Συγκεκριμένα, το πλήθος των κακόβουλων αλληλεπιδράσεων ήταν μικρότερο από το προηγούμενο σενάριο, παρά την εντονότερη κακόβουλη δραστηριότητα. Η επίδοση αυτή ήρθε ως αποτέλεσμα της εντυπωσιακής απόδοσης της αμυντικής λειτουργίας με τον πολύ μεγάλο αριθμό των προσπαθειών για κακόβουλες αλληλεπιδράσεις που απορρίφθηκαν. Μάλιστα η αμυντική λειτουργία του μοντέλου τα πήγε καλά από πολύ νωρίς (περίπου από την 100<sup>η</sup> αλληλεπίδραση και μετά).

Τα αποτελέσματα της προσομοίωσης του 6<sup>ου</sup> σεναρίου για το μοντέλο PeerTrust φαίνονται στην εικόνα 6-14 .



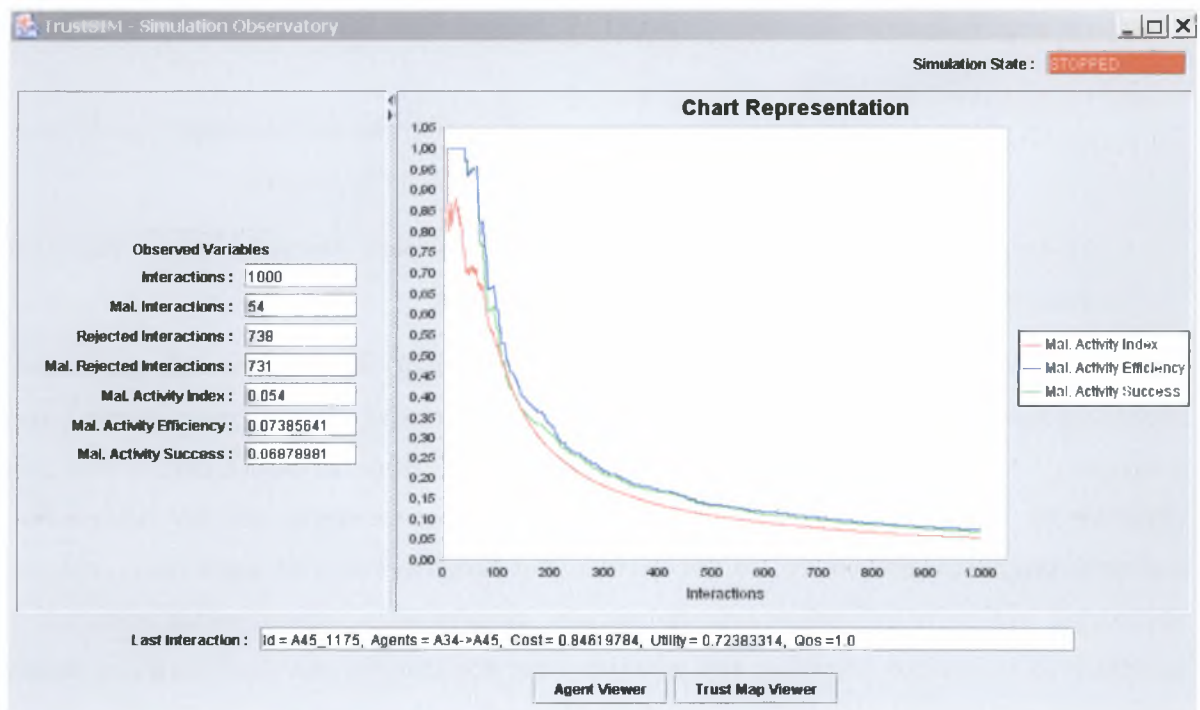
Εικόνα 6-14 : PeerTrust - Αποτελέσματα 6<sup>ου</sup> σεναρίου

Σε σύγκριση με το βασικό μοντέλο, η αμυντική λειτουργία του PeerTrust είναι σημαντικά καλύτερη σ' αυτό το περιβάλλον. Όλες οι παράμετροι έδωσαν καλύτερες (χαμηλότερες) τιμές. Υπήρξε ένα μικρό μόνο τίμημα με την απόρριψη ενός μικρού σχετικά αριθμού (συνολικά 10) προσπαθειών για καλοπροαίρετες αλληλεπιδράσεις, όταν το βασικό μοντέλο δεν παρουσίασε κανένα πρόβλημα στον τομέα αυτό.

- **Σενάριο 7** : Αντιμετώπιση έντονα κακόβουλων συμπεριφορών με χρήση κοινωνικών μηχανισμών και αυστηρά κριτήρια επιλογής αλληλεπιδράσεων

### Παρατηρήσεις - Σχολιασμός

Με τη χρήση υψηλότερου κατωφλίου εμπιστοσύνης το μοντέλο PeerTrust δεν παρουσίασε καμιά διαφοροποίηση στις επιδόσεις του. Όπως φαίνεται και από τα αποτελέσματα της προσομοίωσης (εικόνα 6-15) όλοι οι δείκτες κυμάνθηκαν στα ίδια επίπεδα με την προσομοίωση του 6<sup>ου</sup> σεναρίου. Η μόνη αισθητή διαφορά είναι ο σημαντικά μικρότερος αριθμός των αλληλεπιδράσεων που απορρίφθηκαν. Η αμυντική λειτουργία του μοντέλου φαίνεται να αποθάρρυνε την εκδήλωση κακόβουλων ενεργειών. Οι μικρές αποκλίσεις στις τιμές των μεταβλητών παρακολούθησης είναι μέσα στα φυσιολογικά όρια και οφείλονται στην αυτόνομη και τυχαία δραστηριοποίηση των πρακτόρων.



**Εικόνα 6-15 : PeerTrust - Αποτελέσματα 7<sup>ου</sup> σεναρίου**

Σε σύγκριση με το βασικό μοντέλο, το PeerTrust εμφανίζει σημαντικά μικρότερο αριθμό κακόβουλων αλληλεπιδράσεων. Αυτός είναι και ο λόγος που οι δείκτες κακόβουλης συμπεριφοράς είναι ελαφρά μειωμένοι, υποδηλώνοντας μια σχετικά καλύτερη αμυντική λειτουργία.

## 6.4 Συμπεράσματα

Από τις προσομοιώσεις των δύο μοντέλων εμπιστοσύνης προκύπτουν ενδιαφέροντα συμπεράσματα. Συγκεκριμένα :

- Δεν υπάρχει καθολική υπεροχή κάποιου από τα δύο μοντέλα αφού καθένα έχει δυνατά και αδύνατα σημεία.
- Το βασικό μοντέλο πλεονεκτεί :
  - Στη διαμόρφωση κλίματος εμπιστοσύνης σε κοινωνίες με κακόβουλα μέλη, χωρίς όμως τη χρήση κοινωνικών μηχανισμών
  - Στην εγκυρότητα του μηχανισμού εκτίμησης της αξιοπιστίας
- Το μοντέλο PeerTrust πλεονεκτεί :
  - Στη διαμόρφωση κλίματος εμπιστοσύνης τόσο σε κοινωνίες χωρίς κακόβουλα μέλη όσο και σε κοινωνίες με κακόβουλα μέλη, με τη χρήση όμως, κοινωνικών μηχανισμών
  - Στην αμυντική λειτουργία, περιορίζοντας πιο αποτελεσματικά τις κακόβουλες συμπεριφορές.
- Το μοντέλο PeerTrust εμφανίζει τα παρακάτω προβλήματα :
  - Είναι πολύ βαρύ υπολογιστικά
  - Η αμυντική λειτουργία μπορεί να παρεμποδίσει κάποιο ποσοστό καλοπροαίρετης δραστηριότητας
  - Ο υπολογισμός της εμπιστοσύνης μπορεί να οδηγήσει στη διαμόρφωση πολύ χαμηλών τιμών εμπιστοσύνης ακόμη και για τα έντιμα μέλη της κοινωνίας
- Η ενεργοποίηση των κοινωνικών μηχανισμών επέδρασε θετικά και στα δύο μοντέλα βελτιώνοντας σημαντικά την αποτελεσματικότητά τους

Τα παραπάνω συμπεράσματα αφορούν τη συγκριτική αξιολόγηση των μοντέλων που προσομοιώθηκαν. Πέραν αυτών όμως προέκυψε και ένα γενικότερο και πολύ πιο σημαντικό, κατά τη γνώμη μας, συμπέρασμα. Ότι το περιβάλλον του προσομοιωτή μπορεί να αποτελέσει ένα πολύ χρήσιμο εργαλείο για τη μελέτη των χαρακτηριστικών μοντέλων εμπιστοσύνης και την αξιολόγηση της απόδοσής τους. Όπως φάνηκε από τα πειραματικά αποτελέσματα των προσομοιώσεων, εκτός από τη δυνατότητα πειραματικής επαλήθευσης θεωρητικών εκτιμήσεων, ο προσομοιωτής επιτρέπει στους μελετητές να εντοπίσουν αδυναμίες των μοντέλων τους που ίσως να μην είναι ορατές σε θεωρητικό πεδίο και οι οποίες αποκαλύπτονται μέσα από τη σύγκριση με άλλα μοντέλα.

Οι προσομοιώσεις που διενεργήθηκαν και παρουσιάστηκαν στο κεφάλαιο αυτό αποτελούν ένα πολύ μικρό δείγμα των διαφορετικών προσομοιώσεων που μπορεί να υποστηρίξει ο προσομοιωτής. Ενδεικτικά αναφέρουμε ότι θα μπορούσαμε να δοκιμάσουμε πολλά επιπλέον σενάρια για τη μελέτη της επίδρασης, στη συμπεριφορά των μοντέλων, του κάθε κοινωνικού μηχανισμού ατομικά, όπως επίσης

και της επίδρασης των παραμέτρων που παρέμειναν σταθερές κατά τις διενεργηθείσες προσομοιώσεις. Οι συνδυασμοί είναι υπερβολικά πολλοί για να παρουσιαστούν σ' αυτήν την εργασία. Επιφυλασσόμαστε να παρουσιάσουμε περισσότερους σε μελλοντικές εργασίες, μετά και την κωδικοποίηση περισσότερων μοντέλων εμπιστοσύνης.



# ΚΕΦΑΛΑΙΟ

## 7. Συμπεράσματα

Η διατριβή αυτή ασχολήθηκε με το πρόβλημα της πειραματικής μελέτης και αξιολόγησης των μοντέλων εμπιστοσύνης. Η συγκεκριμένη ερευνητική περιοχή δεν έχει λάβει ακόμη τη δέουσα προσοχή, παρά το εύρος και τη σπουδαιότητα των υπολογιστικών εφαρμογών που χρησιμοποιούν την εμπιστοσύνη για την υποστήριξη της λειτουργικότητάς τους. Η έρευνα που διεξάγεται στο χώρο της εμπιστοσύνης φαίνεται ότι επικεντρώνεται σχεδόν αποκλειστικά στην ανάπτυξη νέων μοντέλων, αγνοώντας, προς τα παρόν, την ανάγκη ύπαρξης ενός κοινού πλαισίου για την αξιολόγησή τους. Μέσα από την δουλειά αυτή προσπαθήσαμε να καλύψουμε αυτό το κενό, προτείνοντας μια κοινή πλατφόρμα αξιολόγησης μοντέλων εμπιστοσύνης. Η πρόταση αυτή ευελπιστούμε πως θα βοηθήσει στην κατεύθυνση της διεξαγωγής μιας ευρύτερης συζήτησης για την ανάπτυξη νέων μοντέλων εμπιστοσύνης και την αποτίμησή τους.

Στο κεφάλαιο αυτό συνοψίζεται η συνεισφορά της διατριβής, καταγράφονται οι σημαντικότερες δυσκολίες που αντιμετωπίσαμε κατά την εκπόνησή της και παρουσιάζονται οι κύριες κατευθύνσεις για μελλοντική δουλειά.

### 7.1 Συνεισφορά

Η κύρια συνεισφορά της διατριβής είναι η ανάπτυξη μιας κοινής πλατφόρμας για τη μελέτη και αξιολόγηση μοντέλων εμπιστοσύνης. Η πλατφόρμα προσφέρει ένα παραμετροποιημένο περιβάλλον προσομοίωσης, που επιτρέπει τον πειραματισμό με μοντέλα εμπιστοσύνης σε διάφορες συνθήκες (μέγεθος κοινωνίας, σύνθεση πληθυσμού, χρησιμοποίηση κοινωνικών μηχανισμών κλπ), οι οποίες μπορεί να είναι και διαφορετικές απ' αυτές για τις οποίες αρχικά σχεδιάστηκαν. Μέσω της πλατφόρμας είναι δυνατή η παρακολούθηση της συμπεριφοράς των προσομοιούμενων μοντέλων και η καταγραφή της απόδοσής τους, με τη βοήθεια ενός κατάλληλου γραφικού περιβάλλοντος που απεικονίζει τις μεταβλητές της προσομοίωσης. Τα αποτελέσματα της προσομοίωσης είναι διαθέσιμα on-line (κατά τη διάρκεια της προσομοίωσης), αλλά και σε μορφή αρχείων δεδομένων, ώστε να είναι δυνατή η αξιοποίησή τους και από άλλα εργαλεία (π.χ. στατιστικής επεξεργασίας). Η πλατφόρμα δοκιμάστηκε στην πράξη με την εκτέλεση μιας σειράς προσομοιώσεων πάνω σε δύο μοντέλα εμπιστοσύνης. Τα αποτελέσματα των προσομοιώσεων έδωσαν μια σαφή εικόνα για την αποδοτικότητα των συγκεκριμένων μοντέλων, υποδεικνύοντας τα δυνατά και αδύναμα σημεία του κάθε μοντέλου. Με τον τρόπο αυτό τεκμηριώθηκε ότι η πλατφόρμα είναι λειτουργική και χρήσιμη.

Μια δεύτερη σημαντική συνεισφορά της διατριβής είναι ανάπτυξη μιας σειράς μετρικών για την αξιολόγηση της απόδοσης των προσομοιούμενων μοντέλων εμπιστοσύνης. Οι μετρικές είναι σχεδιασμένες ώστε να χρησιμοποιούν γενικές μεταβλητές που υπάρχουν σε όλα τα μοντέλα

εμπιστοσύνης και να αξιολογούν την αποτελεσματικότητα με την οποία ένα μοντέλο επιτυγχάνει τους γενικούς του στόχους. Με τον τρόπο αυτό οι μετρικές :

- είναι αντικειμενικές
- μπορούν να αξιολογήσουν οποιοδήποτε μοντέλο εμπιστοσύνης
- επιτρέπουν τη σύγκριση των αξιολογούμενων μοντέλων

Η ανάπτυξη των μετρικών αποτελεί μια πρώτη προσπάθεια καθιέρωσης κάποιων κοινών κριτηρίων αξιολόγησης όλων των μοντέλων εμπιστοσύνης. Έτσι δημιουργείται ένα κοινό σύνολο γενικών κριτηρίων αξιολόγησης που γεφυρώνει τις εξειδικευμένες προσεγγίσεις αξιολόγησης που έχουν εμφανιστεί στο παρελθόν και εξομαλύνει την υφιστάμενη ανομοιομορφία και υποκειμενικότητα που επικρατεί στο χώρο της αξιολόγησης των μοντέλων εμπιστοσύνης.

Τρίτη κύρια συνεισφορά της διατριβής είναι ότι προσφέρει, έμμεσα, τη δυνατότητα να τεκμηριώνεται η εφαρμοσιμότητα των μοντέλων εμπιστοσύνης. Ο προσομοιωτής μας είναι ένα πραγματικό πολυπρακτορικό περιβάλλον που βασίζεται στην πλατφόρμα του JADE. Στην ουσία ένα μοντέλο που έχει κωδικοποιηθεί για να εισαχθεί στον προσομοιωτή είναι μια κανονική πολυπρακτορική εφαρμογή. Αυτό πρακτικά σημαίνει ότι η πλατφόρμα παρέχει ένα πολύ ισχυρό τεκμήριο εφαρμοσιμότητας των προσομοιούμενων μοντέλων :

*Εάν ένα μοντέλο μπορεί να αναπαρασταθεί και να προσομοιωθεί στην προτεινόμενη πλατφόρμα αξιολόγησης, τότε είναι εφαρμόσιμο.*

Η ελλιπής ή ασαφής περιγραφή μοντέλων εμπιστοσύνης είναι ένα αρκετά συνηθισμένο φαινόμενο. Οι συγγραφείς, τις περισσότερες φορές, εστιάζουν την προσοχή τους και προτείνουν λύσεις μόνο για μια υπο-περιοχή του προβλήματος της διαχείρισης εμπιστοσύνης. Έτσι όμως, δημιουργείται η αμφιβολία για το κατά πόσο ένα μοντέλο είναι εφαρμόσιμο και πρακτικό. Ο προσομοιωτής μας μπορεί να βοηθήσει να απαντηθεί αυτό το ερώτημα.

Πέρα από την κύρια συνεισφορά της διατριβής αυτής, στους τρεις άξονες που συζητήθηκαν παραπάνω, υπάρχει και μια παράπλευρη συνεισφορά. Η μελέτη των υφιστάμενων μοντέλων εμπιστοσύνης, μας οδήγησε στον προσδιορισμό και τη συστηματική καταγραφή ενός συνόλου μηχανισμών και διαδικασιών, που είναι απαραίτητα ώστε να είναι ένα μοντέλο εμπιστοσύνης πλήρες και υλοποιήσιμο. Η καταγραφή αυτή μπορεί να λειτουργήσει ως μια λίστα προδιαγραφών πάνω στην οποία μπορούν να βασιστούν τα επερχόμενα μοντέλα εμπιστοσύνης προκειμένου να θεωρούνται πλήρη και υλοποιήσιμα.

## 7.2 Δυσκολίες

Για μπορέσει να ολοκληρωθεί η διατριβή χρειάστηκε να αντιμετωπιστούν κάποιες σημαντικές δυσκολίες. Η πρωταρχική και κυριότερη δυσκολία αφορούσε την έλλειψη άλλων σχετικών εργασιών

που θα μπορούσαν να χρησιμοποιηθούν ως οδηγός για την επίλυση των θεμελιωδών ζητημάτων. Η έλλειψη προηγούμενης μελέτης για τα βασικά δομικά στοιχεία των μοντέλων εμπιστοσύνης, όπως επίσης και η έλλειψη κοινών κριτηρίων και μεθοδολογιών αξιολόγησης των μοντέλων εμπιστοσύνης κατέστησε αναγκαία την διεξοδική μελέτη ενός ταχέως αναπτυσσόμενου και ποικιλόμορφου χώρου από τον οποίο απουσιάζει η προτυποποίηση.

Στο κομμάτι της υλοποίησης, η επιλογή να δώσουμε στο περιβάλλον προσομοίωσης τα χαρακτηριστικά ενός πραγματικού πολυπρακτορικού περιβάλλοντος, αύξησε δραματικά την πολυπλοκότητα του κώδικα και των συνθηκών αποσφαλμάτωσης. Έτσι, έπρεπε να αντιμετωπίσουμε την εγγενή πολυπλοκότητα των πολυπρακτορικών συστημάτων που πηγάζει από την παράλληλη εξέλιξη γεγονότων (concurrency) και την αυτόνομη δραστηριοποίηση των πρακτόρων.

### 7.3 Μελλοντική δουλειά

Η ενασχόλησή μας με το πεδίο αυτό μας οδήγησε στον εντοπισμό αρκετών ευκαιριών για περαιτέρω έρευνα. Οι κατευθύνσεις στις οποίες μπορεί να διεξαχθεί μελλοντική δουλειά αφορούν αφενός την αξιοποίηση της πλατφόρμας για την υποστήριξη θεωρητικών αναζητήσεων στο χώρο των μοντέλων εμπιστοσύνης και των κοινωνιών που βασίζονται στην εμπιστοσύνη και αφετέρου την υλοποίηση λειτουργικών βελτιώσεων στο περιβάλλον προσομοίωσης της πλατφόρμας. Ας δούμε τις σημαντικότερες απ' αυτές τις κατευθύνσεις.

Μια πρώτη δυνατότητα μελλοντικής δουλειάς είναι η κωδικοποίηση επιπλέον μοντέλων εμπιστοσύνης και η ενσωμάτωσή τους στην βιβλιοθήκη των διαθέσιμων μοντέλων της πλατφόρμας. Ως εργαλείο, η πλατφόρμα μπορεί να φανεί χρήσιμη και σε άλλες κατηγορίες ερευνητών (πλην των δημιουργών μοντέλων εμπιστοσύνης), όπως για παράδειγμα στους κοινωνιολόγους. Είναι σημαντικό λοιπόν, να είναι διαθέσιμος ένας ικανοποιητικός αριθμός μοντέλων, ώστε να μπορούν και οι μη εξειδικευμένοι χρήστες να διεξάγουν τις μελέτες τους, χωρίς να χρειάζεται να γράψουν κώδικα.

Έχοντας κωδικοποιήσει ένα σημαντικό αριθμό από τα υφιστάμενα μοντέλα της βιβλιογραφίας, επόμενη ενδιαφέρουσα προοπτική αποτελεί η εκπόνηση μιας διεξοδικής μελέτης και σύγκρισης των κωδικοποιημένων μοντέλων, ανά κατηγορία ή και συνολικά, με την εκτέλεση προσομοιώσεων σε διάφορες συνθήκες, στο πνεύμα των δοκιμών που παρουσιάστηκαν στο 5<sup>ο</sup> κεφάλαιο της διατριβής. Η συγκριτική αξιολόγηση μοντέλων εμπιστοσύνης απουσιάζει, μέχρι στιγμής, από τη βιβλιογραφία και θεωρούμε πως θα τύχει ιδιαίτερης προσοχής από την ερευνητική κοινότητα στα επόμενα χρόνια.

Στην παρούσα έκδοση της πλατφόρμας, η εισαγωγή ενός νέου μοντέλου εμπιστοσύνης απαιτεί γνώσεις προγραμματισμού (Java). Θα ήταν εξαιρετικά επιθυμητό εάν μπορούσαμε να περιγράψουμε ένα μοντέλο σε κάποια γλώσσα υψηλού επιπέδου, απλή στο συντακτικό της, η οποία δεν θα απαιτούσε ιδιαίτερες γνώσεις προγραμματισμού. Μια πρώτη, πολύ πρώιμη μορφή αυτής της γλώσσας αποτελεί το συντακτικό των αρχείων διαμόρφωσης μοντέλων, που παρουσιάστηκε στο 4<sup>ο</sup> κεφάλαιο και στο



παράρτημα Β. Σε μια άλλη εκδοχή θα μπορούσε ίσως να δημιουργηθεί ένα μετα-μοντέλο περιγραφής μοντέλων εμπιστοσύνης. Η ευκολία εισαγωγής νέων μοντέλων στην πλατφόρμα είναι προφανές ότι θα ενισχύσει σημαντικά την αποδοχή της από την ευρύτερη ερευνητική κοινότητα.

Μια άλλη σημαντική προοπτική μελλοντικής δουλειάς είναι η ανάπτυξη νέων μετρικών αξιολόγησης. Υπάρχουν πτυχές της απόδοσης των μοντέλων, που στην παρούσα έκδοση των μετρικών δεν αξιολογούνται, όπως για παράδειγμα η καθυστέρηση στη διαμόρφωση της εμπιστοσύνης, οι απαιτήσεις σε χωρητικότητα και υπολογιστική δύναμη κλπ. Η ανάπτυξη νέων μετρικών προς την κατεύθυνση αυτών των χαρακτηριστικών θα προσφέρει επιπλέον δυνατότητες στη πλατφόρμα, βοηθώντας στην πληρέστερη μελέτη των μοντέλων. Παράλληλα, μπορεί να διερευνηθεί η δυνατότητα βελτίωσης των υφιστάμενων μετρικών. Στην τρέχουσα μορφή τους, οι προτεινόμενες μετρικές που σχετίζονται με την κακόβουλη συμπεριφορά επηρεάζονται άμεσα από τη σύνθεση της κοινωνίας και την αναλογία καλοπροαίρετων – κακόβουλων πρακτόρων. Μια πιθανή βελτίωση θα ήταν η ενσωμάτωση στο τύπο των μετρικών κάποιου παράγοντα που θα εξουδετέρωνε αυτή την επίδραση και θα καθιστούσε τις μετρικές ακόμη πιο αντικειμενικές.

Στην παρούσα έκδοση της πλατφόρμας αξιολόγησης, οι πράκτορες που απαρτίζουν την κοινωνία της προσομοίωσης, λειτουργούν εντελώς αυτόνομα και δραστηριοποιούνται με εντελώς τυχαία σειρά. Αυτό έχει ως αποτέλεσμα να μην μπορεί να παραχθεί η ίδια αλληλουχία γεγονότων σε δύο διαφορετικές προσομοιώσεις και συνεπώς να μην μπορεί να επαναληφθεί επακριβώς ένα πείραμα. Θα ήταν ενδιαφέρον να διερευνηθεί εάν α) μπορούμε να δημιουργήσουμε τις κατάλληλες συνθήκες ώστε να είναι εφικτή η ακριβής επανάληψη μιας σειράς γεγονότων σε διαδοχικές προσομοιώσεις<sup>28</sup> και β) κατά πόσο τα αποτελέσματα που θα προκύψουν θα είναι αξιόπιστα.

Μια επιμέρους προοπτική μελλοντικής δουλειάς είναι και η αξιοποίηση των αποτελεσμάτων των μελετών που θα διεξαχθούν στα υφιστάμενα μοντέλα εμπιστοσύνης, με στόχο τη σύνθεση ενός υπερ-μοντέλου που θα ενσωματώνει τα καλύτερα από τα χαρακτηριστικά που θα έχουν καταγραφεί. Θα ήταν επίσης εξαιρετικά ενδιαφέρον να δούμε (με τη βοήθεια της πλατφόρμας) εάν ένα τέτοιο μοντέλο θα συμπεριφερόταν με βέλτιστο τρόπο σε όλες τις συνθήκες.

Τέλος, υπάρχει αρκετή δουλειά που μπορεί και χρειάζεται να γίνει στην κατεύθυνση της βελτίωσης του περιβάλλοντος του προσομοιωτή. Ενδεικτικά αναφέρουμε :

- Μετατροπή του παρατηρητηρίου της προσομοίωσης (Observatory) σε ανεξάρτητη εφαρμογή που θα διαβάζει και θα απεικονίζει τα δεδομένα της προσομοίωσης από τα αρχεία εξόδου της προσομοίωσης στο δίσκο. Με τον τρόπο αυτό θα είναι εφικτή η παρακολούθηση μιας προσομοίωσης και από τρίτους απομακρυσμένους (πάνω στο web) παρατηρητές.

<sup>28</sup> Αυτό θα μπορούσε για παράδειγμα να επιτευχθεί ρυθμίζοντας τους πράκτορες ώστε να δραστηριοποιούνται βάσει κάποιου συγκεκριμένου σεναρίου.

- Βελτίωση του GUI του προσομοιωτή με ενσωμάτωση οπτικών απεικονίσεων των συνθηκών που διαμορφώνονται στην προσομοιούμενη κοινωνία (φιλίες, αξιοπιστία πρακτόρων, διαδρομές που ακολουθούν οι συστάσεις κλπ)
- Δημιουργία οδηγού (wizard) για την κατασκευή user-defined μετρικών χρησιμοποιώντας λειτουργικές παραμέτρους των μοντέλων και μεταβλητές της προσομοίωσης.

## 7.4 Γενικά συμπεράσματα

Η διατριβή αυτή ασχολήθηκε με το πρόβλημα της αξιολόγησης μοντέλων εμπιστοσύνης, προτείνοντας μια ολοκληρωμένη πλατφόρμα που επιτρέπει την υλοποίηση, τη μελέτη και την αξιολόγηση μοντέλων εμπιστοσύνης, μέσα από παραμετροποιημένες προσομοιώσεις πάνω σε κοινωνίες πρακτόρων. Τα πειραματικά αποτελέσματα που προέκυψαν από τις δοκιμές της πλατφόρμας έδειξαν ότι, μέσω της πλατφόρμας, η αξιολόγηση μοντέλων εμπιστοσύνης :

- είναι εφικτή με ένα ενιαίο και αντικειμενικό τρόπο
- μπορεί να οδηγήσει (με κατάλληλη χρήση γενικών μετρικών) στη σύγκριση των μοντέλων
- μπορεί να βοηθήσει στη βελτίωση των μοντέλων (μέσα από την εντοπισμό των καλύτερων τεχνικών διαχείρισης της εμπιστοσύνης, αλλά και των αδυναμιών του κάθε μοντέλου)

Η προτεινόμενη πλατφόρμα έχει δύο σημαντικά πλεονεκτήματα :

1. Το περιβάλλον προσομοίωσης είναι ένα πραγματικό πολυπρακτορικό περιβάλλον στο οποίο οι συμμετέχοντες πράκτορες λειτουργούν εντελώς αυτόνομα βάσει της λογικής του προσομοιούμενου μοντέλου
2. Η κωδικοποίηση ενός μοντέλου και η εισαγωγή του στην πλατφόρμα είναι μια σχετικά απλή προγραμματιστική διαδικασία που απαιτεί βασική γνώση της Java και μια πολύ μικρή εξοικείωση με το χρησιμοποιούμενο από την πλατφόρμα API του JADE.

Κλείνοντας, ελπίζουμε ότι η διατριβή αυτή θα συνεισφέρει στην ενίσχυση του ενδιαφέροντος της επιστημονικής κοινότητας για την αξιολόγηση των μοντέλων εμπιστοσύνης και να συμβάλει ουσιαστικά στη συζήτηση που έχει ξεκινήσει στη συγκεκριμένη ερευνητική περιοχή. Στην κατεύθυνση αυτή, αναμένουμε με ιδιαίτερο ενδιαφέρον και τη δημοσιοποίηση των αποτελεσμάτων του project *Agent Reputation and Trust (ART) Testbed* [20], προκειμένου να διερευνήσουμε την δυνατότητα αλληλοσυμπλήρωσης και σύγκλισης σε μια καθολικά αποδεκτή προσέγγιση. Θεωρούμε ότι η ανάπτυξη αξιόπιστων μεθοδολογιών και συστημάτων αξιολόγησης μοντέλων εμπιστοσύνης αποτελεί απαραίτητη προϋπόθεση για την ανάπτυξη νέων, βελτιωμένων, μοντέλων εμπιστοσύνης.



## Βιβλιογραφία

1. S. Marsh, "*Formalizing Trust as a Computational Concept*", PhD thesis, University of Stirling, U.K., 1994.
2. D. Gambetta. "*Can We Trust Trust?*", In D. Gambetta, editor, *Trust: Making and Breaking Cooperative Relations*, pages 213-238. Basil Blackwell. Oxford, 1990.
3. D.H. McKnight and N.L. Chervany, "*The Meanings of Trust*", Technical Report MISRC, Working Paper Series 96-04, University of Minnesota, Management Information Systems, Research Center, 1996.
4. N.Luhmann, "*Trust and Power*", Wiley, Chichester, 1979.
5. A.Josang, C.Keser, T.Dimitrakos, "*Can We Manage Trust ?*", In the Proceedings of the 3<sup>rd</sup> International Conference on Trust Management, (iTrust), Paris, May 2005.
6. L.Capra, "*Engineering Human trust in Mobile System Collaborations*", In Proceedings of the 12th International Symposium on Foundations of Software Engineering, pages 107–116, Newport Beach, CA, USA, November 2004. ACM Press.
7. T.Grandison, M.Sloman, "*A Survey of Trust in Internet Applications*", IEEE Communications Surveys, 2000
8. C.Castelfranchi, R.Falcone, "*Principles of Trust for MAS: Cognitive Anatomy, Social Importance and Quantification*", In: Demazeau, Y. (ed.), Proceedings of the Third International Conference on Multi-Agent Systems, IEEE Computer Society, Los Alamitos, pp. 72-79, 1998.
9. M. Blaze, J. Feigenbaum, and J. Lacy, "*Decentralized Trust Management*", IEEE Conf. Security and Privacy, Oakland California, 1996
10. L.Ra smusson, S.Janssen. "*Simulated Social Control for Secure Internet Commerce*", In Catherine Meadows, editor, Proceedings of the 1996 New Security Paradigms Workshop, ACM, 1996
11. A.Josan g, "*The right type of trust for distributed systems*", Proceedings of the 1996 workshop on New security paradigms, 1996
12. Q.Z hang, T.Yu, K.Irwin, "*A Classification Scheme for Trust Functions in Reputation-Based Trust Management*", In ISWC'04, 3rd International Semantic Web Conference, Workshop on "Trust, Security, and Reputation on the Semantic Web", Hiroshima, Japan, 2004
13. Z. Liang, W.Shi, "*Performance Evaluation of Rating Aggregation Algorithms in Reputation Systems*", IEEE Computer Society, CollaborateCom 2005 conference in San Jose, CA, December 2005
14. M.A.Patton, A.Josang, "*Technologies for Trust in Electronic Commerce*", Electronic Commerce Research, 4:9-21, 2004

15. Z. Liang, W. Shi, "PET; A Personalized Trust Model with Reputation and Risk Evaluation for P2P Resource Sharing", Proceedings of the 38<sup>th</sup> Hawaii International Conference on System Sciences, 2005
16. M. Li, H. Liu, L. Cao, J. Yu, Y. Li, Q. Qian, W. Jin, "Semantics and Formalizations of Mission-Aware behaviour Trust Model for Grids", Grid and Cooperative Computing, Second International Workshop, GCC 2003, Shanghai, China, December 2003
17. M. Blaze, J. Feigenbaum, J. Ionnidis, A. Keromytis, "The Role of Trust management in Distributed Systems Security", In Secure Internet Programming: Security Issues for Mobile and Distributed Objects. New York, NY, USA, Springer-Verlag:185 – 210, 1999
18. eBay , <http://www.eBay.com>, 2002.
19. BizRate , <http://www.bizrate.com>, 2002.
20. A RT Testbed, Agent Reputation and Trust Testbed, <http://www.lips.utexas.edu/art-testbed>, 2005.
21. K. Fullam, T. Klos, G. Muller, J. Sabater, A. Schlosser, Z. Topol, K. S. Barber, J. S. Rosenschein, L. Vercouter, and M. Voss, "A Specification of the Agent Reputation and Trust (ART) Testbed", In Proc. AAMAS, 2005.
22. Schlos ser, M. Voss, and L. Bruckner, "Comparing and Evaluating Metrics for Reputation Systems by Simulation", Proc. IAT2004 Workshop on Reputation in Agent Societies, 2004.
23. Z hengqiang Liang and Weisong Shi, "Performance Evaluation of Different Recommendations Aggregation Schemes in Reputation Systems", In Proceedings of the first IEEE International Conference on Collaborative Computing: Networking, Applications, and Worksharing (CollaborateCom '05), San Jose, December 19-21, 2005
24. Josh Boyd, "In Community We Trust : Online Security Communication at eBay", In Journal of Computer-Mediated Communication 7 (3), 2002
25. Paul Resnick, Richard Zeckhauser, "Trust Among Strangers in Internet Transactions : Empirical Analysis of eBay's Reputation System", In M.R. Baye, editor, The Economics of the Internet and E-commerce, volume II of Advances in Applied Microeconomics, Elsevier Sience, 2002
26. Richard J. Zeck hauser and W. Kip. Viscusi, "Risk within Reason", 1990, Science, 248(May 4th), 559–564
27. R. Dechter, & D. Michie December, "Induction of plans", 1984, Tech. rept. TIRM-84-006, The Turing Institute.
28. L. Xi ong and L. Liu, "Peer Trust: A trust mechanism for an open peer-to-peer information system", Technical Report GIT-CC-02-29, Georgia Institute of Technology, College of Computing, 2002

29. Lik Mui, Mojdeh Mohtashemi, Ari Halberstadt, "A Computational Model of Trust and Reputation", Proceedings of the 35th Hawaii International Conference on System Sciences, 2002
30. Ostrom E. Ostrom, "A Behavioral Approach to the Rational-Choice Theory of Collective Action", 1998, American Political Science Review, 92(1), pp. 1-22
31. D. Gambetta, "*Trust: Making and Breaking Cooperative Relations*", Oxford: Basil Blackwell, 1988
32. L. Mui, M. Mohtashemi, C. Ang, P. Szolovits, A. Halberstadt, "*Bayesian Ratings in Distributed Systems: Theories, Models and Simulations*", MIT LCS Memorandum, 2001
33. D. Heckerman, "*A Tutorial on Learning with Bayesian Networks*", Technical Report MSR-TR-95-06, Microsoft Research, 1996
34. Abdul-Rahman and S. Hailes, "*Supporting Trust in Virtual Communities*", Hawaii International Conference on System Sciences, 2000
35. F. Azzedin, M. Maheswaran, "*Evolving and Managing Trust in Grid Computing Systems*", Proceedings of the 2002 IEEE Canadian Conference on Electrical Computer Engineering, 2002
36. B. Yu, M. P. Singh, "*A Social Mechanism of Reputation management in Electronic Communities*", 2000
37. B. Yu, M. Venkatraman, M. P. Singh, "*An Adaptive social network for information access : Theoretical and experimental results*", Applied Artificial Intelligence, 2000
38. B. Yu, M. P. Singh, "*A multiagent referral system for expertise location*", In working notes of the AAAI Workshop on Intelligent Information Systems, pages 66-69, 1999
39. Yao Wang, Julita Vassileva, "*Trust and Reputation Model in Peer-to-peer Networks*", Proc. of IEEE Conference on P2P Computing, 2003
40. Mina xi Gupta, Paul Judge, Mostafa Ammar, "*A Reputation System for Peer-to-peer Networks*", NOSSDAV '03: Proceedings of the 13th international workshop on Network and operating systems support for digital audio and video (2003), pp. 144-152, 2003
41. Sarvapali D. Ramchurn, Nicholas R. Jennings, Carles Sierra, Lluís Godó, "*A Computational Trust Model for Multi-Agent Interactions based on Confidence and Reputation*", In Proceedings of 6th International Workshop of Deception, Fraud and Trust in Agent Societies, pp. 69-75, Melbourne, Australia, 2003
42. Yao Wang, Julita Vassileva, "*Bayesian Network-Based Trust Model in Peer-to-Peer Networks*", Proceedings of the IEEE / WIC International Conference on Web Intelligence (WI'03), Halifax, Canada, pages 372-378, October 2003

43. Paul Resnick, Richard Zeckhauser, "*Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System*", In *The Economics of the Internet and E-Commerce*. Michael R. Baye, editor. Volume 11 of *Advances in Applied Microeconomics*. Amsterdam, Elsevier Science. [online]
44. Jade, <http://jade.tilab.com/>
45. Swarm, <http://www.swarm.org/>
46. Wilensky U., NetLogo. <http://ccl.northwestern.edu/netlogo/>, Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, 1999.
47. Repast, <http://repast.sourceforge.net/>
48. Mason, <http://cs.gmu.edu/~eclab/projects/mason>
49. Ascape, <http://brook.edu/es/dynamics/models/ascape/>

## Παράρτημα Α :

### Οι παράμετροι προσομοίωσης του TrustSIM

Οι παράμετροι που ρυθμίζουν τα χαρακτηριστικά και την πορεία της προσομοίωσης, ανάλογα με τη λειτουργικότητα και την περιοχή της προσομοίωσης που επηρεάζουν, μπορούν να ομαδοποιηθούν σε τρεις κατηγορίες:

- i. Τις παραμέτρους προσομοιούμενου μοντέλου
- ii. Τις παραμέτρους περιβάλλοντος προσομοιωτή
- iii. Τις παραμέτρους σεναρίου προσομοίωσης

Αναλυτικά οι παράμετροι ανά κατηγορία είναι (εμφανίζονται με τα τεχνικά τους ονόματα):

- **Παράμετροι προσομοιούμενου μοντέλου**

#### BEHAVIOUR\_CLASSES

Δομή δεδομένων με τα ονόματα των κλάσεων (Java classes) που υλοποιούν τους μηχανισμούς του προσομοιούμενου μοντέλου. Σε κάθε μηχανισμό-διαδικασία αντιστοιχεί μια JADE συμπεριφορά που προστίθεται στο σύνολο των συμπεριφορών κάθε πράκτορα εμπιστοσύνης. Η δομή περιέχει αντιστοιχίσεις της μορφής (όνομα συμπεριφοράς, όνομα κλάσης).

Τύπος : DataStore

#### SOCIAL\_MECHANISMS

Τα ονόματα των κοινωνικών μηχανισμών (συστάσεις, φήμη κλπ) που χρησιμοποιεί το προσομοιούμενο μοντέλο.

Τύπος : ArrayList (of Strings)

#### INITIAL\_TRUST\_VALUE

Η αρχική τιμή εμπιστοσύνης που αποδίδει ένας πράκτορας σε ένα άγνωστο πράκτορα (για τον οποίο δεν έχει καμιά πληροφορία).

Τύπος : Float

#### MIN\_TRUST\_VALUE, MAX\_TRUST\_VALUE

Η ελάχιστη και η μέγιστη τιμή που μπορεί να πάρει η εμπιστοσύνη ενός πράκτορα.

Τύπος : Float

#### TRUST\_THRESHOLD

Η τιμή εμπιστοσύνης η οποία ορίζει το όριο της αξιοπιστίας. Ένας πράκτορας θεωρείται αξιόπιστος όταν και μόνο όταν η τιμή της εμπιστοσύνης του είναι μεγαλύτερη ή ίση απ' αυτό το όριο.



Τύπος : Float

#### INTERESTS\_HANDLING

Λογικός διακόπτης που υποδεικνύει εάν θα γίνεται διαχείριση ενδιαφερόντων, δηλ. εάν οι πράκτορες θα έχουν ενδιαφέροντα τα οποία θα λαμβάνουν υπόψη κατά την αναζήτηση αλληλεπιδράσεων.

Τύπος : Boolean

#### SITUATION\_HANDLING

Λογικός διακόπτης που υποδεικνύει το αν κατά την διεξαγωγή των αλληλεπιδράσεων θα λαμβάνεται υπόψη η κατάσταση, δηλ. τα ιδιαίτερα χαρακτηριστικά της κάθε αλληλεπίδρασης. Στην τρέχουσα έκδοση του TrustSIM, τα χαρακτηριστικά που ορίζουν την κατάσταση μιας αλληλεπίδρασης είναι το κόστος (cost) και η χρησιμότητά της (utility).

Τύπος : Boolean

- **Παράμετροι σεναρίου προσομοίωσης**

#### TRUST\_MODEL

Το όνομα του προσομοιούμενου μοντέλου, όπως αυτό είναι δηλωμένο μέσα στο αρχείο ρύθμισης μοντέλου (\*.mcf file).

Τύπος : String

#### INTEREST\_VECTOR\_SIZE

Το εύρος του διανύσματος ενδιαφερόντων των πρακτόρων εμπιστοσύνης. Προσδιορίζει το μέγιστο πλήθος των διαφορετικών ενδιαφερόντων που μπορούν να έχει ένας πράκτορας.

Τύπος : Integer

#### MAL\_ACTIVITY\_RATE

Ο ρυθμός εκτέλεσης κακόβουλων (η ανεπιθύμητων) ενεργειών από τους κακόβουλους πράκτορες. Κυμαίνεται από 0.0 (καθόλου κακόβουλες ενέργειες) έως 1.0 (πάντα κακόβουλες ενέργειες).

Τύπος : Float

#### NUM\_OF\_AGENTS

Το πλήθος των πρακτόρων εμπιστοσύνης (καλοί + κακόβουλοι) που θα απαρτίσουν την προσομοιούμενη κοινωνία.

Τύπος : Integer

#### NUM\_OF\_MAL\_AGENTS

Το πλήθος των κακόβουλων πρακτόρων της προσομοιούμενης κοινωνίας

Τύπος : Integer

#### OPTIMISTIC\_AGENTS, PESSIMISTIC\_AGENTS, REALISTIC\_AGENTS

Το ποσοστό μέσα στην κοινωνία των αισιόδοξων, των απαισιόδοξων και των ρεαλιστών πρακτόρων.

Τύπος : Integer

#### MAX\_NUM\_OF\_FRIENDS

Ο μέγιστος αριθμός φίλων που επιτρέπεται να έχει ένας πράκτορας. Στην περίπτωση που υπάρχουν πολλοί που πληρούν τα προσόντα για να θεωρούνται φίλοι, θα πρέπει με κάποιο τρόπο να γίνεται επιλογή ώστε να μην υπερβαίνεται αυτό το μέγιστο όριο φίλων. Η συγκεκριμένη παράμετρος έχει νόημα όταν είναι ενεργοποιημένος ο κοινωνικός μηχανισμός της διαχείρισης φίλων.

Τύπος : Integer

#### MAX\_NUM\_OF\_SUPPLIED\_RECOMMENDATIONS

Ο μέγιστος αριθμός παρεχόμενων συστάσεων. Όταν ζητηθεί από ένα πράκτορα να δώσει συστάσεις για ένα σύνθετο θέμα με πολλές εναλλακτικές (π.χ. για το ποιους πράκτορες θεωρεί καλύτερους σε μια συγκεκριμένη ειδικότητα), αυτός θα πρέπει να περιορίσει το πλήθος των συστάσεων που θα δώσει, ώστε να μην ξεπεραστεί το όριο αυτό.

Τύπος : Integer

#### STYLE\_OF\_QOS\_VALUES

Ένα λεκτικό που προσδιορίζει το στυλ των τιμών για το QoS που ανταλλάσσουν οι πράκτορες κατά τη διεξαγωγή των αλληλεπιδράσεων. Υπάρχουν δύο επιλογές α) δύο συγκεκριμένες τιμές μια για τα καλά QoS και μια για τα κακόβουλα και β) συνεχείς τιμές (πραγματικές) από μια περιοχή τιμών.

Τύπος : String

#### MALICIOUS\_QOS

Η τιμή που αντιπροσωπεύει τα κακόβουλα QoS (δηλ. τις κακές υπηρεσίες ή τις εξαπατήσεις). Η παράμετρος αυτή χρησιμοποιείται όταν έχει επιλεγεί το δυαδικό στυλ τιμών για τα QoS.

Τύπος : Float

#### WELL\_QOS

Η τιμή που αντιπροσωπεύει τα καλά QoS (δηλ. τις καλές υπηρεσίες). Η παράμετρος αυτή χρησιμοποιείται όταν έχει επιλεγεί το δυαδικό στυλ τιμών για τα QoS.

Τύπος : Float

#### MIN\_MALICIOUS\_QOS , MAX\_MALICIOUS\_QOS

Η ελάχιστη και μέγιστη τιμή που μπορεί να πάρει ένα κακόβουλο QoS. Οι παράμετρος αυτοί χρησιμοποιούνται όταν έχει επιλεγεί το συνεχές στυλ τιμών για τα QoS.

Τύπος : Float

#### MIN\_WELL\_QOS, MAX\_WELL\_QOS

Η ελάχιστη και μέγιστη τιμή που μπορεί να πάρει ένα καλό QoS. Οι παράμετρος αυτοί χρησιμοποιούνται όταν έχει επιλεγεί το συνεχές στυλ τιμών για τα QoS.

Τύπος : Float

#### MAX\_NUM\_OF\_INTERACTIONS

Το μέγιστο πλήθος αλληλεπιδράσεων που θα λάβουν χώρα κατά την προσομοίωση. Μετά την ολοκλήρωση του συγκεκριμένου αριθμού αλληλεπιδράσεων η προσομοίωση τερματίζεται αυτόματα.

Τύπος : Integer

#### MAX\_WAITING\_TIME

Ο μέγιστος χρόνος αναμονής (σε milliseconds) για την άφιξη απάντησης σε μια πρόσκληση (για αλληλεπίδραση ή παροχή συστάσεων). Όσο μεγαλύτερη είναι η κοινωνία τόσο μεγαλύτερος πρέπει να είναι ο χρόνος αυτός αφού η CPU μοιράζεται σε περισσότερες processes και η ταχύτητα λειτουργίας των πρακτόρων μειώνεται.

Τύπος : Integer

#### MEMORY\_DEPTH

Το μέγεθος της μνήμης των πρακτόρων. Η παράμετρος αυτή εξομοιώνει το γνώρισμα της λήθης που εμφανίζουν οι πραγματικές κοινωνίες. Έτσι, ένας πράκτορας δε μπορεί να "θυμάται" (δηλ. να κρατά στις εσωτερικές του δομές) όλες τις πληροφορίες που συλλέγει, αλλά μόνο τις πιο πρόσφατες.

Τύπος : Integer

### • Παράμετροι περιβάλλοντος προσομοιωτή

#### CHART\_DATA\_SERIES\_MAX\_SIZE

Το μέγιστο πλήθος στοιχείων που απεικονίζονται ταυτόχρονα στον άξονα X του γραφήματος των μεταβλητών παρακολούθησης. Με την παράμετρο αυτή ουσιαστικά ρυθμίζεται η πυκνότητα της απεικονιζόμενης πληροφορίας στο γράφημα.

Τύπος : Integer

#### CHART\_HORIZONTAL\_AXIS\_VARIABLE

Το όνομα της μεταβλητής παρακολούθησης που αναπαρίσταται στον άξονα X του γραφήματος των μεταβλητών παρακολούθησης.

Τύπος : String

#### CHARTED\_VARIABLES

Τα ονόματα των μεταβλητών παρακολούθησης που αναπαρίστανται στον άξονα Y του γραφήματος του Observatory window.

Τύπος : ArrayList (of Strings)

#### DEBUGGING\_INFO

Λογικός διακόπτης που καθορίζει εάν θα παρέχονται τεχνικές πληροφορίες που βοηθούν στον έλεγχο σφαλμάτων.

Τύπος : Boolean

#### EVENT\_DISPLAYING\_LEVEL

Το επίπεδο λεπτομέρειας των πληροφοριών που αναφέρει κάθε πράκτορα εμπιστοσύνης και οι οποίες εμφανίζονται στην κονσόλα. Είναι διαθέσιμα 4 επίπεδα λεπτομέρειας (από το λιγότερο προς το περισσότερο λεπτομερειακό) :

- α) Καμιά πληροφορία
- β) Κάθε αλλαγή κατάστασης του πράκτορα
- γ) Κάθε διακριτό συμβάν στη ζωή του πράκτορα
- δ) Κάθε διακριτό βήμα κατά την εκτέλεση κάθε συμπεριφοράς

Τύπος : Integer

#### EVENT\_LOGGING\_LEVEL

Το επίπεδο λεπτομέρειας των πληροφοριών που αναφέρει κάθε πράκτορα εμπιστοσύνης και οι οποίες καταγράφονται στο δίσκο. Είναι διαθέσιμα τα ίδια 4 επίπεδα λεπτομέρειας της προηγούμενης παραμέτρου.

Τύπος : Integer

#### FOCUSED\_AGENT

Το όνομα του πράκτορα πάνω στον οποίο εστιάζεται η παρακολούθηση των συμβάντων. Εάν έχει καθοριστεί τέτοιος πράκτορας, μόνο οι δικές του αναφορές εμφανίζονται στην κονσόλα και καταγράφονται στο δίσκο.

Τύπος : String

#### OBSERVED\_VARIABLES

Τα ονόματα των μεταβλητών που θα παρακολουθούνται στο παρατηρητήριο της προσομοίωσης.

Τύπος : ArrayList (of Strings)

#### OUTPUT\_FILENAME

Το όνομα των αρχείων εξόδου (\*.out, \*.log, \*.err)

Τύπος : String

#### OUTPUT\_FILEPATH

Η διαδρομή των αρχείων εξόδου

Τύπος : String

#### REPORT\_VALUES

Τα ονόματα των στοιχείων της προσομοίωσης των οποίων τις μεταβολές θα πρέπει να στέλνουν οι πράκτορες στον κεντρικό πράκτορα προκειμένου να γίνει η ενημέρωση των μεταβλητών παρακολούθησης. Τέτοια στοιχεία είναι για παράδειγμα τα χαρακτηριστικά της τελευταίας ολοκληρωμένης συναλλαγής η τιμή της τελευταίας αξιολόγησης εμπιστοσύνης κλπ. Τα στοιχεία αυτά είναι κατά κάποιο τρόπο τα πρωτογενή υλικά τα οποία παράγουν τις τιμές των μεταβλητών παρακολούθησης.

Τύπος : ArrayList (of Strings)

## Παράρτημα Β :

### Τα αρχεία ρύθμισης μοντέλου (model configuration files)

Σ' ένα αρχείο ρύθμισης μοντέλου καταγράφονται όλες οι ρυθμίσεις των παραμέτρων της κατηγορίας "Παράμετροι προσομοιούμενου μοντέλου". Οι παράμετροι αυτές προσδιορίζουν τα ιδιαίτερα λειτουργικά χαρακτηριστικά ενός μοντέλου εμπιστοσύνης και χρησιμοποιούνται από το TrustSIM για την επιλογή των κατάλληλων διαδικασιών που προσομοιώνουν με μεγαλύτερη πιστότητα το συγκεκριμένο μοντέλο.

Τα περιεχόμενα ενός αρχείου ρύθμισης μοντέλων έχουν τη μορφή απλού κειμένου που ακολουθεί ένα συντακτικό παρόμοιο με αυτό των αρχείων \*.ini των windows. Δηλ. υπάρχουν ρυθμίσεις εκφρασμένες σαν προτάσεις της μορφής <μεταβλητή> = <τιμή>, οι οποίες οργανώνονται σε ενότητες (sections) ανάλογα με την πληροφορία που παρέχουν ή τη λειτουργική περιοχή την οποία ρυθμίζουν. Ένα παράδειγμα αρχείου ρύθμισης μοντέλου φαίνεται στην εικόνα Β-1

```
[Model Name]
Peer Trust

[Social Mechanisms]

[Operational Parameters]
Interests Handling = Yes
Situation Handling = No

[Behaviour Classes]
HandleInfoRequest = PT_HandleInfoRequestBehaviour
EvaluateAgent = PT_EvaluateAgentBehaviour
UpdateTrustData = PT_UpdateTrustDataBehaviour
```

Εικόνα Β-1 : Παράδειγμα αρχείου ρύθμισης μοντέλου

Πιο συγκεκριμένα, σε ένα αρχείο ρύθμισης μοντέλου μπορούμε να διακρίνουμε μέχρι και 4 διαφορετικές ενότητες ρυθμίσεων. Η αρχή κάθε ενότητας σηματοδοτείται από το όνομά της περιστοιχισμένο από αγκύλες (π.χ. [Social Mechanisms]) ενώ το τέλος της υποδεικνύεται από την αρχή τη επόμενης ενότητας. Η σειρά εμφάνισης των ενοτήτων μέσα στο αρχείο δεν έχει σημασία, ενώ μια ενότητα επιτρέπεται να εμφανιστεί και περισσότερες από μια φορές μέσα στο αρχείο. Τα ονόματα των ενοτήτων που αναγνωρίζονται από το TrustSIM, καθώς και το είδος των ρυθμίσεων που περιλαμβάνουν, φαίνονται στον παρακάτω πίνακα :

Ενότητα	Περιεχόμενο - Ρυθμίσεις
Model Name	Το όνομα του μοντέλου με το οποίο θα εμφανίζεται το μοντέλο μέσα στη λίστα των διαθέσιμων για προσομοίωση μοντέλων του TrustSIM
Social Mechanisms	Οι κοινωνικοί μηχανισμών που χρησιμοποιεί το μοντέλο. Τα τεχνικά ονόματα των διαθέσιμων μηχανισμών είναι : FRIENDS_MANAGEMENT RECOMMENDATION_MANAGEMENT REPUTATION_MANAGEMENT GOSSIP
Operational Parameters	Ρυθμίσεις των παραμέτρων λειτουργίας του μοντέλου όπως π.χ. το εύρος τιμών της αναπαράστασης της εμπιστοσύνης ή το αν λαμβάνονται υπόψη οι περιστάσεις (situation /context) κλπ.
Behaviour Classes	Τα ονόματα των Java κλάσεων (αρχεία *.class) που υλοποιούν τις διάφορες συμπεριφορές που πρέπει να επιδεικνύουν οι πράκτορες εμπιστοσύνης της προσομοιούμενης κοινωνίας.

## Παράρτημα Γ :

### Τα μηνύματα επικοινωνίας των πρακτόρων του TrustSIM

Οι πράκτορες του TrustSIM επικοινωνούν μεταξύ τους ανταλλάσσοντας μηνύματα τύπου ACL. Τα χαρακτηριστικά και ο σκοπός των μηνυμάτων αυτών συνοψίζονται στον παρακάτω πίνακα :

Από	Προς	FIPA Performative	Περιεχόμενο	Σκοπός
Trust agent	Trust agent	PROPOSE	Ένα record με τα χαρακτηριστικά της περίπτωσης	Προσκαλεί ένα πράκτορα σε αλληλεπίδραση. Στο περιεχόμενο του μηνύματος τοποθετούνται τα χαρακτηριστικά της περίπτωσης της αλληλεπίδρασης (κατηγορία, κόστος, ωφελιμότητα)
Trust agent	Trust agent	ACCEPT_PROPOSAL REJECT_PROPOSAL	To string "Interaction"	Ενημερώνει για την αποδοχή ή την απόρριψη μιας πρότασης για αλληλεπίδραση
Trust agent	Trust agent	INFORM	Η τιμή του QoS της αλληλεπίδρασης	Παρέχει σε ένα πράκτορα την αξιολόγηση της αλληλεπίδρασης που έλαβε χώρα.
Trust agent	Trust agent	PROPAGATE	Μια λίστα με συστάσεις	Παρέχει σε κάποιο πράκτορα μια απρόκλητη σύσταση (υπό τη μορφή κουσομπολιού)
Trust agent	Trust agent	QUERY_REF	Ένα record με τα στοιχεία της αίτησης για παροχή συστάσεων	Ζητά από ένα πράκτορα συστάσεις. Το είδος της αιτούμενης σύστασης περιέχεται σε ένα record μαζί με τα υπόλοιπα στοιχεία της αιτούμενης σύστασης. Οι συστάσεις που υποστηρίζονται από την παρούσα έκδοση του TrustSIM είναι : <ul style="list-style-type: none"><li>- Γενική σύσταση για ένα συγκεκριμένο πράκτορα</li><li>- Ειδική σύσταση για μια συγκεκριμένη ειδικευση ενός συγκεκριμένου πράκτορα</li><li>- Σύσταση για το ποιο πράκτορες είναι καλοί σε μια συγκεκριμένη ειδικευση.</li></ul>



Trust agent	Trust agent	INFORM_REF	Μια λίστα με συστάσεις	Παρέχει σε κάποιο πράκτορα μια σύσταση ως απάντηση σε προηγούμενη αίτηση για παροχή σύστασης.
Trust agent	Central agent	INFORM	Το string της περιγραφής κάποιου συμβάντος	Ενημερώνει τον κεντρικό πράκτορα για κάποιο συμβάν. Το συμβάν προωθείται στο simulation center για να καταγραφεί στο αρχείο συμβάντων (*.log file). Μεταξύ των συμβάντων που αποστέλλονται είναι και οι μεταβολές της κατάστασης του πράκτορα.
Trust agent	Central agent	INFORM	Μια λίστα με δεδομένα από την προσομοίωση	Στέλνει στον κεντρικό πράκτορα δεδομένα της προσομοίωσης που επηρεάζουν τις μεταβλητές παρακολούθησης (π.χ. το ποσοστό της αξιολόγησης εμπιστοσύνης ή φήμης, τα στοιχεία μιας αλληλεπίδρασης κλπ). Επίσης μέσω τέτοιων μηνυμάτων αποστέλλονται στοιχεία για την εσωτερική του κατάσταση του πράκτορα.
Central agent	Trust agent	REQUEST	Ένα από τα strings : "Internal State Reporting" "No Internal State Reporting"	Ζητά από ένα πράκτορα εμπιστοσύνης να αρχίσει ή να σταματήσει την αποστολή στοιχείων για την εσωτερική του κατάσταση. Η αποστολή των στοιχείων αυτών θα γίνεται μόνο όταν λαμβάνει χώρα κάποιο συμβάν που μεταβάλλει την εσωτερική κατάσταση του πράκτορα.
Central agent	Trust agent	INFORM	Ένα από τα strings : "Agent Start : A"" "Agent End : A""	Ενημερώνει ένα πράκτορα για την γέννηση ή το θάνατο ενός άλλου πράκτορα (μέλους της κοινωνίας). Η ενημέρωση αυτή είναι απαραίτητη προκειμένου ο πράκτορας να ενημερώσει τον ατομικό χάρτη των ζωντανών πρακτόρων που διατηρεί στις εσωτερικές του δομές.

Σε κάθε μήνυμα που χρειάζεται απάντηση (π.χ. μήνυμα πρόσκλησης ή αίτησης) τοποθετείται ένας μοναδικός κωδικός συνομιλίας (conversation id). Ο κωδικός αυτός τοποθετείται από τον απαντώντα πράκτορα και στην απάντηση που στέλνει. Με τον τρόπο αυτό εξασφαλίζεται ότι οι πράκτορες αντιστοιχίζουν σωστά τα μηνύματα που στέλνουν με τις απαντήσεις που δέχονται.

