П О Λ Υ Τ Ε Χ Ν Ε Ι Ο  Κ Ρ Η Τ Η Σ

Σ Χ Ο Λ Η  Μ Η Χ Α Ν Ι Κ Ω Ν  Π Α Ρ Α Γ Ω Γ Η Σ  Κ Α Ι  Δ Ι Ο Ι Κ Η Σ Η Σ


# Μ Ε Τ Α Π Τ Υ Χ Ι Α Κ Η  Ε Ρ Γ Α Σ Ι Α  Μ Ε  Θ Ε Μ Α

«Προσέγγιση του προβλήματος ‘το δίλημμα του φυλακισμένου’ με χρήση μεθευρετικών αλγορίθμων εμπνευσμένων απ’ τη φύση»

Εκπόνηση εργασίας: Μπούτσης Παναγιώτης

Επιβλέπων καθηγητής: Μαρινάκης Ιωάννης

“Research is creating new
knowledge”
Neil Armstrong

Many thanks to my brother
And my family

# Περιεχόμενα

## Πίνακας Σχημάτων

# ΚΕΦΑΛΑΙΟ 1 – ΕΙΣΑΓΩΓΗ ΚΑΙ ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

## 1.1 Εισαγωγή

Καθώς ο Δεύτερος Παγκόσμιος Πόλεμος έφτανε στο τέλος του, ο John von Neumann, ένας απ' τους μεγαλύτερους μαθηματικούς εκείνης της εποχής, αν όχι ο μεγαλύτερος, ήταν απασχολημένος με δύο ανεξάρτητα πνευματικά ζητήματα τα όποια όμως θα διαμόρφωναν το υπόλοιπο του 20$^{ου}$ αιώνα. Αυτά ήταν η θεωρία παιγνίων (game theory) και οι αλγόριθμοι. Το 1944 (16 χρόνια μετά το θεώρημα min-max) δημοσίευσε, μαζί με τον Oscar Morgenstern, το "Games and Economics Behavior", θεμελιώνοντας όχι μόνο την θεωρία παιγνίων αλλά και την θεωρία χρησιμοτήτων, καθώς και την μικροοικονομία. Δύο χρόνια μετά έγραψε μια αναφορά σχετικά με το EDVAC ( Electronic Discrete Variable Automatic Computer) εγκαινιάζοντας την εποχή των ψηφιακών υπολογιστών και του λογισμικού τους, καθώς και των αλγορίθμων.

Θα μπορούσε όμως ο von Neumann να προβλέψει πως οι δίδυμες δημιουργίες του θα σύγκλιναν μισό αιώνα αργότερα; Σίγουρα ήταν αρκετά μπροστά από τους σύγχρονους του όσον αφορά την αντίληψη του πως οι υπολογισμοί είναι δυναμικοί, πανταχού παρόντες και εμπλέκονται με την κοινωνία σχεδόν οργανικά. [1]

Επιστρέφοντας στο σήμερα, και μέσα από τη συγκεκριμένη εργασία, θα προσπαθήσουμε να γεφυρώσουμε λίγο ακόμα αυτές τις δύο έννοιες. Από την μία έχουμε την θεωρία παιγνίων και από την άλλη την αχανή ομάδα των αλγορίθμων. Η αποτελεσματική υπολογισιμότητα έχει αναδειχθεί ως ένα πολύ επιθυμητό χαρακτηριστικό των προβλέψεων συμπεριφοράς και μέσα από τα υπολογιστικά μοντέλα αντανακλά τις πραγματικές πτυχές τις αγοράς. Σκόπος μας είναι μέσα από το πρόβλημα που θα προσπαθήσουμε να λύσουμε, να βγάλουμε κάποια συμπεράσματα όσον αφορά την πρόβλεψη των αποφάσεων.

Ένα από τα συνηθέστερα μοντέλα λήψης αποφάσεων και ένα από τα κλασσικότερα προβλήματα της θεωρίας παιγνίων αποτελεί το δίλημμα του φυλακισμένου (prisoner's dilemma). Το πρόβλημα μελετήθηκε ιδιαίτερα την δεκαετία του 80', αλλά περισσότερο εξελιγμένες τεχνικές προσπαθούν να το προσεγγίσουν μέχρι και σήμερα. Το δίλημμα του φυλακισμένου αποτελεί ένα παίγνιο. Ένα παίγνιο το οποίο περιλαμβάνει δύο παίκτες (φυλακισμένους), που πρέπει να κάνουν μια επιλογή ανάμεσα στις δύο εναλλακτικές (συνεργασία – άρνηση) εις βάρος του αντιπάλου και προς όφελος τους. Αν οι παίκτες επαναλάβουν το παίγνιο πάνω από μιά φορά, τότε το πρόβλημα μετατρέπεται στο επαναλαμβανόμενο πρόβλημα του διλήμματος του φυλακισμένου (Iterated Prisoner's Dilemma).

Στόχος της εργασίας είναι να μελετηθεί το επαναληπτικό δίλημμα του φυλακισμένου απ' την οπτική γωνία των εξελικτικών αλγορίθμων που είναι εμπνευσμένοι από τη φύση. Πιο συγκεκριμένα θα αναλύσουμε, θα σχεδιάσουμε και θα εφαρμόσουμε τον αλγόριθμο βελτιστοποίησης σμήνους σωματιδίων και γνωστές παραλλαγές του, έτσι ώστε να προσεγγίσουμε το συγκεκριμένο πρόβλημα. Αυτοί οι αλγόριθμοι αποτελούν μεθόδους επίλυσης που συνδυάζουν διαδικασίες τοπικής αναζήτησης και υψηλότερου επιπέδου στρατηγικές για να δημιουργήσουν μια διαδικασία που είναι ικανή να ξεφύγει από κάποιο τοπικό ελάχιστο. Με αυτό τον τρόπο θα έχουμε βάλει ένα λιθαράκι για το γεφύρωμα αυτών των δύο εννοιών, της θεωρίας παιγνίων και των αλγορίθμων. Οι εμπνευσμένοι απ' τη φύση αλγόριθμοι ενδείκνυνται εφαρμογής σε τομείς όπως: διαχείριση εφοδιαστικής αλυσίδας (προβλήματα δρομολόγησης οχημάτων, χωροθέτησης εγκαταστάσεων, ελέγχου αποθεμάτων, κ.α..), πρόβλεψη ζήτησης, μάρκετινγκ, παραγωγή, λήψη αποφάσεων, κ.α..

## 1.2 Ιστορική Αναδρομή

Το 1838 ένας Γάλλος οικονομολόγος, ο Augustin Cournot, σκέφτηκε να αναλύσει τις μονοπωλιακές καταστάσεις με τρόπο παρόμοιο με τις σύγχρονες μεθόδους τις θεωρίας παιγνίων. Το όνομα του πήρε και το αντίστοιχο μοντέλο (Cournot). Πενήντα χρόνια μετά, ο Άγγλος οικονομολόγος Francis Edgeworth ασχολήθηκε με την εφαρμογή των μαθηματικών στις κοινωνικές επιστήμες. Το 1913 ο Γερμανός μαθηματικός Ernest Zermelo απέδειξε ότι το σκάκι έχει λύση από οποιαδήποτε κατάσταση, γεγονός που ήταν αρκετά προοδευτικό για την εποχή του και την επιστήμη της θεωρίας παιγνίων. Σημαντική εξέλιξη στον τομέα έφερε ο John von Neumann, που απέδειξε ότι μια σημαντική κατηγορία παιγνίων, τα παίγνια μηδενικού αθροίσματος, έχουν πάντα λύση.

Το 1944 οι Neumann και Oskar Morgenstern εξέδωσαν το βιβλίο “Theory of Games & Economic Behavior”. Μέσα απ’ αυτό όρισαν αξιωματικά την θεωρία της χρησιμότητας (utility theory), ανέλυσαν διεξοδικά τις βέλτιστες λύσεις στα παίγνια μηδενικού αθροίσματος και εισήγαγαν μια νέα κατηγορία παιγνίων, τα συνεργατικά παίγνια (cooperative games).

Το 1950 ο John Nash εισήγαγε την έννοια της ισορροπίας, η οποία είναι η πιο ευρέως χρησιμοποιούμενη έννοια στη σύγχρονη θεωρία παιγνίων, παίρνοντας έτσι και το όνομα του. Η έννοια της ισορροπίας Nash εφαρμόζεται και στα παίγνια μη-μηδενικού αθροίσματος. Η εργασία του Nash μπορεί να θεωρηθεί ότι αποτελεί επέκταση της εργασίας του Cournot. Γι αυτήν του την έρευνα βραβεύτηκε και με το βραβείο Nobel στα οικονομικά το 1994. Ουσιαστικά, ισορροπία Nash σε ένα παίγνιο είναι μια κατάσταση από την οποία δεν συμφέρει κανέναν παίκτη να ξεφύγει μεμονωμένα.

Το 1965 ο Reinhard Selten γενίκευσε τις ιδέες του Nash στα δυναμικά παίγνια, δηλαδή σε παίγνια που εξελίσσονται στην πορεία του χρόνου. Σε συνέχεια αυτού, το 1967-1968 ο John Harsanyi γενίκευσε τις ιδέες του Nash σε παίγνια μη-πλήρους πληροφόρησης σχετικά με τις προτιμήσεις και τις αποφάσεις των άλλων παικτών. Οι Selten και Harsanyi μοιράστηκαν, μαζί με τον John Nash, το βραβείο Nobel στα οικονομικά το 1994.

Από την δεκαετία του 60’ και μετέπειτα, η θεωρία παιγνίων είχε αλματώδη ανάπτυξη και άρχισε να εφαρμόζεται σε όλους τους τομείς και τις πολιτικές επιστήμες, ενώ πληθώρα ερευνητικών πειραμάτων ξεκίνησαν προσπαθώντας να βρουν λύση σε όλο και περισσότερα προβλήματα.

Τη δεκαετία του 70’ άρχισε να εφαρμόζεται και στον κλάδο της βιολογίας σαν αποτέλεσμα της εργασίας του John Maynard Smith σχετικά με την έννοια της «εξελικτικά σταθερή στρατηγική – evolutionary stable strategy» [6]

Στα τέλη της δεκαετίας του 1990 η θεωρία παιγνίων εφαρμόστηκε στον σχεδιασμό δημοπρασιών. Πάνω σε αυτό ασχολήθηκαν διάφοροι επιστήμονες για την κατανομή δικαιωμάτων χρήσης του ηλεκτρομαγνητικού φάσματος στη βιομηχανία των κινητών τηλεπικοινωνιών.

Το 2005 ο Αμερικάνος επιστήμονας Tomas Schelling και ο Γερμανός θεωρητικός παιγνίων Robert Aumann κέρδισαν το βραβείο Νόμπελ για τις Οικονομικές επιστήμες επειδή εμπλούτισαν την αντίληψη μας σχετικά με τις έννοιες του ανταγωνισμού και της συνεργασίας μέσω της παιγνιοθεωρητικής ανάλυσης. Τους τελευταίους ακολούθησαν το 2007 οι Roger Myerson, Leionid Hurwicz και Eric Maskin για "τη θεμελίωση της θεωρίας σχεδιασμού μηχανισμών". [7]

## 1.3 Θεωρία Παιγνίων

### 1.3.1 Τί είναι η Θεωρία Παγνιών

Η θεωρία παιγνίων ασχολείται με το πώς τα ορθολογικά άτομα λαμβάνουν αποφάσεις όταν είναι αμοιβαία αλληλοεξαρτώμενα. Τα τελευταία χρόνια η θεωρία αυτή εφαρμόζεται όλο και περισσότερο σε διάφορους κλάδους της οικονομίας. Συχνά αυτή η σύνθεση βελτιώνει σημαντικά την κατανόηση των οικονομικών θεμάτων και οδηγεί σε σημαντικές νέες ιδέες που αναπτύσσονται. Σε πολλές περιπτώσεις, η εφαρμογή της θεωρίας παιγνίων έχει μετατρέψει τον τρόπο που οι οικονομολόγοι σκέφτονται τόσο σε μικροοικονομικά όσο και σε μακροοικονομικά προβλήματα. Αυτό αποδεικνύεται από την συχνή χρήση των νέων ιδεών όταν η θεωρία παιγνίων εφαρμόζεται σε διάφορους κλάδους της οικονομίας. Για παράδειγμα, είναι πλέον διαδεδομένο για τους οικονομολόγους να αναφέρονται στην Νέα Βιομηχανία και την Νέα Διεθνή Οικονομία. Αυτές οι δύο περιοχές έχουν αναπτυχθεί ακριβώς επειδή η θεωρία παιγνίων έχει εφαρμοστεί σε παραδοσιακές πτυχές των οικονομικών.

Τα παίγνια είναι μια μέθοδος ανάλυσης προβλημάτων που έχουν σχέση με τον τρόπο λήψης αποφάσεων σε καταστάσεις σύγκρουσης και συνεργασίας. Ο παίκτης μπορεί να είναι ένα πρόσωπο, μια οργάνωση, ένα κράτος ή ένας συνασπισμός. Ως αντικείμενο έρευνας μπορούν να θεωρηθούν διάφορα προβλήματα πολιτικής, ψυχολογικής, κοινωνικής και οικονομικής μορφής.

Για την λύση αυτών, θεωρείται προηγουμένως απαραίτητη η ανάλυση καταστάσεων, όπου δύο ή περισσότεροι δράστες (παίκτες) βρίσκονται αντιμέτωποι και ακολουθούν συνεργατικές στρατηγικές. Κάθε παίκτης προσπαθεί να χρησιμοποιήσει όλα τα μέσα που διαθέτει, για να εμποδίσει τον αντίπαλό του να αποκτήσει πλεονεκτήματα που θα περιορίσουν τα κέρδη του. Επομένως, οι ενέργειες του εξαρτώνται άμεσα από τη θέση (στρατηγική) που θα επιλέξει ο αντίπαλος.

Η θεωρία παιγνίων χαρακτηρίζεται από τρία πράγματα: Τον ατομικισμό, τον ορθολογισμό και την αμοιβαία αλληλοεξάρτηση.

## 1.3.2 Ατομικισμός

Είναι σύνηθες να διαχωρίζουμε την θεωρία παιγνίων σε δύο ξεχωριστές κατηγορίες: Στην συνεργατική και στην μη-συνεργατική θεωρία παιγνίων. Για να κυριολεκτήσουμε, ο προηγούμενος ορισμός της θεωρίας παιγνίων αναφέρεται μόνο για την μη-συνεργατική θεωρία παιγνίων. Στην μη-συνεργατική θεωρία παιγνίων, τα άτομα ή οι παίκτες ενός παιγνίου δεν είναι σε θέση να εισέλθουν σε δεσμευτικές και εκτελεστικές συμφωνίες ο ένας με τον άλλον. Δεδομένης αυτής της υπόθεσης τα μη συνεργατικά παίγνια είναι εγγενώς ατομικιστικά.

Απ' την άλλη μεριά, η συνεργατική θεωρία παιγνίων αναλύει καταστάσεις εφόσον οι συμφωνίες αυτές είναι δυνατές να συμβούν. Το επίκεντρο της συνεργατικής θεωρίας παιγνίων είναι το πως οι ομάδες των ατόμων ή παικτών δεσμεύονται ο ένας στον άλλον για να διατυπώσουν ορθολογικές αποφάσεις.

Η διάκριση αυτή δεν σημαίνει ότι η μη-συνεργατική θεωρία παιγνίων δεν επιτρέπει στα άτομα να συνεργάζονται μεταξύ τους. Ωστόσο αυτό σημαίνει πως κάτι τέτοιο θα συμβεί μόνο αν τα άτομα αντιλαμβάνονται μια τέτοια συνεργασία να είναι για το δικό τους συμφέρον. Από αυτή την άποψη τα άτομα δεν συνεργάζονται επειδή πρέπει αλλά επειδή επιλέγουν οικειοθελώς να το πράξουν.

Το σίγουρο είναι πως, σε πολλές περιπτώσεις, το να θεωρήσουμε ένα παίγνιο συνεργατικό ή μη δεν είναι απόλυτα αποσαφηνισμένο. Για παράδειγμα, σε πολλές περιπτώσεις με πολύπλοκους οργανισμούς, όπως είναι οι επιχειρήσεις, οι κυβερνήσεις και οι χώρες θεωρούνται πως ενεργούν ως μεμονωμένοι ιδύνοντες. Σαφώς αυτό είναι μια ακραία απλούστευση που αγνοεί πως λαμβάνονται οι αποφάσεις μέσα σε αυτούς τους οργανισμούς. Η αξία όμως αυτής της απλούστευσης είναι πως κάνει τα προκύπτοντα μοντέλα περισσότερο προσιτά. [3]

## 1.3.3 Ορθολογισμός

Το δεύτερο χαρακτηριστικό της θεωρίας παιγνίων είναι η υπόθεση πως τα άτομα δρουν απόλυτα λογικά. Αυτό σημαίνει πως τα άτομα προσπαθούν να ενεργούν για το δικό τους καλό. Αυτό προϋποθέτει πως τα άτομα είναι σε θέση να καθορίσουν, τουλάχιστον πιθανολογικά, το αποτέλεσμα των ενεργειών τους και να έχουν προτιμήσεις σε αυτά τα αποτελέσματα. Όπως και με τον ατομικισμό, αυτό το χαρακτηριστικό κυριαρχεί στην νεοκλασσική οικονομία και η αιτιολόγηση έχει επιχειρηθεί σε μια σειρά από τρόπους.

Ο πρώτος λόγος είναι για να υποστηρίξει πως τα άτομα είναι πράγματι λογικά. Ωστόσο, δεδομένης της πολυπλοκότητας των πολλών αποφάσεων και το ποσό των πληροφοριών που συχνά πρέπει να αναλυθούν, δεν φαίνεται και πολύ ρεαλιστικός. Πράγματι, στοιχεία από πολλές πειραματικές μελέτες υποδηλώνουν ότι τα άτομα δεν είναι πλήρως ορθολογικά, και για να επιλύσουν σύνθετα προβλήματα θεσπίζουν απλοϊκούς κανόνες που είναι γενικά αναντίστοιχοι.

Μία δεύτερη αιτιολόγηση για τον ορθολογισμό είναι ότι λόγω της διαδικασίας της φυσικής επιλογής, η οικονομία συγκλίνει τελικά για πλήρως ορθολογικά αποτελέσματα. Από αυτή την οπτική, η υπόθεση της ορθολογικότητας είναι να συνάδει επικεντρώνοντας στην μακροχρόνια ισορροπία της οικονομίας. Για παράδειγμα, υποστηρίζεται ότι, εάν οι επιχειρήσεις υπο-βελτιστοποιηθούν τότε η διαδικασία διαγωνισμού, θα τις αναγκάσει τελικά να εγκαταλείψουν τον κλάδο. Το αποτέλεσμα αυτού είναι πως σε μακροχρόνιες ισορροπίες στις επιχειρήσεις, η βελτιστοποίηση θα πρέπει να είναι πλήρως ορθολογική.

Ωστόσο, ένα πρόβλημα με αυτό το συλλογισμό έιναι ότι, αν και μία τέτοια εξελικτική διεργασία/διαδικασία μπορεί να θεωρηθεί σχετική για συναγωνισμό μεταξύ εταιρειών, δεν είναι πάντα ξεκάθαρο πώς εφαρμόζεται σε άλλες περιστάσεις. Για παράδειγμα, δεν φαίνεται να υπάρχει εξελικτική διαδικασία στην οποία καταναλωτές με χρήση της λογικής να μπορούν να εξαλείψουν καταναλωτές χωρίς λογική. Χωρίς μια τέτοια διαδικασία επιλογής, η οικονομία δε θα συγκλίνει απαραίτητα στο λογικό αποτέλεσμα. Η τελική αιτιολόγηση για τον ορθολογισμό είναι ότι δεν αποσκοπεί στο να περιγράψει το πώς τα άτομα λύνουν σύνθετες αποφάσεις στην πραγματικότητα, απλά γίνεται η μοναδική υπόθεση ότι τα άτομα δρουν σαν να έκαναν πλήρη χρήση της λογικής.

Για άλλη μια φορά, χρησιμοποιείται η υπόθεση του ορθολογισμού για να κάνει τα μοντέλα πιο κατανοήσιμα. Όπως σημειώνεται από τον Friedman (1953), όλες οι θεωρίες πρέπει να περιλαμβάνουν κάποια απλοποίηση, καθώς καμία θεωρία δε μπορεί να περιέχει όλα τα πιθανά χαρακτηριστικά της πραγματικότητας. Σύμφωνα με αυτή τη θετική μεθοδολογία, η υπόθεση του ορθολογισμού δε θα πρέπει να απορριφθεί απλά και μόνο επειδή θεωρείται ότι δεν είναι ρεαλιστική. Και αυτό, γιατί όλες οι απλοποιητικές υποθέσεις έιναι απαραίτητα μη ρεαλιστικές. Αντί αυτού, ο ορθολογισμός θα πρέπει να απορρίπτεται μόνο όταν τα αποτελέσματα που βασίζονται σε αυτή την υπόθεση αποδεικνύονται ότι δεν είναι βοηθητικά. Αυτό θα είναι αληθές είτε εάν η θεωρία δεν επιφέρει σχετικές προβλέψεις είτε εάν οι προβλέψεις αυτές προκύπτουν λανθασμένες από εμπειρικά στοιχεία.

Με αυτή τη μεθοδολογία, μία θεωρία θα πρέπει να κρίνεται σύμφωνα με τη χρησιμότητά της και όχι με βάση τον υποτιθέμενο ρεαλισμό των υποθέσεών της. Γενικά, δίνεται η εντύπωση ότι η θεωρία παιγνίων που βασίζεται στον ορθολογισμό μπορεί να είναι εξαιρετικά χρήσιμη στο να μας βοηθήσει να κατανοήσουμε ένα ευρύ φάσμα οικονομικών θεμάτων. Αυτό, όμως, δε σημαίνει ότι η αποχή από τον πλήρη ορθολογισμό δε θα προσφέρει εξίσου χρήσιμες πληροφορίες και προβλέψεις. [3]

## 1.3.4 Αμοιβαία Αλληλοεξάρτηση

Το τελευταίο χαρακτηριστικό της θεωρίας παιγνίων είναι ότι θεωρεί καταστάσεις όπου τα άτομα είναι αμοιβαία αλληλοεξαρτώμενα. Σε αυτή την κατάσταση, η άρτια/καλή κατάσταση κάθε ενός ατόμου σε ένα παίγνιο καθορίζεται, τουλάχιστον εν μέρει, από τις ενέργειες των υπολοίπων παικτών στο παίγνιο.

Συγκεκριμένα με την αμοιβαία αλληλοεξάρτηση, τα άτομα ή διαγωνιζόμενοι μπορούν πλέον να έχουν το κίνητρο να δράσουν στρστηγικά. Λαμβάνοντας αποφάσεις με στρατηγικό τρόπο, τα άτομα αποβλέπουν στο να αντιληφθούν την επίδραση που θα έχουν οι δικές τους ενέργειες στην συμπεριφορά των άλλων. Με δεδομένη αυτή την προσδοκία, κάθε άτομο καθορίζει την ιδανική απάντηση του για να πετύχει το πιο επιθυμητό αποτέλεσμα.

Σε αντίθεση με τον ατομικισμό και τον ορθολογισμό, αυτό το χαρακτηριστικό αμοιβαίας αλληλοεξάρτησης συναντάται λιγότερο συχνά στη νεοκλασσική οικονομική. Για παράδειγμα, στη Γενική Θεωρία Ισορροπίας, όλα τα άτομα θεωρούνται ότι είναι ατομιστές. Αυτό διαβεβαιώνει ότι οι δράσεις των ατόμων, θεωρούμενες μεμονωμένες, δεν έχουν καμία επίδραση στα αποτελέσματα των αγορών ή στην άρτια κατάσταση των άλλων. Αυτό σημαίνει ότι κανένα άτομο δε μπορεί να βελτιώσει τη θέση του ή την κατάσταση του δίχως να επιβαρύνει τη θέση ή κατάσταση κάποιου άλλου. Αντίθετα, μόλις εισάγεται η αλληλοεξάρτηση, έτσι ώστε η καλή κατάσταση ενός ατόμου να εξαρτάται από τις ενέργειες των άλλων, υπάρχει η πιθανότητα αποτυχίας της αγοράς. Σε αυτη την κατάσταση, τουλάχιστον ένα άτομο μπορεί να καλυτερεύσει τη κατάσταση του χωρίς να επιβαρύνει την κατάσταση οποιουδήποτε άλλου ατόμου. Η πιθανότητα για μια τέτοια αναποτελεσματικότητα έχει επιβεβαιωθεί σε πολυάριθμες οικονομολογικές εφαρμογές της θεωρίας παιγνίων. [3]

# 1.4 Παίγνια

## 1.4.1 Παίγνια μη-σταθερού αθροίσματος

Από την θεωρία γνωρίζουμε πως τα παίγνια χωρίζονται σε παίγνια μηδενικού αθροίσματος, παίγνια γενικού αθροίσματος και παίγνια μη-σταθερού αθροίσματος. Δεδομένου πως το δίλημμα του φυλακισμένου εντάσσεται στην τελευταία κατηγορία, θα αναλύσουμε περισσότερο αυτήν.

Τα περισσότερα πρότυπα παιγνίων που εμφανίζονται σε οικονομικές και άλλες κοινωνικές περιπτώσεις δεν είναι σταθερού ή μηδενικού αθροίσματος, διότι είναι σπάνιο φαινόμενο να υπάρχουν π.χ. οικονομικοί ανταγωνιστές που έχουν ολική σύγκρουση συμφερόντων. Στα παίγνια αυτά παρατηρείται το φαινόμενο να ποικίλει το άθροισμα των αμοιβών ή απωλειών των παικτών. Έτσι, αν $a_{ij}$ είναι η αμοιβή του παίκτη Α και $b_{ij}$ η αμοιβή του παίκτη Β για την επιλογή στρατηγικής i από τον πρώτο και j από τον δεύτερο, έχουμε ότι η ποσότητα $a_{ij} + b_{ij}$

είναι διαφορετική για διαφορετικά ζεύγη στρατηγικών (i,j), σε αντίθεση προς τα παίγνια μηδενικού αθροίσματος, όπου $a_{ij} = - b_{ij}$ για κάθε (i,j) και τα παίγνια σταθερού αθροίσματος όπου $b_{ij} = c - a_{ij}$ για κάθε (i,j). Για τον λόγο αυτό τα παίγνια αυτά ονομάζονται παίγνια μη-σταθερού αθροίσματος. Στα παίγνια αυτά είμαστε αναγκασμένοι να παρουσιάζουμε ρητώς τις αμοιβές αμφοτέρων των παικτών, αφού δεν είναι δυνατόν να συμπεράνουμε τις αμοιβές του ενός από τις αμοιβές του άλλου.

Γενικά, σε ένα παίγνιο με μη-σταθερό και άρα γενικό άθροισμα, υποθέτουμε ότι οι παίκτες είναι πλήρως πληροφορημένοι, γνωρίζουν δηλαδή όλες τις πληροφορίες του παιγνίου. Επειδή τα συμφέροντα των παικτών δεν είναι αναγκαστικά τελείως αντίθετα, έχουμε ουσιαστικά ανταγωνιστές και όχι αντιπάλους. Επίσης, οι παίκτες θα θεωρηθούν ορθολογιστές. Αυτό σημαίνει ότι έκαστος επιδιώκει το ατομικό του συμφέρον. Είναι επίσης εγωκεντρικός αλλά όχι κακεντρεχής. Οι επιλογές δεν γίνονται με σκοπό να προκαλέσουν απώλειες στον ανταγωνιστή του εάν ο ίδιος δεν επιτυγχάνει κάποιο κέρδος απ αυτήν την διαδικασία. Από την άλλη πλευρά η έννοια της αλληλεγγύης του είναι τελείως ξένη. Το συλλογικό συμφέρον δεν τον αγγίζει. Για την επιλογή της τελικής του δράσης, συγκρίνει τις αμοιβές που αφορούν αποκλειστικά αυτόν τον ίδιο. Για το λόγο αυτό, ο ορθολογιστής παίκτης δεν επιλέγει στρατηγική με βάση το αμοιβαίο συμφέρον αλλά μόνον με βάση το ατομικό. Οι ενδεχόμενες τάσεις φιλαλληλίας του παίκτη θα πρέπει να εκφραστούν αριθμητικά στην μήτρα αμοιβής του παιγνίου. Όταν οι προτιμήσεις για τις εκβάσεις του παιγνίου έχουν καταχωρηθεί μ’ αυτόν τον τρόπο, ο αλτρουιστής παίκτης θα μεταμορφωθεί στον εγωκεντρικό ορθολογιστή, που επιδιώκει το ατομικό του συμφέρον.

Στα αυστηρώς ανταγωνιστικά παίγνια δεν είναι δυνατό να επιτευχθεί συμφωνία για συνεργασία μεταξύ των παικτών που να αποβαίνει σε κοινό όφελος. Στα παίγνια μη-σταθερού αθροίσματος, ενώ η συνεργασία μεταξύ των παικτών ενός τέτοιου παιγνίου δύναται να αποβεί εις όφελος και των δύο, συχνά παρατηρείται το φαινόμενο της απαγόρευσης μιας τέτοιας συνεργασίας, π.χ. λόγω νομοθεσίας ή διεθνών συνθηκών περί συντεχνιών ή εναρμονισμένης τιμολογιακής τακτικής. Άλλες πάλι φορές, μια δεσμευτική συμφωνία μεταξύ των παικτών αποτρέπεται λόγω της έλλειψης αμοιβαίας εμπιστοσύνης ή και επικοινωνίας. Γι αυτό το λόγο είμαστε υποχρεωμένοι να ορίσουμε ακριβώς στους κανόνες του παιγνίου αν η συνεργασία, μέσω προκαταρκτικής επικοινωνίας των παικτών, επιτρέπεται ή όχι. Γενικά, θα μπορούσαμε να τα χωρίσουμε σε 2 υποκατηγορίες:

- Μη-συνεργατικά παίγνια ή παίγνια άνευ συνεργασίας (non cooperative games), όπου η επικοινωνία των παικτών πριν την έναρξη του παιγνίου δεν επιτρέπεται και συνεπώς οποιαδήποτε σύμπραξη των παικτών πριν από την έναρξη απαγορεύεται
- Συνεργατικά παίγνια ή παίγνια συνεργασίας (cooperative games), όπου οι παίκτες έχουν το δικαίωμα επικοινωνίας πριν την έναρξη της παρτίδας και συνεπώς επιτρέπεται να προβούν σε όλες τις μορφές σύμπραξης και συνεργασίας που θα είναι και δεσμευτικές

Στα παίγνια μη-σταθερού αθροίσματος συναντούμε στοιχεία συνεργασίας αλλά και ανταγωνισμού. Τα συμφέροντα των παικτών μπορεί να είναι αντίθετα σε κάποια σημεία και συμπληρωματικά σε κάποια άλλα. Στις άλλες δύο κατηγορίες, μηδενικού και σταθερού αθροίσματος, οι παίκτες δεν έχουν τίποτα κοινό στις επιθυμίες τους. Από την άλλη πλευρά, στα πλήρως συνεργατικά παίγνια οι παίκτες έχουν μόνο κοινά συμφέροντα και καμιά αντίθεση. Τα περισσότερα παίγνια μη-σταθερού αθροίσματος διαμορφώνουν τον χώρο μεταξύ των δύο αυτών ακραίων περιπτώσεων.

Στη καθημερινή ζωή συναντούμε πιο συχνά παίγνια που συνδυάζουν τόσο στοιχεία συνεργασίας όσο και ανταγωνισμού και είναι πιο ενδιαφέροντα και πιο περίπλοκα απ' αυτά που είναι πλήρως ανταγωνιστικά ή πλήρως συνεργατικά.

Ο παραπάνω διαχωρισμός σε συνεργατικά και μη συνεργατικά παίγνια, όπως έχει οριστεί, αφορά τους κανόνες που διέπουν το παίξιμο του παιγνίου και ότι την ταυτόχρονη ύπαρξη ανταγωνιστικών και συνεργατικών στοιχείων.

Συνοψίζοντας, μια επιλογή στρατηγικών από κάθε παίκτη, σ' ένα παίγνιο με μη-σταθερό άθροισμα, αντιστοιχεί σε ένα μη-συνεργατικό σημείο ισορροπίας εάν κανένας εκ των παικτών δεν έχει την διάθεση να αλλάξει την επιλογή στρατηγικής του, εάν δηλαδή κανείς από τους παίκτες δεν κερδίζει κάτι παραπάνω αν αλλάξει την στρατηγική του μόνο αυτός. [2]

Τα μη συνεργατικά σημεία ισορροπίας λέγονται και σημεία ισορροπίας Nash.

Το πρόβλημα του διλήμματος των φυλακισμένων αποτελεί το κλασσικότερο παράδειγμα προβλήματος μη-σταθερού αθροίσματος και θα περιγραφεί αναλυτικά στο επόμενο κεφάλαιο.

### 1.4.2 Ισορροπία Nash

Η έννοια της ισορροπίας Nash αποτελέι γενίκευση της ανάλογης έννοιας των παιγνίων μηδενικού αθροίσματος στα παίγνια μη-σταθερού αθροίσματος. Ουσιαστικά σημαίνει πως και οι δύο παίκτες θα παραμείνουν ικανοποιημένοι με την επιλογή στρατηγικής τους και την έκβαση του παιγνίου, εφόσον έχει επιτευχθεί η μεγιστοποίηση της αμοιβής ή η ελαχιστοποίηση της απώλειας. Στο παραπάνω θεωρούμε δεδομένο πως και ο δεύτερο παίκτης θα ενεργούσε με σκοπό την επίτευξη αντίστοιχων στόχων για τον εαυτό του.

Ο μαθηματικός ορισμός της ισορροπίας Nash έχει ως εξής:

Ένα ζεύγος στρατηγικών (k,l), που αντιστοιχεί στον στοίχο k και στην στήλη l, αποτελεί ένα μη-συνεργατικό σημείο ισορροπίας ή σημείο ισορροπίας Nash για το δι-μητρικό παίγνιο με μήτρες απώλειας A και B, εάν οι ακόλουθης ανισότητες ισχύουν για όλα τα i=1,…,m και όλα τα j=1,…,n αντιστοίχως:

$$a_{kl} \leq a_{il}$$

$$b_{kl} \leq b_{kj}$$

Η τιμή $a_0 = a_{kl}$ είναι η απώλεια Nash του παίκτη στοίχου, ενώ η τιμή $b_0 = b_{kl}$ είναι η απώλεια Nash του παίκτη στήλης.

Αν οι A και B είναι οι μήτρες αμοιβής, οι φορές των παραπάνω ανισοτήτων αντιστρέφονται και μιλάμε πλέον για αμοιβές Nash. (2)

# ΚΕΦΑΛΑΙΟ 2 – ΤΟ ΔΙΛΗΜΜΑ ΤΟΥ ΦΥΛΑΚΙΣΜΕΝΟΥ

## 2.1 Εισαγωγή στο Πρόβλημα

### 2.1.1 Εισαγωγή

Το παίγνιο των φυλακισμένων παρουσιάστηκε πρώτη φορά από τους Dresher και Flood το 1950 στα πλαίσια της πειραματικής θεωρίας παιγνίων και έλαβε την τελική του μορφή και ονομασία από τον Tucker. Όμως η κατάσταση που μορφοποιεί είναι ενδεχομένως γνωστή από αρχαιοτάτων χρόνων.

Ο Θουκυδίδης αναφέρει ότι όταν, κατά την εκστρατεία των Αθηναίων στη Σικελία κάποιος ή κάποιοι έσπασαν τις στήλες του Ερμή, το οποίο θεωρήθηκε ως πράξικοπηματική πράξη των ολιγαρχικών έναντι των δημοκρατικών, με αποτέλεσμα να αρχίσουν να συλλαμβάνονται και να φυλακίζονται πολίτες. Κανείς δεν γνώριζε τους ενόχους οπότε προκυρήχτηκαν δημόσια, σύμφωνα με τον Θουκυδίδη, μεγάλες αμοιβές για να ανακαλύψουν τους ενόχους. Οποιοσδήποτε πολίτης, είτε ξένος είτε δούλος ήξερε κάποια ιεροσυλία, έπρεπε να την καταγγείλει χωρίς φόβο για το άτομο του. Οι ανακριτές χωρίς να ελέγχουν το ποιόν των καταδοτών, αποδέχονταν φιλύποπτα όλες τις καταγγελίες, και δίνοντας πίστη σε αχρείους ανθρώπους, έπιαναν και έριχναν στην φυλακή ευϋπόληπτους πολίτες, Έτσι δεκάδες πολιτών, ανάμεσα τους και πολλοί αθώοι, κατέληξαν στη φυλακή.

Το γεγονός ότι η ομολογία του ενός δύναται να επιφέρει την καταδίκη αθώου η μη-αποδεδειγμένα ενόχου έχει οδηγήσει πολλές χώρες ανα τον κόσμο, στην εισαγωγή σχετικής νομοθεσίας. Για παράδειγμα, την Σουηδία υπάρχει πρόσφατη δικαστική απόφαση αθώωσης δύο υπόπτων για ανθρωποκτονία διότι αμφότεροι ομολογούσαν ότι το φονικό όργανο το κρατούσε ο άλλος.

Το παίγνιο του φυλακισμένου διευκρινίζει ότι, όταν οι παίκτες επιλέξουν τις στρατηγικές τους με ατομικά ορθολογικά κριτήρια, καταλήγουν σε μία έκβαση που για έναν ουδέτερο παρατηρητή δεν φαίνεται να είναι η καλύτερη δυνατή για το κοινό τους συμφέρον. Η ατομική ορθολογική συμπεριφορά από τη μία και η έλλειψη επικοινωνίας από την άλλη δεν φαίνεται να επιφέρουν το «βέλτιστο» αποτέλεσμα για τους δύο παίκτες.

Το συγκεκριμένο παίγνιο είναι ενδιαφέρον γιατί εξηγεί τον λόγο για τον οποίο δύο ανταγωνιστές αποτυγχάνουν συχνά να συνεργαστούν, έστω και αν μια τέτοια συνεργασία θα απέβαινε σε αμοιβαίο όφελος. Η μη συνεργατική ισορροπία ή ισορροπία Nash αναδεικνύει την μεγάλη δαφορά μεταξύ ατομικού και συλλογικού συμφέροντος. Το παράδοξο, ότι η ατομική ορθολογική δράση δεν αποφέρει το βέλτιστο αποτέλεσμα για τους παίκτες, έχει μεγάλη επίδραση στις κοινωνικές επιστήμες, καθώς και σε πολλά προβλήματα της σύγχρονης κοινωνίας, όπως ο εξοπλιστικός ανταγωνισμός, η ατμοσφαιρική ρύπανση, η υπερεκμετάλευση των φυσικών πόρων κ.α.

### 2.1.2 Περιγραφή του Παιγνίου

Το δίλημμα του φυλακισμένου, όπως αναφέρθηκε στο 1$^{ο}$ κεφάλαιο, είναι ένα μη-μηδενικό παίγνιο δύο παικτών. Ο όρος «μη-μηδενικό» υποδεικνύει πως όταν το όφελος συσσωρεύεται στον έναν παίκτη αυτό δεν σημαίνει πως υποχρεωτικά θα συσσωρεύεται μία αντίστοιχη τιμωρία στον άλλο παίκτη. Το δίλημμα του φυλακισμένου μπορεί να οριστεί και ως «μη-συνεργατικό», για να υποδείξει πως δεν επιτρέπεται καμιά επικοινωνία μεταξύ των παικτών εκτός της δράσης του παιγνίου. Παρ' όλα αυτά όμως, ο όρος «μη-συνεργατικό» μπορεί να θεωρηθεί παραπλανητικός, επειδή η ανάλυση της «συνεργασίας» και οι αναδυόμενες συμπεριφορές συνεργασίας αποτελούν ουσιαστικά τα αντικείμενα της έρευνας γύρω από το παίγνιο.

Στην πιο ουσιώδη μορφή του, το παίγνιο έχει να κάνει με μια κατάσταση που πρέπει να αντιμετωπιστεί απο δύο παίκτες, οι οποίοι δύναται να επιλέξουν μεταξύ δύο πιθανών επιλογών: η πρώτη μπορεί να χαρακτηριστεί ως συνεργατική και η δεύτερη ως εγωιστική. Αν και οι δύο παίκτες αποφασίσουν να συνεργαστούν, μπορούν να κερδίσουν μια μεγάλη αμοιβή. Απ την άλλη μεριά, αν και οι δύο αποφασίσουν να δράσουν εγωιστικά, θα λάβουν μια μικρή αμοιβή. Αν, όμως, μόνο ο ένας από τους δύο συνεργαστεί, τότε αυτός θα λάβει μια πολύ μικρή αμοιβή – ή και καθόλου – ενώ ο άλλος θα λάβει την μέγιστη αμοιβή.

Πιο συγκεκριμένα, η ιστορία που περιγράφει κατατοπιστικά το παιγνιο έχει ως εξής: Οι δύο ύποπτοι συλλαμβάνονται από την αστυνομία και τοποθετούνται σε διαφορά δωμάτια για να ανακριθούν. Οι ανακριτές τους προσφέρουν την ίδια συμφωνία. Αν ο ένας απ' αυτούς ομολογήσει και ο άλλος όχι, τότε ο πρώτος θα αφαιθεί ελεύθερος και ο δεύτερος θα εκτίσει χρόνια φυλάκισης. Αν και οι δύο ομολογήσουν τότε η ποινή τους θα είναι μικρότερη. Οι ύποπτοι γνωρίζουν επίσης πως ο ανακριτής δεν έχει πολλά στοιχεία στα χέρια του για να τους καταδικάσει, έτσι, αν και οι δύο παραμείνουν σιωπηλοί, η ποινή τους δεν θα είναι τόσο σοβαρή. Γνωρίζοντας αυτά τα δεδομένα, ποιά είναι η σοφότερη επιλογή τους; Να εμπιστευθούν ο ένας τον άλλον κρατώντας μυστικό την ενοχή τους ή να καταδώσουν ο ένας τον άλλον; Κάπως έτσι προκύπτει πως ο κάθε παίκτης έχει 2 επιλογές, είτε να συνεργαστεί (cooperate - C) είτε να αποστατήσει (defect - D). Οι αμοιβές που θα λάβουν αντίστοιχα φαίνονται στον παρακάτω πίνακα.

Στον πίνακα το κάθε κελί μπορεί να περιέχει 2 τιμές. Η αριστερή τιμή είναι του παίκτη γραμμής και η δεξιά τιμή είναι του παίκτη στήλης. Αν θέσουμε ως R (reward) την αμοιβή για συνεργατική συμπεριφορά και P (punishment) την αμοιβή αν και οι δύο δράσουν εγωιστικά, τότε T (temptation) θα είναι η αμοιβή του παίκτη που θα δράσει εγωιστικά όταν ο άλλος θα συνεργαστεί και S (sucker) θα είναι η αμοιβή εκείνου που συνεργάστηκε όταν ό άλλος έδρασε εγωιστικά.

| ΠΙΝΑΚΑ Α | **Α παίκτης Συνεργασία** | **Α παίκτης Αποστασί α** |
|---|---|---|
| **Β παίκτης Συνεργασία** | R,R | S,T |
| **Β παίκτης Αποστασία** | T,S | P,P |

Οι απαιτήσεις του παζλ θα είναι ως εξής:

$$T > R > P > S \qquad (2.1)$$

Επίσης, η διπλή αμοιβή για την συνεργατική συμπεριφορά πρέπει να είναι μεγαλύτερη από τον μέσο όρο των αμοιβών T και S:

$$2 * R > S + T \qquad (2.2)$$

Το δίλημμα μπορεί να τεθεί και ως ελαχιστοποιήση μιας ποινής φυλάκισης αλλά και ως μεγιστοποίηση ανταμοιβής για την συνεργασια του ατόμου.

Ένα στιγμιότυπο των παραπάνω σχέσεων εμφανίζεται στον παρακάτω πίνακα. Σύμφωνα με αυτόν, αν ένας λογικός παικτης θα επιλέξει αποστασία ανεξάρτητα από αυτό που θα επιλέξει ο αντίπαλος (5>3 και1>0), τότε και οι 2 παίκτες θα κερδίζουν από έαν πόντο. Αν όμως συνεργαστούν θα λάβουν 3 πόντους. Δυστυχώς αυτό το παίγνιο θα παικτεί 1 φορά οπότε οι παίκτες δεν μπορούν να αντισταθούν στον πειρασμό να αποστατήσουν οπότε η κοινή συνεργασία είναι δύσκολο να προκύψει. [11][13]

| ΠΙΝΑΚΑ Α | Β παίκτης Συνεργασία | Β παίκτης Αποστασί α |
|---|---|---|
| Α παίκτης Συνεργασία | 3,3 | 0,5 |
| Α παίκτης Αποστασία | 5,0 | 1,1 |

Το δίλημμα εμφανίζεται όταν κάποιος υποθέτει ότι και οι δύο παίκτες νοιάζονται μόνο για να μεγιστοποιήσουν το κέρδος τους. Κάθε παίκτης έχει δύο στρατηγικές επιλογές, είτε να συνεργαστεί, είτε να αποστατήσει. Για παράδειγμα το καλύτερο αποτέλεσμα για τον παίκτη Α είναι να παραμείνει σιωπηλός (να αποστατήσει δηλαδή) και ο παίκτης Β να έχει ομολογήσει (να συνεργαστεί). Το επόμενο καλύτερο αποτέλεσμα για τον Α είναι να ομολογήσουν και οι δύο (κοινή συνεργασία), ενώ το χειρότερο σενάριο είναι να αποστατήσει ο Β ενώ ο Α θα έχει συνεργαστεί. Το αντίστοιχο ισχύει και για τον παίκτη Β. Είναι, λοιπόν, φανερό πως, οτιδήποτε και να σκοπεύει να κάνει ο Β, ο παίκτης Α θα πρέπει να επιλέξει την δεύτερη στρατηγική (να αποστατήσει δηλαδή), αφού έτσι θα έχει καλύτερα αποτελέσματα. Ομοίως ισχύει και για τον Β παίκτη, ο οποίος θα προτιμήσει και αυτός να αποστατήσει. Σε αυτό το σημείο υπάρχει το δίλημμα, αφού από τον πίνακα φαίνεται πως οι παίκτες θα αποκομίσουν μεγαλύτερο όφελος αν και οι δύο επιλέξουν να ομολογήσουν

και να συνεργαστούν μεταξύ τους από το να τα ομολογήσουν όλα. . Έτσι, η καλύτερη στρατηγική για τον καθένα ξεχωριστά παράγει ένα αποτέλεσμα που δεν είναι καλό για την ομάδα, κρίνοντας όμως από τα ατομικά κίνητρα να υπονομεύουν το κοινό συμφέρον.

Παζλ, όπως το παραπάνω πρόβλημα, επινοήθηκαν και προτάθηκαν από τους Merril Flood και Melvin Dresher το 1950 σαν κομμάτι ερευνών στην θεωρία παιγνίων και τις πιθανές εφαρμογές τους στην διεθνή πυρηνική στρατηγική της εταιρίας RAND. Το όνομα “prisoner’s dilemma” διαλέχτηκε από τον Albert Tucker, που ήθελε να κάνει την ιδέα των Flood και Dresher περισσότερο προσβάσιμη στους ψυχολόγους. Φαντάστηκε, λοιπόν, δύο εγκληματίες που έχουν συλληφθεί για μια επίθεση και έχουν τοποθετηθεί σε δύο διαφορετικά και απομονωμένα κελιά. Ο κάθε κακοποιός μπορεί να επιλέξει να συνεργαστεί με τον συνεργό του αρνούμενος την οποιαδήποτε εμπλοκή στο έγκλημα ή να αποστατήσει, ομολογώντας στην αστυνομία και προδίδοντας τον συνεργό του. Εφόσον ο τελικός στόχος είναι να μειωθεί ο χρόνος παραμονής στη φυλακή και όχι να αυξηθεί το κέρδος, οι παραπάνω εξισώσεις αντιστρέφονται.

Αγνοώντας τους ηθικούς φραγμούς, η πιο βολική συμπεριφορά είναι η εγωιστική. Αν ο ένας εκ των δύο συνεργαστεί, η περισσότερη κερδοφόρα πράξη για τον άλλον είναι να εκμεταλλευτεί την καλή πίστη και να γίνει εγωιστής. Απ' την άλλη μεριά, αν ο ένας παίκτης δράσει εγωιστικά τότε το καλύτερο για τον δεύτερο είναι να δράσει κι αυτός το ίδιο, μειώνοντας την ζημιά. Αν όμως το παζλ συνεχίζεται επαναληπτικά και επ' αόριστον τότε η υπόθεση γίνεται ακόμα πιο ενδιαφέρουσα. Σε ένα τρέξιμο με πολλές επαναλήψεις η πιο βολική στρατηγική είναι να έρθουν σε συμφωνία οι δυό τους και να ενστερνιστούν μια στρατηγική συνεργασίας. Αυτή η έκδοση του παζλ καλείται "iterated prisoner's dilemma" και ορίστηκε από τότε που ορίστηκε και το ίδιο το παίγνιο, ωστόσο το ενδιαφέρον επιταχύνθηκε μετά τις δημοσιεύσεις του Robert Axelrod [8]

### 2.1.3 Επαναληπτικό Δίλημμα του Φυλακισμένου

Το 1979 ο Axelrod οργάνωσε το πρώτο σημαντικό τουρνουά με επαναληπτικό δίλημμα του φυλακισμένου προσελκύοντας στρατηγικές που είχαν δημοσιευθεί μέχρι τότε στον τομέα. Οι 14 συμμετοχές είχαν να ανταγωνιστούν τον 15$^{o}$ με τον οποίο η πιθανότητα συνεργασίας ή αποστασίας ήταν η ίδια. Η κάθε στρατηγική έπαιξε με κάθε άλλη για πάνω από 200 κινήσεις. Ο νικητής του τουρνουά ήταν ο Anatol Rapoport, γεννημένος στην Ρωσία, Αμερικανός πολίτης, μαθηματικός και ψυχολόγος. Η στρατηγική του λεγόταν "tit for tat" – οφθαλμός αντί οφθαλμού. Η στρατηγική περιλάμβανε συνεργασία πάντα στην πρώτη κίνηση και μετά απ' αυτό μίμηση της προηγούμενης κίνησης του αντιπάλου, οποιαδήποτε κι αν ήταν αυτή. Μετέπειτα ανάλυση πάνω στην στρατηγική έδειξε πως ήταν εύρωστη γιατί ποτέ δεν αποστατούσε πρώτη και ποτέ δεν λάμβανε υπόψιν της σαν πλεονέκτημα κινήσεις που συνέβαιναν σε πάνω από μια επαναλήψεις. Σε ένα δεύτερο τουρνουά, ο Axelrod συγκέντρωσε 62 στρατηγικές και πάλι η νικήτρια στρατηγική ήταν αυτή του οφθαλμού αντί οφθαλμού.

Το επαναληπτικό δίλημμα του φυλακισμένου μελετήθηκε έντονα στην κοινότητα των μελετητών των εξελικτικών αλγορίθμων. Ο ίδιος ο Axelrod προσπάθησε να αναπτύξει έναν πληθυσμό στρατηγικών στην αρχή του 1980 χρησιμοποιώντας έναν γενετικό αλγόριθμο και αναπαριστώντας τις στρατηγικές σαν πίνακες από bit. Αργότερα, οι μηχανές πεπερασμένων καταστάσεων έγιναν αρκετά κοινές στο να αναπαριστούν τις στρατηγικές, εφόσον μπορούν και συμπεριφέρονται σαν σύνθετα μοντέλα Markov. Μετέπειτα έρευνες χρησιμοποίησαν και άλλες προσεγγίσεις, όπως είναι ο εξελικτικός προγραμματισμός και νευρωνικά δίκτυα. [9]

Οι πιο δημοφιλείς «γνωστές στρατηγικές είναι οι ακόλουθες:

- AC (always coop): Πάντα συνεργασία.
- AD (always defect): Πάντα αποστασία.
- BA (Brainless alternation): εναλλάσσει συνεργασία με αποστασία
- E-BA (Evil Brainless alternation): εναλλάσσει αποστασία με συνεργασία
- G (Grudger): συνεργασία μέχρι να πετύχει αποστασία. Μετά αποστατεί μέχρι το τέλος του παιχνιδιού.

- SG (Soft Grudger): συνεργασία μέχρι ο αντίπαλος να αποστατήσει. Σε αυτή την περίπτωση συμπεριφέρεται με 4 αποστασίες στη σειρά. Έπειτα συνεχίζει με 2 συνεχόμενες συνεργασίες.
- TFT (Tit for tat): συνεργασία αρχικά και έπειτα ότι κάνει ο αντίπαλος.
- E-TFT (Evil tit for tat): αποστασία αρχικά και έπειτα ότι κάνει ο αντίπαλος.
- A-TFT (Anti tit for tat): συνεργασία αρχικά και έπειτα το αντίθετο του αντιπάλου του.
- TF2T (Tit for two tat): συνεργασία μέχρι ο αντίπαλος να κάνει 2 αποστασίες ο αντίπαλος. Έπειτα αποστατεί και θα ξανασυνεργαστεί όταν ο αντίπαλος κάνεις 2 συνεργασίες στη σειρά.
- 5TM (five is too much): ίδια λογική με την κλασσική TFT, απλά αν ο αντίπαλος συνεργαστεί 5 φορές στη σειρά τότε αποστατούμε 2 φορές.
- GC (General Cooperator): ξεκινάει με λογική TFT απλά αν και οι 2 αποστατήσουν τότε για τον επόμενο γύρο το ξεχνάει και ξανα συνεργάζεται.
- P (Pavlov): Ξεκινάει με συνεργασία. Μετά επαναλαμβάνει την προηγούμενη κίνηση αν ήταν κερδοφόρα αλλιώς την αλλάζει.
- 2TT (Two to trust): ο παίκτης συνεργάζεται εκτός κι αν ο αντίπαλος αποστατήσει οπότε ξεκινάει να αποστατεί. Χρειάζεται 2 συνεχόμενες κινήσεις συνεργασίας για να ξεκινήσει να συνεργάζεται ξανά. Είναι γνωστή και ως Two to forgive.
- E-2TT (Evil two to trust): ο παίκτης αποστατεί εκτός κι αν ο αντίπαλος συνεργαστεί για 2 συνεχόμενες φορές στη σειρά. Αν ο αντίπαλος αποστατήσει θέλει ξανά 2 συνεχόμενες συνεργασίες για να ξεκινήσει ξανά να συνεργάζεται.
- 2TB (Two to betray): συνεργασία εκτός κι αν ο αντίπαλος αποστατήσει 2 φορές στη σειρά. Σε αυτή την περίπτωση ξεκινάει να αποστατεί. Μόλις ο αντίπαλος συνεργαστεί τότε ξεκινάει ξανά η συνεργασία.
- E-2TB (Evil two to betray): Ο παίκτης αποστατεί μέχρι να συναντήσει συνεργασία. Στη συνέχεια συνεργάζεται μέχρι ο αντίπαλος να αποστατήσει 2 φορές στη σειρά.
- TTP (three then punish): Συνεργασία μέχρι ο αντίπαλος να αποστατήσει 3 φορές στη σειρά. Μόλις συμβεί αυτό ο παίκτης αποστατεί 3 φορές. Μετά απ αυτό αν ο αντίπαλος συνεργαστεί, συνεργάζεται και ο παίκτης αλλιώς συνεχίζει να αποστατεί.
- T/D (Trust/Distrust): Ο παίκτης εσωτερικά βαθμολογεί την πίστη του αντιπάλου του σε μια κλίμακα από το 1 μέχρι το 4. Όταν ο αντίπαλος συνεργάζεται, η πίστη αυξάνεται, όταν αποστατεί μειώνεται. Οπότε για 1 ή 2 αποστατεί και για 3 ή 4 συνεργάζεται.
- EGET (Evil Good Evil Trust): Ο παίκτης ξεκινά με μια αλληλουχία συνεργασία-αποστασία-συνεργασία. Αν ο αντίπαλος απαντήσει με συνεργασία στις 3 αυτές κινήσεις θεωρείται ως έμπιστος και συνεχίζει να συνεργάζεται.

### 2.1.4 MiniMax διαδικασία

Κατά το παίξιμο του παιγνίου, υποθέτουμε στις περισσότερες περιπτώσεις πως ο αντίπαλος είναι απόλυτα λογικός και χρησιμοποιεί μια στρατηγική παρόμοια με την δική μας. Αυτό οδήγησε τον John Von Neumann να αναπτύξει την διάσημη MiniMax διαδικασία το 1928. Αρκετές βελτιώσεις και βελτιστοποιήσεις προτάθηκαν τα επόμενα χρόνια και μερικές παραλλαγές προσπάθησαν να λάβουν υπόψιν τους την εξαπάτηση ή την παρορμητικότητα του παίκτη για να βελτιωθεί η εμπειρία του παιγνίου. Ωστόσο, ο θεμελιώδης στόχος είναι να μεγιστοποιηθεί η αμοιβή απέναντι σε έναν αντίπαλο που αναζητά τον ίδιο στόχο.

Στα παίγνια μηδενικού αθροίσματος, όπου η αμοιβή ενός παίκτη παίρνεται από τον άλλον, η ιδέα του να μεγιστοποιήσουμε την αμοιβή και να μειώσουμε τον κίνδυνο μπορεί να θεωρηθεί ισοδύναμη. Σε περισσότερο περίπλοκα παίγνια μη-μηδενικού αρθοίσματος μπορεί αυτοί οι στόχοι να είναι περισσότεροι από έναν και αντικρουόμενοι. Ακόμα, σε αυτές τις περιπτώσεις δεν ψάχνουν όλοι οι παίκτες να μειώσουν τον μέγιστο κίνδυνο που μπορεί να προκαλέσει ο αντίπαλος. Καθώς μερικοί παίκτες είναι πρόθυμοι να πάρουν ρίσκο απ' το αντίθετο. Για παράδειγμα, ένας παίκτης που παίζει πάντα εγωιστικά δεν μπορεί να νικηθεί, και το πιθανότερο είναι να μαζέψει μια πολύ μικρή αμοιβή σε κάθε παίγνιο του.

Επίσης πρέπει να θεωρηθεί πιθανό πως κάποιοι παίκτες μπορεί να έχουν παράλογα κίνητρα και να κινηθούν προς ένα συγκεκριμένο στόχο. Αυτό είναι εν μέρη σωστό σε ένα παίγνιο με άμεσες ηθικές επιπτώσεις, όπως είναι το δίλημμα του φυλακισμένου.

Το άμεσο συμπέρασμα είναι πως, σε σύνθετα σενάρια, είναι λογικό να σκεφτούμε πως ο αντίπαλος μας πιθανόν να χρησιμοποιεί μια στρατηγική διαφορετική από την δική μας. Ακόμα, ο αλγόριμος MiniMax δεν είναι άμεσα χρησιμοποιήσιμος. Η γνώση της πρόθεσης και της στρατηγικής του αντιπάλου είναι πολύ σημαντικά στοιχεία για να παίξουμε αποτελεσματικά. Η ιδέα μοντελοποίησης του αντιπάλου προτάθηκε από τότε που το παίγνιο άρχισε να παίζεται στους υπολογιστές στις αρχές του 50'. Αρκετοί αλγόριθμοι έχουν προταθεί από τότε που ενσωματώνουν όλη αυτή τη γνώση αλλά κανένας απ αυτούς δεν έχει αποδειχθεί απόλυτα αποτελεσματικός.

## 2.2 Προσέγγιση του προβλήματος με εξελικτικούς αλγορίθμους

### 2.2.1 Προσαρμοστική μοντελοποίηση

Η συγκεκριμένη προσέγγιση είναι βασισμένη στην έννοια της προσαρμοστικής μοντελοποίησης του αντιπάλου. Δηλαδή, ο παίκτης χωρίζεται σε δύο συνδυασμένα υποσυστήματα: ένα υποσύστημα αντιγραφέα ο οποίος είναι ικανός να διαμορφώσει την στρατηγική του αντιπάλου σε FSM, και ένα υποσύστημα που αντιπροσωπεύει τον παίκτη και είναι σε θέση να εκμεταλευτεί το μοντέλο έτσι ώστε να βρεί την καταλληλότερη επόμενη κίνηση.

Ο στόχος του υποσυστήματος του αντιγραφέα είναι να επινοήσει ένα μοντέλο ικανό να μαντέψει μελλοντικές κινήσεις. Παρακολουθεί το παίγνιο, καταγράφει την ακολουθία των κινήσεων που πραγματοποιούνται από τον ίδιο τον παίκτη και τις αντιδράσεις του αντιπάλου. Τέτοια πληροφορία χρησιμοποιείται τελικά για να δημιουργήσει ένα εσωτερικό μοντέλο της στρατηγικής. Το μοντέλο ενημερώνεται όταν είναι διαθέσιμα τα νέα δεδομένα, δηλαδή στην προκειμένη περίπτωση μετά την ολοκλήρωση ενός γύρου.

Το υποσύστημα βασίζεται σε έναν εξελικτικό αλγόριθμο, ενώ παράλληλα καλλιεργείται ένας πληθυσμός διαφορετικών υποψήφιων μοντέλων. Οι εξελικτικοί αλγόριθμοι έχουν αποδειχθεί πως είναι ικανοί να αντιμετωπίσουν δύσκολα προβλήματα, επιδεικνύοντας την ικανότητα τους να προσαρμόζονται στα μεταβαλλόμενα σενάρια, οπότε αποτελούν ένα ιδανικό εργαλείο για την συγκεκριμένη εργασία.

Σχετικές παραδοχές που έγιναν είναι οι ακόλουθες: καταρχάς το εσωτερικό μοντέλο τροποποιείται συνέχεια και είναι πιθανό να αλλάζει μετά από κάθε γύρο. Η ποικιλία του εσωτερικού πληθυσμού πρέπει να διατηρηθεί για να αποφευχθεί πρόωρη σύγκλιση. Αυτό το είδος προβλήματος, θα παρεμποδίσει την μελλοντική προσαρμογή. Ακόμα, σε κάθε βήμα ο αριθμός G των γενεών του εξελικτικού αλγορίθμου θα πρέπει να παραμένει αρκετά χαμηλός. Πράγματι, τρέχοντας έναν περιορισμένο αριθμό γενεών μπορεί να αυξηθεί η γενική ανταπόκριση του παίκτη.

Το υποσύστημα που αντιπροσωπεύει τον παίκτη εκμεταλεύεται το μοντέλο του αντιπάλου έτσι ώστε να βρει την βέλτιστη κίνηση. Θεωρεί τις πιθανές εκβάσεις του παιγνίου, επιλέγει το βέλτιστο, μιας και παίζει την πρώτη κίνηση του, και την εκτελεί. Ο υπολογισμός πρέπει να επαναλαμβάνεται σε κάθε στροφή.

Όσο το μοντέλο του αντιπάλου είναι απόλυτα ντετερμινιστικό, το υποσύστημα χρησιμοποιεί μια πολύ απλή προσέγγιση δύναμης εξαναγκασμού: αναπαράγει όλες τις δυνατές $2^n$ ακολουθίες των N κινήσεων και επιλέγει αυτήν που οδηγεί στο καλύτερο αποτέλεσμα. Είναι προφανές πως, εφόσον το μοντέλο του αντιπάλου ενημερώνεται μετά την καταγραφή της κίνησης, η συνεχής πρόβλεψη μπορεί να αποδειχθεί ανακριβής.

Το βάθος της ανάλυσης από την δύναμη εξαναγκασμού N είναι η μοναδική παράμετρος του υποσυστήματος που αντιπροσωπεύει τον παίκτη. Αυτό επηρεάζει άμεσα το πλήθος των πόρων που χρειάζονται για να επιλεγεί μια κίνηση. Εφόσον το μοντέλο του αντιπάλου είναι προσεγγιστικό και μπορεί να αλλάξει, το να υπολογίζουμε μεγάλες ακολουθίες κινήσεων δεν είναι χρήσιμο.

### 2.2.2 Προσέγγιση Laran

Οι Elio Piccolo και Giovanni Squillero προσπάθησαν να προσεγγίσουν την επαναληπτική διαδικασία του διλήμματος του φυλακισμένου μοντελοποιώντας έναν προσαρμοζόμενο αντίπαλο. Η προτεινόμενη προσέγγιση ονομάστηκε Laran, από το θεό του πολέμου στη μυθολογία των Ετρούσκων [8]

Ο Laran αντιπροσωπεύει μια στρατηγική σαν μια μηχανή του Moore, δηλαδή σαν ένα FSM του οποίου η τιμή εξόδου καθορίζεται μόνο από την τρέχουσα κατάστασή του και όχι από την τιμή της εισόδου του. Άλλες εργασίες πάνω στο δίλημμα του φυλακισμένου χρησιμοποιούν FSM στα οποία η έξοδος είναι μία συνάρτηση της τρέχουσας καταστασης και της εισόδου, τα λεγόμενα μηχανές Mealy· αυτό αποτελεί μία ισοδύναμη αναπαράσταση.


*Figure 1- Μηχανή Moore*

Στο παραπάνω σχήμα φαίνεται μία μηχανη Moore που περιγράφει τη στρατηγική Tit for tat. Υπάρχουν δύο καταστάσεις που αριθμούνται με «0» και «1». Η έξοδος της κατάστασης διακρίνεται από το χρώμα του πλαισίου της: ανοιχτό σημαίνει ότι ο παίκτης συνεργάζεται, το σκούρο ότι αποστατεί. Οι μεταβάσεις κατάστασης παρουσιάζονται με βέλη ονομασμένα σύμφωνα με την κίνηση του αντιπάλου: «c» για συνεργασία και «d» για αποστασία. Η εναρκτήρια κατάσταση επισημαίνεται με ένα βέλος.

Όπως είναι αναμενόμενο, ο εξελικτικός αλγόριθμος που είναι ενσωματωμένος στον αντιγραφέα θυμίζει την πρωτοποριακή δουλειά του Fogel στην επιλογή τελεστών. Πολλές λεπτομέρειες, ωστόσο, έχουν μεταλλαχθεί από διαφορετικούς, πιο πρόσφατους εξελικτικούς αλγορίθμους.

Ο εξελικτικός αλγόριθμος χρησιμοποιεί μια προσέγγιση σταθερής κατάστασης, με ένα πληθυσμό ‘Ι’ ατόμων. Ο αρχικός πληθυσμός παράγεται τυχαία. Στη συνέχεια, σε κάθε γενεά ‘Ο’ γονείς επιλέγονται τυχαία και σε κάθε έναν από αυτούς αντιστοιχίζεται ένας γενετικός τελεστής. Ο γενετικός τελεστής επιλέγεται τυχαία από μία ομάδα τεσσάρων: ‘άλλαξε αρχική κατάσταση’, ‘πρόσθεσε κόμβο’, ‘δίαγραψε κόμβο’ και ‘τροποποίησε μετάβαση’. Δεν έχουν συμπεριληφθεί τελεστές επανασυνδυασμού.

Ο τελεστής ‘άλλαξε αρχική κατάσταση’ κλωνοποιεί τον γονέα και στη συνέχεια τροποποιεί την αρχική κατάσταση του FSM επιλέγοντας τυχαία ένα νέο κόμβο. Ο τελεστής ‘πρόσθεσε κόμβο’ κλωνοποιεί τον γονέα και έπειτα προσθέτει ένα νέο κόμβο και τον συνδέει τυχαία με τους υπάρχοντες. Ο τελεστής ‘διάγραψε κόμβο’ επιλέγει έναν τυχαίο κόμβο και δημιουργεί δύο νέα άτομα χωρίς αυτόν: στον πρώτο, ο πρόγονος συνδέεται με τον διάδοχο στην περίπτωση συνεργασίας· στον δεύτερο, ο πρόγονος συνδέεται με τον διάδοχο στην περίπτωση αποστασίας. Ο τελεστής ‘τροποποίησε μετάβαση’ κλωνοποιεί τον γονέα, και έπειτα τον συνδέει τυχαία με έναν υπάρχοντα κόμβο. Όλοι οι γενετικοί τελεστές έχουν ίσες πιθανότητες.

Οι γονείς επιλέγονται μέσα από ένα τουρνουά μεγέθους δύο: δύο άτομα ξεχωρίζονται τυχαία, συγκρίνονται και το ένα επιλέγεται σαν ο γονέας. Για να αποφευχθεί η εξόγκωση των διανυσμάτων, χρησιμοποιείται μία οπή καταλληλότητας. Με πιθανότητα p, το κριτήριο για την επιλογή στα τουρνουά δεν είναι η τιμή της αντικειμενικής συνάρτησης αλλά το μέγεθος των ατόμων: το μικρότερο άτομο επιλέγεται για αναπαραγωγή, όχι το πιο ιδανικό. Το όνομα ‘οπή καταλληλότητας’ προέρχεται από το γεγονός ότι πρόκειται για μία τρύπα/οπή/κενό στο μοίρασμα (distribution) των πιθανοτήτων που καθορίζει την επιλογή.

Η επιλογή επιβίωσης είναι μια καθαρά ντετερμινιστική διαδικασία. Τα χειρότερα μέλη αφαιρούνται από τον πληθυσμό μέχρι το μέγεθος του πληθυσμού να επανέλθει σε ‘*I*’.

Όπως αναφέρθηκε παραπάνω, ο στόχος του υποσυστήματος αντιγραφέα είναι να εφεύρει ένα μοντέλο ικανό να εξηγήσει παρελθοντικές κινήσεις και να μαντέψει μελλοντικές. Για να ικανοποιήσει και τις δύο αυτές απαιτήσεις, το μοντέλο του αντιπάλου θα πρέπει να μπορεί να παίξει σχετικά καλά. Ένα απλό FSM που μπορεί να αναπαράγει όλες τις προηγούμενες κινήσεις του αντιπάλου ίσως να μην έχει ικανότητα πρόβλεψης. Συνεπώς, είναι απαραίτητο να ληφθεί υπόψη η δύναμη των μοντέλων, και όχι μόνο η ικανότητά τους να αναπαράγουν προηγούμενες κινήσεις.

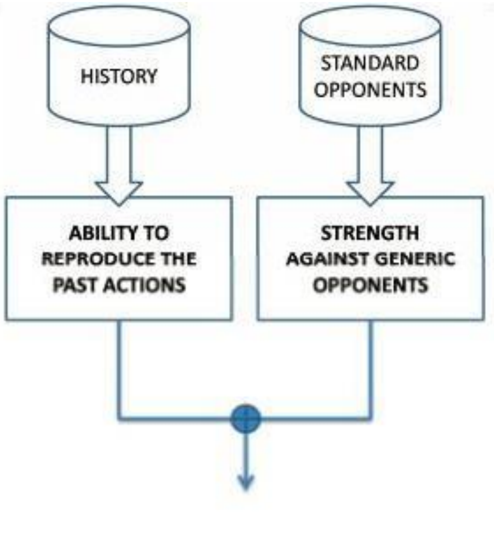



*Figure 2- Εξελικτικός Αλγόριθμος*

.

Για να πετύχει αυτό το στόχο, ο εξελικτικός αλγόριθμος χωρίζει τον υπολογισμό του κάθε υποψηφίου μοντέλου σε δύο βήματα, όπως φαίνεται στο παραπάνω σχήμα: πρώτα, υπολογίζεται η ικανότητα να επαναληφθούν οι προηγούμενες ενέργειες του παιχνιδιού. Το υποψήφιο μοντέλο λαμβάνει ένα σκορ που δηλώνει το πόσες φορές η στρατηγική καθόρισε σωστά τις ενέργειες που έχουν παιχτεί στο παιχνίδι. Στη συνέχεια, ο υποψήφιος αντιμετωπίζει μια σειρά προκαθορισμένων αντιπάλων. Ο στόχος του δεύτερου υπολογισμού είναι να μετρηθεί η δύναμη της υποψήφιας στρατηγικής.

Η σειρά των καθορισμένων εχθρών δεν τροποποιείται ποτέ, ούτε λαμβάνει μέρος στην εξελικτική διαδικασία. Χρησιμοποιείται απλά για να δώσει ένα ποσοτικοποιημένο υπολογισμό της δύναμης του υποψηφίου. Θεωρητικά, θα μπορούσε να αντικατασταθεί από μία στατική ανάλυση ή κάποιο άλλο κατάλληλο εργαλείο.

Στην προτεινόμενη ενσωμάτωση η σημαντικότητα της πρώτης συνεισφοράς υπερβαίνει τη σημαντικότητα της δεύτερης, γί'αυτό οι στρατηγικές πρέπει εξαρχής να έχουν πλήρη συνοχή με το ιστορικό.

Με το πέρας της έρευνας, όπως είχαν προβλέψει, το μοντέλο Laran αποδείχθηκε ικανό να καθοδηγήσει σωστά το μοντέλο και μακροπρόθεσμα να ξεπεράσει τις αντίπαλες στρατηγικές πρότυπα. Κινούμενοι προς αυτή την κατεύθυνση, ενδιαφέρον θα έχει να εφαρμοστεί το μοντέλο προσαρμοσμένου αντιπάλου σε περιπτώσεις που ο αντίπαλος δεν είναι μοντελοποιημένος με FSM και δεν μπορεί να μοντελοποιηθεί κατ' αυτόν τον τρόπο.

### 2.2.3 Προσέγγιση με γενετικό αλγόριθμο

Οι Wang Tao, Chen Zhi-Gang, Deng Xiao-Heng και Zhagn Jin, δημιούργησαν ένα ελεύθερης κλίμακας δίκτυο με βάση το BA μοντέλο. Το μοντέλο BA (Barabasi – Albert) είναι ένας αλγόριθμος για παραγωγή τυχαίων δικτύων ελεύθερης κλίμακας, που χρησιμοποιεί ένα προνομιακό μηχανισμό προσάρτησης. Παράλληλα έκαναν χρήση ενός γενετικού αλγορίθμου για να μελετήσουν πως τα άτομα μπορούν να εξελιχθούν σε πολύπλοκα δίκτυα παίζοντας το παίγνιο του διλήμματος.

Κάθε κόμβος στο δίκτυο είναι άμεσα συνδεδεμένος με άλλους κόμβους που παίζουν το παίγνιο του διλήμματος και μπορεί να θυμηθεί τις τελευταίες L κινήσεις. Για παράδειγμα, αν L=3, τότε χρησιμοποιούν το bit 1 για να καθορίσουν την συνεργασία και το Bit 0 για να καθορίσουν την αποστασία. Οπότε χρησιμοποιούν 2L bits για να σημειώσουν τα τελευταία L παιχνίδια. Για παράδειγμα, αν L=3 και $H_{ij}$=100111 καθορίζονται τα τελευταία 3 παίγνια. Πιο συγκεκριμένα σημαίνει πως ο i συνεργάστηκε και ο j αποστάτησε, ο i αποστάτησε και ο j συνεργάστηκε, ο i συνεργάστηκε και ο j συνεργάστηκε. Αυτό ουσιαστικά σημαίνει 3 επαναληπτικά παιχνίδια με

64 πιθανά διαφορετικά σενάρια. Αυτές οι συμβολοσειρές από bit μπορούν να μετατραπούν σε δεκαδικούς αριθμούς από το 0 μέχρι το 63. Για κάθε διαφορετικό συνδυασμό, οι κόμβοι μπορούν να επιλέξουν μια στρατηγική (αποστασία 0 – συνεργασία 1) για τον επόμενο γύρο. Έπειτα ξεκινούν με ένα 64bit χρωμοσώμα, όπου η κάθε γενναιά του χρωμοσώματος μπορεί να είναι 1 ή 0, καθορίζοντας την στρατηγική του ατόμου σε ένα λεπτομερές ιστορικό. Για παράδειγμα ο κόμβος i έχει χρωμοσώμα $C_i$ =100110111….Η πρώτη γενναιά (Αρ. 0) μεταφράζεται σε μια ιστορικότητα 000000 και παράλληλα σε έναν συνεργάσιμο κόμβο για το επόμενο παιχνίδι. Ο δεύτερος κόμβος (Αρ. 1) σημαίνει πως η ιστορικότητα θα είναι 000001 και η τιμή 0 θα σημαίνει πως ο κόμβος θα αποστατήσει.

Στο ξεκίνημα, για το μήκος της μνήμης L, τυχαία άτομα που αρχικοποιούν το χρωμοσώμα (μήκος $2^{2L}$) τίθενται σε 0 ή σε 1. Στα πρώτα L παιχνίδια, τα άτομα επιλέγουν τυχαία αν θα συνεργαστούν ή θα αποστατήσουν. Μετά κάθε άτομο επαναλαμβάνει $2^{2L+2}$ παιχνίδια με όλους τους γείτονες του, βασισμένο στο χρωμόσωμα και την ιστορικότητα των παιχνιδιών. Για παράδειγμα αν το L=3 πρέπει να υπάρχουν 64 γενναιές και 256 επαναληπτικά παιχνίδια. Κάθε κόμβος προσθέτει την αμοιβή του στο συνολικό μετά από κάθε παιχνίδι. Μετά από ολα τα παιχνίδια, ο κάθε κόμβος διαιρεί τις συνολικές αμοιβές του με το βαθμό που του αντιστοιχεί για να φτάσει στην συνάρτηση καταλληλότητας. Κάθε κόμβος επιλέγει δύο άλλους κόμβους της γειτονιάς του σαν γονείς (περιλαμβάνοντας τον εαυτό του) που έχουν την μεγαλύτερη καταλληλότητα. Μετά από εφαρμογή διασταύρωσης και μετάλαξης, το χρωμόσωμα γονέας εξελίσσεται σε δύο χρωμοσώματα παιδιά. Μετά ένα απ' αυτά επιλέγεται σαν το χρωμόσωμα του κόμβου και η επόμενη γενναιά (επανάληψη) ξεκινά. Ως εκ τούτου, τα επιτυχημένα άτομα έχουν περισσότερες πιθανότητες να διατηρούν τις στρατηγικές τους και στους απογόνους.

Με το πέρας της έρευνας τους κατέληξαν πως αυξάνοντας την μνήμη δεν μπορεί να αυξηθεί το επίπεδο συνεργασίας και επίσης δεν μπορούν να μειωθούν οι εξελιγμένες στρατηγικές του συστήματος για ένα σταθερό όριο συνεργασίας. Αυτό συμβαίνει διότι όταν η μνήμη αυξάνεται, η μεγάλη αμοιβή που λαμβάνεται από την αποστασία υπάρχει ακόμα στο χρωμόσωμα των απογόνων και τα χαρακτηριστικά γονίδια χρειάζονται περισσότερους γύρους-γενναιές για να εξελιχθούν. Ωστόσο, παρά τα υψηλά κέρδη που προέρχονται από την αποστασία, η συνεργασία που προκύπτει από το σύστημα είναι πάνω από το 50%. Τα στατιστικά αποτελέσματα δείχνουν πως μετά από μερικές γενναιές, σε κάποιο συγκεκριμένο ιστορικό των παιχνιδιών, κάποια γονίδια τείνουν συγκεκριμένα στο 0 ή στο 1 αντίστοιχα. Τα περισσότερα απ' αυτά τείνουν στο 1 και συνεργάζονται.

Παρατηρήθηκε επίσης πως άλλοι κόμβοι έπαιρναν υψηλή αμοιβή ακολουθώντας στρατηγικές ταλάντωσης. Όταν το σύστημα φτάνει σε ένα σταθερό επίπεδο συνεργασίας, μόνο λίγες γενναιές χρησιμοποιούνται σε μεγάλο βαθμό. Όταν η αμοιβή της αποστασίας είναι χαμηλή, οι περισσότεροι κόμβοι διαλέγουν γονίδια συνεργασίας. Όταν η αμοιβή της αποστασίας είναι υψηλή, τότε ο γειτονικός κόμβος διαλέγει γονίδια που ταλαντώνονται μεταξύ αποστασίας και συνεργασίας. Επιπρόσθετα, παρατηρήθηκε πως σε διαφορετικά δίκτυα απ' ότι τα κλασσικά ετερογενή, η αποστασία δεν προέκυπτε μόνο σε κόμβους μικρού ή μεγάλου βαθμού, αλλά εξίσου στους περισσότερους κόμβους διαφορετικού βαθμού.

Γενικά οι Γενετικοί αλγόριθμοι είναι ένα είδος κατανεμημένου αλγορίθμου που εμπνεύστηκαν από την βιολογική εξέλιξη, ενώ τα εξελικτικά παίγνια μελετώνται απ' την οπτική γωνία της μακροσκοπικής συμπεριφοράς όπως είναι η συνεργασία η αυτό-οργάνωση που εμφανίζεται σε έναν πληθυσμό που αποτελείται από ανεξάρτητα, εγωιστικά άτομα. Συνδυάζοντας τους γενετικούς αλγορίθμους με τα εξελικτικά παίγνια, τα ανεξάρτητα άτομα διαλέγουν διαφορετικές στρατηγικές με τους διαφορετικούς τους γείτονες. Αυτό είναι περισσότερο σύμφωνο με την πραγματική κατάσταση όσον αφορά τις ομάδες ατόμων όπως είναι οι βιολογικές ομάδες ή κοινωνικές ομάδες κτλ. [12]

## 2.3 Εφαρμογή στην καθημερινή ζωή

Σαν ένα εφαρμοσμένο παράδειγμα του παιγνίου του διλήμματος του φυλακισμένου μπορούμε να αναφέρουμε το πρόβλημα των Internet Service Providers (ISPs) που απαιτείται να στείλουν την κίνηση, ή πιο απλά, πακέτα δεδομένων ο ένας στον άλλον.

Πιο συγκεκριμένα, στην δρομολόγηση της κίνησης που προέρχεται από έναν ISP με προορισμό έναν διαφορετικό ISP, η επιλογή της διαδρομής που πραγματοποιείται από τον πρώτο ISP επηρεάζει αυτόματα και τον φόρτο στον προορισμό του δεύτερου ISP.

Ας θεωρήσουμε, λοιπόν, πως οι ISPs, όπως φαίνονται στο παρακάτω σχήμα, έχουν ο καθένας το δικό τους ξεχωριστί δίκτυο (περιοχές ISP1 και ISP2). Τα 2 δίκτυα μπορούν να ανταλλάξουν την κίνηση μέσω 2 σημείων διαμετακόμισης, που αποκαλούνται σημεία ανταλλαγής κίνησης (peering points). Στην συγκεκριμένη περίπτωση τα απεικονίζουμε με C και S αντίστοιχα.

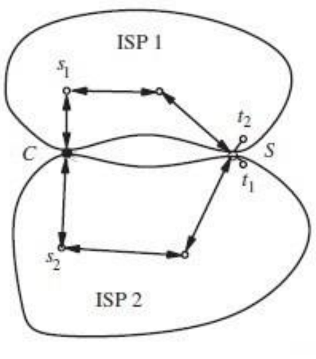



*Figure 3- ISP routing*

Στο παραπάνω αυτό σχεδιάγραμμα, φαίνονται επίσης 2 ζευγάρια σημείων προέλευσης και προορισμού $s_i$, $t_i$, το καθένα απ' αυτά περνάει ανάμεσα στους τομείς (domains). Υποθέτουμε ότι ο ISP 1 χρειάζεται να στείλει κίνηση από το σημείο $s_1$ του δικού του τομέα στο $t_1$ που ανήκει στον τομέα του ISP 2. Έτσι, ο ISP 1 έχει 2 επιλογές για να στείλει τα πακέτα, αντίστοιχες με τα 2 σημεία ανταλλαγής κίνησης (peering points). Ο ISP ουσιαστικά συμπεριφέρεται 'εγωιστικά' και προσπαθεί να ελαχιστοποιήσει τα δικά του κόστη, στέλνοντας τα πακέτα στο κοντινότερο σημείο ανταλλαγής κίνησης (C), δεδομένου ότι ο ISP με τον κόμβο προορισμού πρέπει να δρομολογήσει τα πακέτα ανεξάρτητα από το που εισέρχονται στον τομέα του. Το σημείο ανταλλαγής κίνησης C είναι πιο κοντά, οπότε, χρησιμοποιώντας αυτό το σημείο ανταλλαγής κίνησης, ο ISP 1 λαμβάνει 1 πόντο κόστους (με το να στείλει τα δεδομένα έναν κόμβο παραπέρα). Ενώ αν χρησιμοποιήσει το σημείο ανταλλαγής κίνησης S, τότε λαμβάνει κόστος 2 πόντων.

Εδώ, θα σημειωθεί πως ο μακρύτερος προορισμός S είναι περισσότερο άμεσος στη διαδρομή του προορισμού $t_1$. Ως εκ τούτου, η δρομολόγηση μέσω του S έχει ως αποτέλεσμα το συνολικό μικρότερο μονοπάτι. Το μήκος της διαδρομής μέσω του C είναι 4 ενώ μέσα από το S είναι 2, εφόσον ο προορισμός είναι πολύ κοντά στο S

Η κατάσταση που περιγράφηκε για την διαχείριση της κίνησης του ISP 1 από το $s_1$ στο $t_1$ είναι κατά έναν τρόπο ανάλογη των επιλογών των φυλακισμένων στο παίγνιο του διλήμματος τους φυλακισμένου. Για να ταυτιστεί απόλυτα η κατάσταση με το δίλημμα του φυλακισμένου, μπορούμε να υποθέσουμε πως συμμετρικά και ο δεύτερος ISP χρειάζεται να στείλει δεδομένα από το $s_2$ που ανήκει στον τομέα του, στο σημείο $t_2$ που ανήκει στον τομέα του ISP 1. Οι 2 επιλογές του δεύτερου ISP οδηγούν σε παίγνιο με πίνακα κόστους πανομοιότυπο με εκείνο που πειργράφει και το παίγνιο του διλήμματος του φυλακισμένου, με το C να αντιστοιχεί σε συνεργασία και το S σε αποστασία. [1]

# ΚΕΦΑΛΑΙΟ 3 – ΕΞΕΛΙΚΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ

## 3.1 Μεθευρετικοί και Εξελικτικοί Αλγόριθμοι

### 3.1.1 Εισαγωγή

Οι εξελικτικοί αλγόριθμοι (evolutionary algorithms – EA) είναι στοχαστικές μέθοδοι αναζήτησης, οι οποίες έχουν εφαρμοσθεί σε ποικίλα προβλήματα αναζήτησης, βελτιστοποίησης και μηχανικής μάθησης. Η βασική ιδέα των εξελικτικών αλγορίθμων είναι πως οι πιθανές λύσεις ομαδοποιούνται σε έναν πληθυσμό και επεξεργάζονται με ανταγωνιστικό τρόπο, ώστε να καταλήξουν σε μία βέλτιστη λύση. Στο παρακάτω σχήμα παρουσιάζεται μία ταξινόμηση των μεθόδων αναζήτησης.

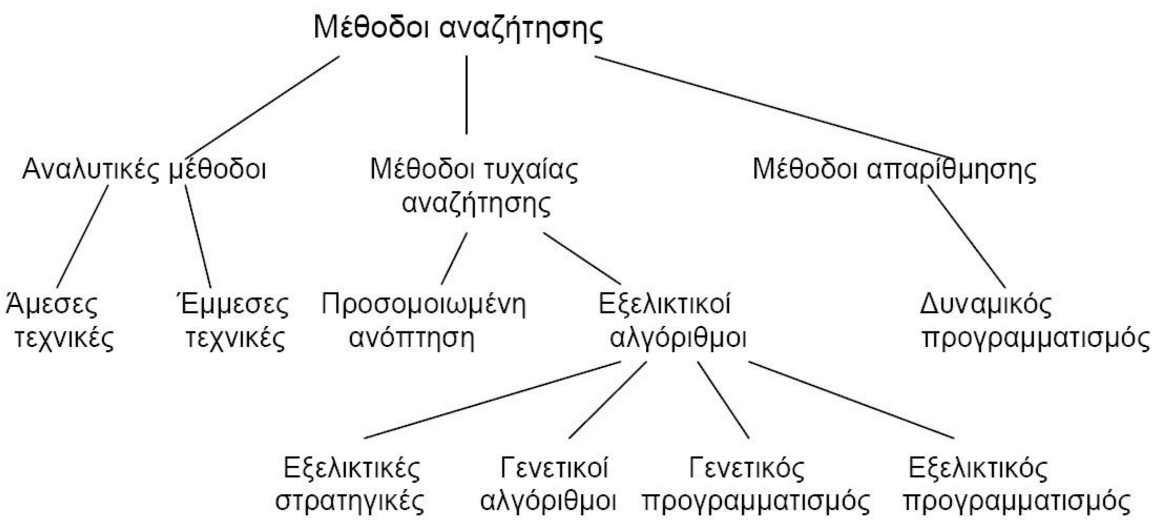



*Figure 4- Μέθοδοι αναζήτησης*

Στον τομέα της τεχνητής νοημοσύνης, ένας εξελικτικός αλγορίθμος (EA) είναι ένα υποσύνολο του εξελικτικού υπολογισμού, ένας γενικός μεθευριστικός αλγόριθμος βελτιστοποίησης βασισμένος στον πληθυσμό. Ένας EA χρησιμοποιεί ορισμένους μηχανισμούς εμπνευσμένους από τη βιολογική εξέλιξη: στην αναπαραγωγή, την μετάλλαξη, τον ανασυνδυασμό και την επιλογή. Οι υποψήφιες λύσεις στο πρόβλημα βελτιστοποίησης παίζουν το ρόλο των ατόμων σε έναν πληθυσμό και η συνάρτηση καταλληλότητας καθορίζει το περιβάλλον μέσα στο οποίο «ζουν» οι λύσεις. Η εξέλιξη του πληθυσμού, στη συνέχεια, λαμβάνει χώρα μετά την επανειλημμένη εφαρμογή των παραπάνω φορέων.

Η τεχνητή εξέλιξη (artificial evolution – AE) περιγράφει μια διαδικασία που βασίζεται σε επιμέρους εξελικτικούς αλγορίθμους. Οι EA είναι τα επιμέρους συστατικά που συμμετέχουν σε μια τεχνητή εξέλιξη.

Οι εξελικτικοί αλγόριθμοι συχνά έχουν πολύ καλές προσεγγιστικές λύσεις σε όλα τα είδη των προβλημάτων, επειδή στην ιδανική περίπτωση δεν πραγματοποιείται καμία υπόθεση για το υποκείμενο πεδίο λειτουργίας. Αυτό γενικότερα φαίνεται από τις επιτυχίες σε τομείς αρκετά διαφορετικούς, όπως η μηχανική, η τέχνη, η βιολογία, η οικονομία, το μάρκετινγκ, η γενετική, η ρομποτική, οι κοινωνικές επιστήμες, η φυσική, η πολιτική και η χημεία.

Εκτός από τη χρήση τους ως μαθηματικό εργαλείο, οι εξελικτικοί αλγόριθμοι και οι εξελικτικοί υπολογισμοί έχουν χρησιμοποιηθεί σαν πειραματικό πλαίσιο μέσα στο οποίο γίνεται προσπάθεια για επικύρωση εκείνων των θεωριών όσον αφορούν τη βιολογική εξέλιξη και τη φυσική επιλογή, κυρίως μέσα από την εργασία στον τομέα της τεχνητής ζωής. Τεχνικές των εξελικτικών αλγορίθμων που εφαρμόζονται για την μοντελοποίηση της βιολογικής εξέλιξης είναι γενικά περιορισμένες σε εξερευνήσεις των μικρο-εξελικτικών διεργασιών. Ωστόσο, κάποιες προσομοιώσεις σε υπολογιστή, όπως τα Tierra και Avida, προσπαθούν να μοντελοποιήσουν την μικρο-εξελικτική δυναμική.

Στις περισσότερες από τις πραγματικές εφαρμογές των εξελικτικών αλγορίθμων, η υπολογιστική πολυπλοκότητα αποτελεί απαγορευτικό παράγοντα. Η προσεγγιστική καταλληλότητα (fitness approximation) είναι μία από τις τεχνικές για την αντιμετώπιση αυτού του προβλήματος. [9]

### 3.1.2 Βασικές Έννοιες

Έτσι όπως γνωρίζουμε απ' την ιστορία του αντικειμένου, έχουμε αρκετές εκδοχές εξελικτικών αλγορίθμων. Το σίγουρο είναι πως η βασική ιδέα πίσω απ' αυτές τις τεχνικές είναι κοινή. Ξεκινάμε από έναν αρχικό πληθυσμό ατόμων, στον οποίο η πίεση του περιβάλλοντος προκαλεί την φυσική επιλογή που ουσιαστικά σημαίνει την επιβίωση του καταλληλότερου. Αυτό έχει ως αποτέλεσμα την αύξηση καταλληλότητας (fitness) του πληθυσμού. Είναι δυνατόν, λοιπόν, να δημιουργηθούν τυχαία σύνολα υποψηφίων λύσεων, εάν έχουμε μια συνάρτηση ποιότητας που πρέπει να μεγιστοποιηθεί. Έπειτα, η συνάρτηση ποιότητας μπορεί να εφαρμοστεί σαν μια γενική μέτρηση της καταλληλότητας (όσο υψηλότερη η μέτρηση, τόσο το καλύτερο).

Βασιζόμενοι σε αυτή την έννοια της καταλληλότητας, κάποιες από τις καλύτερες υποψήφιες λύσεις επιλέγονται έτσι ώστε να δημιουργήσουν την επόμενη γενιά, εφαρμόζοντας επανασυνδυασμό (rexombination) και/ή μετάλλαξη (mutation) σε αυτές. Πιο συγκεκριμένα, ο επανασυνδυασμός είναι ένας τελεστής ο οποίος εφαρμόζεται σε δύο ή περισσότερες επιλεγμένες υποψήφιες λύσεις (αποκαλούμενες γονείς) και η μετάλλαξη εφαρμόζεται σε μια υποψήφια λύση και καταλήγει σε μια νέα υποψήφια λύση. Εφαρμόζοντας επανασυνδυασμό και/ή μετάλλαξη οδηγούμαστε σε ένα νέο σύνολο υποψήφιων λύσεων (απόγονοι – offspring) που συναγωνίζεται, βασιζόμενο στην καταλληλότητα, τα παλιά σύνολα λύσεων για μια θέση στην επόμενη γενιά. Αυτή η διαδικασία μπορεί να επαναλαμβάνεται μέχρι μια υποψήφια λύση με επαρκή καταλληλότητα να βρεθεί, ή μέχρι να φτάσει η διαδικασία στο υπολογιστικό όριο ενός προηγούμενου συνόλου.

Σκοπός των παραπάνω είναι με συνδυασμένη εφαρμογή επανασυνδυασμών, μεταλλάξεων και επιλογών των καταλληλότερων λύσεων να οδηγούμαστε στην βελτίωση των τιμών καταλληλότητας σε διαδοχικούς πληθυσμούς. Στην περίπτωση που μια διαδικασία είναι προς βελτιστοποίηση, επιτυγχάνεται προσέγγιση των βέλτιστων δυνατών τιμών, οι οποίες λαμβάνουν όλο και κοντινότερες τιμές κατά την εξέλιξη της διαδικασίας. Εναλλακτικά, εξέλιξη παρατηρείται συχνά σαν μια διαδικασία προσαρμογής. Από αυτή την οπτική γωνία, η καταλληλότητα δεν φαίνεται σαν μια αντικειμενική συνάρτηση που πρέπει να βελτιστοποιηθεί, αλλά σαν μια έκφραση των περιβαλλοντικών απαιτήσεων. Βελτιστοποιώντας αυτές τις απαιτήσεις όλο και περισσότερο, προκύπτει μια αυξημένη βιωσιμότητα, ανακλώμενη σε ένα μεγαλύτερο αριθμό απογόνων. Η εξελικτική διαδικασία κάνει τον πληθυσμό να προσαρμόζεται στο περιβάλλον όλο και καλύτερα.

Αξίζει να σημειωθεί ότι πολλά στοιχεία μιας τέτοιας εξελικτικής διαδικασίας είναι στοχαστικά. Κατά τη διάρκεια της επιλογής, τα άτομα με τις καλύτερες συνθήκες καταλληλότητας έχουν περισσότερες ευκαιρίες να επιλεχθούν σε σχέση με τα άτομα που υστερούν σε επίπεδα καταλληλότητας, αλλά τυπικά ακόμα και τα αδύναμα άτομα έχουν την δυνατότητα να γίνουν γονείς ή ακόμα να επιβιώσουν. Για τον επανασυνδυασμό των ατόμων η επιλογή των τμημάτων που θα επανασυνδυαστούν είναι τυχαία. Παρομοίως για την μετάλλαξη, τα κομμάτια που θα μεταλλαχθούν εντός μίας υποψήφιας λύσης, καθώς και τα νέα κομμάτια που θα τα αντικαταστήσουν, επιλέγονται τυχαία. [10]

### 3.1.3 Επιλογή του ιδανικού Αλγορίθμου

Κατά την σχεδίαση ενός μεθευρετικού αλγορίθμου, πρέπει πάντα να λαμβάνονται υπόψι η σε πλάτος αναζήτηση του χώρου λύσεων, η οποία ονομάζεται διάχυση της αναζήτησης (diversification) και η σε βάθος αναζήτηση των καλύτερων λύσεων που ονομάζεται εντατικοποίηση της αναζήτησης (intensification).

Η εντατικοποίηση απαιτείται για να βρίσκονται οι πιο ελπιδοφόρες περιοχές αναζήτησης γύρω από κάποιες καλές λύσεις και από την άλλη πλευρά, η διάχυση μιας αναζήτησης είναι επίσης απαραίτητη για να δοκιμάζονται νέες, ανεξερεύνητες περιοχές. Ο στόχος είναι να μην υπάρχουν περιοχές που δεν έχουν εξερευνηθεί καθόλου έτσι ώστε να μην περιοριστεί ο αλγόριθμος μόνο σε κάποιες περιοχές του χώρου αναζήτησης.

Θα μπορούσαμε να χωρίσουμε τους μεθευρετικούς αλγορίθμου στις παρακάτω κατηγορίες βάση κριτηρίων που οφείλονται στη δομή τους.

- Αλγόριθμοι εμπνευσμένοι από τη φύση ή όχι. Οι γενετικοί αλγόριθμοι όπως και οι αλγόριθμοι τεχνητού ανοσοποιητικού συστήματος που έχουν εμπνευστεί από διαδικασίες της βιολογίας, αποτελούν το κλασσικότερο παράδειγμα. Επίσης έχουμε τους αλγορίθμους των μυρμηγκιών, των μελισσών και του σμήνους σωματιδίων που βασίζονται στην αλληλεπίδραση σμήνους εντόμων ή γενικότερα σωματιδίων.
- Αλγόριθμοι που χρησιμοποιούν μνήμη από προηγούμενες επαναλήψεις και αλγόριθμοι που δεν θυμούνται τις λύσεις των προηγούμενων επαναλήψεων. Κλασσικά παραδείγματα της πρώτης κατηγορίας είναι η μέθοδος της Περιορισμένης αναζήτησης που χρησιμοποιεί μνήμη μικρής, μεσαίας ή και μακράς διάρκειας, ενώ της δεύτερης είναι η Προσομοιωμένη ανόπτηση και η Διαδικασία Άπληστης Τυχαιοποιημένης Προσαρμοστικής Αναζήτησης.
- Ντετερμινιστικοί ή στοχαστικοί αλγόριθμοι. Με τους πρώτους επιλύονται προβλήματα βελτιστοποίησης με τη χρήση ντετερμινιστικών αποφάσεων όπως τοπική αναζήτηση, και περιορισμένη αναζήτηση. Απ’ την άλλη μεριά στους στοχαστικούς υπάρχει ένας αριθμός από τυχαιοποιημένους κανόνες που χρησιμοποιούνται όπως γενετικοί αλγόριθμοι, αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων, κα.
- Αλγόριθμοι βασισμένοι στον πληθυσμό λύσεων και αλγόριθμοι βασισμένη σε μία λύση. Οι τελευταίοι, βασίζονται σε μία λύση όπως είναι η τοπική αναζήτηση και η προσομοιωμένη ανόπτηση και διαχειρίζονται αυτή την λύση κατά την διάρκεια των επαναλήψεων. Αντίθετα στους αλγορίθμους που βασίζονται σε πληθυσμό λύσεων, χρησιμοποιούν ένα μεγάλο πληθυσμό λύσεων κατά τη διάρκεια των επαναλήψεων. Χαρακτηριστικά παραδείγματα είναι ξανά οι γενετικοί αλγόριθμοι όπως και ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων.
- Τέλος οι επαναληπτικοί αλλά και άπληστοι αλγόριθμοι. Στους επαναληπτικούς ξεκινάμε μια ολοκληρωμένη λύση ή έναν πληθυσμό λύσεων η οποία τροποποιείται κατά τη διάρκεια των επαναλήψεων ενώ αντίθετα στους άπληστους αλγορίθμους ξεκινάμε από μια κενή λύση η οποία συμπληρώνεται βηματικά.

Πολύ σημαντικό σημείο κατά την σχεδίαση του αλγορίθμου παίζει και η αναπαράσταση. Αν θα είναι με χρήση πραγματικών τιμών (real time representation) ή θα είναι δυαδική που είναι κατάλληλη για διακριτά προβλήματα που μπορούν να πάρουν μόνο δύο δυνατές τιμές όπως για παράδειγμα το πρόβλημα του διλήμματος του φυλακισμένου που μπορεί να λάβει δύο εναλλακτικές τιμές – 1 για Συνεργασία και 0 για Αποστασία. Σε άλλες περιπτώσεις πρέπει να χρησιμοποιηθούν διακριτές τιμές (discrete values representation), κατάλληλη επιλογή για ακέραια προβλήματα που μπορούν να πάρουν τιμές στο σύνολο των ακεραίων. [4]

Η επίλυση ενός προβλήματος βελτιστοποίησης και ιδιαίτερα συνδυαστικής βελτιστοποίησης γίνεται ολοένα και δυσκολότερη όσο αυξάνει το μέγεθος του προβλήματος και, πολλές φορές, το να προσπαθούμε να βρούμε την ολικά βέλτιστη λύση σε λογικό χρόνο είναι πρακτικά αδύνατο. Για να επιλυθούν προβλήματα αυτή της μορφής, συχνά καταφεύγουμε σε διαφορετικές τεχνικές που μας οδηγούν σε μια όχι βέλτιστη, αλλά αρκετά ικανοποιητική, λύση. Μια λύση ενός ευρετικού αλγορίθμου γίνεται αποδεκτή αν ικανοποιεί κάποια κριτήρια όπως η ποιότητα της λύσης,δηλαδή η λογική πάνω στην οποία στηρίζονται οι κανόνες του ευρετικού αλγορίθμου που χρησιμοποιήθηκαν για να οδηγηθούμε στη λύση. Για κάθε πρόβλημα βελτιστοποίησης δεν υπάρχει μονάχα ένας ευρετικός αλγόριθμος που να δίνει τη βέλτιστη λύση, αλλά έχουν αναπτυχθεί αρκετοί αλγόριθμοι, οι οποίοι συγκρινόμενοι μεταξύ τους, οδηγούν σε ολοένα και καλύτερες λύσεις.

Το πιο σημαντικό σημείο είναι εκείνο της ποιότητας της λύσης. Σε μερικά προβλήματα, όπως αναφέρθηκε, είναι αδύνατο να βρεθεί η βέλτιστη λύση για κάποιο πρόβλημα σε ικανοποιητικό χρόνο. Ένας απλός τρόπος για να επιβεβαιωθεί η ποιότητα της λύσης που προκύπτει απ’ τον αλγόριθμο είναι να δημιουργήσουμε μικρότερα παραδείγματα από αυτό που θέλουμε να λύσουμε, τα οποία μπορούμε να τα λύσουμε με κάποια ακριβή μέθοδο, και να δούμε πόσο κοντά στο βέλτιστο είναι η λύση που παίρνουμε με τη χρήση του ευρετικού αλγορίθμου. Ένας άλλος τρόπος είναι η δημιουργία ενός φράγματος αποδεκτής λύσης με την επίλυση ενός χαλαρωμένου προβλήματος, για παράδειγμα με τη διαδικασία διακλάδωσης και οριοθέτησης (branch and bound), και έτσι όλες οι λύσεις που θα πάρουμε με τη χρήση του ευρετικού αλγορίθμου και των οποίων η τιμή τους δεν θα παραβιάζει την τιμή του φράγματος θα είναι ικανοποιητικές.

Οι ευρετικοί αλγόριθμοι δουλεύουν γρήγορα και αποτελεσματικά. Η ποιότητα της λύσης, όμως, που παρέχουν είναι ένα εντελώς διαφορετικό θέμα. Πριν από την εμφάνιση της μεθόδου ανάλυσης των προσεγγιστικών αλγορίθμων η απόδοση ενός ευρετικού αλγορίθμου κρινόταν από δοκιμαστικά τρεξίματα σε ένα σύνολο από παραδείγματα προβλημάτων αναφοράς και συγκρινόταν με την απόδοση άλλων αλγορίθμων που επέλυαν το ίδιο πρόβλημα. Αυτή η μέθοδος είχε πάντοτε κάποια προβλήματα. Αρχικά, το σύνολο των προβλημάτων που δοκιμαζόταν ήταν ένα μικρό δείγμα αλλά συνήθως όχι αντιπροσωπευτικό όλων των παραδειγμάτων ενός συγκεκριμένου προβλήματος. Επίσης, μια καλή βελτίωση της λύσης σε κάποιο παράδειγμα σε σύγκριση με κάποιον άλλο αλγόριθμο δεν σημαίνει και απαραίτητα μια γενική βελτίωση της λύσης σε οποιοδήποτε παράδειγμα. Είναι ξεκάθαρο ότι χρειάζεται μια ανάλυση για το πόσο καλή είναι η λύση που μπορεί να μας δώσει ένας αλγόριθμος απληστίας και γενικότερα ένας ευρετικός αλγόριθμος.

Η δημιουργία μιας αρχικής λύσης είναι μια πολύ σημαντική διαδικασία για την αποτελεσματικότητα του αλγορίθμου. Αν η αρχική λύσης είναι καλή μπορεί να οδηγήσεις πιο γρήγορα τον αλγόριθμο σε σύγκλιση. Από την άλλη μεριά μια πολύ κακή αρχική λύση θεωρητικά θα αργήσει να συγκλίνει σε κάποια καλή λύση. Συνήθως υπάρχουν δύο τρόποι για να δημιουργηθεί μια αρχική λύση . Ο πρώτος τρόπος είναι με τη χρήση μιας τυχαιοποιημένης διαδικασίας κατάλληλα προσαρμοσμένης στο πρόβλημα που επιλύουμε και ο δεύτερος τρόπος με τη χρήση κάποιου αλγορίθμου απληστίας. Ο τελευταίος αλγόριθμος προσπαθεί να οδηγήσει σε μια εφικτή λύση του προβλήματος, αλλά πολλές φορές χρειάζεται πάρα πολύ μεγάλο χρόνο γιατί είναι μυωπικός αλγόριθμος, δηλαδή βλέπει μόνο μπροστά. Από τη μέθοδο που θα χρησιμοποιήσουμε στη φάση της τοπικής αναζήτησης εξαρτάται και ο τρόπος που θα επιλέξουμε για να δημιουργήσουμε τις αρχικές λύσεις. Έτσι, για παράδειγμα, αν έχουμε ένα συνεχές πρόβλημα είναι καλύτερο να δημιουργήσουμε τις αρχικές λύσεις με τυχαίο τρόπο. Αντίθετα, σε ένα πρόβλημα συνδυαστικής βελτιστοποίησης είναι πιο αποτελεσματικό να δημιουργήσουμε τις αρχικές λύσεις με ένα αλγόριθμο απληστίας. Επίσης, αν έχουμε ένα αλγόριθμο που χρησιμοποιεί ένα πληθυσμό από λύσεις για να βελτιώσει τις λύσεις του κατά τη διάρκεια των επαναλήψεων, τότε είτε θα δημιουργήσουμε τις αρχικές λύσεις με τυχαίο τρόπο είτε θα πρέπει να χρησιμοποιήσουμε πολλούς διαφορετικούς αλγορίθμους απληστίας. [4]

### 3.1.4 Μεθευρετικοί Αλγόριθμοι

Η μέθοδος τοπικής αναζήτησης που χρησιμοποιούμε συνήθως θα συγκλίνει γύρω από κάποιο τοπικό ελάχιστο. Επιπλέον, ο αλγόριθμος είναι πολύ ευαίσθητος στην αρχική λύση, δηλαδή αν η αρχική λύση είναι πολύ καλή και βρίσκεται πολύ κοντά σε κάποιο τοπικό ελάχιστο, τότε με τη διαδικασία της τοπικής αναζήτησης θα βρεθεί πολύ εύκολα το τοπικό ελάχιστο, αλλά δεν ξέρουμε αν μετά ο αλγόριθμος παγιδευτεί σε αυτό το τοπικό ελάχιστο και δεν μπορεί να βελτιώσει άλλο τη λύση. Για να επιλυθεί αυτό το πρόβλημα έχουν προταθεί πολλοί αλγόριθμοι που βοηθούν την αναζήτηση να ξεφύγει από κάποιο τοπικό ελάχιστο. Οι αλγόριθμοι αυτοί ονομάζονται μεθευρετικοί αλγόριθμοι. [4]

Αν χωρίσουμε τους αλγορίθμους σε 2 ομάδες, θα μπορούσαμε να πούμε πως έχουμε εκείνους που χρησιμοποιούν μία λύση και κάνουν αναζήτηση στη γειτονία της λύσης και εκείνους τους αλγόριθμους που έχουν έναν πληθυσμό από λύσεις και προσπαθούν να κάνουν αναζήτηση σε όλο το χώρο των λύσεων.

Στην περίπτωση που ο αλγόριθμος κάνει αναζήτηση γειτονιάς, έχει το σημαντικό πλεονέκτημα να εκμεταλεύεται (exploitation) και να εντατικοποιεί (Intensification) την περιοχή των λύσεων. Οι αλγόριθμοι που έχουν έναν πληθυσμό λύσεων, αντίθετα, έχουν πλεονεκτήματα στην εξερεύνηση (exploration) και στη διάχυση (diversification) σε όλη την περιοχή τών λύσεων. Η ουσία είναι να αναζητούμε πάντα τις στρατηγικές που συνδυάζουν τους δύο παραπάνω τύπους αλγορίθμων.

Σε συνέχεια των παραπάνω, θα μπορούσαμε να χωρίσουμε τους μεθευρετικούς αλγόριθμους που επικεντρώνονται στην αναζήτηση γύρω από κάποιο σημείο που έχει βρεθεί τοπικό ελάχιστο σε τέσσερις βασικές κατηγορίες.

Η πιο χαρακτηριστική κατηγορία είναι οι αλγόριθμοι πολυεναρκτήριας τοπικής αναζήτησης (multistart local search), επαναληπτικής τοπικής αναζήτησης (iterated local search) και διαδικασίας άπληστης τυχαιοποιημένης προσαρμοστικής αναζήτησης (greedy randomized adaptive search procedure (GRASP)). Όλες οι παραπάνω έχουν το κοινό πως ξεκινούν από διαφορετικές αρχικές λύσεις.

Δύο ακόμα χαρακτηριστικοί αλγόριθμοι είναι η προσομοιωμένη ανόπτηση (simulated annealing) και η περιορισμένη αναζήτηση (tabu search). Αυτοί ανήκουν στην κατηγορία αλγορίθμων που δέχονται γειτονικές κινήσεις χωρίς να βελτιώνουν ουσιαστικά τη λύση. Σε αυτές τις μεθόδους, μία κίνηση που δεν βελτιώνει τη λύση μπορεί να γίνει υπό κάποιες συνθήκες αποδεκτή. Με αυτό τον τρόπο μπορεί σε μία από τις επόμενες κινήσεις να ξεφύγουμε από το τοπικό ελάχιστο και να οδηγηθούμε σε κάποιο επόμενο τοπικό ελάχιστο το οποίο να είναι καλύτερο από το τρέχον.

Επόμενη κατηγορία αποτελούν ο αλγόριθμος μεταβλητής γειτονιάς αναζήτησης (variable neighborhood Search (VNSI)) και ο αλγόριθμος επέκτασης της γειτονιάς αναζήτησης (Expanding Neighborhood Search (ENSI)). Αυτοί οι δύο αλγόριθμοι έχουν το πλεονέκτημα πως, όταν κολλήσουν σε ένα τοπικό ελάχιστο, αλλάζουν τον αλγόριθμο που χρησιμοποιούν για την αναζήτηση σε γειτονικά σημεία του χώρου λύσεων.

Τέλος, έχουμε τον αλγόριθμο καθοδηγούμενης τοπικής αναζήτησης, στον οποίο αλλάζει η αντικειμενική συνάρτηση ή κάποια από τα δεδομένα του προβλήματος έτσι ώστε να οδηγηθούμε στην καλύτερη λύση. [4]

## 3.2 Αλγόριθμοι εμπνευσμένοι απ’ τη φύση

### 3.2.1 Εισαγωγή

Όπως αναφέρθηκε στις προηγούμενες παραγράφους οι γενετικοί αλγόριθμοι είναι διαδικασίες αναζήτησης που βασίζονται στους μηχανισμούς της φυσικής επιλογής και της γενετικής.

Με αφορμή τους γενετικούς αλγορίθμους, τα τελευταία χρόνια έχουν αναπτυχθεί στο χώρο των εφαρμοσμένων μαθηματικών και της πληροφορικής, ποικίλες μέθοδοι βελτιστοποίησης που έχουν εμπνευστεί από μηχανισμούς που συναντούμε στη φύση σε καθημερινή βάση. Παράλληλα έχουμε και μια άλλη ομάδα αλγορίθμων που έχουν προκύψει από την μελέτη του ανοσοποιητικού συστήματος του ανθρώπου.

Παραδείγματα τέτοιων αλγορίθμων αποτελούν: η Βελτιστοποίηση Αποικίας Μυρμηγκιών (Ant Colony Optimization), η Βελτιστοποίηση Σμήνους Σωματιδίων (Particle Swarm Optimization), η Αναζήτηση της Μουσικής Αρμονίας (Harmony Search Algorithm), η Επιλογή Κλώνων (Clone Selection Algorithm), η Βελτιστοποίηση Σμήνους Πυγολαμπίδων (Glowworm Swarm Based Optimization Algorithm), η Προσομοίωση Βατράχων μέσω Αλμάτων (Shuffled Frog Leaping Algorithm), η Προσομοίωση της Διαδικασίας Εύρεσης Τροφής (foraging behaviour) κ.α.

Οι πρώτοι δύο είναι αυτοί που αποτελούν τα χαρακτηριστικότερα και δημοφιλέστερα παραδείγματα της κατηγορίας αλλά τα τελευταία χρόνια πραγματοποιούνται πολλές προσεγγίσεις προβλήματων με τη χρήση διαφόρων εκ των παραπάνω αλγορίθμων.

Όλοι οι αλγόριθμοι της κατηγορίας είναι συστήματα που μιμούνται τη αντίστοιχη συμπεριφορά που συναντάται στη φύση. Για παράδειγμα τα μυρμήγκια αναπτύσσουν μια τεχνική για να βρουν τη συντομότερη διαδρομή από τη φωλιά τους προς την πηγή της τροφής τους και αντίθετα. Ξεκινούν την αναζήτηση της τροφής γύρω από την πηγή με τυχαίο τρόπο και καθώς κινούνται αφήνουν μια ποσότητα μιας ουσίας που ονομάζεται φερομόνη και με αυτό τον τρόπο 'σημειώνουν' το μονοπάτι που έχουν διανύσει. Η ποσότητα της φερομόνης στο κάθε μονοπάτι εξαρτάται από την απόσταση, την ποιότητα και την ποσότητα της τροφής που βρέθηκε. Αυτό έχει ως αποτέλεσμα το επόμενο μυρμήγκι που θα φύγει από τη φωλιά του να είναι πολύ πιθανό να ακολουθήσει τη φερομόνη που θα υπάρχει σε κάποιο μονοπάτι, αφήνοντας μια δική τους ποσότητα φερομόνης σε αυτό το μονοπάτι. Καθώς η ποσότητα φερομόνης στο συγκεκριμένο όλο και θα αυξάνεται, όλο και περισσότερα μυρμήγκια ακολουθούν τελικά το ίδιο μονοπάτι. Παράλληλα καθώς περνάει η ώρα, η φερομόνη ελαττώνεται από τα υπόλοιπα μη επιλέξιμα μονοπάτια, με αποτέλεσμα όσο περνάει η ώρα τα μυρμήγκια να ακολουθούν τελικά το ίδιο μονοπάτι, που είναι και η βέλτιστη ή τουλάχιστον η σχεδόν βέλτιστη λύση.

Μοντελοποιώντας την παραπάνω συμπεριφορά των μυρμηγκιών σε αλγόριθμο δημιουργούμε ένα πρόβλημα εύρεσης μονοπατιού ελαχίστου κόστους. Κάθε μυρμήγκι αποτελεί και μια λύση για το πρόβλημα. Ο αλγόριθμος θα αποτελείται από έναν αριθμό επαναλήψεων όπου σε κάθε επανάληψη το κάθε μυρμήγκι ξεκινάει και κατασκευάζει μια λύση με βάση πάντα την εμπειρία που έχει αποκτήσει από τις λύσεις που δημιούργησαν τα προηγούμενα μυρμήγκια.

Παρόμοια φιλοσοφία ακολουθούν και οι υπόλοιποι εμπνευσμένοι από τη φύση αλγόριθμοι που αναφέρθηκαν παραπάνω. Στην παρούσα εργασία θα εστιάσουμε στην μοντελοποίηση του Αλγορίθμου Βελτιστοποιήσης Σμήνους Σωματιδίων (PSO) και εναλλακτικών του, για να δημιουργήσουμε έναν ανταγωνιστικό αντίπαλο έναντι των γνωστών στρατηγικών όσον αφορά το παίγνιο του διλήμματος του φυλακισμένου.

### 3.2.2 Νοημοσύνη Σμήνους

Οι περισσότεροι από τους αλγορίθμους της προηγούμενης παραγράφου στηρίζονται στη μοντελοποίηση συστημάτων που βασίζονται στη συνεργασία μεταξύ ατόμων από ένα πληθυσμό και στις περισσότερες περιπτώσεις τα άτομα του πληθυσμού δεν έχουν πλήρη γνώση του προς επίλυση προβλήματος.

Οι εκάστοτε πληροφορίες που έχει στη διάθεση του το κάθε μέλος του πληθυσμού, χρησιμοποιούνται αρχικά για να εκτελέσει την ενέργεια ή κίνηση του και στη συνέχεια για να ενημερώσει και να μοιραστεί την γνώση που απέκτησε στα υπόλοιπα μέλη του πληθυσμού.

Συνήθως το σύνολο του πληυθσμού ονομάζεται σμήνος (swarn) και όλες αυτές οι μέθοδοι ονομάζονται αλγόριθμοι νοημοσύνης σμήνους (swarm intelligence).

Όπως αναφέρθηκε παραπάνω, το μέλος του πληθυσμού, για παράδειγμα το έντομο, δεν χρειάζεται να γνωρίζει την συνολική διαδικασία που πρέπει να ακολουθηθεί για την ολοκλήρωση μιας εργασίας ούτε το στάδιο το οποίο βρίσκεται η διαδικασία. Η μοναδική γνώση που απαιτείται είναι αυτή που το καθιστά ικανό να φέρει εις πέρας την εργασία που του έχει ανατεθεί. Η ικανότητα των εντόμων να έχουν μια συλλογική συμπεριφορά και να ολοκληρώνουν μια εργασία χωρίς να ξέρουν όλα τα στάδια ονομάζεται στιγμεργία (stigmergy). Ο όρος stigmergy προέρχεται από τις 2 ελληνικές λέξεις, στίγμα και έργο. Σύμφωνα με αυτήν, το κάθε άτομο του πληθυσμού παρατηρεί κάποια σημάδια και αυτά τα σημάδια το οδηγούν για να ολοκληρώσει την εργασία που του έχει ανατεθεί.

Σε αυτά τα αποκεντρωμένα συστήματα εργασιών, κάθε μέλος της ομάδας έχει μόνο μια τοπική γνώση για το περιβάλλον του και δεν έχει την ικανότητα να καταλάβει το συνολικό πρόβλημα που πρέπει να επιλυθεί. [4]

### 3.2.3 Αλγόριθμος Βελτιστοποίησης Σμήνους Σωματιδίων

Ο Αλγόριθμος Βελτιστοποίησης Σμήνους Σωματιδίων (Particle Swarm Optimization - PSO) είναι ένας απ’ τους πιο δημοφιλείς αλγόριθμους που προτάθηκε από τους Kennedy και Eberhard για να προσομοιώσει την κοινωνική συμπεριφορά κάποιων ομάδων όπως το πέταγμα των πουλιών σε μορφή σμήνους, και την ομαδική κίνηση των ψαριών.

Αρχικά η ιδέα ήταν να προσομοιωθεί γραφικά η κίνηση που κάνει ένα σμήνος πουλιών, να βρεθούν και να αναλυθούν οι κανόνες που οδηγούν τα πουλιά να κινούνται ως ένα σμήνος και πως ξαφνικά αλλάζουν κατεύθυνση χωρίς να καταστρέφεται ο σχηματισμός που είχαν αρχικά.

Ο αλγόριθμος χρησιμοποιεί τη φυσική κίνηση των ατόμων ενός πληθυσμού μέσα στο σμήνος και έχει πολυ ευέλικτο και καλά ισορροπημένο μηχανισμό που προσαρμόζεται στις ολικές και τοπικές ικανότητες εξερεύνησης των ατόμων του σμήνους. Οι αλλαγές ενός ατόμου μέσα στο σμήνος επηρεάζονται από την εμπειρία και τις γνώσεις των γειτονικών του σωματιδίων και ως εκ τούτου ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων μπορεί να θεωρηθεί ως ένας συμβιωτικός συνεργατικός αλγόριθμος.

Έχει αποδειχτεί πως η συγκεκριμένη μέθοδος είναι πολύ χρήσιμη σε ένα μεγάλο σύνολο προβλημάτων εξαιτίας της εύκολης υλοποίησης της και των πολύ καλών αποτελεσμάτων που έχει δώσει σε προβλήματα που έχουν συνεχείς μεταβλητές αλλά τελευταία και με εφαρμογές σε προβλήματα συνδυαστικής βελτιστοποίησης.

Σε σύγκριση με τους γενετικούς, ο PSO εμπεριέχει μνήμη, πράγμα που σημαίνει πως η γνώση από προηγούμενες καλές λύσεις κληρονομείται στις επόμενες γενιές και δεν χάνεται από επανάληψη σε επανάληψη. Επίσης παρατηρείται μια μεγάλη συνεργασία ανάμεσα στα σωματίδια του σμήνους αφού τα μέλη της ομάδας συνεργάζονται μεταξύ τους στην κατασκευή των λύσεων.

Για την δημιουργία του αρχικού σμήνους, δημιουργείται ένας αρχικός αριθμός απο σωματίδια που το κάθε ένα έχει μια συγκεκριμένη θέση στο χώρο λύσεων και κινείται με συγκεκριμένη ταχύτητα. Αντιστοιχίζοντας με έναν γενετικό αλγόριθμο θα είχαμε πως το σμήνος (swarm) αντιστοιχεί στον πληθυσμό (population) και το σωματίδιο (particle) αντιστοιχεί σε ένα άτομο (individual).

Οι συγκεκριμένες λύσεις του προβλήματος αντιπροσωπεύονται με τις θέσεις των σωματιδίων. Δεδομένου αυτού, αν N είναι το μέγεθος του πληθυσμού και n ο αριθμός των διαστάσεων, $x_{ij}$ θα είναι οι λύσεις με i=1,2,…,N και j=1,2,…n. Η απόδοση της θέσης-λύσης εκτιμάται από την προκαθορισμένη συνάρτηση καταλληλότητας $f(x_{ij})$ (fitnsess function). Επίσης η ταχύτητα $u_{ij}$ αντιπροσωπεύει τις αλλαγες που θα γίνουν έτσι ώστε να κινηθεί το σωματίδιο από μια θέση σε μια άλλη. Η κατεύθυνση που θα κινηθεί το σωματίδιο υπολογίζεται από την δυναμική αλληλεπίδραση της δικής του εμπειρίας αλλά και της εμπειρίας ολόκληρου του σμήνους.

Κατά την κίνηση τους τα σωματίδια, μπορούν είτε να επιλέξουν να ακολουθήσουν μια δική τους διαδρομή, είτε να κινηθούν προς την βέλτιστη θέση που είχαν κατά τη διάρκεια των επαναλήψεων $pbest_{ij}$, είτε να κινηθούν προς την θέση που έχει το βέλιστο σωματίδιο στον πληθυσμό ($gbest_j$)

Οι ταχύτητες και οι θέσεις των σωματιδίων όταν έχουμε συνεχείς μεταβλητές, υπολογίζονται βάση των παρακάτων εξισώσεων:

$$u_{ij}(t+1) = u_{ij}(t) + c_1 rand_1(pbest_{ij} - x_{ij}(t)) + c_2 rand_2(gbest_j - x_{ij}(t)) \quad (3.1)$$

$$x_{ij}(t+1) = x_{ij}(t) + u_{ij}(t+1) \qquad (3.2)$$

Όπου t είναι ο μετρητής των επαναλήψεων, $c_1$ και $c_2$ είναι οι μεταβλητές επιτάχυνσης, $rand_1$ και $rand_2$ είναι δύο τυχαίες μεταβλητές στο διάστημα [0,1]. Οι μεταβλητές επιτάχυνσης ελέγχουν το πόσο μακριά μπορεί να κινηθεί το σωματίδιο κατα τη διάρκεια μιας απλής επανάληψης, Χαμηλές τιμές επιτρέπουν στα σωματίδια να βρεθούν πολύ μακριά από τη στοχευμένη περιοχή πριν βρεθούν κοντά σε τοπικό ελάχιστον, ενώ πολύ υψηλές τιμές έχουν ως αποτέλεσμα να κάνουν μια απότομη κίνηση προς τις στοχευμένες περιοχές. Τυπικά και οι δύο αυτές μεταβλητές παίρνουν την τιμή 2. Αυτός ο αλγόριθμος ονομάζεται και gbest – PSO γιατί η γειτονιά αναζήτησης που έχουν τα σωματίδια αντιστοιχεί σε ολόκληρο το σμήνος.

Η βέλτιστη θέση $pbest_{ij}$ ενός σωματιδίου στο σμήνος υπολογίζεται από την ακόλουθη εξίσωση:

Σε περίπτωση προβλήματος ελαχιστοποίησης:

$$pbest_{ij}\begin{cases} x_{ij}(t+1), & \varepsilon\acute{\alpha}\nu\ f(x_{ij}(t+1)) < f(x_{ij}(t)) \\ pbest_{ij}, & \alpha\lambda\lambda\iota\acute{\omega}\varsigma \end{cases} \quad (3.3)$$

Σε περίπτωση προβλήματος μεγιστοποίησης:

$$pbest_{ij}\begin{cases} x_{ij}(t+1), & \varepsilon\acute{\alpha}\nu\ f(x_{ij}(t+1)) > f(x_{ij}(t)) \\ pbest_{ij}, & \alpha\lambda\lambda\iota\acute{\omega}\varsigma \end{cases} \quad (3.4)$$

Η βέλτιστη θέση όλου του σμήνους τη χρονική στιγμή t υπολογίζεται από την εξίσωση:

$$gbest_j \in \{pbest_{1j}, pbest_{2j}, \dots, pbest_{Nj} | f(gbest_j)\} = \min\{f(pbest_{1j}), f(pbest_{2j}), \dots, f(pbest_{Nj})\} \quad (3.5)$$

Ο υπολογισμός της ταχύτητας σε μια χρονική στιγμή t+1 αποτελείται από τρείς όρους. Αρχικά από την προηγούμενη ταχύτητα ($u_{ij}(t)$) που στην ουσία χρησιμοποιείται ως μνήμη των προηγούμενων κατευθύνσεων που είχε χρησιμοποιήσει το σωματίδιο κατά την αμέσως προηγούμενη χρονική στιγμή. Αυτή η μνήμη εμποδίζει το σωματίδιο να αλλάξει δραματικά την κατεύθυνσή του κατά τη διάρκεια μιας επανάληψης.

Ο δεύτερος όρος ($c_1 rand_1(pbest_{ij} - x_{ij}(t))$) ο οποίος αναφέρεται και ως γνωστικός όρος (cognitive component) χρησιμοποιείται για να ποσοτικοποιήσει την απόδοση κάθε σωματιδίου σε σχέση με τις αποδόσεις του κατά το παρελθόν. Στην ουσία αυτός ο όρος έχει το ϱόλο της επαναφοράς του σωματιδίου σε προηγούμενες καταστάσεις που ήταν πολύ αποδοτικές για το ίδιο το σωματίδιο.

Ο τρίτος όρος ($c_2 rand_2(gbest_j - x_{ij}(t))$) ο οποίος αναφέρεται και ως κοινωνικός όρος (social component) χρησιμοποιείται για να ποσοτικοποιήσει την απόδοση κάθε σωματιδίου σε σχέση με τις αποδόσεις ολόκληρου του σμήνους ή μιας γειτονιάς σωματιδίων. Στην ουσία αυτός ο όρος έχει το ϱόλο της αναζήτησης από το σωματίδιο του καλύτερου μέλους του σμήνους.

Όταν αρχικοποιούμε τα σωματίδια είναι πολύ πιθανό να αφήσουμε κάποια μέϱη του χώρου λύσεων ακάλυπτα και αυτό να οδηγήσει είτε σε αργή σύγκλιση προς το ϐέλτιστο είτε σε παγίδευση του αλγορίθμου σε τοπικό ελάχιστο. Ο συνηθέστερος τρόπος είναι να αρχικοποιήσουμε τις ϑέσεις των σωματιδίων με τυχαίο τρόπο στο χώρο λύσεων. Έτσι, ανάλογα με το πρόβλημα που ϑέλουμε να επιλύσουμε πρέπει να προσπαθούμε να αρχικοποιούμε κατάλληλα τον πληθυσμό ώστε να καλύπτουμε όσο το δυνατόν μεγαλύτερο μέρος του χώρου αναζήτησης. Αν θέσουμε ως $x_{min}$ την ελάχιστην τιμή και $x_{max}$ την μέγιστη, τότε όλες οι τιμές του διανύσματος του κάθε σωματιδίου μπορούν να υπολογιστούν από την σχέση:

$$x_{ij}(0) = x_{min} + r_i(x_{max} - x_{min}) \quad (3.6)$$

Με το $r_i$ να είναι ένας τυχαίος αριθμός στο (0,1).

Αντίθετα με τις αρχικές θέσεις των σωματιδίων που είναι πολύ σημαντικό να οριστούν σωστά οι αρχικές ταχύτητες σχεδόν πάντα παίρνουν την τιμή 0. Οι αρχικές τιμές του pbest για κάθε σωματίδιο τίθενται ίσες με τις αρχικές θέσεις των σωματιδίων. Ο αλγόριθμος τερματίζεται όταν ένα από τα κλασικά κριτήρια που ισχύουν σε όλους τους εξελικτικούς αλγορίθμους ικανοποιηθεί. Δηλαδή είτε αν έχουμε φθάσει το μέγιστο αριθμό επαναλήψεων, είτε αν η βέλτιστη λύση βρίσκεται μέσα σε κάποιο προκαθορισμένο κατώφλι, είτε αν μετά από ένα αριθμό επαναλήψεων δεν έχει εμφανιστεί κάποια βελτίωση στη λύση κτλ.

Οι παράγοντες επιτάχυνσης $c_1$ και $c_2$ μαζί με τα τυχαία διανύσματα $rand_1$ και $rand_2$ ελέγχουν την τυχαία επιρροή του γνωστικού και κοινωνικού παράγοντα στη συνολική ταχύτητα του σωματιδίου. Αν το $c_1 = c_2 = 0$ τα σωματίδια πετούν στην κατεύθυνση που τους δίνει η ταχύτητα τους. Εάν το $c_1 > 0$ και το $c_2 = 0$ τότε το κάθε μέλος του πληθυσμού επηρεάζεται μόνο από τις προηγούμενες κινήσεις του και κινείται ανεξάρτητα από τα άλλα σωματίδια του σμήνους.

Στην αντίθετη περίπτωση όπου $c_2 > 0$ και το $c_1 = 0$ ολόκληρο το σμήνος κυνηγάει ένα σωματίδιο, το βέλτιστο. Ο στόχος είναι να βρεθεί η κατάλληλη ισορροπία μεταξύ του $c_1$ και του $c_2$. Συνήθως το $c_1 = c_2$, όπου σημαίνει ότι το σωματίδιο το έλκουν και οι δύο παράγοντες ισόποσα. Για παράδειγμα, αν θεωρήσουμε ότι το $c_{1,min}$, $c_{1,max}$, $c_{2,min}$, $c_{2,max}$ είναι οι μέγιστες και οι ελάχιστες τιμές που θα μπορούσαν να πάρουν τα $c_1$ και $c_2$ αντίστοιχα, τότε:

$$c_1 = c_{1,min} + \frac{c_{1,max} - c_{1,min}}{iter_{max}} * t \qquad (3.7)$$

$$c_2 = c_{2,min} + \frac{c_{2,max} - c_{2,min}}{iter_{max}} * t \qquad (3.8)$$

Όπου όπου t είναι ο αριθμός της τρέχουσας επανάληψης και $iter_{max}$ είναι ο μέγιστος αριθμός επαναλήψεων. Με αυτό τον τρόπο αρχικά οι τιμές των $c_1$ και $c_2$ είναι μικρές ενώ στη συνέχεια καθώς προχωρούν οι επαναλήψεις οι τιμές αυξάνονται και πλησιάζουν τη μέγιστη τιμή. Άρα αρχικά τα σωματίδια επηρεάζονται περισσότερο από τον πρώτο παράγοντα της εξίσωσης και στη συνέχεια αρχίζουν να επηρεάζονται από το δεύτερο και τον τρίτο παράγοντα. Αυτό γίνεται για να υπάρχει μεγαλύτερη ελεύθερια κίνησης μέσα στο χώρο των σωματιδίων κατά τις πρώτες επαναλήψεις ώστε να εντοπιστεί γρηγορότερα το ελάχιστο. [4]

Σε μορφή ψευδοκώδικα ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων θα έχει ως εξής:

**Αλγόριθμος** Βελτιστοποίηση Σμήνους Σωματιδίων
**Αρχικοποίηση**
**Επιλογή** του αριθμού των σμηνών

**Επιλογή** του αριθμού των σωματιδίων σε κάθε σμήνος
**Αρχικοποίηση** της θέσης και της ταχύτητας κάθε σωματιδίου
**Υπολογισμός** του αρχικού κόστους του κάθε σωματιδίου
**Εύρεση** Βέλτιστου σωματιδίου ολόκληρου του σμήνους
**Εύρεση** Βέλτιστης λύσης κάθε σωματιδίου
Κύρια Φάση
**Εκτέλεση έως ότου** δεν έχει φθάσει ο μέγιστος αριθμός επαναλήψεων
    Υπολογισμός της ταχύτητας του κάθε σωματιδίου
    Υπολογισμός της νέας θέσης του κάθε σωματιδίου
    Υπολογισμός της νέας συνάρτησης ποιότητας του κάθε σωματιδίου
    Ενημέρωση της βέλτιστης λύσης του κάθε σωματιδίου
    Εύρεση του βέλτιστου σωματιδίου ολόκληρου του σμήνους
**Τέλος εκτέλεσης**
**Επιστροφή** βέλτιστου σωματιδίου (βέλτιστης λύσης).

### 3.2.4 Σχετικές προσεγγίσεις

Σε σχετική τους δημοσίευση οι Yuce Tekol και Adnam Acan, ανέλυσαν την προσέγγιση τους σχετικά με την ανάπτυξη αξιόπιστων στρατηγικών για το παίγνιο του διλλήματος του φυλακισμένου κάνοντας χρήση Βελτιστοποίησης Αποικίας Μυρμηγκιών (Ant Colony Optimization). Μέσα από την εργασία τους παρουσίασαν πως τα μυρμήγια μπορούν να παίξουν το επαναληπτικό δίλημμα του φυλακισμένου και παρά το γεγονός πως χρειάζονται μεγαλύτερους χρόνους, φαίνεται πως τα αποτελέσματα τους είναι καλύτερα από εκείνα των γενετικών αλγορίθμων. Την επιτυχία της ανεπτυγμένης στρατηγικής την σύγκριναν με εκείνες που αναπτύχθηκαν χρησιμοποιώντας Γενετικούς αλγορίθμους μέσω πειραματικών αξιολογήσεων. [13]

Τα τελευταία χρόνια έχουν πραγματοποιηθεί κι άλλες παρόμοιες δημοσιεύσεις όπως η ‘Evolving Group Strategies for IPD’ από Philip Hingston, η ‘Evolving Strategies for the Prisoner’s Dilemma’ από Andrew Errity, η ‘Evolution, Neural Networks, Games and Intelligence’ από Kumar Chellapilla, η ‘Engineering Design of Stategies for Winning Iterated Prisoner’s Dilemma Competitions’ από Jiawei Li και Philip Hingston, κ.α.. [5][15][16][17]

Η δημοσίευση όμως που μας ενέπνευσε και απετέλεσε την αφορμή και την πηγή γι’ αυτήν την εργασία ήταν η ‘Particle Swarm Optimization Approaches Coevolve Strategies for the Iterated Prisoner’s Dilemma’ από Nelis Franken και Andries P. Engelbrecht. [14]

Στην δημοσίευση αυτή παρουσιάζουν και ερευνούν μια εφαρμογή από εξελικτικές τεχνικές εκπαίδευσης βασισμένες στην βελτιστοποίηση σμήνους σωματιδίων με τελικό σκοπό να εξελίξουν τις στρατηγικές ενός παίκτη για το μη-μηδενικό παίγνιο του επαναληπτικού διλήμματος των φυλακισμένων. Για την ακρίβεια αναλύουν τρεις διαφορετικούς τρόπους προσέγγισης: Ο πρώτος είναι ένα Νευρωνικό δίκτυο το οποίο προσπαθεί να προβλέψει την επόμενη ενέργεια, ο δεύτερος είναι προσέγγιση δυαδικού PSO που το κάθε σωματίδιο αναπαριστά μια ολοκληρωμένη στρατηγική παιχνιδιού και τέλος μια νέα προσέγγιση που εκμεταλεύεται την συμμετρική κατασκευή των αντρωπογενών στρατηγικών. Οι τρεις αυτές προσεγγίσεις αξιολογούνται και συγκρίνονται πειραματικά η μία με την άλλη αλλά επίσης και σε σχέση με τις ήδη γνωστές στρατηγικές. Η απόδοση αυτών μελετάται σε καθαρό αλλά και σε θορυβώδες περιβάλλον. Τα αποτελέσματα δείχνουν πως το NN συμπεριφέρθηκε σωστά αλλά μπορεί να αναπτύξει αδύναμες στρατηγικές που μπορεί να οδηγήσουν σε κακά αποτελέσματα. Σε αντίθεση ο PSO δεν παρουσίασε το ίδιο ελάττωμα αλλά οδηγούσε σε μια συνολική κατάσταση ισορροπίας κατά την οποία οι στρατηγικές μπορούν να εκμεταλευτούν τον πληθυσμό αλλά ποτέ δεν κυριασχούν. Τέλος η συμμετρική προσέγγιση δεν ήταν τόσο πετυχημένη όσο ο δυαδικός PSO. Γενικά κατέληξαν πως οι τεχνικές PSO ήταν επιτυχείς στο να παράγουν μια ποικιλία στρατηγικών που μπορούν να χρησιμοποιηθούν για το επαναληπτικό δίλημμα του φυλακισμένου, αναπαράγοντας και βελτιώνοντας τους υπάρχοντες εξελικτικούς πληθυσμούς.

Στην παρούσα εργασία θα κάνουμε χρήση της ιδέας του binary PSO έτσι ώστε να προετοιμάσουμε με την σειρά μας τους κατάλληλους εξελικτικούς πληθυσμούς που θα έχουν την δυνατότητα να ανταγωνιστούν τις υπάρχοντες γνωστές στρατηγικές του χώρου. [14]

# ΚΕΦΑΛΑΙΟ 4 – ΑΛΓΟΡΙΘΜΟΣ / ΥΛΟΠΟΙΗΣΗ

## 4.1 Περιγραφή απαιτήσεων

Όπως σε όλα τα προβλήματα, το πρώτο βήμα είναι να αναγνωρίσουμε την ανάγκη και να καθοριστεί το πρόβλημα. Ο στόχος μας διαμορφώθηκε στο να δημιουργήσουμε έναν αλγόριθμο βασισμένο σε δυαδικό PSO, ο οποίος θα έχει την δυνατότητα να παίξει ανταγωνιστικά έναντι γνωστών στρατηγικών του επαναληπτικού διλήμματος του φυλακισμένου. Σκοπός μας δεν είναι μόνο να συγκεντρώνει πάντα το μεγαλύτερο σκορ σε σχέση με τους αντιπάλους του, αλλά παράλληλα να μελετάται και η συμπεριφορά του (συνεργατική ή μη) και το πως αυτή εξελίσσεται σύμφωνα, κάθε φορά, με τον αντίστοιχο αντίπαλο. Αναφορικά, η όλη υλοποίηση έχει πραγματοποιηθεί σε περιβάλλον **Matlab**.

Γενικά, για την μορφοποίηση ενός προβλήματος βελτιστοποίησης περιλαμβάνονται τα ακόλουθα τρία βασικά σύνολα στοιχείων:

- Μεταβλητές απόφασης (decision variables). Οι μεταβλητές απόφασης είναι οι άγνωστοι που πρέπει να καθοριστούν από την επίλυση του προτύπου. Οι παράμετροι αποτελούν τις μεταβλητές ελέγχου του συστήματος.
- Περιορισμοί (constraints). Για την μέτρηση των φυσικών περιορισμών του συστήματος, το πρότυπο πρέπει να περιέχει περιορισμούς που να περιορίζουν τις μεταβλητές απόφασης στις εφικτές (επιτρεπτές) τιμές του. Ένα πρόβλημα κατατάσσεται σε διαφορετικές κατηγορίες βάσει της φύσης των περιορισμών του (γραμμικοί, ακέραιοι, μη γραμμικοί κ.τ.λ.).
- Αντικειμενική συνάρτηση (objective function). Η αντικειμενική συνάρτηση καθορίζει το μέτρο της αποτελεσματικότητας του συστήματος ως μαθηματική συνάρτηση των μεταβλητών απόφασής του. Για παράδειγμα, αν ο αντικειμενικός στόχος του συστήματος είναι η μεγιστοποίηση του συνολικού κέρδους, η αντικειμενική συνάρτηση πρέπει να καθορίζει το κέρδος σε όρους των μεταβλητών απόφασης. Γενικά η βέλτιστη (optimal) λύση ενός μοντέλου επιτυγχάνεται όταν οι αντίστοιχες τιμές των μεταβλητών απόφασης οδηγούν στην βέλτιστη τιμή της αντικειμενικής συνάρτησης ικανοποιώντας ταυτόχρονα, όλους τους περιορισμούς. [2]

Όπως αναφέρθηκε και στο προηγούμενο κεφάλαιο (§3.2.3), η μέθοδος της βελτιστοποίησης σμήνους σωματιδίων έχει αποδειχθεί πολύ χρήσιμη σε ποικίλα προβλήματα, αλλά ιδιαίτερα σε προβλήματα που έχουν συνεχείς μεταβλητές. Όταν, όμως, έχουμε διακριτές μεταβλητές και όχι συνεχείς, όπως έχουμε στο συγκεκριμένο πρόβλημα, δεν μπορεί να χρησιμοποιηθεί η εξίσωση (3.2) για τον υπολογισμό της θέσης του σωματιδίου.

Το κάθε σωματίδιο θα πρέπει να κινείται σε μια περιορισμένη περιοχή όπου μπορεί να πάρει μόνο 2 τιμές – 0 ή 1. Έτσι, οι τροχιές των σωματιδίων καθορίζονται ως οι αλλαγές στην πιθανότητα και το $u_{ij}$ μετράει την πιθανότητα του κάθε ενός σωματιδίου να πάρει την τιμή 1 στην θέση i. Εάν η ταχύτητα είναι μεγάλη, τότε το σωματίδιο θα είναι πιθανότερο να πάρει την τιμή 1, ενώ αν είναι χαμηλή θα είναι πιθανότερο να πάρει την τιμή 0.

Η παραπάνω διακριτή προσέγγιση του αλγορίθμου βελτιστοποίησης σμήνους σωματιδίων προτάθηκε από τους Kennedy και Eberhard και θα την χρησιμοποιήσουμε σε αυτή την εργασία για να παρακάμψουμε το πρόβλημα που παρουσιάζεται: [4]

Για να υπολογίσουμε τις ταχύτητες των σωματιδίων, κάνουμε χρήση σιγμοειδούς συνάρτησης, η οποία μετατρέπει την ταχύτητα σε κανονικοποιημένες τιμές στο διάστημα [0,1]:

$$sig(u_{ij}) = \frac{1}{1+e^{-u_{ij}}} \qquad (4.1)$$

Οι ταχύτητες υπολογίζονται με τον ίδιο τύπο από την σχέση (3.1). Οι θέσεις όμως, των σωματιδίων υπολογίζονται από τον παρακάτω τύπο:

$$s_{ij}(t+1)\begin{cases} 1, & \varepsilon\acute{\alpha}\nu\ rand < sig(u_{ij}) \\ 0, & \varepsilon\acute{\alpha}\nu\ rand \geq sig(u_{ij}) \end{cases} \qquad (4.2)$$

Έχοντας λύσει το πρόβλημα της μετατροπής στο δυαδικό σύστημα, θα αναλύσουμε παρακάτω την ροή και τον σχεδιασμό του αλγορίθμου, καθώς και την ανάλυση των μεγεθών, όπως και των αντίστοιχων διανυσμάτων που θα πρέπει να χρησιμοποιηθούν στο Matlab.

## 4.2 Ανάλυση/Υλοποίηση

Στην συγκεκριμένη υλοποίηση, θεωρήσαμε πως το κάθε σωματίδιο μέσα στο σμήνος περιγράφεται από ένα διάνυσμα, στο οποίο η κάθε θέση μπορεί να πάρει μόνο 2 διακριτές τιμές - 1 και 0. Το 1 αντιστοιχεί σε συνεργασία και το 0 σε αποστασία. Συνεργασία όπως αναφέραμε και στο κεφάλαιο 2 (§2.1.2) σημαίνει πως ο ύποπτος ομολόγησε ενώ αντίθετα αποστάσια ή μη-συνεργασία σημαίνει πως ο ύποπτος παρέμεινε σιωπηλός.

Ο πίνακας που ακολουθεί είναι ο πίνακας αμοιβών που χρησιμοποιήθηκε. Η αριστερή τιμή αντιστοιχεί στον παίκτη γραμμής και η δεξιά τιμή είναι του παίκτη στήλης. Επίσης τα νούμερα που έχουμε επιλέξει είναι τέτοια, ώστε να επαληθεύονται οι εξισώσεις 5>3>1>0 (σχεση 2.1) και 2*3>0+5 (σχέση 2.2).

| ΠΙΝΑΚΑ Α | Β παίκτης Συνεργασία | Β παίκτης Αποστασί α |
|---|---|---|
| **Α παίκτης Συνεργασία** | 3,3 | 0,5 |
| **Α παίκτης Αποστασία** | 5,0 | 1,1 |

Ο παίκτης Α θα λάβει επαναληπτικές αποφάσεις, σε αντίκρουση, πάντα, με τις επαναληπτικές αποφάσεις του παίκτη Β.

Μεταφράζοντας το παραπάνω σε αλγόριθμο, ο παίκτης j θα έχει ένα διάνυσμα decision(i) που σε κάθε θέση i θα εμπεριέχει την απόφαση του παίκτη j για την συγκεκριμένη επανάληψη. Το σμήνος, λοιπόν, θα αποτελείται από σωματίδια decision, τα οποία ουσιαστικά αποτελούν μια αλληλουχία κινήσεων ή αλλιώς την στρατηγική του εκάστοτε παίκτη. Τα διανύσματα decision είναι εκείνα τα οποία θα περάσουν μέσα από τον binary PSO για να επεξεργαστούν και με το πέρας τον επαναλήψεων να διαμορφώνουν την καταλληλότερη και πιο ανταγωνιστική στρατηγική που θα ανταγωνιστεί τους «γνωστούς» αντιπάλους.

Τα διανύσματα-στρατηγικές decision ξεκινούν να παίζουν μεταξύ τους. Ο παίκτης j=1 θα παίξει με τους παίκτες j=2,3,...,N όπου N το σύνολο των παικτών. Αντίστοιχα, και ο 2$^{ος}$ θα παίξει με όλους τους υπόλοιπους και ούτω καθ’ εξής.

Από τα μεταξύ τους παιχνίδια, ο κάθε παίκτης, ανάλογα με την στρατηγική του και με τον αντίπαλο που θα έχει να αντιμετωπίσει, διαμορφώνει ένα διάνυσμα αμοιβών (payoff). Τα payoff των παικτών συγκρίνονται μεταξύ τους, διαμορφώνοντας τα διανύσματα pbest και gbest. Το pbest περιέχει την καλύτερη στρατηγική του πάικτη, με το αντίστοιχο payoff, σε όλη την διάρκεια των επαναλήψεων ενώ το gbest περιέχει το καλύτερο pbest του καλύτερου παίκτη όλου του πλήθους.

Στη συνέχεια, στα διανύσματα decision, pbest, gbest εφαρμόζουμε την σιγμοειδή συνάρτηση και μετατρέπουμε τις τιμές από δυαδικές σε συνεχείς. Έτσι, μας δίνεται η δυνατότητα να εφαρμόσουμε τους τύπους (3.1) και (3.2) που αφορούν την Βελτιστοποίηση Σμήνους Σωματιδίων. Μόλις πραγματοποιηθεί ο υπολογισμός των μεγεθών της θέσης και της ταχύτητας, επαναφέρουμε το διάνυσμα decision σε δυαδικές τιμές και επαναλαμβάνουμε τα παιχνίδια με τους υπόλοιπους παίκτες.

Μόλις ολοκληρωθεί η παραπάνω διαδικασία, ο καλύτερος παίκτης (διάνυσμα gbest) ξεκινάει να αντιμετωπίζει αντιπάλους που ακολουθούν συγκεκριμένες στρατηγικές, όπως αυτές έχουν προκύψει μέσα από τη βιβλιογραφία, όπως και random αντιπάλους, που πάντα είναι πιο ανταγωνιστικοί γιατί είναι σχετικά απρόβλεπτοι και η νίκη συνήθως αποδίδεται στις καλύτερες πιθανότητες που έχει ο PSO αλγόριθμος.

Στο τέλος αυτής της διαδικασίας, συγκρίνονται οι αμοιβές που έλαβε ο καλύτερος PSO παίκτης σε σχέση με τους αντιπάλους του που ακολουθούσαν γνωστές στρατηγικές. Οι πιο «δημοφιλείς» αντίπαλοι αναφέρονται στο 2$^{ο}$ κεφάλαιο (§2.1.3) αλλά στην παρούσα εργασία διαλέγουμε να αντιμετωπίσουμε τους πιο ανταγωνιστικούς και να δούμε την απόδοση της PSO προσέγγισης. Πιο συγκεκριμένα, οι Tit_for_tat και Pavlov είναι οι πρώτοι αντίπαλοι που πρέπει να ανταγωνιστεί ο διαμορφωμένος αλγόριθμος, καθώς απ' τα αποτελέσματα άλλων σχετικών εργασιών, προκύπτει ότι συγκεντρώνουν τις καλύτερες αμοιβές και παρουσιάζουν τις καλύτερες αποδόσεις απ' τους υπόλοιπους. Επίσης, δεν θα μπορούσε να απουσιάζει, όπως αναφέρθηκε παραπάνω, ένας καθαρά random παίκτης. Τέλος, συμπεριλαμβάνονται και οι αντίπαλοι Evil_tit_for_tat και two_to_trust. Αυτό συνέβη διότι ο πρώτος διαθέτει παρόμοια ανταγωνιστικά στοιχεία όπως ο tit_for_tat και ο δεύτερος διότι, όπως θα δείξουμε στο επόμενο κεφάλαιο, αποτέλεσε έναν απ' τους αντιπάλους που δυσκόλεψε τον PSO.

Συνοψίζοντας τα παραπάνω, ο ψευδοκώδικας που διέπει τις σχέσεις και τις συσχετίσεις και όλα τα βήματα των επαναλήψεων είναι ο ακόλουθος:

1. Δημιουργώ το πλήθους σωματιδίων N – σύνολο N παικτών στην προκειμένη περίπτωση.
2. Επαναλαμβάνω για L επαναλήψεις
   A. Ορίζω σε κάθε παίκτη τυχαία θέση και αποθηκεύω το προσωπικό του καλύτερο στο χώρο λύσεων (διάνυσμα pbest).
   B. Για κάθε παίκτη στο χώρο των λύσεων:
      i. Παίζει με κάθε άλλο παίχτη στο χώρο των λύσεων:
         a. Ξεκινάει από τυχαία θέση
         b. Ανταλλάζει 150 αποφάσεις συνεργασία ή μη (150 αποφάσεις καθορίζουν την στρατηγική του κάθε παίχτη – διάνυσμα decision )
         c. Ορίζεται η ανταμοιβή σύμφωνα με τον πίνακα A ( διάνυσμα payoff).
      ii. Συνεχίζω να αποθηκεύω το συνολικό payoff του κάθε παίχτη
   C. Για κάθε παίκτη στο χώρο των λύσεων:
      i. Μετατρέπω τα 0-1 με σιγμοειδή*
      ii. Συγκρίνω μέσω των τύπων PSO με το προσωπικό καλύτερο του παίχτη
      iii. Συγκρίνω με το συνολικό καλύτερο όλου του χώρου λύσεων (διανυσμα gbest).
      iv. Ανανεώνω την ταχύτητα
      v. Ανανεώνω την θέση
      vi. Μετατρέπω ξανά σε 0-1

3. Καθορίζεται, τελικά, αυτός που έχει την καλύτερη απόδοση αθροιστικά από το σύνολο των παιχτών.
4. Ο καλύτερος παίζει αρκετές επαναλήψεις Μ απέναντι σε γνωστές στρατηγικές ( – totally random – Pavlov – tit for tat – Evil tit for tat – Two to trust)
5. Επαναλαμβάνω ξανά για W επαναλήψεις.

## 4.3 Παραλλαγές PSO

### Παραλλαγή βάρους αδράνειας w

Ανάλογα με την φύση του προβλήματος, έχουν προταθεί αρκετές παραλλαγές του αρχικού PSO αλγορίθμου, που αποσκοπούν στο να βελτιστοποιήσουν τις λύσεις ενός συγκεκριμένου προβλήματος. Διερευνητικά, θα δοκιμασούμε κάποιες απ’ αυτές και θα εξετάσουμε πως επιδρούν στον PSO που κατασκευάσαμε παραπάνω.

Η πιο χαρακτηριστική βελτίωση χρησιμοποιεί ένα βάρος αδράνειας w. Με την εισαγωγή αυτού του βάρους, προσπαθούμε να αυξήσουμε την ικανότητα διασποράς της αναζήτησης και να εντατικοποιήσουμε την λύση. Το βάρος αδράνειας ελέγχει την επιρροή των προηγούμενων ταχυτήτων ενός σωματιδίου στην τρέχουσα ταχύτητα. Το w χρησιμοποιείται, επίσης, για να ελέγχει την σύγκλιση του αλγορίθμου.

$$u_{ij}(t+1) = wu_{ij}(t) + c_1 rand_1(pbest_{ij} - x_{ij}(t)) + c_2 rand_2(gbest_j - x_{ij}(t)) \quad (4.3)$$

Για την επιλογή του w θα πρέπει να γίνει εκτενέστερη μελέτη, καθώς το w επηρεάζει τη σύγκλιση του αλγορίθμου σε τοπικό ή ολικό ελάχιστο και την ταχύτητα σύγκλισης του αλγορίθμου. Αν επιλεγεί μια πολύ μικρή σταθερή τιμή, μειώνονται οι ικανότητες διασποράς των λύσεων, με κίνδυνο να παγιδευτεί ο αλγόριθμος σε κάποιο τοπικό ελάχιστο. Αν, αντίθετα, επιλεγεί μια πολύ μεγάλη σταθερή τιμή, τότε μειώνονται οι ικανότητες εντατικοποίησης της αναζήτησης σε κάποιο καλό σημείο, με κίνδυνο να προσπεραστεί το βέλτιστο χωρίς να εντοπιστεί.

Βάση των παραπάνω, ο ενδεικτικότερος τρόπος επιλογής του w είναι η εκκίνηση της τιμής του από ένα υψηλό επίπεδο και, κατά τη διάρκεια των επαναλήψεων, η μείωση της, επιτρέποντας έτσι στον αλγόριθμο να προχωρήσει και σε ανεξερεύνητες περιοχές. Ο τύπος που θα μας καθορίζει το w είναι ο ακόλουθος:

$$w = w_{max} - \frac{w_{max} - w_{min}}{iter_{max}} * t \quad (4.4)$$

Όπου $w_{max}$, $w_{min}$ είναι η μέγιστη και η ελάχιστη τιμή που μπορεί να λάβει το βάρος αδράνειας και $iter_{max}$ είναι ο μέγιστος αριθμός επαναλήψεων. [4]

## Παραλλαγή παράγοντα περιορισμού ταχύτητας

Αυτή η παραλλαγή έχει σκοπό να εξασφαλίσει πως δεν θα πραγματοποιηθεί μια ραγδαία άυξηση της ταχύτητας με αποτέλεσμα να παραμεριστεί κάποιο πιθανό βέλτιστο. Έτσι αυτή η παραλλαγή εισάγει έναν παράγοντα περιορισμού (constriction factor) στις ταχύτητες των σωματιδίων.

Η νέα ταχύτητα ενός σωματιδίου υπολογίζεται από μια παραλλαγή της σχέσης (3.1):

$$u_{ij}(t+1) = X(u_{ij}(t) + c_1 rand_1 \left(pbest_{ij} - x_{ij}(t)\right) + c_2 rand_2 \left(gbest_j - x_{ij}(t)\right)) \quad (4.5)$$

Όπου:

$$X = \frac{2}{\left|2-c-\sqrt{c^2-4c}\right|} \; και \; c = c_1 + c_2, c > 4 \quad (4.6)$$

Έχει προταθεί, επίσης, μια εναλλακτική της παραπάνω σχέσης, όπου ο αριθμητής του κλάσματος υπολογισμού πολλαπλασιάζεται με έναν παράγοντα $k \in (0,1]$.

Έτσι, η εξίσωση (4.6) μετατρέπεται στην παρακάτω:

$$X = \frac{2k}{\left|2-c-\sqrt{c^2-4c}\right|} \; και \; c = c_1 + c_2, c > 4 \quad (4.7)$$

Όπου αν το k πάρει μια τιμή κοντά στο 0, που σημαίνει πως και το X θα τείνει στο μηδεν, τότε ο αλγόριθμος αναπτύσσει μεγάλες ικανότητες εντατικοποίησης της αναζήτησης γύρω από κάποιο πιθανό καλό σημείο. Αντίθετα αν το k πλησιάζει τη μονάδα τότε ο αλγόριθμος μπορεί να αναπτύξει μεγαλύτερες ικανότητες αναζήτησης στο χώρο των λύσεων. [4]

## Παραλλαγή γνωστικού-κοινωνικού μοντέλου

Αξίζει να σημειωθεί ότι για τον υπολογισμό της ταχύτητας υπάρχει και η παραλλαγή γνωστικού-κοινωνικού μοντέλου, στην οποία δε χρησιμοποιούνται όλοι οι παράγοντες που περιγράψαμε παραπάνω.

Το πρώτο μοντέλο έχει ονομαστεί γνωστικό-μοντέλο (cognition only model) και δεν χρησιμοποιεί καθόλου τον παράγοντα κοινωνικότητας της ταχύτητας. Σε αυτή την περίπτωση, η σχέση (3.1) απλοποιείται σε:

$$u_{ij}(t+1) = u_{ij}(t) + c_1 rand_1 \left(pbest_{ij} - x_{ij}(t)\right) \quad (4.8)$$

Το δεύτερο μοντέλο έχει ονομαστεί κοινωνικό-μόνο μοντέλο (social-only model) και δεν χρησιμοποιεί καθόλου το δεύτερο παράγοντα της σχέσης (3.1). Σε αυτή την περίπτωση η σχέση απλοποιείται σε:

$$u_{ij}(t+1) = u_{ij}(t) + c_2 rand_2 \left(gbest_{ij} - x_{ij}(t)\right) \tag{4.9}$$

# ΚΕΦΑΛΑΙΟ 5 – ΑΠΟΤΕΛΕΣΜΑΤΑ / ΣΥΜΠΕΡΑΣΜΑΤΑ

## 5.1 Καθορισμός του στόχου

Όπως έχει αναφερθεί στα προηγούμενα κεφάλαια, στο παίγνιο ενός γύρου (one-shot prisoner's dilemma), ο κάθε παίκτης παίρνει μια μοναδική απόφαση για την έκβαση του παιγνίου. Κατ' αυτό το σενάριο η καλύτερη στρατηγική είναι να αποστατήσει και να μείνει σιωπηλός, καθώς είτε θα πάρει την μέγιστη αμοιβή '5' είτε θα έρθει ισοπαλία με τον αντίπαλό του και θα λάβει από '1' μονάδα. Όταν όμως ξεκινάμε να παίζουμε επαναληπτικά ο στόχος είναι να «χτίσουμε» μια σχέση εμπιστοσύνης μεταξύ των 2 παικτών έτσι ώστε και οι δύο να αποκομίζουν την ιδανική αμοιβή και να είναι και οι δύο ικανοποιημένοι.

Στην συγκεκριμένη εργασία προσεγγίζουμε το παίγνιο δίνοντας αμοιβές στους παίκτες για τις αποφάσεις τους. Αν όμως σε ένα ρεαλιστικό σενάριο αυτό μεταφραζόταν αντίστροφα σε χρόνια φυλάκισης, η αντίληψη μας για το πρόβλημα θα ήταν διαφορετική. Πιο συγκεκριμένα, αν το πρόβλημα άλλαζε κι από πρόβλημα μεγιστοποίησης των αμοιβών γινόταν ελαχιστοποίησης των απωλειών τότε οι σχέσεις (2.1) και (2.2) θα αντιστρέφονταν σε

$$T < R < P < S \qquad (5.1)$$

Επίσης, η διπλή αμοιβή για την συνεργατική συμπεριφορά θα ήταν μικρότερη από τον μέσο όρο των αμοιβών T και S:

$$2 * R < S + T \qquad (5.2)$$

Αντιστοιχίζοντας με αριθμούς θα είχαμε πως το T=0 χρόνια φυλάκισης, R=2 χρόνια φυλάκισης, P=10 χρόνια φυλάκισης και S=30 χρόνια φυλάκισης. Έτσι, για παράδειγμα, θα είχαμε τα παρακάτω πιθανά ενδεχόμενα:

- Αν και οι 2 ομολογήσουν και συνεργαστούν, τότε και οι 2 έχουν πολύ μειωμένη ποινή της τάξης των 2 ετών έκαστος.
- Αν και οι 2 παραμείνουν σιωπηλοί και αποστατήσουν, τότε και οι δύο θα έχουν την δεύτερη χειρότερη ποινή της τάξεως των 10 ετών έκαστος.
- Αν τώρα ο $1^{ος}$ συνεργαστεί και ο $2^{ος}$ αποστατήσει και παραμείνει σιωπηλός, τότε αυτός που ομολόγησε θα έχει πιαστεί κορόιδο και θα λάβει όλη την ποινή δηλαδή 30 χρόνια.
- Αντίστοιχα αν ο $1^{ος}$ αποστατήσει και ο $2^{ος}$ συνεργαστεί τότε θα λάβει ο $1^{ος}$ το εισητήριο για την ελευθερία του.

Αναφέρουμε την παραπάνω εναλλακτική προσέγγιση για να δείξουμε, πως αν κρίνουμε την ρεαλιστική κατάσταση, τότε δεν αποτελεί ξεκάθαρη επιλογή και για τους 2 παίκτες να επιλέγουν πάντα την αποστασία και να έχουν μια ποινή της τάξης των 10 ετών φυλάκισης. Αν δίναμε την δυνατότητα επικοινωνίας στους παίκτες πριν πάρουν την απόφαση τους και δεδομένης της εμπιστοσύνης μεταξύ τους, τότε είμαστε σίγουροι πως και οι 2 θα επέλεγαν να ομολογήσουν από κοινού το έγκλημα τους και σε 2 χρόνια να είναι ελεύθεροι.

Οπότε μπορεί η αποστασία να φαίνεται ιδανική όταν προσεγγίζουμε το δίλημμα καθαρά μαθηματικά, αλλά η συνεργασία είναι αυτή που δίνει το μέγιστο κοινό όφελος και στους δύο παίκτες. Γι αυτό το λόγο, αποφύγαμε να συμπεριλάβουμε στις συγκρίσεις τις στρατηγικές AD (Always Defect) και AC (Always Cooperate). Στόχος μας, λοιπόν, απ' τη μία είναι να είμαστε ανταγωνιστικοί απέναντι στους αντιπάλους μας, μεγιστοποιώντας τις αμοιβές μας μέσα στα ρεαλιστικά πλαίσια κι απ' την άλλη να δημιουργούμε πάντα συνεργατική προσέγγιση.

## 5.2 Αποτελέσματα βασικού PSO

<u>1ο τρέξιμο με χαμηλό αριθμό επαναλήψεων:</u>

Σε πρώτη φάση τρέξαμε τον αλγόριθμο για N=4 όπου N ο αριθμός των PSO παικτών, L=2 όπου L ο αριθμός των επαναλήψεων, M=5 όπου M ο αριθμός των αντίπαλων παικτών που παίζουν κάνοντας χρήση συγκεκριμένων στρατηγικών και W=20, όπου W ο αριθμός των συνολικών επαναλήψεων που επαναλαμβάνουμε την διαδικασία. Να σημειωθεί επίσης πως τα $c_1$ και $c_2$ τα αντιστοιχήσαμε με την τιμή 2 καθώς καταλήξαμε σε αυτήν την τιμή πειραματικά. Στα παρακάτω αποτελέσματα κρατάμε σταθερό το W. Το W προστέθηκε για να εξαλειφθεί ο παράγοντας «σύμπτωσης» στα συγκριτικά αποτελέσματα Στις παρακάτω γραφικές παραστάσεις (με χρήση scatter plot), εμφανίζεται με κόκκινο σταυρό η επίδοση του PSO παίκτη και με πράσινο κύκλο η επίδοση του αντίστοιχου αντιπάλου. Στον οριζόντιο άξονα απεικονίζονται οι W επαναλήψεις. Τα W μεταξύ τους θεωρούνται ανεξάρτητες διαδικασίες και δεν συγκρίνονται το ένα με το άλλο. Στον κάθετο άξονα εμφανίζεται το άθροισμα των αμοιβών για κάθε παίκτη. Τα ποσά του αθροίσματος από σχήμα σε σχήμα μπορεί να διαφέρουν κι αυτό οφείλεται στα μεγέθη των επαναλήψεων. Αυτό όμως που μας ενδιαφέρει είναι η ποσοστιαία διαφορά της επίδοσης του ενός παίκτη απ' τον άλλον, για σταθερό W. Για παράδειγμα για W=7, στο παρακάτω σχήμα, ο PSO παίκτης έχει μαζέψει περίπου 1680 πόντους ενώ ο random παίκτης έχει μαζέψει περίπου 1720

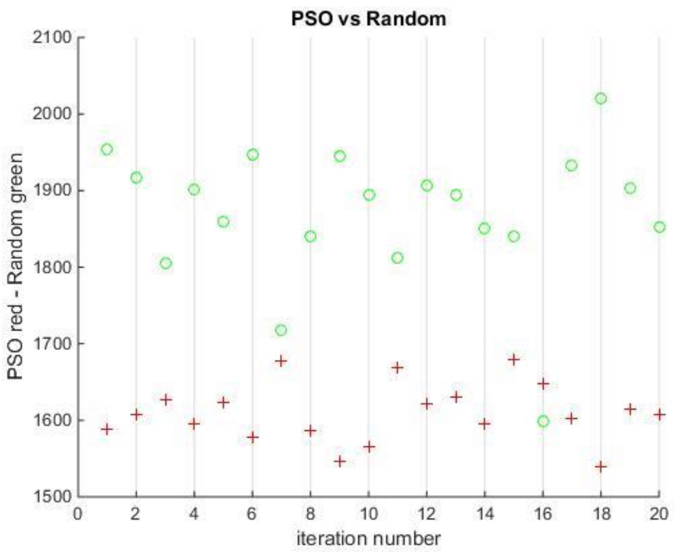



Figure 5-- PSO vs Random 1st level

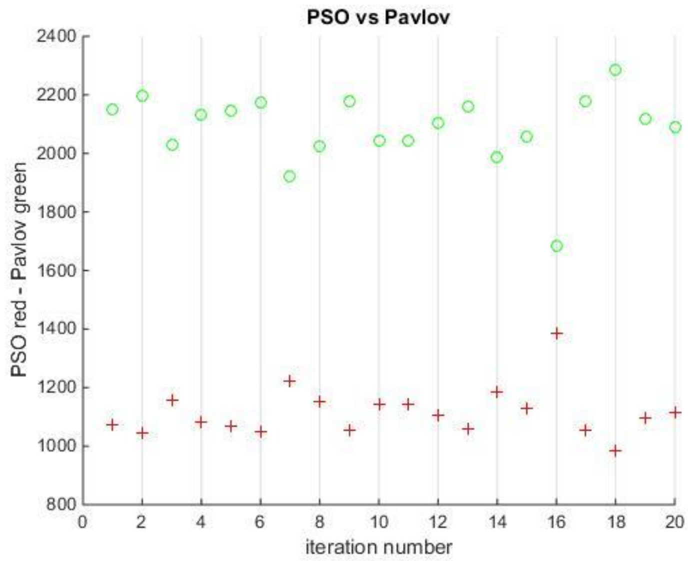



Figure 6- PSO vs Pavlov 1st level

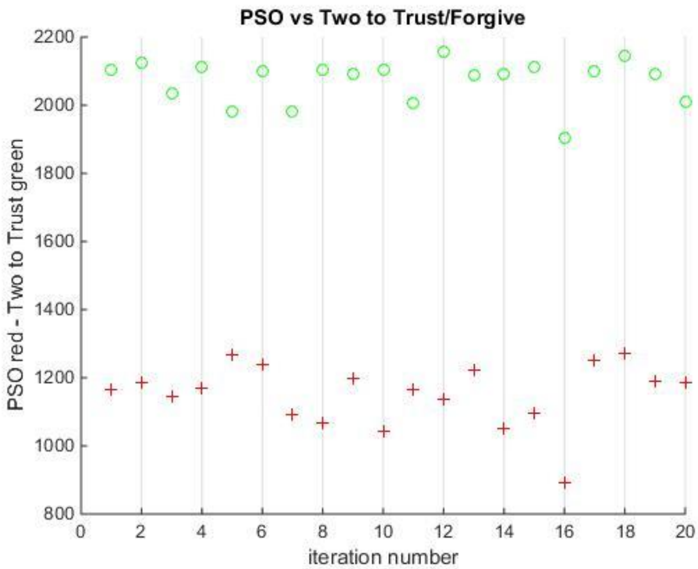



Figure 7- PSO vs Twott 1st level

Ενώ στους τρεις πρώτους παρατηρούμε πως υπάρχει μεγάλη διαφορά μεταξύ των επιδόσεων, στους επόμενους δύο (tft και evil tft) βλέπουμε πολύ κοντινές επιδόσεις, με τον PSO, βέβαια, να έρχεται δεύτερος στις περισσότερες επαναλήψεις ή στην καλύτερη ισόπαλος:

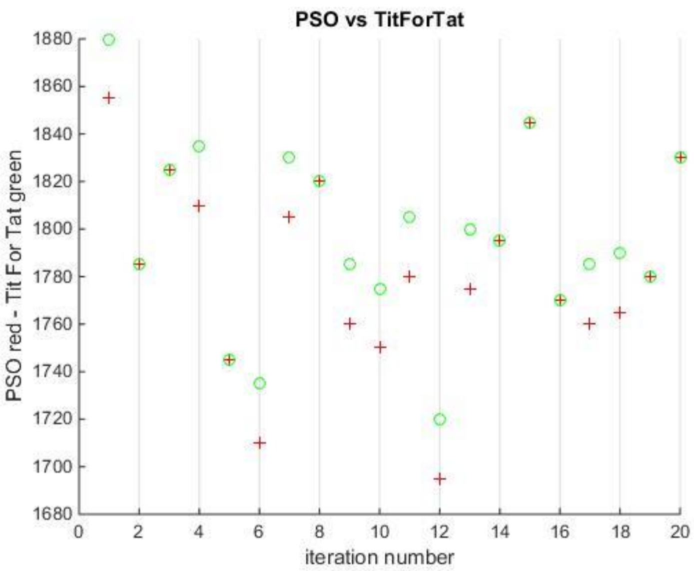



*Figure 8- PSO vs TFT 1st level*

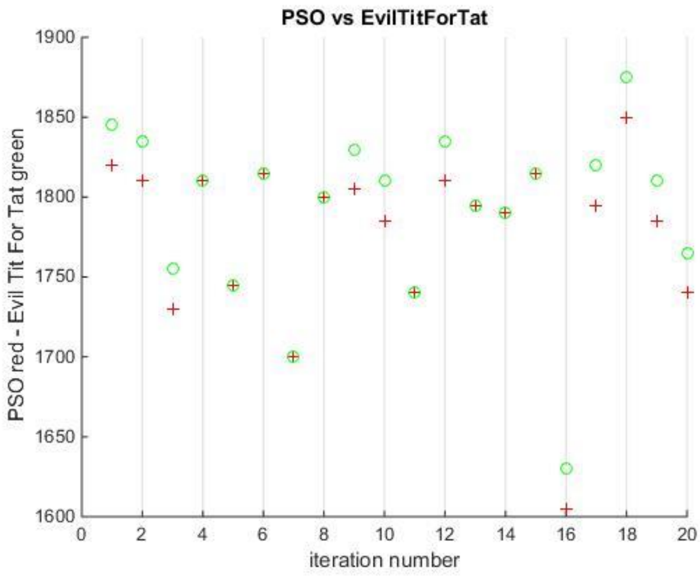



*Figure 9- PSO vs E-TFT 1st level*

Η παραπάνω συμπεριφορά οφείλεται στο γεγονός πως και οι δύο αυτοί αλγόριθμοι ενεργούν πάντα προσαρμοστικά πάνω στον αντίπαλό τους.

Σε επόμενο βήμα και για να συγκρίνουμε κατα πόσον βελτιώνεται ο αλγόριθμος αυξάνοντας τις επαναλήψεις, εισαγάγαμε την ποσοστιαία διαφορά του PSO με τον εκάστοτε αντίπαλο. Πιο συγκεκριμένα:

$$Perc_{opponent}^{z} = \frac{payoff_{PSO} - payoff_{opponent}}{payoff_{opponent}} \qquad (5.3)$$

όπου z αποτελεί δείκτη, που αντιστοιχεί σε 1=Low για μικρό αριθμό επαναλήψεων, 2=Medium για μεσαίο αριθμό επαναλήψεων, κτλ.

Έτσι για τις παραπάνω περιπτώσεις τα ποσοστά που διαμορφώθηκαν είναι τα ακόλουθα:

$Perc_{random}^{1}$= - 12,3%

$Perc_{pavlov}^{1}$= - 41,2%

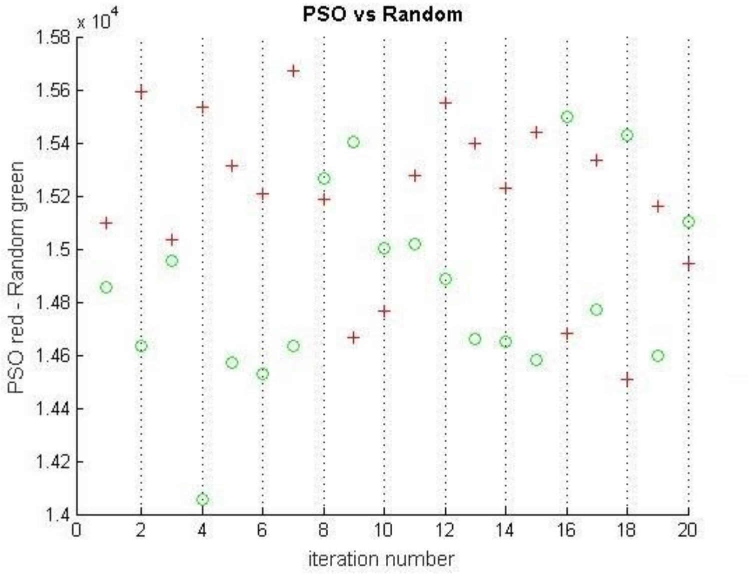
$Perc_{twott}^{1}$= - 47,8%

$Perc_{tft}^{1}$= - 0,88%

$Perc_{e-tft}^{1}$= - 1,02%

2$^{ο}$ τρέξιμο με υψηλότερο αριθμό επαναλήψεων:

Αυξήσαμε των αριθμό των παικτών σε N=20, τον αριθμό των επαναλήψεων σε L=50 και τον αριθμό των αντίπαλων παικτών σε M=50. Όπως αναφέρθηκε και παραπάνω, τα W είναι ανεξάρτητα μεταξύ τους και είναι απλά ενδεικτικά οπότε τα κρατήσαμε σταθερά σε W=20. Παρακάτω παραθέτουμε τα αντίστοιχα scatter plot γι' αυτές τις τιμές:



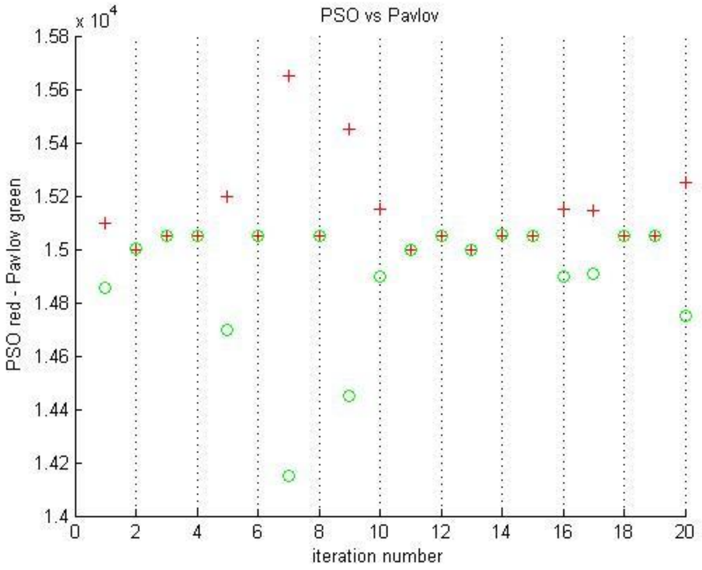




*Figure 10- PSO vs Random 2nd level*

*Figure 11- PSO vs Pavlov 2nd level*

Όπως παρατηρούμε στο 1$^{ο}$ σχήμα, έχει σημειωθεί ήδη αρκετή βελτίωση των επιδόσεων του PSO παίκτη έναντι ενός random, αλλά ακόμα διαπιστώνουμε πως σε μερικές παρτίδες ο PSO έρχεται δεύτερος.

Με αντίπαλο pavlov διαπιστώνουμε πως έχει ανατραπεί εξ' ολοκλήρου η αρχική επίδοση αλλά παρ' όλα αυτά σε 12 από τι 20 δοκιμές φέραμε ισοπαλία με τον αντίπαλο. Αντίθετα με αντίπαλο two to trust διαπιστώνουμε πως η κατάσταση χειροτερεύει όσο αυξάνονται οι επαναλήψεις όπως φαίνεται και παρακάτω. Αυτό οφείλεται στο γεγονός πως η συγκεκριμένη στρατηγική παρουσιάζει μεγάλο ποσοστό αποστασιών (μηδενικά - 0), το οποίο της δίνει πολύ καλές επιδόσεις εναντίον άλλων παικτών. Αλλά αν τα επίπεδα των συνολικών αμοιβών μιας τέτοιας στρατηγικής είναι τόσο χαμηλά, τότε σε ρεαλιστικές συνθήκες, μπορούν να αποβούν καταστροφικά.

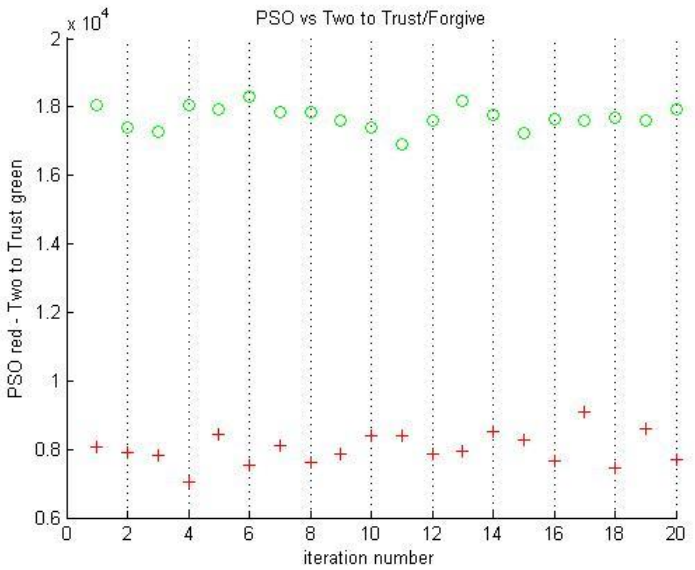



*Figure 12- PSO vs TwoTT 2nd level*

Η στρατηγική του ‘two to trust to forgive’ αποδείκτηκε πολύ ανταγωνιστική έναντι της PSO. Φάνηκε πως μεγαλώνοντας τον αριθμό των επαναλήψεων, ζημιωνόταν ο PSO αντί να βελτιώνεται. Με την ολοκλήρωση και των παρακάτω επαναλήψεων θα αναφέρουμε και τους λόγους που συνέβη κάτι τέτοιο.

Τέλος, γι αυτή την σειρά επαναλήψεων έχουμε τα scatter plot των στρατηγικών tit for tat και evil tit for tat όπως επισυνάπτονται παρακάτω:

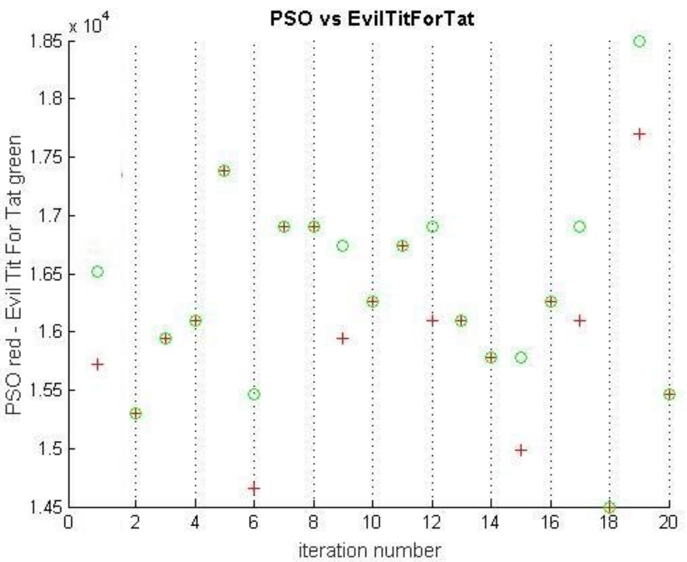



*Figure 13- PSO vs E-TFT 2nd level*

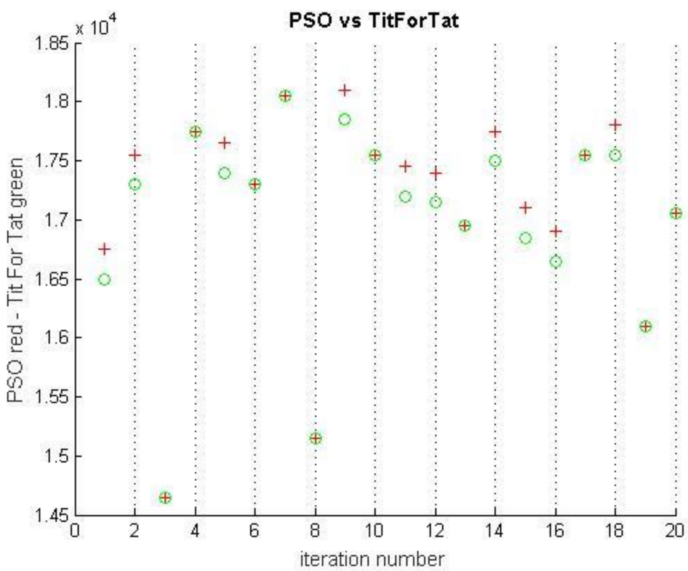



*Figure 14- PSO vs TFT 2nd level*

Βάση του τύπου 5.3, τα νέα ποσοστά των αποδόσεων είναι τα ακόλουθα:

$Perc^2_{random}$= 2,55%

$Perc^2_{pavlov}$= 1,64%

$Perc^2_{twott}$= - 55,3%

$Perc^2_{tft}$= 0.8%

$Perc^2_{e-tft}$= - 1,8%

Απ' τις παραπάνω γραφικές αλλά και τα ποσοστά παρατηρούμε πως η κατάσταση ανατράπηκε εντελώς όσον αφορά τους αντιπάλους random και Pavlov. Η επιδόση του PSO αυξήθηκε και όσον αφορά τον tft. Εντυπωσιακό, βέβαια, ήταν το γεγονός πως ο e-tft δεν επέφερε καλύτερα απότελεσματα.

<u>3ο τρέξιμο με υψηλότερο αριθμό επαναλήψεων:</u>

Συνεχίζοντας, αυξήσαμε αρκετά τον αριθμό των επαναλήψεων για να εξετάσουμε κατά αν και πόσο βελτιώνονται ακόμα οι επιδόσεις του PSO. Πιο συγκεκριμένα, ανεβάσαμε το N σε 400, το L σε 150 και τον αριθμό των αντίπαλων παικτών σε M=100. Όπως αναφέρθηκε και παραπάνω, τα W είναι ανεξάρτητα μεταξύ τους και είναι απλά ενδεικτικά, οπότε τα κρατήσαμε σταθερά σε W=20. Παρακάτω παραθέτουμε τα αντίστοιχα scatter plot γι' αυτές τις τιμές:

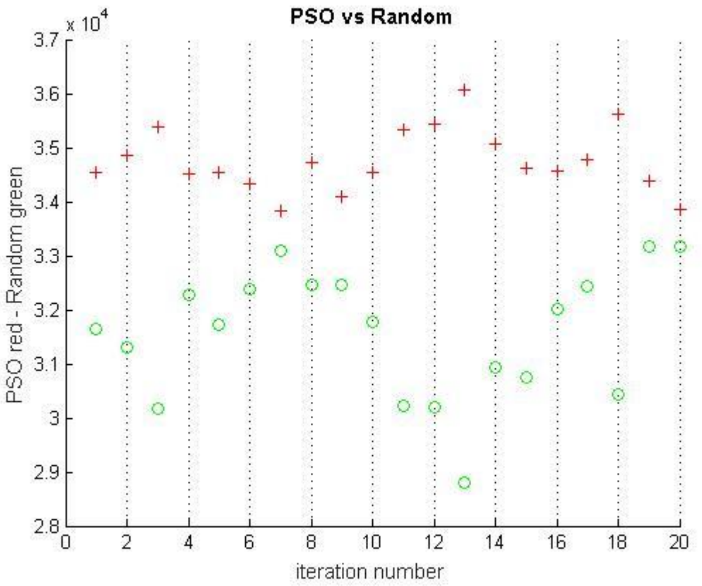



*Figure 15- PSO vs Random 3rd level*

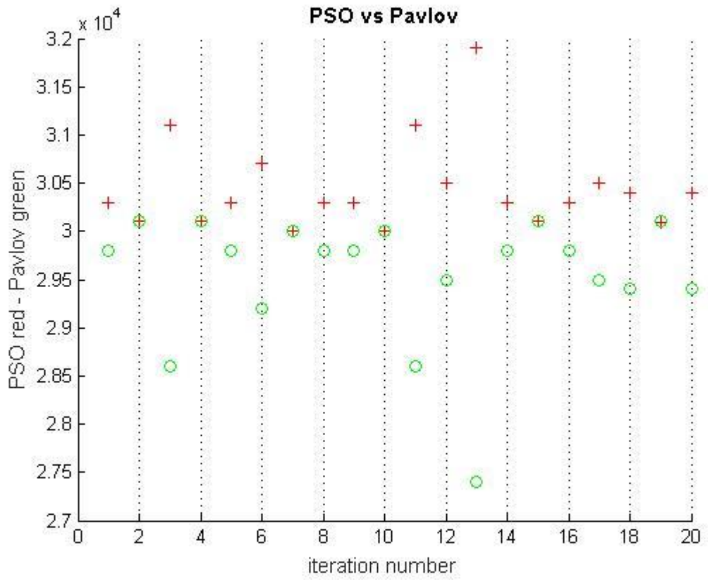



*Figure 16- PSO vs Pavlov 3rd level*

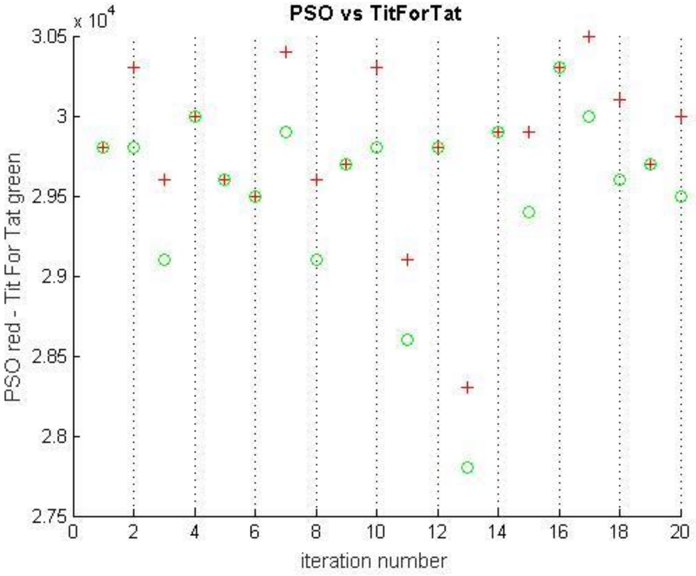



*Figure 17- PSO vs TFT 3rd  level*

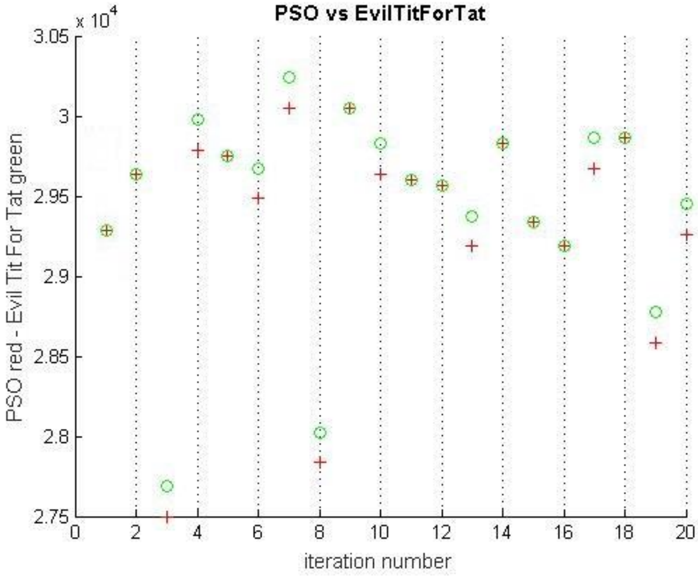



*Figure 18- PSO vs E-TFT 3rd  level*

Στα παραπάνω σχήματα προτιμήσαμε να μην συμπεριλάβουμε εκ νέου του ‘two trust to forgive’, δεδομένου ότι είχε ξανά σχετική απόδοση όπως και νωρίτερα. Για την συγκεκριμένη στρατηγική διαπιστώθηκε πως ο PSO δεν ήταν ικανός να προβλέψει την κίνηση του αντιπάλου μετά από 2 γύρους και όχι έναν. Όπως αναφέραμε και παραπάνω, στόχος που είχαμε θέσει ήταν να οδηγήσουμε τους παίκτες σε κοινές συνεργασίες και αν είναι δυνατό να έχουμε και καλύτερες αμοιβές απ’ αυτούς. Στην στρατηγική ‘two trust to forgive’, επειδή ο παίκτης αποφασίζει να συνεργαστεί μετά από 2 επαναλαμβανόμενες συνεργασίες, το αποτέλεσμα ήταν να έχει μεγάλο αριθμό αποστασιών (~120/150). Αντίθετα επειδή ο PSO δεν μπορούσε με την συνεργασία του να επηρεάσει τον αντίπαλο, ζημιωνόταν όλο και περισσότερο.

Βάση του τύπου 5.3, τα νέα ποσοστά των αποδόσεων είναι τα ακόλουθα:

$Perc^{3}_{random}$= 10,1%

$Perc^{3}_{pavlov}$= 3,8%

$Perc^{3}_{tft}$= 1,44%

$Perc^{3}_{e-tft}$= - 0,35%

Συνολικά, το καλύτερο ποσοστό ο PSO το παρουσίασε έναντι random αντιπάλων. Το γεγονός όμως που δεν ήταν πλήρως θετικό σε αυτή την περίπτωση είναι πως δεν κράτησε σε υψηλά επίπεδα τις συνολικές συνεργασίες του. Δηλαδή, ουσιστικά 'ξεγέλασε' τους αντιπαλους του, με ένα ποσοστό συνεργασιών ~45-55%. Το ποσοστό αυτό, επαληθεύεται κι απ' τις παραπάνω γραφικές παραστάσεις, όπου οι συνολικές αμοιβές του PSO έναντι random, είναι σχετικά πολύ υψηλές σε σχέση με των υπολοίπων.

Μικρή βελτίωση παρουσίασαν και οι Pavlov με tft, ενώ τελικά ο etft, ανέκαμψε μειώνοντας τις ζημιές του και φέρνοντας σχεδόν ισοπαλία. Αξιοσημείωτο στους τρεις τελευταίους αντιπάλους είναι το γεγονός πως ο PSO κατάφερε να μεταφέρει το 'πνεύμα' συνεργασίας στους αντιπάλους του. Κατά μέσο όρο ~100/150 συνεργασίες έναντι Pavlov και ~110/150 συνεργασίες έναντι tft και etft. Αυτό συνέβη γιατί και οι τρεις παραπάνω στρατηγικές βασίζονται στην προσαρμοστικότητα. Αυτός άλλωστε ήταν και ο λόγος που εξ αρχής οι tft και etft δεν διέφεραν και πολύ με την τυχαία και 'κακή' πρώτη στρατηγική μας. Όσο βελτιωνόταν η στρατηγική του PSO έτσι βελτιώνονταν και αυτοί.

Εδώ να αναφέρουμε, πως πραγματοποιήθηκαν και περαιτέρω δοκιμές με μεγαλύτερο ή και πολύ μεγαλύτερο αριθμό επαναλήψεων, αλλά φάνηκε πως τα ποσοστά παρέμεναν στα ίδια επίπεδα. Θεωρούμε, λοιπόν, πως ο PSO αλγόριθμος ολοκληρώθηκε, πετυχαίνοντας να μεγιστοποιήσει, όσο ήταν δυνατό, τα κέρδη του και παράλληλα να μεταφέρει και το 'πνεύμα' συνεργασίας στους αντιπάλους του.

## 5.3 PSO με θόρυβο

Συνεχίζοντας τον πειραματισμό και την ανάλυση, αποφασίσαμε να εξετάσουμε και αποτελέσματα PSO, όταν στο περιβάλλον εισαχθεί θόρυβος. Έτσι διαλέξαμε να εισάγουμε ένα ποσοστό 10%, όπου ο PSO λαμβάνει αντίθετη απόφαση από την ενδεικνυόμενη.

Παραθέτουμε τα σχετικά scatter plot καθώς και τα σχόλια μας:

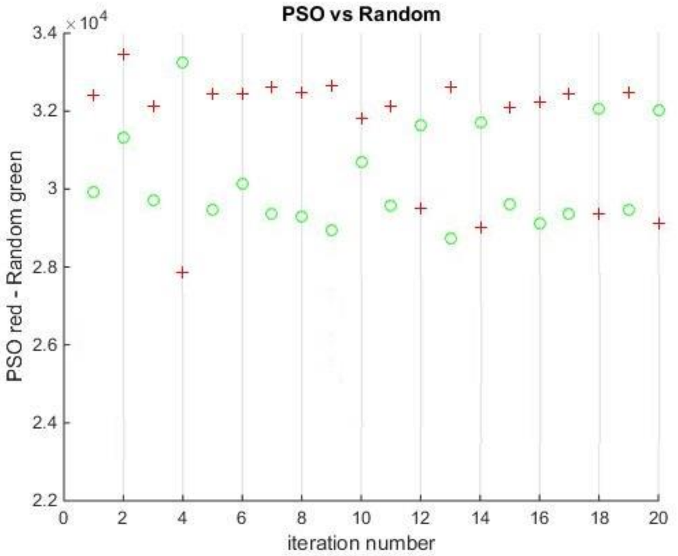



*Figure 19- PSO vs Random noise level*

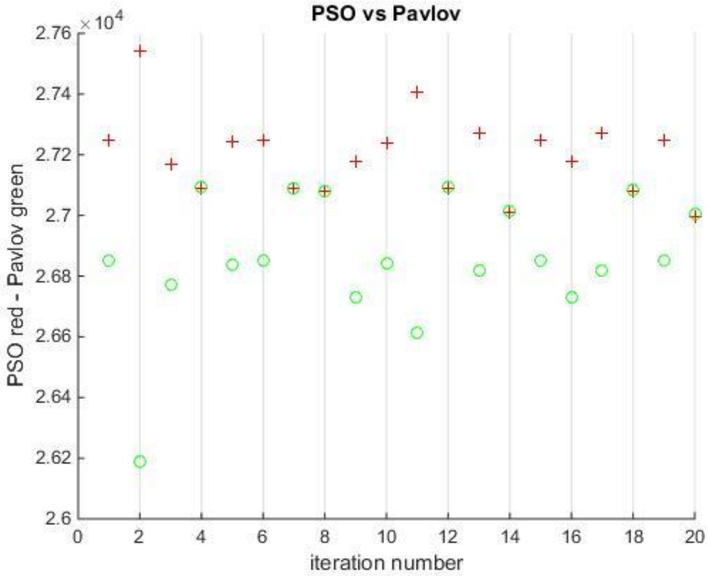



*Figure 20- PSO vs Pavlov noise level*

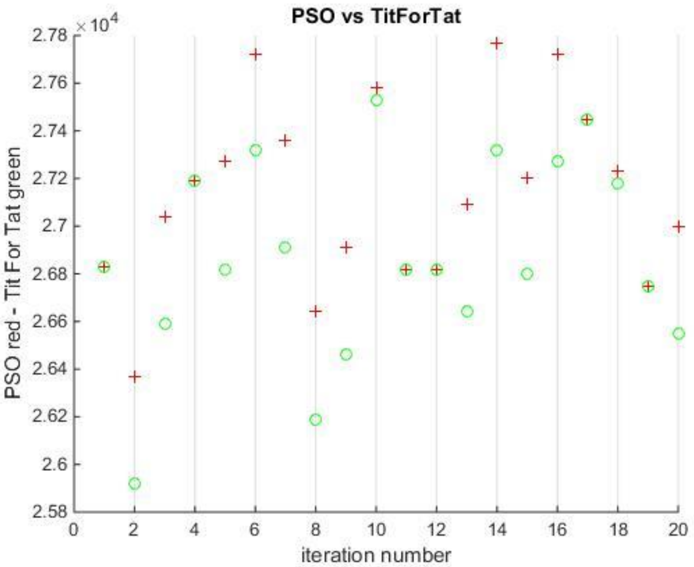



*Figure 21- PSO vs TFT noise level*

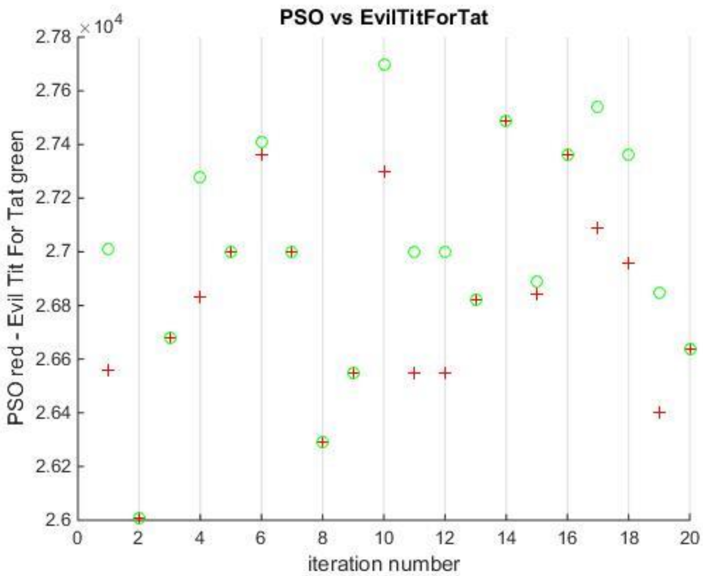



*Figure 22- PSO vs E-TFT noise level*

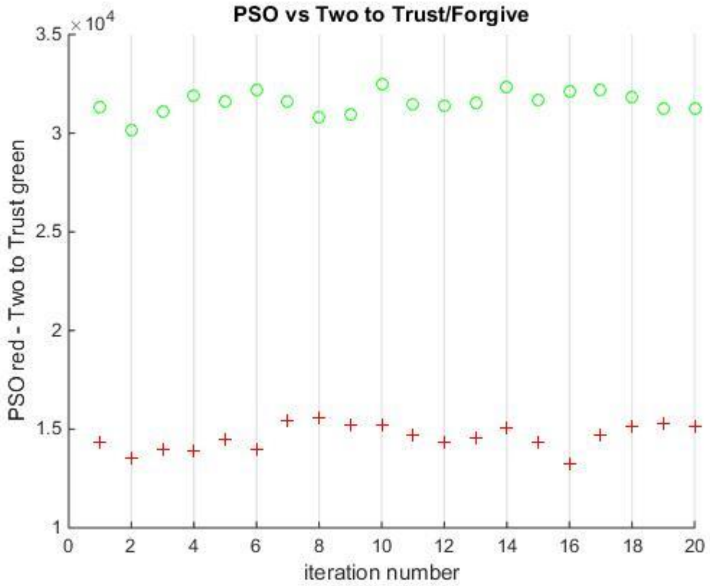



*Figure 23- PSO vs TwoTT noise level*

Εξετάζοντας τα παραπάνω σχήματα, παρατηρούμε πως τα αποτελέσματα δεν επηρεάστηκαν ιδιαίτερα. Αξιοσημείωτο όμως είναι πως το ποσοστό του θορύβου φαίνεται να έχει περάσει στα ποσοστά επιδόσεων του PSO με αρνητικό τρόπο.

Πιο συγκεκριμένα και βάση του τύπου 5.3, τα νέα ποσοστά των αποδόσεων είναι τα ακόλουθα:

$Perc_{random}^{3+Noise}$= 8,84%

$Perc_{pavlov}^{3+Noise}$= 1,87%

$Perc_{tft}^{3+Noise}$= 0.92%

$Perc_{e-tft}^{3+Noise}$= - 0,56%

$Perc_{twott}^{3+Noise}$= - 54,7%

Συμπεραίνουμε, λοιπόν, πως ο PSO επηρεάζεται αρνητικά όταν αναπτύσεται σε περιβάλλον θορύβου.

## 5.4 PSO με βάρος αδράνειας w

Στο κεφάλαιο 4 (§4.3) αναφέραμε την παραλλαγή του PSO όπου σε κάποια προβλήματα έχει επιτευχθεί καλύτερη απόδοση με την εισαγωγή ενός βάρους αδράνειας w. Παρακάτω παραθέτουμε σχετικά scatter plot μετά την εισαγωγή του σχετικού βάρους w.

Οι δοκιμές πραγματοποιήθηκαν για διάφορους αριθμούς επαναλήψεων, αλλά για να έχουμε συγκριτικά αποτελέσματα, παραθέτουμε τα αποτελέσματα για N=400, L=150 και M=100.

Επίσης, να αναφερθεί εδώ, πως για την επιλογή του w, έγινε χρήση της σχέσης (4.4). Για την επιλογή των $w_{max}$ και $w_{min}$ έπρεπε να πραγματοποιηθούν δοκιμαστικά τρεξίματα της διαδικασίας έτσι ώστε να προκύψουν τα ιδανικά μεγέθη. Δεδομένου ότι το πρόβλημα είναι διακριτό και όχι συνεχές, οι μεταβολές ήταν σχετικά μικρές με την αλλαγή αυτών των μεγεθών. Τελικώς επιλέχθηκαν για $w_{max}$=20 και $w_{min}$=1 και $iter_{max}$=150*N, όπου 150 είναι οι γύροι που παίζει ένας αντίπαλος με τον άλλον και N είναι ο αριθμός των παικτών. Όπως σημειώθηκε και στο προηγούμενο κεφάλαιο, Το w ξεκινάει από μια μεγάλη τιμή και με το πέρασμα των επαναλήψεων το t τείνει να γίνει ίσο με τον παρονομαστή και το w να πάρει τελικώς την τιμή $w_{min}$. Στην προκειμένη περίπτωση, εφόσον το $w_{min}$ = 1, ο αλγόριθμος μετά από συγκεκριμένο αριθμό επαναλήψεων καταλήγει στον κλασσικό PSO.

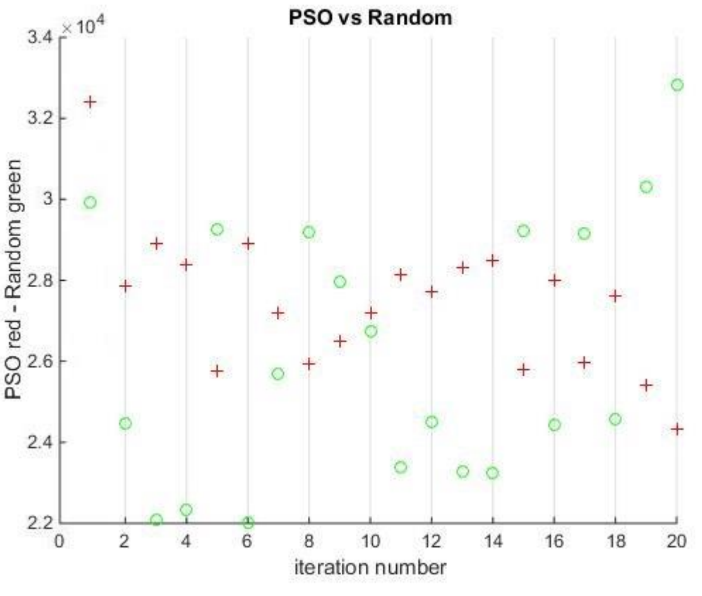



Figure 24- PSO vs Random ‘w’ level

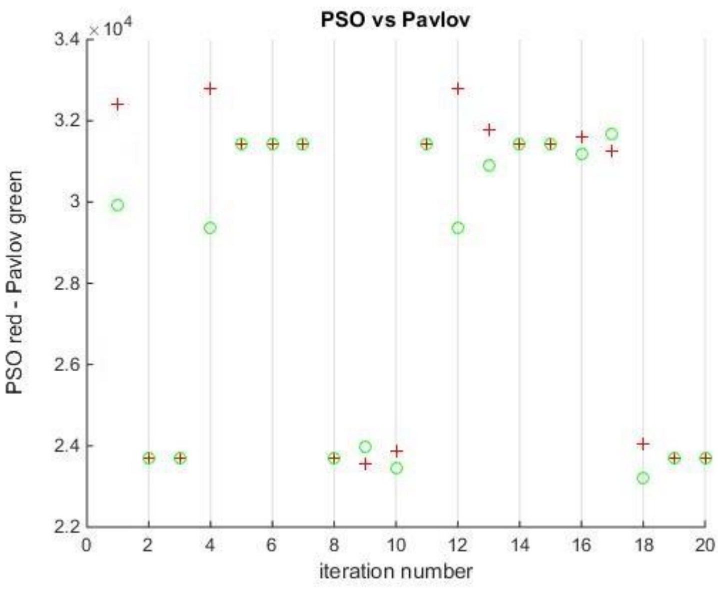



Figure 25- PSO vs Pavlov ‘w’ level

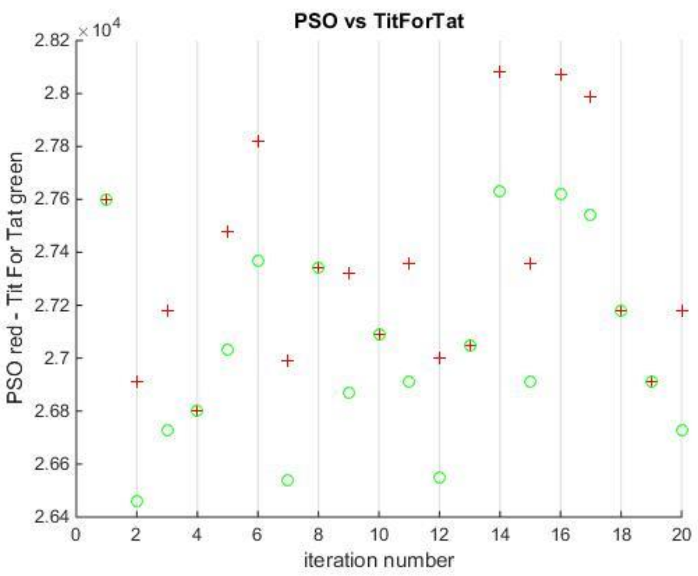



*Figure 26- PSO vs TFT 'w' level*

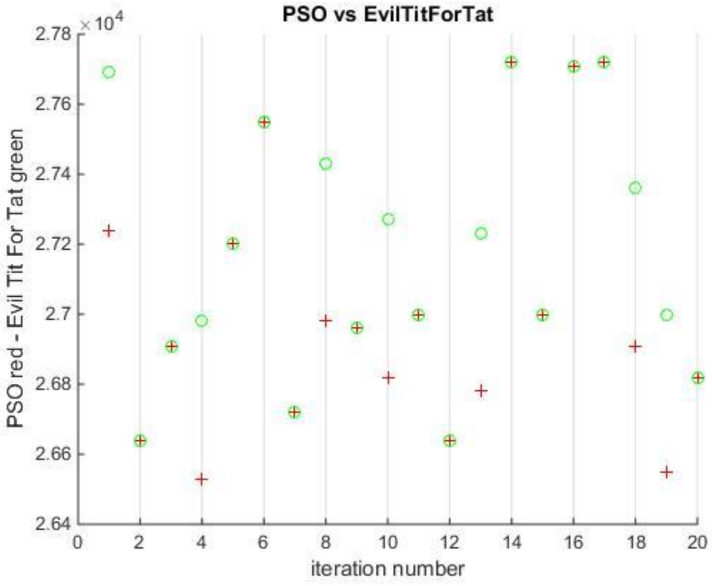



*Figure 27- PSO vs E-TFT 'w' level*

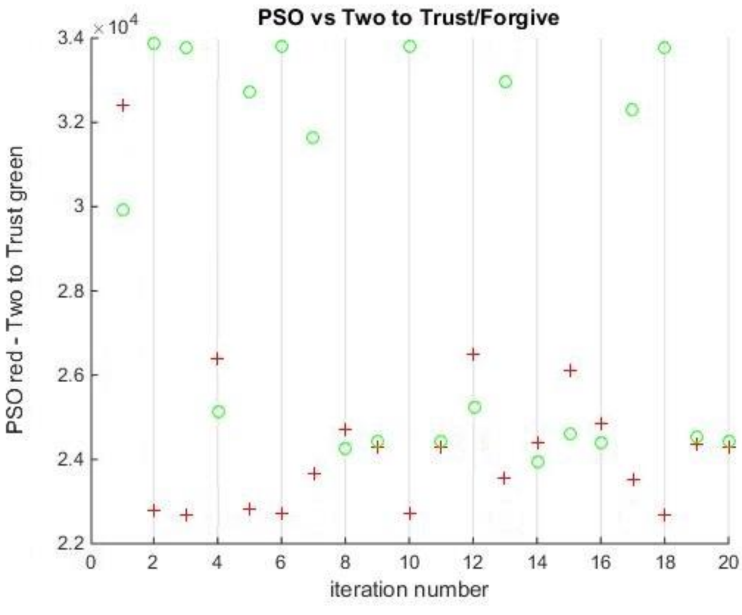



*Figure 28- PSO vs TwoTT 'w' level*

Τα συνολικά σχόλια με το πέρας των πειραματισμών ήταν αρνητικά καθώς ο PSO χειροτέρεψε τις επιδόσεις του. Επίσης φάνηκε, πως στις περιπτώσεις του random και του Pavlov αντιπάλου, χάθηκε η ισορροπία αποτελεσμάτων που υπήρχε μεταξύ των παιχνιδιών. Φαίνεται πως με την προσθήκη του βάρους αδράνειας σε άλλα παιχνίδια έχανε με αρκετή διαφορά και σε άλλα υπερτερούσε με το ίδιο ή και μεγαλύτερο άθροισμα αμοιβών.

Τα ποσοστά των επιδόσεων βάση της σχέσης 5.3 ήταν τα ακόλουθα:

$Perc_{random}^{3-w}$= 5,9%

$Perc_{pavlov}^{3-w}$= 1,77%

$Perc_{tft}^{3-w}$= 1,17%

$Perc_{e-tft}^{3-w}$= - 0,56%

$Perc_{twott}^{3-w}$= - 12,74%

Εξετάζοντας και τις παραπάνω ποσοστιαίες επιδόσεις, παρατηρούμε πως τα ποσοστά έχουν εμφανώς μειωθεί αλλά όπως αναφέραμε παραπάνω έχει χαθεί η ισορροπία και η ευστάθια του συστήματος καθώς στον κλασσικό PSO διαπιστώσαμε πως η στρατηγική ήταν σταθερά νικητήρια.

Οι tft και etft φαίνεται να έχουν διατηρήσει την επιδοσή τους με τον etft να φέρνει ξανά σχεδόν ισοπαλία με τον PSO.

Το μόνο αξιοσημείωτο με αυτήν την προσθήκη, ήταν η βελτίωση του PSO έναντι της ‘two trust to forgive). Σταθερά αρνητικό το ποσοστό, αλλά τουλάχιστον σημειώθηκε βελτίωση, την οποία δεν είχαμε συναντήσει σε οποιαδήποτε άλλη δοκιμή.

## 5.5 PSO με περιορισμό ταχύτητας.

Στο κεφάλαιο 4 (§4.3) έγινε αναφορά και της παραλλαγής του PSO όπου ένας περιοριστικός παράγοντας προστίθεται στον βασικό τύπο έτσι ώστε να μην αφήσει την ταχύτητα να ξεφύγει και να χαθούν πιθανά βέλτιστα. Αποφασίσαμε να εισάγουμε και αυτή την παραλλαγή του βασικού PSO έτσι ώστε να δούμε και να αναλύσουμε τα σχετικά της αποτελέσματα για το συγκεκριμένο παίγνιο.

Να σημειωθεί εδώ πως επειδή παρατηρήθηκε ότι η συγκεκριμένη παραλλαγή συγκλινει σχετικά πιο αργά απ’ τις προηγούμενες, τρέξαμε την διαδικασία για παραπάνω επαναλήψεις. Ενδεικτικά παραθέτουμε τα scatter plot για N=200, L=200 και M=400. Σε αντίθεση με τις προηγούμενες εκδοχές, εδώ διαπιστώσαμε πως αυξάνοντας τον αριθμό των γύρων που παίζει ο παίκτης με τους αντιπάλους (M) αυξάνεται και η επίδοση του αλγορίθμου. Επίσης, δεδομένης της συνθήκης $c = c_1 + c_2, c > 4$, επιλέξαμε για $c_1$=$c_2$=3.

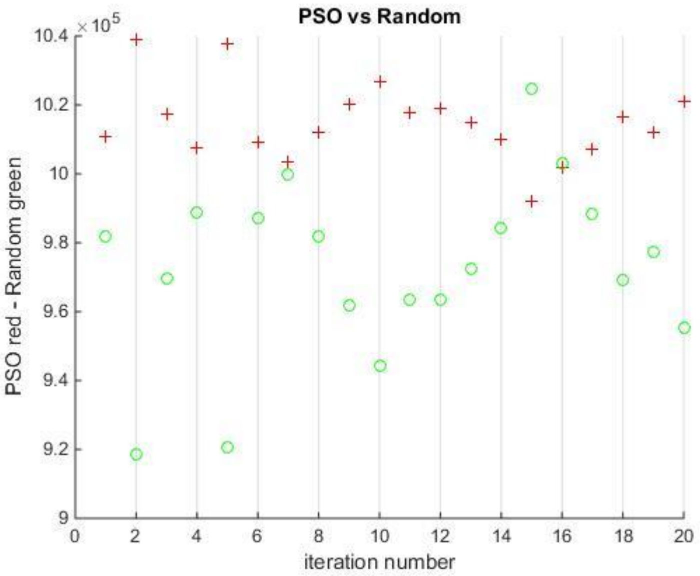



*Figure 29- PSO vs Random 'Veloc. Restr.' Level*

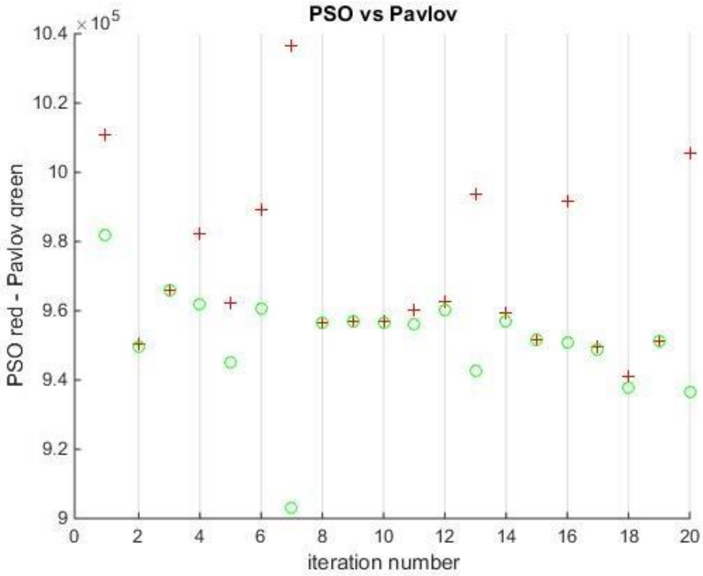



*Figure 30 - PSO vs Pavlov 'Veloc. Restr.' level*

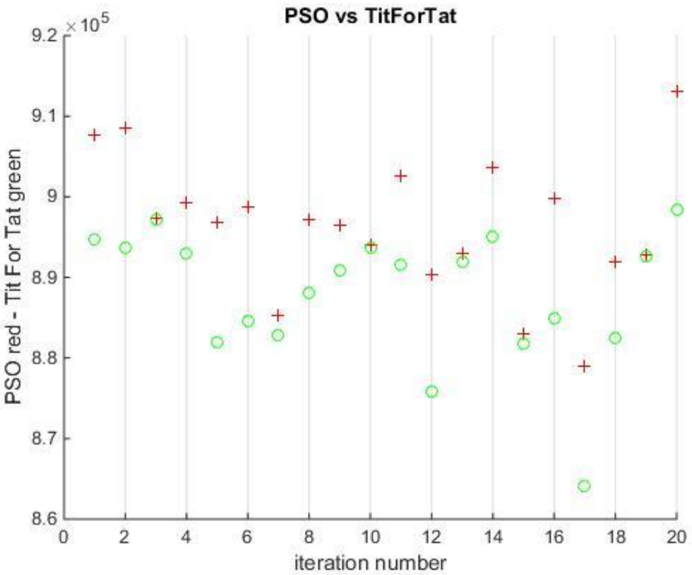



*Figure 31- PSO vs TFT 'Veloc. Restr.' Level*

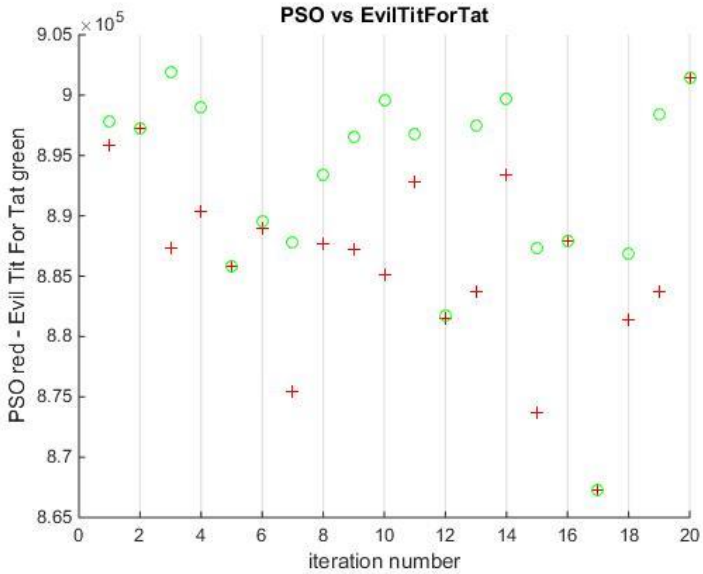



*Figure 32- PSO vs E-TFT 'Veloc. Restr.' level*

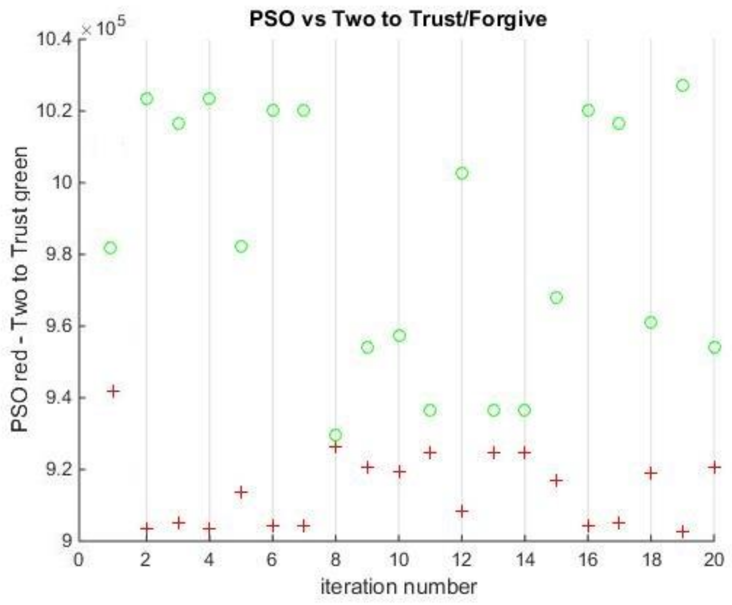



*Figure 33- PSO vs TwoTT 'Veloc. Restr.' level*

Επίσης τα αντίστοιχα ποσοστά βάση της σχέσης (5.3) είναι τα ακόλουθα:

$Perc_{random}^{3-Vr}$= 4,48%

$Perc_{pavlov}^{3-Vr}$= 2,04%

$Perc_{tft}^{3-Vr}$= 1,03%

$Perc_{e-tft}^{3-Vr}$= - 0,64%

$Perc_{twott}^{3-Vr}$= - 6,72%

Από τις παραπάνω γραφικές αλλά και τις ποσοστιαίες αποδόσεις διαπιστώνουμε πως ούτε αυτή η παραλλαγή έφερε καλύτερα αποτελέσματα από τον βασικό PSO. Συγκριτικά με την παραλλαγή του βάρους αδράνειας, τα ποσοστά Pavlov και tft έχουν βελτιωθεί και παράλληλα δείχνουν και μεγαλύτερη ευστάθεια συγκρίνοντας τις διαφορετικές διαδικασίες μεταξύ τους. Επίσης η επίδοση έναντι του random αντιπάλου έχει μειωθεί αλλά συγκρίνοντας τις γραφικές με την παραλλαγή του βάρους αδράνειας διαπιστώνουμε πως στον περιορισμό της ταχύτητας ο PSO χάνει μόνο σε μία εκ των 20 διαδικασιών, οπότε θα μπορούσαμε κι εδώ να πούμε πως με τον περιορισμό της ταχύτητας έχουμε πιο ισορροπημένα αποτελέσματα.

Συνολικά, και έχοντας ολοκληρώσει όλους τους ελέγχους μας, την καλύτερη επίδοση, όπως παρατηρήσαμε, την σημείωσε ο βασικός PSO. Αύξηση των επιδόσεων παρατηρήθηκε περισσότερο όταν αυξανόταν ο συντελεστης Μ, που αναφέρεται στα παιχνίδια έναντι αντίπαλων στρατηγικών. Αξιοσημείωτα ήταν βέβαια και τα αποτελέσματα που προέκυψαν με την προσθήκη του θορύβου, καθώς το ποσοστό (10%) που προστέθηκε, φάνηκε να περνάει στον PSO και να επηρεάζει αρνητικά τις επιδόσεις του. Τέλος, από τις παραλαγές που δοκιμάστηκαν, διαπιστώθηκε πως εκείνη με τον περιορισμό της ταχύτητας είχε πιο ισορροπημένα αποτελέσματα. Παράλληλα διατήρησε σε καλό βαθμό το ποσοστό συνεργασιών όπως είχαν σημειωθεί και παραπάνω για τον βασικό PSO.

# Βιβλιογραφία


(1) Noam Nisan, Tim Roughgarden, Eva Tardos, Vijay V. Varirani – Algorithmic game theory – Βιβλίο – Cambridge – 2007

(2) Αθανάσιος Μυγδαλάς – Θεωρία Παιγνίων και Προγραμματισμός Ισορροπίας – Βιβλίο – Χανιά – 2007

(3) Graham Romp – Game Theory Introduction and Applications – Βιβλίο – Oxford – 1997

(4) Ι. Μαρινάκης, Μ. Μαρινάκη, Ν. Ματσατσίνης, Κ. Ζοπουνίδης – Μεθευρετικοί και Εξελικτικοί Αλγόριθμοι σε προβλήματα Διοικητικής Επιστήμης – Βιβλίο – Αθήνα – 2011

(5) Andrew Errity – Evolving Strategies for the Prisoner's Dilemma –Technical Manual – Faculty of Computing and Mathematical Science Dep. – Dublin – 2003

(6) http//:el.wikipedia.org/wiki/θεωρία παιγνίων

(7) Παναγοπούλου Ν. Παναγιώτα – Αλγοριθμική και εξελικτική Θεωρία Παιγνίων – Διδακτορικό Δίπλωμα, Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής Πολυτεχνικής Σχολής Πανεπιστημίου Πατρών – 2009

(8) Elio Piccolo, Giovanni Squillero – Adaptive Opponent Modeling for the IPD – IEEE Congress on Evolutionary Computation – Ιούνιος 2011

(9) Kotler Phlilip – Marketing management: Analysis, planning, implementation and control – Βιβλίο – Ιανουάριος 1994

(10) A.E. Eiben, J. E. Smith – Introduction to Evolutionary Computing – Βιβλίο – Springer-Verlag Berlin Heidelberg – 2003

(11) Yuce Tekol – To cooperate or to defect: That's the Prisoner's Dilemma – IEEE Computer Society Looking Forward Magazine – 2003

(12) Wang Tao, Chen Zhi-Gang, Deng Xiao-Heng και Zhagn Jin – Properties of Prisoner's Dilemma Game Based on Genetic Algorithm – IEEE Congress on Evolutionary Computation – Δεκέμβριος 2003

(13) Yuce Tekol, Adnan Acan – Ants can Play Prisoner's Dilemma – IEEE Congress on Evolutionary Computation – Ιούνιος 2011

(14) Nelis Franken, Andries P. Engelbrecht – Particle Swarm Optimization Approaches to Coevolve Strategies for the Iterated Prisoner's Dilemma – IEEE Transaction on Evolutionary Computation Vol .9 No. 6 – Δεκέμβριος 2005

(15) Philip Hingston – Evolving Group Strategies for IPD – IEEE Congress on Evolutionary Computation – Ιούλιος 2010

(16) Kumar Chellapilla – Evolution, Neural Networks, Games and Intelligence – Proceedings of the IEEE Vol .87 Issue .9 – Σεπτέμβριος 1999

(17) Jiawei Li,Philip Hingston – Engineering Design of Stategies for Winning Iterated Prisoner's Dilemma Competitions – Computational Intelligence and AI in Games, IEEE Transactions on Vol .3 Issue .4 – Αύγουστος 2011