



ΣΤΡΑΤΙΩΤΙΚΗ ΣΧΟΛΗ ΕΥΕΛΠΙΔΩΝ
Τμήμα Στρατιωτικών Επιστημών

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΔΙΔΡΥΜΑΤΙΚΟ ΔΙΑΤΜΗΜΑΤΙΚΟ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΑΚΑΔΗΜΑΪΚΟΥ ΕΤΟΥΣ 2018-19
ΕΦΑΡΜΟΣΜΕΝΗ ΕΠΙΧΕΙΡΗΣΙΑΚΗ
ΕΡΕΥΝΑ & ΑΝΑΛΥΣΗ)

(ΠΔ 96 /2015/ΦΕΚ 163Α'/20.08.2014)



ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
Σχολή Μηχανικών Παραγωγής & Διοίκησης

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

ΣΤΡΑΤΗΓΙΚΕΣ ΕΞΕΛΙΞΗΣ και ΣΤΟΧΑΣΤΙΚΕΣ ΜΕΘΟΔΟΙ
ΑΝΑΖΗΤΗΣΗΣ ΣΕ ΠΡΟΒΛΗΜΑΤΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ ΜΕΓΑΛΗΣ
ΚΛΙΜΑΚΑΣ ΠΟΛΥ-ΠΡΑΚΤΟΡΙΚΩΝ ΡΟΜΠΟΤΙΚΩΝ ΣΜΗΝΩΝ

Διατριβή που υπεβλήθη για την μερική ικανοποίηση των απαιτήσεων
για την απόκτηση Μεταπτυχιακού Διπλώματος Ειδίκευσης

Υπό:

ΓΕΩΡΓΙΟΣ ΚΑΝΤΡΕΒΑΣ

A.M.: 2017018025

ΙΟΥΝΙΟΣ 2019

Η Μεταπτυχιακή Διατριβή του ΓΕΩΡΓΙΟΥ ΚΑΝΤΡΕΒΑ εγκρίνεται:

ΤΡΙΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Δάρας Ιω. Νικόλαος (Επιβλέπων)

Παπαδάκης Νικόλαος

Τσαφάρκης Στέλιος

ΣΕΛΙΔΑ ΣΚΟΠΙΜΑ ΚΕΝΗ

© Copyright υπό

Έτος 2019

Αφιερώνω την προσπάθειά μου στον Πατέρα μου.

ΣΕΛΙΔΑ ΣΚΟΠΙΜΑ ΚΕΝΗ

Θερμές ευχαριστίες στον επιβλέποντα καθηγητή και Κοσμήτορα
κ . Δάρα Ιω. Νικόλαο

Δομή Εργασίας (αφορά μια αντιπροσωπευτική δομή η οποία θα εξειδικεύεται ανά περίπτωση)

Format εργασιών:

Γραμματοσειρά: Garamond: 14

Line spacing: At least 16 pt Spacing before 0.6 pt

ΣΕΛΙΔΑ ΣΚΟΠΙΜΑ ΚΕΝΗ

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ.....	1
---------------	---

ΕΙΣΑΓΩΓΗ.....	2
ΚΕΦΑΛΑΙΟ 1	
Διαμόρφωση του προβλήματος μέσω μαθηματικών μοντέλων.....	5
ΚΕΦΑΛΑΙΟ 2	
Επίλυση του προβλήματος μέσω μαθηματικών μοντέλων.....	6
ΚΕΦΑΛΑΙΟ 3	
Ισχύουσες διατάξεις δομημένων πολιτικών για πολυπρακτορικά προβλήματα.....	13
ΚΕΦΑΛΑΙΟ 4	
Προσομοιώσεις.....	15
ΕΠΙΛΟΓΟΣ.....	22
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	23

ΠΕΡΙΛΗΨΗ

Η παρούσα μεταπτυχιακή διατριβή έχει ως στόχο να αποδείξει ότι οι εξελικτικές στρατηγικές (Evolution Strategies – ES) αποτελούν ειδική περίπτωση μίας κατηγορίας στοχαστικών αλγορίθμων, οι οποίοι εμφανίζουν αποτελεσματικές ασυμπτωτικές ιδιότητες και ωραία ποσοστά σύγκλισης. Οι Εξελικτικές Στρατηγικές εκτός από μία ευρέως γνωστή εναλλακτική μέθοδος ενισχυτικής μάθησης , αποτελεί και μια ειδική περίπτωση μεθοδολογιών στοχαστικής αναζήτησης που βασίζεται σε μαθηματικά μοντέλα. Η προσέγγιση της παραπάνω μεθόδου (ES) σε προβλήματα πολλαπλών παραγόντων θα παρουσιαστεί σε αυτήν την εργασία μέσω δύο προσομοιώσεων. Θα παρουσιάσουμε το πώς θα προσεγγίσει και έπειτα θα επιτεθεί ένα σμήνος UAV σταθερής πτέρυγας σε μία καλά οργανωμένη αμυντική βάση και δεύτερον , την σύγκρουση αυτού του σμήνους με ένα άλλο εχθρικό σμήνος UAV το οποίο αμύνεται της προαναφερόμενης βάσης.

ΕΙΣΑΓΩΓΗ

Ο αλγόριθμος Evolution Strategies [1] πάνω στον οποίο θα στηριχτούμε, είναι μια ειδική περίπτωση τάξης μοντέλων με βάση τις στοχαστικές μεθόδους αναζήτησης, γνωστή και ως Gradient Based Adaptive Stochastic Search (GASS) [3]. Αυτή η τάξη μεθόδων γενικεύει πολλές στοχαστικές μεθόδους αναζήτησης. Χρησιμοποιώντας τη σύγκλιση των Στρατηγικών Εξέλιξης, επιδεικνύουμε τον τρόπο με τον οποίο ο αλγόριθμος των ES μπορεί να χρησιμοποιηθεί αποτελεσματικά για την επίλυση τόσο των συνεταιριστικών, όσο και των ανταγωνιστικών προβλημάτων μεγάλης κλίμακας πολλών πρακτόρων [2]. Σε αρκετές περιπτώσεις, οι αλγόριθμοι που χρησιμοποιούνται για την επίλυση τέτοιου είδους προβλημάτων βασίζονται σε χειροκίνητους αλγορίθμους. Ένα παράδειγμα είναι ο αλγόριθμος Model Predictive Control (MPC), ο οποίος χρησιμοποιεί ανεξάρτητους ελεγκτές για κάθε παράγοντα του ελάχιστου πράκτορα, οι οποίοι διατηρούν μεταξύ τους κάποιο επίπεδο συνεργασίας και ανταλλαγής πληροφοριών [4], [5]. Αυτοί οι ελεγκτές όμως στηρίζονται σε χειροποίητο σχεδιασμό, έλεγχο, λειτουργία και διαδικασία ανατροφοδότησης. Αυτό σημαίνει ότι απαιτείται η ύπαρξη ειδικών χειριστών και εμπειρογνομώνων [2]. Ακόμα, η χρήση αυτού του είδους αλγορίθμων για την επίλυση πιο σύνθετων προβλημάτων είναι πολύ δύσκολη. Οι εξελικτικοί αλγόριθμοι δοκιμάστηκαν επίσης ως λύση στα προβλήματα πολλών παραγόντων συνήθως με μικρότερα, απλούστερα περιβάλλοντα και πολιτικές χαμηλής πολυπλοκότητας [6], [7]. Πρόσφατα παρουσιάστηκε μια υβριδική προσέγγιση, που συνδυάζει τη χρήση MPC και γενετικών αλγορίθμων για την εύρεση της συνάρτησης κόστους για ένα χειροκίνητο MPC ελεγκτή, ο οποίος χρησιμοποιήθηκε σε ένα σενάριο UAV ρομποτικών σμηνών [8].

Τα δύο σενάρια που θα χρησιμοποιηθούν σ' αυτήν την εργασία έχουν εξεταστεί προηγουμένως σε προσομοιωμένα περιβάλλοντα με μικρότερη πολυπλοκότητα [9], [8]. Θα αξιοποιήσουμε την υπολογιστική αποτελεσματικότητα και την ευελιξία του πρόσφατα αναπτυγμένου προσομοιωτή πολλαπλών πρακτόρων SCRIMMAGE για τα πειράματά μας [10]. Θα δείξουμε επίσης για το ανταγωνιστικό σενάριο πώς συντονίζεται με την πάροδο του χρόνου, μια στρατηγική, ως απάντηση σε έναν εχθρό που μαθαίνει να κάνει το ίδιο συγκρίνοντας την απόδοση του ES έναντι της μεθόδου Cross Entropy [2]. (Video 1)

ΚΕΦΑΛΑΙΟ 1

Διαμόρφωση του προβλήματος μέσω μαθηματικών μοντέλων

Όσον αφορά την αρχική διαμόρφωση του προβλήματος θα θεωρήσουμε μία συνάρτηση των οποίων τις μεταβλητές θα επεξηγήσουμε παρακάτω. Το επόμενο βήμα είναι να εξετάσουμε αν είναι εφικτό να λυθεί το συγκεκριμένο πρόβλημα χρησιμοποιώντας μεθόδους προσαρμοστικής στοχαστικής αναζήτησης Gradient Based και στη συνέχεια θα προσπαθήσουμε να λύσουμε το ίδιο πρόβλημα με τον αλγόριθμο ES, δείχνοντας παράλληλα πώς ο αλγόριθμος ES είναι μια ειδική υπο - περίπτωση αυτών των μεθόδων με πολύ περισσότερα πλεονεκτήματα. Η συνάρτηση έχει ως εξής :

$$\kappa^i = \underset{\kappa \in K}{\operatorname{argmax}} D(\kappa) \quad (1)$$

$$\text{με } K \subset \mathbb{R}^n, D: K \rightarrow \mathbb{R}$$

Το κ θα μπορούσε να είναι να είναι οποιοσδήποτε συνδυασμός μεταβλητών απόφασης του προβλήματός μας, συμπεριλαμβανομένων των βαρών του νευρικού δικτύου, των κερδών PID, των παραμέτρων σχεδιασμού υλικού κ.λπ. που επηρεάζουν το αποτέλεσμα των αποδόσεων D . Για τέτοιου είδους προβλήματα όπως αυτό που εξετάζουμε, το κ αντιπροσωπεύει συνήθως τις παραμέτρους της πολιτικής και το D είναι μια λειτουργία της διαδοχικής εφαρμογής της πολιτικής στο περιβάλλον.

ΚΕΦΑΛΑΙΟ 2

3 Επίλυση του προβλήματος μέσω μαθηματικών μοντέλων

Αρχικά λοιπόν, όπως προαναφέρθηκε θα προσπαθήσουμε να λύσουμε το πρόβλημά μας χρησιμοποιώντας την μέθοδο προσαρμοστικής στοχαστικής αναζήτησης , Gradient Based. Η συγκεκριμένη μέθοδος σκοπεύει στο να μετατρέψει το μη διαφοροποιήσιμο πρόβλημα βελτιστοποίησης (1) σε διαφοροποιήσιμο, καθορίζοντας ένα πιθανοτικό μοντέλο. Θα θεωρήσουμε λοιπόν έστω ότι το μοντέλο αυτό είναι το

$g(\kappa|v)=t(\kappa;v)$, $v \in N$, όπου v είναι μία παράμετρος που εξισώνει την κατανομή πιθανότητας. Οι τιμές της $D(\kappa)$ στο σύνολο τιμών $t(\kappa;v)$

είναι αναμενόμενο ότι θα είναι πάντα μικρότερες από από την μέγιστη τιμή της D .

$$\int_K D(\kappa)t(\kappa;v)d\kappa \leq D(\kappa^i) \quad (2)$$

Ο σκοπός της μεθόδου προσαρμοστικής στοχαστικής αναζήτησης , Gradient Based είναι ότι μπορεί κανείς να εκτελέσει μια αναζήτηση στο χώρο των παραμέτρων της κατανομής N αντί του K , για μια κατανομή που μεγιστοποιεί την προσδοκία στην (2) :

$$v^i = \underset{v \in N}{argmax} \int_K D(\kappa)t(\kappa;v)d\kappa \quad (3)$$

Μεγιστοποιώντας την παραπάνω συνάρτηση (3) σε αντίθεση με την περίπτωση μεγιστοποίησης της (1) αναμένουμε την κατανομή της γύρω από το βέλτιστο κ . Ακόμα η συνάρτηση αυτή μπορεί να γίνει συνεχής και διαφοροποιήσιμη σε σχέση με το v . Σύμφωνα

με τον αλγόριθμο GASS ο οποίος ισχύει για την εκθετική οικογένεια των πιθανοτήτων ισχύει το εξής :

$$t(\kappa; v) = \exp \left\{ v^A \Delta(\kappa) - \chi(\kappa) \right\}, \quad (4)$$

$$\frac{\int \exp(v^A \Delta(\kappa)) d\kappa}{\int d\kappa} \quad \text{και} \quad \Delta(\kappa) \quad \text{είναι η πηγή των στατιστικών των παραμέτρων του}$$

όπου $\chi(\kappa) = \ln \int d\kappa$

νευρωνικού δικτύου.

Παρουσιάζοντας λοιπόν συνοπτικά τον αλγόριθμο προσαρμοστικής στοχαστικής αναζήτησης , Gradient Based , [2] :

- Require:** Learning rate β_r , sample size E_r , initial policy parameters $v_0 = [\mu_0, \sigma_0^2]^A$, smoothing function $G()$, small constant $\lambda > 0$.
1. **For** $r = 0, 1, 2, \dots$ **do**
 2. Sample $\kappa_r^i \sim t(\kappa; v_r), i=1, 2, \dots, E_r$
 3. Compute returns $D(\kappa_r^i)$ for $i=1, 2, \dots, E_r$
 $y_r^i = G(D(\kappa_r^i))$
 4. Compute variance terms $H_r = \widehat{H}_r + \lambda Z$, eq (5),(6)
 5. Calculate normalizer $\zeta = \sum_{i=1}^{E_r} y_r^i$
 6. Update v_{r+1} :
 7. $v_{r+1} \leftarrow v_r + \beta_r \frac{1}{\zeta} H_r^{-1} \sum_{i=1}^{E_r} y_r^i \left(\begin{bmatrix} \tilde{\kappa}_r^i \\ (\tilde{\kappa}_r^i)^2 \end{bmatrix} - \begin{bmatrix} \mu \\ \sigma^2 + \mu^2 \end{bmatrix} \right)$
 8. **End for**

Για κάθε επανάληψη r , ο αλγόριθμος προσαρμοστικής στοχαστικής αναζήτησης, Gradient Based περιλαμβάνει την εκτίμηση των E_r δειγμάτων των παραμέτρων $\kappa_r^i \sim t(\kappa; \nu_r)$, $i = 1, 2, \dots, E_r$. Αυτές οι παράμετροι στη συνέχεια χρησιμοποιούνται για τη δειγματοληψία της συνάρτησης επιστροφής $D(\kappa_r^i)$. Οι αποδόσεις τροφοδοτούνται μέσω μιας συνάρτησης διαμόρφωσης: $G(\cdot): \mathbb{R}^{\hat{L}} \rightarrow \mathbb{R}^L$ και στη συνέχεια χρησιμοποιούνται για τον υπολογισμό μιας ενημέρωσης στις παραμέτρους μοντέλου ν_{r+1} . Η συνάρτηση διαμόρφωσης $G(\cdot)$ απαιτείται να μην μειώνεται και να οριοθετείται από πάνω και κάτω για οριοθετημένες εισόδους, με το κάτω όριο μακριά από το 0. Η αρχική ανάλυση του GASS παίρνει μια πιο γενική μορφή $G_r(\cdot)$ όπου το G μπορεί να αλλάξει σε κάθε επανάληψη. Για απλότητα υποθέτουμε ότι εδώ είναι αμετάβλητη ανά επανάληψη. Το GASS απαιτεί εκτίμηση της διακύμανσης των παραμέτρων δειγματοληψίας:

$$\hat{H}_r = \frac{1}{E_r - 1} \sum_{i=1}^{E_r} \Delta(\kappa_r^i) \Delta(\kappa_r^i)^{\Delta} - \frac{1}{E_r^2 - E_r} \left(\sum_{i=1}^{E_r} \Delta(\kappa_r^i) \right) \left(\sum_{i=1}^{E_r} \Delta(\kappa_r^i) \right)^{\Delta} \quad (5)$$

Στην εργασία αυτή προσεγγίζουμε το H_r με ανεξάρτητους υπολογισμούς της διακύμανσης σε όλες τις παραμέτρους. Θεωρούμε το $\tilde{\kappa}_r^i$ ως ένα κλιμακωτό στοιχείο του κ_r^i . Έχουμε έπειτα, για κάθε στοιχείο $\tilde{\kappa}_r^i$ έναν πίνακα 2×2 , [2] :

$$\hat{H}_r = \frac{1}{E_r - 1} \sum_{i=1}^{E_r} \begin{bmatrix} \tilde{\kappa}_r^i \\ (\tilde{\kappa}_r^i)^2 \end{bmatrix} \begin{bmatrix} \tilde{\kappa}_r^i & (\tilde{\kappa}_r^i)^2 \end{bmatrix} - \frac{1}{E_r^2 - E_r} \left(\sum_{i=1}^{E_r} \begin{bmatrix} \tilde{\kappa}_r^i \\ (\tilde{\kappa}_r^i)^2 \end{bmatrix} \right) \left(\sum_{i=1}^{E_r} \begin{bmatrix} \tilde{\kappa}_r^i & (\tilde{\kappa}_r^i)^2 \end{bmatrix} \right) \quad (6)$$

Όπως συμβαίνει με τους περισσότερους μη κυρτούς αλγόριθμους βελτιστοποίησης, δεν είναι εγγυημένη η επίτευξη του απόλυτου μέγιστου, αλλά με τη χρήση πιθανοτικών μοντέλων και προσεκτικής επιλογής της λειτουργίας διαμόρφωσης θα πρέπει να αποφευχθεί η πρόωμη σύγκλιση σε τοπικό ανώτατο όριο. Ακόμα, το Θεώρημα του αλγορίθμου GASS καθορίζει

επίσης τα ποσοστά σύγκλισης ως προς τον αριθμό των επαναλήψεων r , τον αριθμό των δειγμάτων E_r και τον βαθμό μάθησης β_r .

Υπόθεση 1

- i) The learning rate $\beta_r > 0, \beta_r \rightarrow 0$ as $r \rightarrow \infty, \wedge \sum_{r=0}^{\infty} \beta_r = 0$
- ii) The sample size $E_r = E_0 r^\pi$, where $\pi > 0$; also $\beta_r \wedge E_r$ jointly satisfy $\beta / \sqrt{E_r} = b(r^{-\vartheta})$.
- iii) $\Delta(\kappa)$ is bounded on K .
- iv) If v^i is a local maximum of (3), the Hessian of $\int_K D(\kappa) \tau(\kappa, v) d\kappa$ is continuous and symmetric negative definite in a neighborhood of v^i .

Θεώρημα 1

Let $\beta_r = \beta_0 / r^\beta$ for $0 < \beta < 1$. Let $E_r = E_0 r^{\delta - \beta}$, where $\delta > 2\beta$

Then $\{v_r\}$ converges with rate $b(1/\sqrt{r}^\delta)$

Στην συνέχεια, όπως προαναφέρθηκε, θα εξεταστεί η λύση του προβλήματος μέσω του αλγορίθμου ES [1]. Όπως θα δούμε, ο αλγόριθμος ES είναι μία προσέγγιση του αλγορίθμου GASS. Δηλαδή, αρχικά διαταράσσει τυχαία τις παραμέτρους που προέρχονται μέσω της μεθόδου δειγματοληψίας της Gaussian κατανομής. Στην συνέχεια, υπολογίζει τις αποδόσεις και ενημερώνει τις παραμέτρους. Μόλις υπολογιστούν οι επιστροφές, αποστέλλονται μέσω μιας συνάρτησης $G(\cdot)$.

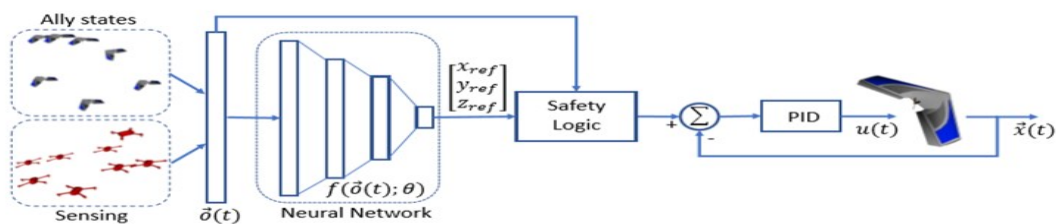
Ο αλγόριθμος ES έχει ως εξής :

Require: Learning rate β_r , noise standard deviation λ , initial policy parameters κ_0 , smoothing function $G()$.

1. **For** $r = 0, 1, \dots$ **Do**
2. *Sample* $\psi_1, \dots, \psi_{E_r} \sim E(0, Z^{r \times r})$
3. *Compute* $y_r^i = G(D(\kappa_r + \lambda \psi_i))$ for $i=1, \dots, E_r$
4. *Update* κ_{r+1} :
5. $\kappa_{r+1} = \kappa_r + \beta_r \frac{1}{E_r \lambda} \sum_{i=1}^{E_r} y_r^i \psi_i$
6. **End for**

Όπως προαναφέρθηκε ο αλγόριθμος ES είναι μία παραδοχή του αλγορίθμου GASS. Ο αλγόριθμος ES μπορεί να προβλέψει βήματα του αλγορίθμου GASS. Αντί του όρου ζ , ο αλγόριθμος ES χρησιμοποιεί τον αριθμό των δειγμάτων E_r . Η μικρή σταθερά στο GASS λ γίνεται ο όρος διακύμανσης στον αλγόριθμο ES. Βλέπουμε ότι το ES λαμβάνει τα ίδια ποσοστά ασυμπτωτικής σύγκλισης που προσφέρονται από την ανάλυση της GASS. Ενώ ο αλγόριθμος GASS είναι μία δευτερεύουσα μέθοδος και το ES είναι μέθοδος πρώτης τάξης, στην πράξη, το ES χρησιμοποιείται κατά προσέγγιση προκειμένου να επιταχυνθεί και να σταθεροποιηθεί ο ρυθμός μάθησης. Παραδείγματα αυτών των μεθόδων είναι τα ADAM, RMSProp, SGD με ορμή, κλπ., τα οποία χρησιμοποιούνται ευρέως στα νευρωνικά δίκτυα. Επομένως, μπορούμε να αντιμετωπίσουμε το ES ως μια προσέγγιση πρώτης τάξης της μεθόδου διακύμανσης που χρησιμοποιεί το GASS. Στα πειράματά μας χρησιμοποιούμε το ADAM [11] για να προσαρμόσουμε τον ρυθμό μάθησης για κάθε παράμετρο. Όπως

αναφέρθηκε και στο [1], όταν χρησιμοποιήσαμε ρυθμούς προσαρμοστικής μάθησης, διαπιστώσαμε μικρή βελτίωση σε σχέση με την προσαρμογή της διακύμανσης της κατανομής δειγματοληψίας. Υποθέτουμε ότι μια μέθοδος πρώτης τάξης με την κατάλληλη προγραμματισμένη εκμάθηση είναι επαρκής για την επίτευξη καλών επιδόσεων κατά τη βελτιστοποίηση των νευρωνικών δικτύων. Για άλλους τύπους παραμέτρων πολιτικής, ωστόσο, η πλήρης επεξεργασία που πραγματοποιεί ο αλγόριθμος του GASS μπορεί να είναι πιο χρήσιμη.



Εικ.2 : Στο παραπάνω διάγραμμα φαίνεται η διαδικασία εφαρμογής της πολιτικής κάθε πράκτορα.

ΚΕΦΑΛΑΙΟ 3

Ισχύουσες διατάξεις δομημένων πολιτικών για πολυπρακτορικά προβλήματα

Στο σημείο που βρισκόμαστε έχουμε πλέον εμβαθύνει σε μεγάλο βαθμό στη σύγκλιση της μεθόδου ES / GASS, και έχουμε πλέον δείξει πώς μπορεί να χρησιμοποιηθεί ο αλγόριθμος ES για τη βελτιστοποίηση μίας πολύπλοκης πολιτικής σε ένα περιβάλλον πολλών παραγόντων μεγάλης κλίμακας. Το επόμενο βήμα είναι να χρησιμοποιούμε το περιβάλλον προσομοίωσης πολλαπλών πρακτόρων SCRIMAGE [10], που μας επιτρέπει να αξιοποιήσουμε μία πικοιλία σεναρίων για προβλήματα πολλαπλών παραγόντων.

Θα πραγματοποιήσουμε την προσομοίωση μέσω της χρήσης αεροσκαφών 6DoF και drone τύπου quadcopter (που φέρουν 4 έλικες) [2] . Παράμετροι όπως αυτός του θορύβου και της δύναμης του ανέμου διαμορφώνονται μέσω του θορύβου Gauss. Άλλες παράμετροι που θα εντάξουμε είναι η παράμετρος μίας ενδεχόμενης πρόσκρουσης των αεροσκαφών είτε

μεταξύ τους , είτε στο έδαφος. Το αποτέλεσμα σ' αυτήν την περίπτωση είναι η καταστροφή του αεροσκάφους. Επιπρόσθετα , στην προσωμοίωσή μας , θα εντάξουμε και την παράμετρο καταστροφής ενός αεροσκάφους από ένα άλλο με τη χρήση οπλισμού. Το ενδεχόμενο προσβολής θα επηρεάζεται ασφαλώς από την απόσταση και το είδος του στόχου. Η τακτική η οποία θα ακολουθήσουμε επιτρέπει τον ανεξάρτητο χειρισμό κάθε αεροσκάφους με αποτέλεσμα να διαμορφώνονται διαφορετικές πιθανόν τακτικές από τις 2 αντίπαλες ομάδες.

-

Εκτός των άλλων, έχει καθοριστεί και ο τομέας της επικοινωνίας μεταξύ των πρακτόρων. Συγκεκριμένα οι φίλιοι μεταξύ τους πράκτορες μπορούν να ανταλλάσσουν πληροφορίες (μέχρι 50 μηνύματα και μόνο σε κοντινές αποστάσεις). Άλλες επιπρόσθετες δυνατότητες των ιπταμένων μέσων είναι ο εντοπισμός μέχρι 5 εχθρικών πρακτόρων. Μέσα στις δυνατότητες των ιπταμένων μέσων βρίσκεται και ο εντοπισμός πληροφοριών σχετικά με τις φίλιες και μη βάσεις.

Όπως προαναφέρθηκε τα δεδομένα (παράμετροι νευρωνικών δικτύων) του θεάτρου επιχειρήσεων είναι τα ίδια για όλους τους πράκτορες. Όμως κάθε πράκτορας θα έχει την δυνατότητα ξεχωριστής αξιολόγησης και αξιοποίησης των παραμέτρων των νευρωνικών δικτύων τους, άρα το αποτέλεσμα δεν θα είναι αναμενόμενο.

Βασικό κομμάτι του όλου εγχειρήματος , είναι να καταφέρουμε να σχεδιάσουμε την πολιτική των νευρωνικών δικτύων με τέτοιο τρόπο ώστε να είναι εύκολη η μετάδοση της από την προσωμοίωση στις πραγματικές επιχειρήσεις. Στην πραγματικότητα όμως, η ασφάλεια παίζει πολύ σημαντικό ρόλο. Έτσι θα χρησιμοποιηθούν ξεχωριστοί ελεγκτές PID για την πλοήγηση κάθε πράκτορα. Ο κάθε ελεγκτής θα παρέχει πληροφορίες σε κάθε πράκτορα σχετικά με το υψόμετρο του, το περιβάλλον και όλες εκείνες τις πληροφορίες που θα αποτρέψουν μία πιθανή συντριβή.

ΚΕΦΑΛΑΙΟ 4

Προσομοιώσεις

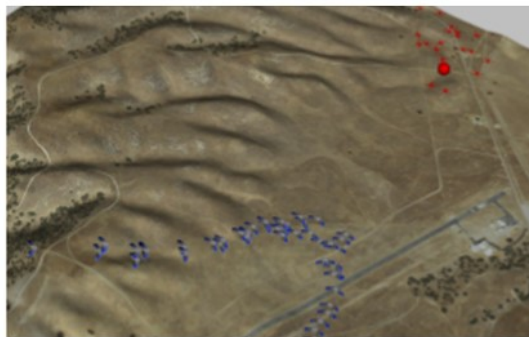
Τα δύο επεισόδια τα οποία θα εξεταστούν σ' αυτήν την εργασία είναι τα εξής [2] :

Στο πρώτο επεισόδιο μία ομάδα 50 μη επανδρωμένων αεροσκαφών σταθερής πτέρυγας έχει αποστολή να επιτεθεί μίας καλά οργανωμένης αμυντικά βάσης την οποία υπερασπίζονται 20 μη επανδρωμένα αεροσκάφη με ειδική διαμόρφωση με 4 έλικες. Στο δεύτερο επεισόδιο 2 ομάδες μη επανδρωμένων αεροσκαφών έχουν ως αποστολή να επιτεθούν η μία στην άλλη. Θα χρησιμοποιήσουμε και στις δύο περιπτώσεις την εξής εξίσωση :

$$D=10 \times (\text{kills}) + 50 \times (\text{collisions with enemy base}) - 1e-5 \times (\text{enemy base at end of episode})$$

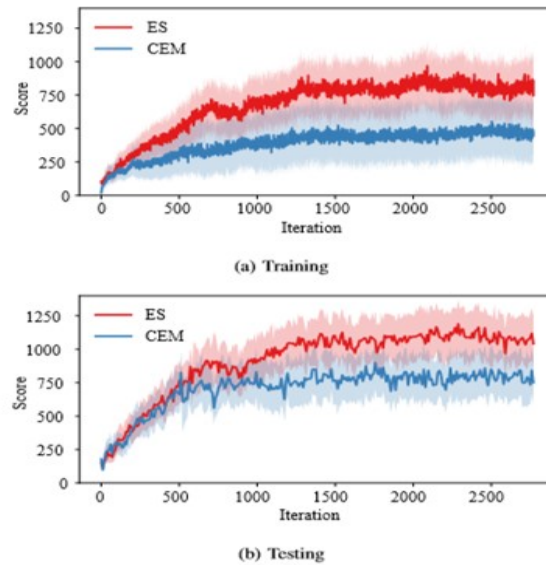
Στην παραπάνω εξίσωση παρατηρούμε ότι συμπεριλαμβάνονται όροι οι οποίοι σχετίζονται με την επίτευξη καταρρίψεων αεροσκαφών του εχθρού , διενέργειας αερομαχιών καθώς επίσης και όροι που αφορούν τυχόν επιθέσεις “αυτοκτονίας” κατά της αμυνόμενης βάσης.

1. Σενάριο επίθεσης σε αμυντική τοποθεσία



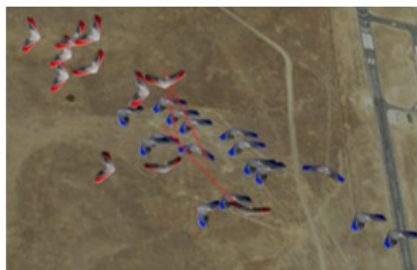
Εικ.3 : Στιγμιότυπο από την επίθεση UAV σταθερής πτέρυγας σε αμυντική βάση.

Στο συγκεκριμένο επεισόδιο που θα εξετάσουμε [2], θα ασχοληθούμε με την επίθεση κατά μίας ισχυρά οχυρωμένης αμυντικής τοποθεσίας. Συγκεκριμένα, 50 μη επανδρωμένα αεροσκάφη σταθερής πτέρυγας θα επιτεθούν μίας βάσης της οποίας θα αμύνονται 20 μη επανδρωμένα αεροσκάφη (φέρουν 4 έλικες το καθένα). Η αρχική τοποθέτηση των δύο ομάδων στο θέατρο επιχειρήσεων που πρόκειται να αναμετρηθούν μεταξύ τους, έγινε με τυχαίο τρόπο. Η ομάδα των 20 μη επανδρωμένων αεροσκαφών αμύνονται με την εξής μέθοδο : Αρχικά μοιράζουν ομοιόμορφα όλο τον χώρο στον οποίο πρόκειται να εξελιχθεί η σύγκρουση. Όταν εντοπίσουν τον πλησιέστερο στο κάθε πράκτορα εχθρό, υπολογίζουν το ύψος του και τον πλησιάζουν προκειμένου να του επιτεθούν . Η όλη διαδικασία επαναλαμβάνεται κάθε φορά, όσες φορές χρειαστεί ανάλογα με τις επιθέσεις. Χρησιμοποιήσαμε $E_r = 300$, $\lambda = 0,02$, ένα χρονικό βήμα 0,1 δευτερολέπτων και συνολικό μήκος επεισοδίου 200 δευτερολέπτων. Για την προσωμοίωσή μας χρησιμοποιήσαμε ένα ειδικό επεξεργαστή τύπου Xeon Phi. Μετά τις δοκιμές που πραγματοποιήσαμε διαπιστώθηκε ότι η πολιτική που χρησιμοποίησε η επιτιθέμενη ομάδα εδράστηκε στην κατά πρώτο λόγο αξιολόγηση της κάθε πληροφορίας που αφορά στον εχθρό και έπειτα στην προσέγγιση της αμυντικής τοποθεσίας. Στην δεύτερη φάση, δηλαδή αυτή της προσέγγισης, τα επιτιθέμενα UAV αναλάμβαναν δύο διαφορετικές αποστολές. Κάποια επιδίδονταν σε αερομαχίες με τα αμυνόμενα αεροσκάφη και κάποια άλλα προσκρούονταν ειθελημένα στην εχθρική τοποθεσία προκειμένου να την καστρέψουν Video 2. Στην σύγκριση του αλγορίθμου της ES με τον αλγόριθμο CEM (cross-entropy method) διαπιστώσαμε την ανωτερότητα του ES. Ο λόγος είναι ότι οι παράμετροι που χρησιμοποιεί ο αλγόριθμος CEM προέρχονται από περιορισμένο δείγμα πληροφοριών , άρα και η εκτίμησή του αλγορίθμου δεν είναι ασφαλής.



Εικ. 4: Τα αποτελέσματα που προκύπτουν ανά επανάληψη επίθεσης στην αμυντική βάση. Η κόκκινη καμπύλη συμβολίζει τον αλγόριθμο ES ενώ η μπλε τον αλγόριθμο στοχαστικής αναζήτησης CEM.

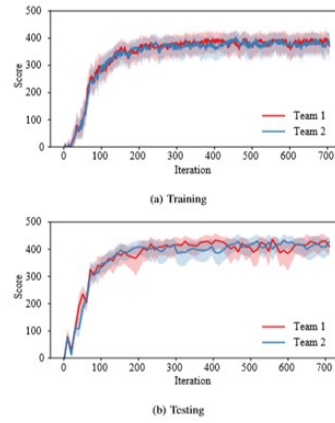
2. Σενάριο αναμέτρησης 2 αντίπαλων ομάδων UAV



Εικ. 5

Στην δεύτερη προσωμοίωση που πραγματοποιήσαμε , εξελίσσεται η σύγκρουση μεταξύ 2 διαφορετικών ομάδων μη επανδρωμένων αεροσκαφών. Στην προσπάθειά τους για εξουδετέρωση του αντιπάλου κάθε ομάδα αναπτύσει δική της τακτική και μεθοδολογία. Στις πάνω από 300 επαναλήψεις προσωμοιώσεων που πραγματοποιήθηκαν η κάθε μία από τις 2 αντίπαλες ομάδες εμφανίζει κάθε φορά ένα διαφορετικό ελάττωμα , οπότε και θα πρέπει η

εταιρη ομάδα να αναπτύξει ένα διαφορετικό τρόπο αντιμετώπισης. Οι βαθμολογίες των 2 ομάδων προέρχονται από τις απώλειες που κατάφεραν στον αντίπαλο μέσω της εκμετάλλευσης των ελαττωμάτων που προαναφέραμε, σε κάθε επανάληψη. Στις δεκάδες επαναλήψεις που πραγματοποιήθηκαν οι 2 ομάδες εκπαιδεύτηκαν και έμαθαν να αντιμετωπίζουν ένα πλήθος διαφορετικών επιθετικών ή αμυντικών τακτικών. Το πιο σημαντικό συμπέρασμα ήταν ότι οι αντιδράσεις και οι επιλογές των 2 ομάδων πλησιάζουν την λεγόμενη ισορροπία Nash. Και οι δύο πλευρές δηλαδή, προσπαθούν να πετύχουν την περισσότερες δυνατές απώλειες και καταρρίψεις αντίπαλων αεροσκαφών. Στο τέλος όμως το αποτέλεσμα είναι ταυτόχρονα καταστροφικό και για τις 2 πλευρές, με τις καταρρίψεις να ισοψιφίζονται σχεδόν. Τέλος, αξίζει να ανφερθεί ότι αν οι δύο πλευρές μπορούσαν να καταχωρίσουν τις διάφορες τακτικές που χρησιμοποιούσαν σε κάθε επανάληψη ώστε να μπορέσουν να τις χρησιμοποιήσουν σε μελλοντικές επαναλήψεις του πειράματός μας, τότε ίσως το αποτέλεσμα να ήταν διαφορετικό και μία από τις δύο πλευρές να είχε τελικά επικρατήσει, κάτι το οποίο έχει επισημανθεί και από την evolutionary computation community [12].



Εικ.6 : Τα αποτελέσματα των καταρρίψεων των ομάδων 1, 2 με κόκκινη και μπλε καμπύλη αντίστοιχα.

Επίλογος

Τα δύο παραπάνω πειράματα μας βοήθησαν να αποδείξουμε ότι ο αλγόριθμος των ES είναι απόλυτα αποτελεσματικός σε προβλήματα με πολλαπλούς παράγοντες. Τα δύο διαφορετικά σενάρια που αναπτύξαμε συμπεριελάμβαναν εκατοντάδες παραμέτρους. Το πρώτο σενάριο ήταν σε επίπεδο συνεργασίας και το δεύτερο σε ανταγωνιστικό επίπεδο μεταξύ δύο αντίπαλων ομάδων μη επανδρωμένων αεροσκαφών. Και στις δύο περιπτώσεις ο αλγόριθμος των ES ανταποκρίθηκε σε πολύ υψηλότερο επίπεδο σε σχέση με άλλες μεθόδους στοχαστικής αναζήτησης. Μελλοντικά, η έρευνα ίσως θα πρέπει να κινηθεί στην συνεργασία των ES με άλλες μεθόδους στοχαστικής αναζήτησης ώστε να πετύχουμε την βελτίωση σύνθετων προβλημάτων, όπως η ελάφρυνση των βαρών του νευρωνικού δικτύου. Τέλος, η έρευνα ίσως μπορεί να επικεντρωθεί σε εξίσου σημαντικές και ενδιαφέρουσες εργασίες όπως η βελτίωση του επιπέδου των παραμέτρων για κάθε πράκτορα ξεχωριστά, κάτι που θα βελτίωνε τις τακτικές της ομάδας σε αντίστοιχα προβλήματα, όπως αυτά με τα οποία ασχοληθήκαμε [2].

Βιβλιογραφία

- [1] T. Salimans, J. Ho, X. Chen, S. Sidor, and I. Sutskever, “Evolution Strategies as a Scalable Alternative to Reinforcement Learning,” ArXiv e-prints, Mar. 2017.
- [2] David D. Fan, Evangelos Theodorou and John Reeder , “Model-Based Stochastic Search for Large Scale Optimization of Multi-Agent UAV Swarms ”.
- [3] J. Hu, “Model-based stochastic search methods,” in Handbook of Simulation Optimization. Springer, 2015, pp. 319– 340.
- [4] J. B. Rawlings and B. T. Stewart, “Coordinating multiple optimizationbased controllers: New opportunities and challenges,” Journal of Process Control, vol. 18, no. 9, pp. 839–845, 2008.
- [5] W. Al-Gherwi, H. Budman, and A. Elkamel, “Robust distributed model predictive control: A review and recent developments,” The Canada Journal of Chemical Engineering, vol. 89, no. 5, pp. 1176–1190, 2011. [Online]. Available: <http://doi.wiley.com/10.1002/cjce.20555>
- [6] G. B. Lamont, J. N. Slear, and K. Melendez, “UAV swarm mission planning and routing using multi-objective evolutionary algorithms,” in IEEE Symposium Computational Intelligence in Multicriteria Decision Making, no. Mcdm, 2007, Conference Proceedings, pp. 10–20.
- [7] A. R. Yu, B. B. Thompson, and R. J. Marks, “Competitive evolution of
- [8] E. Zhou and J. Hu, “Gradient-based adaptive stochastic search for nondifferentiable optimization,” IEEE Transactions on Automatic Control, vol. 59, no. 7, pp. 1818–1832, 2014.
- [9] U. Gaerther, “UAV swarm tactics: an agent-based simulation and Markov process analysis,” 2015. [Online]. Available: <https://calhoun.nps.edu/handle/10945/34665>
- [10] K. J. DeMarco. (2018) SCRIMMAGE multi-agent robotics simulator.[Online]. Available: <http://www.scrimmagesim.org/>
- [11] [D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980, 2014.
- [12] K. O. Stanley and R. Miikkulainen, “Competitive coevolution through evolutionary complexification,” Journal of Artificial Intelligence Research, vol. 21, pp. 63–100, 2004.

Ιστότοποι

Video 1 : https://github.com/ddfan/swarm_evolve

Video 2 : <https://goo.gl/dWvQi7>