5. Υλοποίηση 41

5.1. Εισαγωγή 41

5.2. Συστήματα Παράλληλης Επεξεργασίας 41

5.3. Υλοποίηση των ΣΠΔΕΣ και του Μοντέλου Νησίδων 44
5.3.1. Διάταξη Αναπαραγωγής Επιπέδου 0 του Μοντέλου Νησίδων 47
5.3.2. Διάταξη Αναπαραγωγής Επιπέδου 0 των ΣΠΔΕΣ 48

6. Προσομοιώσεις 51

6.1. Περιγραφή των Προβλημάτων 51
6.1.1. Εκπαίδευση Τεχνητών Νευρωνικών Δικτύων 52
6.1.2. Βελτιστοποίηση Συνάρτησης Rastrigin 53
6.1.3. Πρόβλημα του Περιοδεύοντος Πωλητή 54

6.2. Αποτελέσματα Προσομοιώσεων 56
6.2.1. Εκπαίδευση Τεχνητών Νευρωνικών Δικτύων 57
6.2.2. Βελτιστοποίηση Συνάρτησης Rastrigin 61

Κεφάλαιο 5

Υλοποίηση

5.1. Εισαγωγή

Τα συστήματα πολλών επεξεργαστών παρέχουν την δυνατότητα αύξησης της διαθέσιμης υπολογιστικής δύναμης με λογικό κόστος. Την τελευταία δεκαετία το ενδιαφέρον για τέτοια συστήματα έχει αυξηθεί σημαντικά, με αποτέλεσμα την ανάπτυξη μιας πληθώρας εφαρμογών. Η ιδέα της ανάπτυξης εφαρμογών εξελικτικών αλγορίθμων σε συστήματα παράλληλης επεξεργασίας δεν είναι καινούργια. Την τελευταία πενταετία υπάρχει ένας διαρκώς αυξανόμενος αριθμός εφαρμογών οι οποίες προσπαθούν να εκμεταλλευτούν τις αυξημένες δυνατότητες των συστημάτων πολλών επεξεργαστών

Σε αυτό το κεφάλαιο, μετά την γενική εισαγωγή στα συστήματα παράλληλης επεξεργασίας και ειδικότερα στους transputer και το σύστημα που χρησιμοποιήθηκε, παρουσιάζεται η μέθοδος υλοποίησης των ΣΠΔΕΣ και του μοντέλου των χαΠΓΑ όπου κάθε υπο-πληθυσμός εξελίσσεται χρησιμοποιώντας την ίδια εξελικτική συμπεριφορά. Το δεύτερο μοντέλο θα αναφέρεται ως ***μοντέλο νησίδων***.

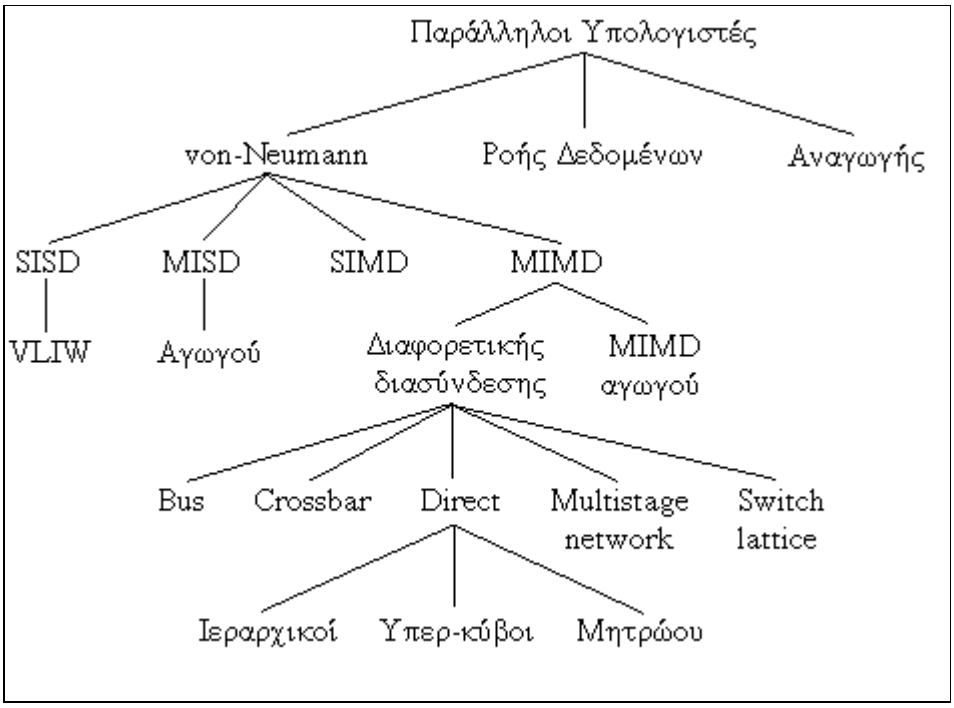
5.2. Συστήματα Παράλληλης Επεξεργασίας

Η χρήση των συστημάτων παράλληλης επεξεργασίας αυξάνεται καθώς τίθενται υπολογιστικά προβλήματα τα οποία είναι δύσκολο να αντιμετωπιστούν σε συστήματα ενός επεξεργαστή και

καθώς οι σειριακοί υπολογιστές πλησιάζουν κάποια φυσικά όρια στην ανάπτυξη της υπολογιστικής τους δύναμης.

Το κύριο εμπόδιο στην γρήγορη εξάπλωση της παράλληλης επεξεργασίας είναι οι τεράστιες επενδύσεις που έχουν γίνει σε λογισμικό για σειριακές μηχανές, η δυσκολία προγραμματισμού των περισσότερων παράλληλων υπολογιστικών συστημάτων και η έλλειψη ενός κοινώς αποδεκτού μοντέλου παραλληλισμού.

Τα κύρια προβλήματα της παράλληλης επεξεργασίας είναι:
- ο τρόπος επικοινωνίας και συγχρονισμού των μονάδων επεξεργασίας,
- ο έλεγχος των ταυτόχρονων αναφορών στην κοινή μνήμη, ή στις μονάδες Ε/Ε
- ο τρόπος ανάθεσης των προγραμμάτων στις μονάδες επεξεργασίας
- ο τρόπος περιγραφής των παράλληλων προγραμμάτων



Σχήμα 5.1. Κύριοι τύποι παράλληλων αρχιτεκτονικών

Στο σχήμα 5.1 παρουσιάζονται οι βασικές αρχιτεκτονικές παράλληλης επεξεργασίας. Υπάρχουν τρεις βασικές προσεγγίσεις: η βασισμένη στην Von-Neumann αρχιτεκτονική, η ροής δεδομένων (data flow), και η προσέγγιση αναγωγής (reduction). Η προσέγγιση Von-Neumann βασίζεται στην ιδέα της εκτέλεσης των εντολών του προγράμματος με βάση μια ακολουθία σε δύο ή περισσότερες μηχανές ενός επεξεργαστή, σε διάφορες τοπολογίες. Στην προσέγγιση ροής δεδομένων οι εντολές εκτελούνται μόλις είναι διαθέσιμοι οι τελεστές της εντολής. Η προσέγγιση αναγωγής περιλαμβάνει την εκτέλεση των εντολών όταν τα αποτελέσματα τους χρειάζονται για άλλους υπολογισμούς.

Η πλειοψηφία των παράλληλων υπολογιστών που αναπτύσσονται σήμερα βασίζεται σε επεξεργαστές που ακολουθούν την αρχή του Von-Neumann. Αυτά τα παράλληλα υπολογιστικά

συστήματα κατηγοριοποιούνται ανάλογα με τον τρόπο παροχής των εντολών και των δεδομένων στις μονάδες επεξεργασίας ως ακολούθως (M. Flynn, 1972):

- *Single Instruction Single Data* (SISD, Μοναδικής Εντολής Μοναδικών Δεδομένων): Μία μονάδα επεξεργασίας εκτελεί ακολουθιακά τις εντολές, πάνω σε μία ακολουθία δεδομένων.
- *Single Instruction Multiple Data* (SIMD, Μοναδικής Εντολής Πολλαπλών Δεδομένων): Οι μονάδες επεξεργασίας εκτελούν μία κοινή ακολουθία εντολών, πάνω σε διαφορετικά δεδομένα.
- *Multiple Instruction Single Data* (MISD, Πολλαπλών Εντολών Μοναδικών Δεδομένων): Οι μονάδες επεξεργασίας διατάσσονται σε μια αλυσιδωτή μορφή. Κάθε επεξεργαστής εκτελεί ανεξάρτητες εντολές, ενώ τα δεδομένα μιας μονάδας επεξεργασίας είναι τα αποτελέσματα της προηγούμενης. Συχνά αναφέρονται και ως μηχανές αγωγού.
- *Multiple Instruction Multiple Data* (MIMD, Πολλαπλών Εντολών Πολλαπλών Δεδομένων): Υποστηρίζονται πολλαπλές ροές εντολών με τεχνική διασωλήνωσης (pipeline) ή με ξεχωριστούς επεξεργαστές. Στην δεύτερη περίπτωση ανήκει και το σύστημα που χρησιμοποιήθηκε στην παρούσα διατριβή. Οι μονάδες επεξεργασίας έχουν την δυνατότητα να εκτελούν ανεξάρτητα, διαφορετικές εντολές με διαφορετικά δεδομένα μία κοινή ακολουθία εντολών, πάνω σε διαφορετικά δεδομένα.

Τα συστήματα πολλών επεξεργαστών προσφέρουν πολλά πλεονεκτήματα:

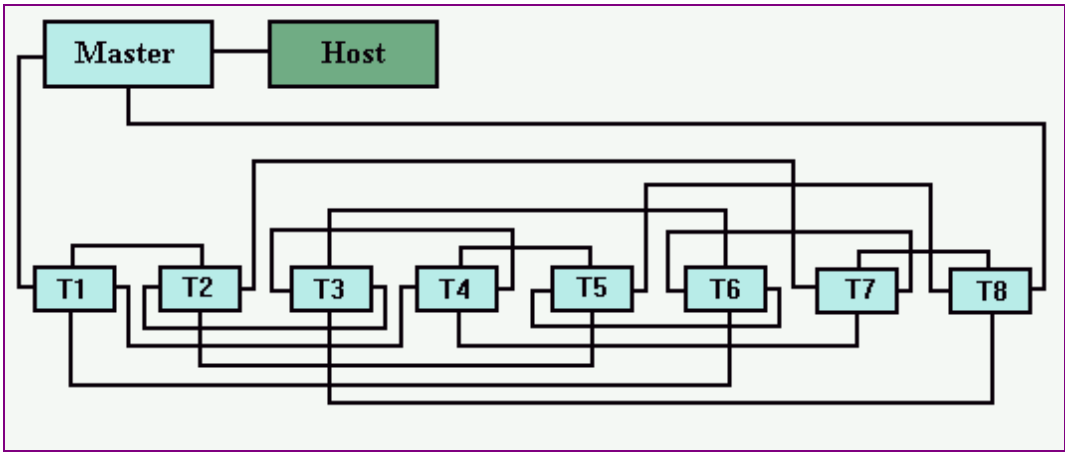
Τα δεδομένα και οι εντολές των επεξεργαστών λαμβάνονται από την κεντρική μνήμη της μηχανής, η οποία συνήθως αποτελείται από αρκετές ανεξάρτητες μονάδες μνήμης. Ανάλογα με τον τρόπο διασύνδεσης των επεξεργαστών με τις μνήμες, τα συστήματα πολλών επεξεργαστών μπορούν να χωριστούν σε δύο κατηγορίες:

- Στενά συνδεδεμένο σύστημα (tightly coupled system), όπου οι μονάδες μνήμης είναι απ'ευθείας προσπελάσιμες από οποιοδήποτε επεξεργαστή (διαμοιραζόμενη μνήμη, Shared memory)
- Χαλαρά συνδεδεμένο σύστημα (loosely coupled system), όπου κάθε μονάδα μνήμης είναι προσπελάσιμη από ένα μόνο επεξεργαστή. Οι επεξεργαστές επικοινωνούν στέλνοντας μηνύματα (message passing). Σ'αυτή την κατηγορία ανήκει και το σύστημα που χρησιμοποιήθηκε στην παρούσα διατριβή.

Σε όλες τις υλοποιήσεις της παρούσας διατριβής χρησιμοποιήθηκε η κάρτα Fast9 της εταιρίας Quintek. Η κάρτα αυτή αποτελείται από ένα δίκτυο εννέα ανεξάρτητων επεξεργαστών (transputers) συνδεδεμένων κατάλληλα μεταξύ τους. Μερικές από αυτές τις συνδέσεις είναι σταθερές, ενώ οι υπόλοιπες υλοποιούνται, είτε με ένα προγραμματιζόμενο ραβδεπαφικό διακόπτη (cross-bar switch), είτε απ' ευθείας με διακόπτες (dip-switches).

Οι transputers της κάρτας Fast9 είναι οι IMS T800 της INMOS, με μήκος λέξης 32-bit και τεχνολογία CMOS. Κάθε transputer έχει τοπική μνήμη (on-chip) 4 Kbytes και μονάδα κινητής υποδιαστολής για αριθμούς 32 bit. Κάθε transputer δουλεύει ταυτόχρονα και ανεξάρτητα από τους άλλους. Έχουν την δυνατότητα να επικοινωνούν με τέσσερις άλλους transputers, μέσω σειριακών συνδέσμων (links), προσφέροντας έτσι υποστήριξη για συστήματα πολλαπλών επεξεργαστών. Η αρχιτεκτονική τους υποστηρίζει το μοίρασμα του χρόνου του επεξεργαστή (time shearing), επιτρέποντας έτσι την ταυτόχρονη (concurrent) εκτέλεση διαδικασιών.

Η κάρτα Fast9 της Quintek (Παράρτημα Δ) μπορεί να προσαρμοστεί σε ένα απλό IBM PC, AT ή συμβατό, καταλαμβάνοντας μια υποδοχή (slot) των 16 bits του διαδρόμου συστήματος. Διαθέτει ένα προγραμματιζόμενο ραβδεπαφικό διακόπτη (crossbar switch) των συνδέσμων, καθώς και κύκλωμα προσαρμογής (host interface) ανάμεσα στην αρχιτεκτονική της INMOS και του διαδρόμου συστήματος του PC. Οι σύνδεσμοι 0 και 2 κάθε transputer είναι μόνιμα συνδεμένοι σχηματίζοντας τοπολογία σωλήνωσης (pipe). Οι υπόλοιπες συνδέσεις υλοποιήθηκαν προγραμματιστικά μέσω του ραβδεπαφικού διακόπτη.



Σχήμα 5.2. Τοπολογία σύνδεσης των transputer

Το σχήμα 5.2 δείχνει την τοπολογία σύνδεσης των transputer που επιλέχθηκε και υλοποιήθηκε. Ο επεξεργαστής του PC αναφέρεται σαν *Host* επεξεργαστής. Ο transputer που συνδέεται απ' ευθείας με τον Host ονομάζεται *Master*, ενώ οι άλλοι οκτώ ονομάζονται *Slaves (T1-T8)*. Το αρχείο διαμόρφωσης (configuration file) δίνεται στο Παράρτημα ΣΤ, όπου υπάρχουν και τα αρχεία διαμόρφωσης των διαφόρων εφαρμογών (σύνδεση των διεργασιών και ανάθεσή τους στους διάφορους transputers).

Η υλοποίηση των αλγόριθμων προγραμματιστικά έγινε με τη γλώσσα υψηλού επιπέδου Parallel C της 3L. Στο παράρτημα Ε δίνεται μια περιληπτική περιγραφή της 3L Parallel C και στο παράρτημα Ζ δίνεται μια περιγραφή των προγραμμάτων και του τρόπου χρήσης τους.

5.3. Υλοποίηση των ΣΠΔΕΣ και του Μοντέλου Νησίδων

Και τα δύο μοντέλα ορίζονται πλήρως αν ορισθούν τα στοιχεία της τετράδας του επίπεδου 1 των ΣΠ (*ΣΠ*^{1}) σύμφωνα με την 3.1. Ακολούθως δίνεται μία περιγραφή των στοιχείων της τετράδας σύμφωνα με τον τρόπο υλοποίησης των δύο μοντέλων που ακολουθήθηκε.

Η διάταξη ***Αναπαραγωγής,*** σύμφωνα με την 3.3, έχει υλοποιηθεί ως ακολούθως και για τα δύο μοντέλα:

- ***εποχή***: χρησιμοποιείται σύγχρονη μετανάστευση με σταθερές τιμές *εποχής* κατά την διάρκεια κάθε πειράματος. Οι τιμές της *εποχής* που χρησιμοποιήθηκαν στα πειράματα, κυμαίνονται από 1 έως 500 γενιές. Πριν την πρώτη μετανάστευση γόνων μεταξύ των υπο-πληθυσμών, αφήνουμε τους υπο-πληθυσμούς να εξελιχθούν για κάποιο αριθμό γενεών. Έτσι κάθε υπο-πληθυσμός έχει την δυνατότητα να ερευνήσει διαφορετικό μέρος του χώρου λύσεων, πριν εμπλουτιστεί με πληροφορίες από τους υπόλοιπους, και τα άτομα μετανάστες έχουν την δυνατότητα να αναπτύξουν σωστά τμήματα κατασκευής.
- ***επιλογή***: το καλύτερο άτομο κάθε υπο-πληθυσμού επιλέγεται προς μετανάστευση
- ***τελεστές***: εκτός της μετανάστευσης, δεν χρησιμοποιούνται τελεστές σε αυτό το επίπεδο
- ***παράμετροι***: ένα μόνο άτομο, από κάθε υπο-πληθυσμό, μεταναστεύει
- ***μηχανισμοί***: κάθε υπο-πληθυσμός στέλνει το καλύτερο άτομό του σε όλους τους υπόλοιπους και λαμβάνει το καλύτερο άτομο όλων των υπολοίπων. Κάθε άτομο που μεταναστεύει αντικαθιστά ένα τυχαίο άτομο στον υπο-πληθυσμό που πηγαίνει. Δεν επιτρέπεται η αντικατάσταση του καλύτερου ατόμου σε κάθε υπο-πληθυσμό.

Τα στοιχεία της διάταξης ***Δομή***, σύμφωνα με την 3.4, είναι ως ακολούθως:

- ***ΜεΠ***: οκτώ υπο-πληθυσμοί χρησιμοποιήθηκαν
- ***ΜεΓ***: οκτώ υπο-πληθυσμοί είναι και το μέγεθος γειτονιάς
- ***ΔοΓ***: όταν το μέγεθος γειτονιάς είναι ίσο με το μέγεθος πληθυσμού, όπως εδώ, δεν είναι δυνατό να οριστούν γειτονιές
- ***λ***: ο αριθμός των επεξεργαστών που είναι διαθέσιμοι είναι οκτώ (worker trasnputers).
- ***τοπολογία***: η τοπολογία είναι στατική. Όλοι οι επεξεργαστές (transputers) είναι άμεσα ή έμμεσα συνδεδεμένοι μεταξύ τους. Κάθε transputer είναι συνδεμένος με τέσσερις άλλους και έχει την δυνατότητα να επικοινωνεί με τους υπόλοιπους, μέσω αυτών που είναι άμεσα συνδεμένος. (σχ. 5.2).
- ***ΑνΠ***: αν και οι αλγόριθμοι υλοποιήθηκαν στην κάρτα Quintek Fast9, επειδή είναι ανεξάρτητοι από την συγκεκριμένη αρχιτεκτονική, θα μπορούσαν να υλοποιηθούν σε οποιοδήποτε παράλληλο (ακόμη και σειριακό) Η/Υ ή δίκτυο Η/Υ. Κάθε υπο-πληθυσμός ανατίθεται σε ένα διαφορετικό επεξεργαστή. Επειδή όλοι οι επεξεργαστές επικοινωνούν μεταξύ τους, κάθε υπο-πληθυσμός είναι δυνατό να ανατεθεί σε οποιοδήποτε επεξεργαστή.

Το κριτήριο τέλους ***τ***, ικανοποιείται όταν έχει συμπληρωθεί ένας αριθμός γενιών, ή έχει βρεθεί το βέλτιστο, ή η καλύτερη λύση δεν βελτιώνεται για κάποιο αριθμό γενιών.

Επειδή οι Συνεργαζόμενοι Πληθυσμοί ορίζονται αναδρομικά, πρέπει να οριστεί και το επίπεδο 0 των δύο μοντέλων (νησίδων και ΣΠΔΕΣ) για να είναι δυνατή κάποια υλοποίηση. Η υλοποίηση των δύο μοντέλων θα ήταν ίδια αν το επίπεδο 0 ($\Sigma\Pi^0$) των μοντέλων οριζόταν με τον ίδιο τρόπο. Έτσι η διαφοροποίηση των ΣΠΔΕΣ και του μοντέλου νησίδων επικεντρώνεται στην διαφοροποίηση της υλοποίησης του επιπέδου 0 για τα δύο μοντέλα

Οι υλοποιήσεις αυτές έχουν σαν στόχο την εδραίωση της άποψης, ότι η απόδοση βελτιώνεται όταν χρησιμοποιούνται υπο-πληθυσμοί με διαφορετικές εξελικτικές συμπεριφορές.

Στην περίπτωση του μοντέλου νησίδων, όπου κάθε υπο-πληθυσμός έχει την ίδια εξελικτική συμπεριφορά, έγινε προσπάθεια να επιλεχθεί η καλύτερη δυνατή διαμόρφωση των $\Sigma\Pi^0$, και συνεπώς η καλύτερη δυνατή εξελικτική συμπεριφορά. Έτσι η σύγκριση με τους ΣΠΔΕΣ γίνεται

γενική, αντικειμενική, πληρέστερη και δεν είναι επηρεασμένη από τις συγκεκριμένες υλοποιήσεις. Στην περίπτωση των ΣΠΔΕΣ, κάθε υπο-πληθυσμός έχει διαφορετική εξελικτική συμπεριφορά, συνεπώς ορίζεται διαφορετικό μοντέλο $\Sigma\Pi^0$ για κάθε υπο-πληθυσμό.

Ο αλγόριθμος που εκτελείται από κάθε υπο-πληθυσμό είναι ως εξής:

```
{
    Initialization
    Repeat {
        Repeat {
            Elitism (Best survives always)
            Reproduction of m-1 offsprings
        } until an_epoch_is_completed
        Send best individual to all other populations
        Receive the best individual from each one of the other populations
    } until termination_criterion
}
```

Για να ορισθεί πλήρως το αρχικό επίπεδο 0 ($\Sigma\Pi^0$) πρέπει να ορισθούν τα στοιχεία της διάταξης της σχέσης 3.2.

Π^0, ο αρχικός πληθυσμός κάθε νησίδας δημιουργείται τυχαία, μετά τον καθορισμό της κωδικοποίησης του εκάστοτε προβλήματος. Σε όλα τα προβλήματα χρησιμοποιήθηκε δυαδική κωδικοποίηση σταθερού μήκους και για τα δύο μοντέλα (νησίδων και ΣΠΔΕΣ). Η δυαδική κωδικοποίηση είναι η πιο κοινή κωδικοποίηση και συνηθίζεται να χρησιμοποιείται από την πλειοψηφία των υλοποιήσεων στον χώρο των ΓΑ ([Gold89a], [Mitc96]). Επειδή οι ΣΠΔΕΣ εστιάζουν στην επίδραση των διαφορετικών εξελικτικών συμπεριφορών, επιλέχθηκε η χρήση της ίδιας κωδικοποίησης σε όλους τους υπο-πληθυσμούς.

Τα στοιχεία της διάταξης *Δομής* είναι ως ακολούθως και είναι τα ίδια και για το μοντέλο νησίδων και για τους ΣΠΔΕΣ.

Το μέγεθος πληθυσμού (*ΜεΠ*) ισούται με τον αριθμό των ατόμων ανά υπο-πληθυσμό. Οι τιμές που χρησιμοποιήθηκαν στα πειράματα κυμαίνονται μεταξύ 50 και 200 ατόμων.

Καθώς οι υπο-πληθυσμοί είναι παν-μικτικοί, το μέγεθος γειτονιάς (*ΜεΓ*) είναι ίσο με το μέγεθος πληθυσμού (*ΜεΠ*) και δεν είναι δυνατό να ορισθούν γειτονιές. Επειδή κάθε υπο-πληθυσμός ανατίθεται σε ένα επεξεργαστή (λ, και *ΑνΠ*), ούτε η τοπολογία έχει φυσική έννοια και δεν ορίζεται.

Το επίκεντρο της έρευνας των ΣΠΔΕΣ είναι η διάταξη *Αναπαραγωγής*, η οποία είναι υπεύθυνη για την εμφάνιση διαφορετικών εξελικτικών συμπεριφορών.

5.3.1. Διάταξη Αναπαραγωγής Επιπέδου 0 του Μοντέλου Νησίδων

Η διάταξη αυτή κάνει δυνατή την εμφάνιση μιας συγκεκριμένης εξελικτικής συμπεριφοράς. Για να υπάρξει αντικειμενική και πλήρης σύγκριση μεταξύ του μοντέλου νησίδων και των ΣΠΔΕΣ, είναι αναγκαία η επιλογή της βέλτιστης διαμόρφωσης των στοιχείων της διάταξης *Αναπαραγωγής*, για κάθε υπο-πληθυσμό του μοντέλου νησίδων.

Εξ ορισμού, η τιμή της *εποχής* είναι 1, αφού με κάθε γενιά παράγεται νέος πληθυσμός.

Το στοιχείο της επιλογής είναι προσανατολισμένο στην επιλογή των καλύτερων ατόμων, με την ελπίδα ότι οι απόγονοί τους θα έχουν ακόμη καλύτερη ποιότητα. Η επιλογή πρέπει να ισορροπεί μεταξύ ανασυνδυασμού και μετάλλαξης, εκμετάλλευσης και διερεύνησης. Δεν είναι δυνατό να δοθούν αυστηρά καθορισμένες κατευθυντήριες γραμμές, για την επιλογή της κατάλληλης μεθόδου για κάθε πρόβλημα. Πρέπει να υπάρχει κατάλληλη ισορροπία μεταξύ της έμφασης που δίνεται στην επιλογή των καλύτερων (selective pressure) και στην γενετική απόκλιση των ατόμων (genetic diversity).

Η μέθοδος που επιλέχθηκε να χρησιμοποιηθεί, είναι η αναλογική επιλογή με τον τροχό της ρουλέτας (roulette wheel, fitness proportionate parent selection), όπου ένας γονέας επιλέγεται με πιθανότητα κατευθείαν ανάλογη της ποιότητάς του, σε συνδυασμό με την μέθοδο γραμμικής δυναμικής κλιμάκωσης των ποιοτήτων. Η αναλογική επιλογή χρησιμοποιείται για να επιτευχθεί η όσο το δυνατό μεγαλύτερη απόκλιση του πληθυσμού κατά την διάρκεια της αναζήτησης. Η κλιμάκωση των ποιοτήτων χρησιμοποιείται επειδή από μόνη της η αναλογική επιλογή δεν ήταν αρκετά αποτελεσματική στην διατήρηση μιας σταθερής πίεσης στην κατεύθυνση της σύγκλισης του αλγόριθμου ([Bäck94b][BäHo91] [dlMT93][GoDe91]).

Υπάρχουν αρκετές δημοσιεύσεις στη βιβλιογραφία των ΓΑ, οι οποίες προσδιορίζουν ποσοτικώς κάποιες πλευρές διαφόρων ανασυνδυαστικών τελεστών, αλλά δεν προσφέρουν καθορισμένη και σαφή καθοδήγηση στην επιλογή του καλύτερου ανασυνδυαστικού τελεστή. Επίσης, υπάρχουν αρκετές δημοσιεύσεις οι οποίες συγκρίνουν πειραματικά την χρησιμότητα και απόδοση διαφόρων τύπων ανασυνδυασμού. Η επιλογή του καταλληλότερου ανασυνδυαστικού τελεστή στηρίχτηκε σε αυτές τις δημοσιεύσεις καθώς επίσης και στα αποτελέσματα κάποιων αρχικών πειραμάτων στα δύο πρώτα προβλήματα (Ενότητες 6.1 και 6.2). Τα αποτελέσματα αυτά οδήγησαν στην επιλογή του προσαρμοστικού ανασυνδυασμού πολλών σημείων (adaptive mluti-point crossover) [SpDJ91], [DJSp92]. Ο αριθμός των σημείων ανασυνδυασμού είναι ίσος με τον αριθμό των γονιδίων στο χρωμόσωμα.

Η μετάλλαξη θεωρείται από πολλούς ερευνητές, κυρίως στο χώρο των ΓΑ, μικρότερης σπουδαιότητας από τον ανασυνδυασμό, λειτουργώντας κυρίως ως τελεστής παρασκηνίου. Από άλλους θεωρείται ότι έχει μεγάλη σημασία και σε μερικές εξελικτικές μεθόδους, η μετάλλαξη είναι η μοναδική πηγή παραλλαγής και εξέλιξης του πληθυσμού. Η ισορροπία μεταξύ ανασυνδυασμού, μετάλλαξης και επιλογής, θεωρείται σπουδαιότερη από την επιλογή μεταξύ ανασυνδυασμού και μετάλλαξης. Αυτές οι παράμετροι δεν είναι δυνατό να βελτιστοποιηθούν ταυτόχρονα. Υπάρχει ένας σημαντικός αριθμός δημοσιεύσεων σε θέματα καθορισμού, ρύθμισης και προσαρμογής των παραμέτρων ([Bäck93][Bram91] [Davi89][Foga89][Mühl92][SCED89][Spea93][StWM90]).

Η προσαρμογή των παραμέτρων του ανασυνδυασμού και της μετάλλαξης και η κλιμάκωση της ποιότητας υλοποιούνται ως ακολούθως. Παρακολουθούνται συνεχώς η πρόωρη σύγκλιση και η υπερβολική απόκλιση του πληθυσμού, μέσω της επεξεργασίας στατιστικών πληροφοριών για τον πληθυσμό. Εάν η επεξεργασία δείξει ότι αναμένεται κάποια τέτοια κατάσταση, τότε προσαρμόζονται οι πιθανότητες των τελεστών και η κλιμάκωση της συνάρτησης ποιότητας έτσι ώστε η κατανομή του πληθυσμού να παραμείνει σε μια λογική κατάσταση. Όταν παρατηρείται πρόωρη σύγκλιση, μειώνεται η πιθανότητα του ανασυνδυασμού και αυξάνεται η πιθανότητα μετάλλαξης. Το αντίθετο συμβαίνει στην περίπτωση της υπερβολικής απόκλισης, δηλαδή αυξάνεται η πιθανότητα του ανασυνδυασμού και μειώνεται η πιθανότητα μετάλλαξης ([BaPK94][KaBP94] [PeKa94][PeKP93]).

Αν και οι κλασσικοί τελεστές των ΓΑ θεωρούνται ο ανασυνδυασμός και η μετάλλαξη, υπάρχουν πολλοί άλλοι τελεστές οι οποίοι έχουν προταθεί και χρησιμοποιηθεί σε πολλές υλοποιήσεις. Ένας από αυτούς, ο τελεστής τοπικής αναρρίχησης χρησιμοποιήθηκε επιτυχώς συνδυαστικά και παραμετρικά προβλήματα βελτιστοποίησης ([Mühl89] [Mühl92] [MüSB91]). Για να επιταχυνθεί η αναζήτηση, χρησιμοποιήσαμε ένα τελεστή ανάλογο του τελεστή αναρρίχησης, με όνομα "φαινοτυπική μετάλλαξη" (phenotype mutation) [PeKP93]. Ο τελεστής εφαρμόζεται μόνο στο καλύτερο άτομο του πληθυσμού της κάθε γενιάς, εκτελώντας μία τοπική αναζήτηση στην γειτονιά του καλύτερου ατόμου, σε μια προσπάθεια να βρεθεί ένα καλύτερο σημείο στον χώρο αναζήτησης. Ο τελεστής αυτός αυξάνει σε μεγάλο βαθμό την σύγκλιση του ΓΑ, ειδικά στην περίπτωση που ο χώρος αναζήτησης δεν έχει ασυνέχειες. Στην περίπτωση του προβλήματος του περιοδεύοντος πωλητή (χώρος αναζήτησης με πολλές ασυνέχειες), ο τελεστής αναρρίχησης σπάνια βρίσκει επιτρεπτά σημεία και έτσι επιλέχθηκε να μην χρησιμοποιηθεί.

Επίσης χρησιμοποιείται και ελιτισμός έτσι ώστε να εξαλειφθεί η πιθανότητα καταστροφής του καλύτερου ατόμου από κάποιο γενετικό τελεστή. Διατηρείται μόνο το καλύτερο άτομο κάθε γενιάς.

5.3.2. Διάταξη Αναπαραγωγής Επιπέδου 0 των ΣΠΔΕΣ

Για να εμφανιστεί η διαφορετική εξελικτική συμπεριφορά κάθε υπο-πληθυσμού των ΣΠΔΕΣ, είναι απαραίτητο να οριστεί διαφορετικά η διάταξη *Αναπαραγωγής* για κάθε υπο-πληθυσμό. Στόχος της υλοποίησης είναι η δημιουργία βάσιμων πειραματικών αποτελεσμάτων τα οποία θα επιβεβαιώνουν ότι δεν είναι μόνο χρήσιμη αλλά και προτιμότερη, η χρήση διαφορετικής εξελικτικής συμπεριφοράς σε κάθε υπο-πληθυσμό. Όταν δηλαδή, κάθε υπο-πληθυσμός χρησιμοποιεί διαφορετική εξελικτική συμπεριφορά, βελτιώνεται η διερεύνηση και εκμετάλλευση του χώρου λύσεων. Στην ουσία, και σύμφωνα με τον ορισμό της εξελικτικής συμπεριφοράς (ενότητα 4.2), η διαφορετική εξελικτική συμπεριφορά κάθε υπο-πληθυσμού, είναι το αποτέλεσμα επιλογής διαφορετικών τελεστών, παραμέτρων και μηχανισμών για κάθε υπο-πληθυσμό, τα οποία είναι τα στοιχεία της διάταξης *Αναπαραγωγής*.

Για να είναι βάσιμος ο έλεγχος της χρησιμότητας και της αποτελεσματικότητας των ΣΠΔΕΣ έναντι του μοντέλου νησίδων, τα στοιχεία της διάταξης *Αναπαραγωγής* για κάθε υπο-πληθυσμό επιλέχθηκαν έτσι ώστε:

α. όταν κάθε διάταξη υλοποιείται από μόνη της σε ένα ΣΓΑ να μην έχει καλύτερη απόδοση από την διάταξη του μοντέλου νησίδων που παρουσιάστηκε στην προηγούμενη ενότητα 5.2.1,

β. ο καθορισμός των τελεστών, παραμέτρων και μηχανισμών για κάθε υπο-πληθυσμό δεν είναι ο καλύτερος δυνατός για τα προβλήματα της δοκιμής, δηλαδή δεν είναι προσανατολισμένος στην αντιμετώπιση και επίλυση των συγκεκριμένων προβλημάτων κάνοντας χρήση ειδικών τελεστών, παραμέτρων, κωδικοποιήσεων και μηχανισμών.

Κάτω από αυτούς τους περιορισμούς η υλοποίηση των ΣΠΔΕΣ **δεν είναι η βέλτιστη**. Έτσι η πιθανή εμφάνιση δυνατοτήτων καλύτερης διερεύνησης και βελτιωμένης απόδοσης των ΣΠΔΕΣ οφείλεται στο συνδυασμό των διαφορετικών εξελικτικών συμπεριφορών και όχι στην καλύτερη παραμετροποίηση των υπο-πληθυσμών.

Επειδή το μέγεθος πληθυσμού (*ΜεΠ*) του μοντέλου του επιπέδου 1 ($ΣΠ^1$) είναι οκτώ, πρέπει να οριστούν οκτώ διαφορετικές διατάξεις *Αναπαραγωγής*.

Εξ'ορισμού η τιμή εποχής είναι ίση με 1 και είναι ίδια με την τιμή του μοντέλου νησίδων

Για να υπάρξουν οι ελάχιστες δυνατές διαφορές από το μοντέλο νησίδων αλλά και για να επιλεγούν διευθετήσεις των διατάξεων οι οποίες δεν έχουν καλύτερη απόδοση από την διευθέτηση της διάταξης του μοντέλου νησίδων, κάποια στοιχεία της διάταξης *Αναπαραγωγής* επιλέχθηκαν να είναι ίδια για κάθε υπο-πληθυσμό και όμοια με την διάταξη του μοντέλου νησίδων. Τα όμοια χαρακτηριστικά είναι:

- αναλογική επιλογή με τον τροχό της ρουλέτας (roulette wheel, fitness proportionate parent selection) σε συνδυασμό με την μέθοδο γραμμικής δυναμικής κλιμάκωσης των ποιοτήτων,
- τελεστής αναρρίχησης "φαινοτυπική μετάλλαξη",
- ελιτισμός, διατηρώντας μόνο το καλύτερο άτομο σε κάθε γενιά,
- προσαρμογή των παραμέτρων του ανασυνδυασμού και της μετάλλαξης (εκτός αν αναφέρεται διαφορετικά) και κλιμάκωση της ποιότητας.

Τα διαφορετικά χαρακτηριστικά των υπο-πληθυσμών περιγράφονται παρακάτω. Για ευκολία αναφοράς οι υπο-πληθυσμοί αριθμούνται από 1 έως 8.

Υπο-πληθυσμός 1: Η διευθέτηση της διάταξης *Αναπαραγωγής* είναι ίδια με την διευθέτηση του μοντέλου νησίδων.

Υπο-πληθυσμός 2: Η διευθέτηση της διάταξης είναι όπως και στον υπο-πληθυσμό 1, εκτός του τελεστή ανασυνδυασμού πολλών σημείων (multi-point crossover). Στην θέση του χρησιμοποιείται ένας τελεστής ανασυνδυασμού ο οποίος χειρίζεται γονίδια και όχι ολόκληρα χρωμοσώματα. Ο τελεστής χρησιμοποιεί ανασυνδυασμό ενός σημείου (one-point crossover) για κάθε γονίδιο, χωρίζοντάς το γονίδιο σε δύο μέρη. Όπως και στον ανασυνδυασμό ενός σημείου, το σημείο κοπής είναι τυχαίο. Ο απόγονος δημιουργείται παίρνοντας το σημαντικότερο μέρος του γονέα με την καλύτερη ποιότητα και το λιγότερο σημαντικό μέρος του γονέα με την χειρότερη ποιότητα. Αυτός ο τελεστής δημιουργεί απογόνους με μικρές διακυμάνσεις της ποιότητας στην περιοχή της ποιότητας του καλύτερου γονέα (τελεστής MostSignificantPart_LeastSignificantPart, MSP_LSP).

Υπο-πληθυσμός 3: Όπως και στον υπο-πληθυσμό 2, δεν χρησιμοποιείται ο τελεστής ανασυνδυασμού πολλών σημείων (multi-point crossover). Στην θέση του

χρησιμοποιείται ένας τελεστής ανασυνδυασμού παρόμοιος με τον τελεστή MSP_LSP του προηγούμενου υπο-πληθυσμού 2. Και αυτός ο τελεστής χειρίζεται γονίδια και όχι ολόκληρα χρωμοσώματα και χρησιμοποιεί ανασυνδυασμό ενός σημείου (one-point crossover) για κάθε γονίδιο, χωρίζοντάς το γονίδιο σε δύο μέρη. Η διαφορά τους είναι στον τρόπο δημιουργίας του απογόνου. Αντίθετα από τον τελεστή MSP_LSP, ο απόγονος δημιουργείται παίρνοντας το λιγότερο σημαντικό μέρος του γονέα με την καλύτερη ποιότητα και το σημαντικότερο μέρος του γονέα με την χειρότερη ποιότητα. Έτσι, αυτός ο τελεστής δημιουργεί απογόνους με μικρές διακυμάνσεις της ποιότητας στην περιοχή της ποιότητας του χειρότερου γονέα (τελεστής LeastSignificantPart_MostSignificantPart, LSP_MSP).

Υπο-πληθυσμός 4: Αντί του τελεστή ανασυνδυασμού πολλών σημείων χρησιμοποιείται ο τελεστής ανασυνδυασμού δύο σημείων (two-point crossover).

Υπο-πληθυσμός 5: Η διευθέτηση της διάταξης είναι όπως και στον υπο-πληθυσμό 1. Η μόνη διαφορά τους βρίσκεται στον αριθμό των σημείων ανασυνδυασμού. Εδώ ο αριθμός των σημείων ανασυνδυασμού είναι ίσος με το ήμισυ του αριθμού των γονιδίων του χρωμοσώματος.

Υπο-πληθυσμός 6: Αντί του τελεστή ανασυνδυασμού πολλών σημείων χρησιμοποιείται ο τελεστής ομοιόμορφου ανασυνδυασμού (uniform crossover) [Sysw89].

Υπο-πληθυσμός 7: Η διευθέτηση της διάταξης είναι όπως και στον υπο-πληθυσμό 1, αλλά η προσαρμογή των πιθανοτήτων εφαρμογής των γενετικών τελεστών ευνοεί την μετάλλαξη. Στην περίπτωση του υπο-πληθυσμού 1, η πρόωρη σύγκλιση ορίζεται όταν το 50% ή περισσότερο του υπο-πληθυσμού έχει ποιότητα κοντά στην ποιότητα του καλύτερου ατόμου. Εδώ το όριο ελαττώνεται στο 8% του υπο-πληθυσμού. Η απόκλιση στον υπο-πληθυσμό 1 ορίζεται όταν μόνο το 10% του υπο-πληθυσμού έχει ποιότητα κοντά στην ποιότητα του καλύτερου ατόμου. Εδώ το όριο μειώνεται στο 3%. Έτσι σ'αυτόν τον υπο-πληθυσμό διατηρείται μια σχετικά υψηλή τιμή της συχνότητας μετάλλαξης, βελτιώνοντας την διερεύνηση του χώρου αναζήτησης και κατευνάζοντας (συνολικά και σε συνεργασία με τους υπόλοιπους υπο-πληθυσμούς) το πρόβλημα της πρόωρης σύγκλισης.

Υπο-πληθυσμός 8: Εδώ χρησιμοποιείται η διευθέτηση της διάταξης του υπο-πληθυσμού 1, αλλά χωρίς προσαρμογή των πιθανοτήτων εφαρμογής των γενετικών τελεστών. Η τιμή της πιθανότητας ανασυνδυασμού διατηρείται σταθερή και μικρή (0.2), ενώ η συχνότητα μετάλλαξης σταθερή και αυξημένη. Ο στόχος της διάταξης είναι η βελτίωση της ικανότητας διερεύνησης, όπως και του προηγούμενου υπο-πληθυσμού.

Ως τελικό σχόλιο της παρουσίασης υλοποίησης των ΣΠΔΕΣ, ας επισημανθεί ότι κατά την δημιουργία της διάταξης κάθε υπο-πληθυσμού, οι επιλογές των στοιχείων δεν είναι οι καλύτερες δυνατές, έτσι ώστε να ελεγχθεί η χρησιμότητα της ιδέας των ΣΠΔΕΣ.

Κεφάλαιο 6

Προσομοιώσεις

Οι γενετικοί αλγόριθμοι έχουν την δυνατότητα να αντιμετωπίσουν με επιτυχία ένα πλήθος διαφορετικών προβλημάτων. Σ'αυτό το κεφάλαιο παρουσιάζονται τα προβλήματα που επιλέχθηκαν για να ελεγχθεί η απόδοση των ΣΠΔΕΣ έναντι των χαΠΓΑ με την ίδια εξελικτική συμπεριφορά σε κάθε υπο-πληθυσμό, καθώς επίσης και τα αποτελέσματα της σύγκρισης

6.1. Περιγραφή των Προβλημάτων

Η επιλογή των προβλημάτων έγινε με το σκεπτικό της δοκιμής των ΣΠΔΕΣ σε αρκετά διαφορετικά προβλήματα με αρκετά διαφορετικούς χώρους λύσεων. Το πρώτο πρόβλημα που δοκιμάστηκε είναι η εκπαίδευση ενός πλήρως συνδεμένου Αναδρομικού Τεχνητού Νευρωνικού Δικτύου (ΑΤΝΔ). Το πρόβλημα είναι πολύ δύσκολο και το ΑΤΝΔ απαιτεί μεγάλο χρόνο εκπαίδευσης. Ο μεγάλος χρόνος εκπαίδευσης που απαιτείται κάνει απαγορευτική την χρήση του στην μελέτη των ΣΠΔΕΣ και την σύγκρισή τους με το μοντέλο νησίδων.

Έτσι οδηγηθήκαμε στην επιλογή της συνάρτησης Rastrigin η οποία είναι μεν αρκετά δύσκολη, αλλά η βελτιστοποίησή της απαιτεί πολύ λιγότερο χρόνο από την εκπαίδευση του ΑΤΝΔ, πράγμα που επιτρέπει την χρήση της στη μελέτη των δύο μοντέλων. Η συνάρτηση Rastrigin είναι μία από τις καθιερωμένες συναρτήσεις ελέγχου της απόδοσης (benchmark) διαφόρων μοντέλων εξελικτικών αλγορίθμων. Πρίν φτάσουμε στην επιλογή της συγκεκριμένης συνάρτησης, δοκιμάστηκαν και άλλες καθιερωμένες συναρτήσεις ελέγχου (π.χ. συνάρτηση Schwefel, συναρτήσεις De Jong, συναρτήσεις παραπλάνησης (deception functions)). Επειδή ήταν

πολύ εύκολο να βρεθεί το ολικό βέλτιστο αυτών των συναρτήσεων, δεν προσφέρονταν για την μελέτη των δύο μοντέλων.

Το τελευταίο πρόβλημα που δοκιμάσαμε, αντίθετα από τα δύο προηγούμενα προβλήματα, είναι ένα συνδυαστικό πρόβλημα, το πρόβλημα του περιοδεύοντος πωλητή, με ασυνεχή χώρο λύσεων. Η θετική απόδοση των ΣΠΔΕΣ στα τρία παραπάνω προβλήματα καθιστά πιθανή την καλή απόδοσή τους σε οποιοδήποτε πρόβλημα και οπωσδήποτε σε ένα ευρύ φάσμα προβλημάτων.

6.1.1. Εκπαίδευση Τεχνητών Νευρωνικών Δικτύων

Τα ΑΤΝΔ, είναι πιο γενικευμένα δίκτυα και υπόσχονται την επίλυση προβλημάτων σχετικών με το χρόνο.

Το ΑΤΝΔ που επιλέχτηκε είναι πλήρως συνδεδεμένο και αποτελείται από σιγμοειδείς νευρώνες. Η τοπολογία του δικτύου φαίνεται στο σχήμα 6.1. Η διάσταση του πίνακα των βαρών W είναι n x (n+m+1), όπου n είναι ο συνολικός αριθμός των νευρώνων και m είναι ο αριθμός των εισόδων. Ο συνολικός αριθμός των βαρών είναι N=n(n+m+1).

Εάν $y_j(t)$ είναι η έξοδος του j νευρώνα την χρονική στιγμή t και $x_i(t)$ είναι η τιμή της i εισόδου την ίδια χρονική στιγμή, τότε η συνολική είσοδος του k νευρώνα την χρονική στιγμή t, δίνεται από το άθροισμα

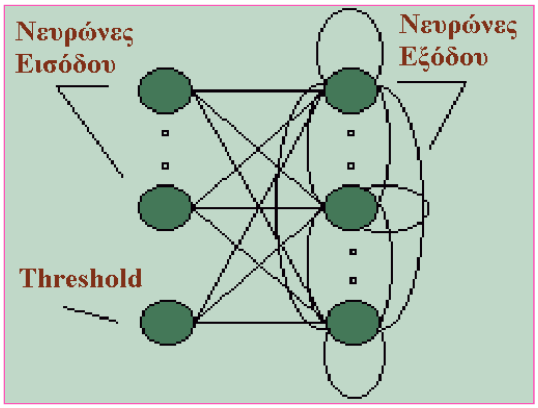
$$S_k(t) = \sum_{j=1}^{n} w_{kj} y_j(t-1) + \sum_{i=n+1}^{n+m+1} w_{ki} x_{in}(t) \tag{6.1}$$

όπου $x_{m+1}(t)$=1, είναι μία επιπλέον είσοδος η οποία ελέγχει το κατώφλι (threshold).

Η έξοδος κάθε νευρώνα την χρονική στιγμή t είναι

$$y_k(t) = f(S_k(t)) \tag{6.2}$$

όπου f είναι η συνάρτηση υπερβολικής εφαπτομένης.



Σχήμα 6.1. Τοπολογία Αναδρομικού Τεχνητού Νευρωνικού Δικτύου

Αν M είναι το σύνολο των νευρώνων για τους οποίους υπάρχει μία επιθυμητή τιμή εξόδου $d_k(t)$ για κάθε χρονική στιγμή, τότε το συνολικό λάθος την χρονική στιγμή t είναι

$$J(t) = \frac{1}{2}\sum_{k}^{M}[d_k(t) - y_k(t)]^2 \tag{6.3}$$

Ο στόχος είναι η προσαρμογή των τιμών των βαρών έτσι ώστε το συνολικό λάθος

$$J_T = \sum_{t=1}^{T} J(t) \tag{6.4}$$

να γίνει μικρότερο από μια προκαθορισμένη τιμή. Τ είναι η περίοδος του επιθυμητού σήματος εξόδου ή ο χρόνος παρουσίασης μιας χρονικής σειράς.

Το μοντέλο του ΑΤΝΔ που αναφέρθηκε, μπορεί να χρησιμοποιηθεί για την δημιουργία οριακών κύκλων, καθώς επίσης και την αποθήκευση χρονοσειρών και τροχιών. Εδώ, το μοντέλο του ΑΤΝΔ, εκπαιδεύεται να παράγει δύο οριακούς κύκλους, με τη μορφή δύο ημιτονοειδών συναρτήσεων διαφορετικού πλάτους και συχνότητας. Το ΑΤΝΔ εκπαιδεύτηκε έτσι ώστε δύο διαφορετικές τιμές εισόδου (0.2 και 0.8) να οδηγούν τους νευρώνες εξόδου σε διαφορετικές ταλαντώσεις. Οι επιθυμητές έξοδοι είναι y1=0.2∗sin(2πt/T) και y2=0.8∗sin(πt/T) αντίστοιχα.

Το ΑΤΝΔ έχει πέντε νευρώνες (35 βάρη), δύο από τους οποίους είναι οι νευρώνες εξόδου. Κάθε βάρος κωδικοποιείται ως ένα σύνολο 16 δυαδικών ψηφίων. Έτσι ο γονότυπος έχει μήκος 25x16=560 δυαδικά ψηφία και ο συνολικός χώρος λύσεων $2^{560}=3.77\text{x}10^{168}$ διαφορετικές τιμές.

Για την εκπαίδευση χρησιμοποιήθηκαν 16 δείγματα ανά περίοδο. Το ΑΤΝΔ εκπαιδεύτηκε μέχρι το συνολικό λάθος να φτάσει στην τιμή 0.003. Είναι ένα αρκετά δύσκολο πρόβλημα που απαιτεί πολύ χρόνο και υπολογιστικές δυνατότητες. Το ΑΤΝΔ μπορεί να χρησιμοποιηθεί στον χώρο της ρομποτικής για την αποθήκευση και αναπαραγωγή πολύπλοκων προτύπων κίνησης, όπως επίσης και στο χώρο της βιολογίας για την μοντελοποίηση αντανακλαστικών αντιδράσεων. Περισσότερα στο θέμα των ΑΤΝΔ μπορούν να βρεθούν στο [PePa94].

6.1.2. Βελτιστοποίηση Συνάρτησης Rastrigin

Το δεύτερο πρόβλημα είναι η ελαχιστοποίηση της τιμής της συνάρτησης *f(x)* που προτάθηκε από τον L. Rastrigin [TöZi89]

$$f(x) = nA + \sum_{1}^{n}(x_i^2 - A\cos(\omega x_i)) \tag{6.5}$$

όπου n=20; A=10; ω=2π; $-5.12 \le x_i \le 5.12$;

Η επιφάνεια της συνάρτησης εξαρτάται από τις μεταβλητές Α και ω, οι οποίες ελέγχουν το πλάτος και την συχνότητα διαμόρφωσης αντίστοιχα. Όταν Α=10, το επιλεγμένο πεδίο ορισμού κυριαρχείται από την διαμόρφωση. Η συνάρτηση είναι συνεχής, με πάρα πολλά τοπικά ελάχιστα. Τα τοπικά ελάχιστα βρίσκονται σε ένα ορθογώνιο πλέγμα με πλάτος 1. Το σχήμα 6.2 δίνει την κατανομή των ελαχίστων για δύο διαστάσεις x_1 και x_2. Το ολικό ελάχιστο βρίσκεται στο σημείο x_i=0, *i*=1...,*n*, όπου *f*(*x*)=0.

Τα σημεία του πλέγματος όπου x_i=0 εκτός μιας συντεταγμένης όπου x_i=1, δίνουν *f*(*x*)=1, που είναι το δεύτερο καλύτερο ελάχιστο. Η τιμή του τοπικών ελαχίστων αυξάνεται όσο αυξάνεται η απόσταση των σημείων από το ολικό ελάχιστο.

8	5	4	5	**8**
5	2	1	2	5
4	1	**0**	1	4
5	2	1	2	5
8	5	4	5	**8**

Σχήμα 6.2. Κατανομή των ελαχίστων για την συνάρτηση Rastrigin

Η συνάρτηση χρησιμοποιείται συχνά για τον έλεγχο της απόδοσης διαφόρων μοντέλων εξελικτικών αλγορίθμων. Χρησιμοποιήθηκε από τους Hoffmeister και Bäck [HoBä90] για να συγκρίνουν την απόδοση των ΓΑ με την μέθοδο Εξελικτικής Στρατηγικής. Οι μέθοδοι που χρησιμοποίησαν δεν βρήκαν το ολικό ελάχιστο. Η καλύτερη τιμή την οποία επέτυχαν (μετά από 250 γενιές με πληθυσμό 50) ήταν f =10. Πολύ καλύτερα αποτελέσματα είχαν οι Mühlenbein κ.α. [MüSB91a][MüSB91b] οι οποίοι βρήκαν το ολικό ελάχιστο χρησιμοποιώντας ΠΓΑ, και ένα τελεστή τοπικής αναρρίχησης. Στον έλεγχο της απόδοσης των ΠΓΑ χρησιμοποιήθηκε επίσης από τον Maresky [Mare94], τους Cantu-Paz και Mejia-Olvera [CaMe94]

6.1.3. Πρόβλημα του Περιοδεύοντος Πωλητή

Το τελευταίο πρόβλημα είναι το γνωστό πρόβλημα του περιοδεύοντος πωλητή (Travelling Salesperson Problem - TSP), όπου πρέπει να βρεθεί ο μικρότερος κύκλος Hamilton σε ένα πλήρη γράφο με N κόμβους. Η Ευκλείδεια απόσταση μεταξύ δύο κόμβων υπολογίζεται από τις συντεταγμένες τους. Πιο απλά το πρόβλημα διατυπώνεται ως εξής. Κάποιος πωλητής πρέπει να επισκεφτεί N πόλεις. Ξεκινώντας από κάποια πόλη πρέπει να επισκεφτεί όλες τις πόλεις ακριβώς μία φορά και μετά να επιστρέψει στην αρχική. Το πρόβλημα είναι η επιλογή της ελάχιστης διαδρομής μεταξύ των N πόλεων.

Η επιλογή του συγκεκριμένου προβλήματος έγινε από την βιβλιοθήκη προβλημάτων TSPLIB έκδοση 1.1, που διατίθεται από τα Πανεπιστήμια Rice, ΗΠΑ και Augsburg, Γερμανία. Υπάρχουν διαθέσιμες τρεις κατηγορίες προβλημάτων.

- Συμμετρικά προβλήματα (Symmetric travelling salesman problems), όπου η απόσταση από κάποιο κόμβο α μέχρι κάποιο κόμβο β είναι ίδια με την απόσταση από τον κόμβο β μέχρι τον α.
- Μή συμμετρικά προβλήματα (Asymmetric travelling salesman problems), όπου η απόσταση από κάποιο κόμβο α μέχρι κάποιο κόμβο β μπορεί να είναι διαφορετική από την απόσταση από τον κόμβο β στον α.
- Προβλήματα δρομολόγησης οχημάτων ορισμένης χωρητικότητας (Capacitated vehicle routing problems), όπου δεδομένου N-1 κόμβων και μιας αποθήκης πρέπει να ικανοποιηθούν οι παραγγελίες όλων των κόμβων από την αποθήκη. Το πρόβλημα είναι η εύρεση της ελάχιστης συνολικής διαδρομής των φορτηγών η οποία ικανοποιεί τις

παραγγελίες των κόμβων χωρίς να παραβιάζεται ο περιορισμός της χωρητικότητας των φορτηγών.

Σκοπός μας δεν είναι η επίλυση του συγκεκριμένου προβλήματος (TSP) αλλά η επιβεβαίωση των αποτελεσμάτων και συμπερασμάτων που προκύπτουν για την απόδοση των ΣΠΔΕΣ από τα προηγούμενα προβλήματα. Επίσης, θα πρέπει να διεξαχθεί ένας σημαντικός αριθμός πειραμάτων για στατιστικούς λόγους. Έτσι, από τα διαθέσιμα προβλήματα της βιβλιοθήκης, επιλέχτηκε ένα μέσης δυσκολίας και μεγέθους με 29 συνολικά πόλεις στην περιοχή της Βαυαρίας στη Γερμανία. Τα στοιχεία παρέχονται στο αρχείο "bays29" της βιβλιοθήκης, όπου αναγράφονται οι ακριβείς αποστάσεις μεταξύ όλων των πόλεων (Παράρτημα Β).

Ο χώρος λύσεων του προβλήματος TSP έχει πολλές ασυνέχειες, κατά συνέπεια ο τελεστής αναρρίχησης σπάνια βρίσκει επιτρεπτά σημεία και έτσι επιλέχθηκε να μην χρησιμοποιηθεί σ'αυτό το συγκεκριμένο πρόβλημα. Αυτή είναι και η μόνη διαφοροποίηση του αλγορίθμου από τα προηγούμενα προβλήματα.

Το πρόβλημα TSP χρησιμοποιείται, όπως και τα προηγούμενα, για να επιβεβαιωθεί ότι η χρήση των ΣΠΔΕΣ προσφέρει δυνατότητες καλύτερης διερεύνησης του χώρου λύσεων έναντι του μοντέλου νησίδων. Στόχος δεν είναι η επίλυση του προβλήματος TSP και έτσι, κρίθηκε αναγκαίο να διατηρηθεί η πρωταρχική συμπεριφορά του αλγόριθμου όπως χρησιμοποιήθηκε στα προηγούμενα προβλήματα. Σύμφωνα με αυτό το σκεπτικό δεν χρησιμοποιήθηκαν κωδικοποίηση, τελεστές ή μηχανισμοί προσανατολισμένοι στην επίλυση του συγκεκριμένου προβλήματος, όπως ανασυνδιασμός ακμών (edge recombination), ανασυνδιασμός σειράς (order crossover), ανασυνδιασμός μερικής ταύτισης (partially matched crossover, PMX), κυκλικός ανασυνδυασμός (cycle crossover), ανασυνδυσμός βασισμένος στην θέση (position based crossover), ανασυνδιασμός μήτρας (matrix crossover), κωδικοποίηση βασισμένη στη θέση των γονιδίων (locus based encoding), πλεονάζουσα γενετική κωδικοποίηση (redundant genetic coding), κλπ. ([BuMo94] [CuFo92] [Davi91] [DzWh94] [HoGL93] [JoSV89] [TKSM94] [MaWh92] [Mühl91] [Mühl92] [MüGK88][Pete90][Tale94][VaJo93][WhSF89])

6.2. Αποτελέσματα Προσομοιώσεων

Όπως ήδη επισημάνθηκε (ενότητα 5.2.2), για να είναι βάσιμος ο έλεγχος της χρησιμότητας και της αποτελεσματικότητας των ΣΠΔΕΣ έναντι του μοντέλου νησίδων, σε κάθε υπο-πληθυσμό επιλέχθηκαν τα στοιχεία της διάταξης *Αναπαραγωγής*, έτσι ώστε όταν κάθε διάταξη υλοποιείται από μόνη της σε ένα ΣΓΑ, να μην έχει καλύτερη απόδοση από την διάταξη του μοντέλου νησίδων (ενότητα 5.2.1).

Έτσι, πριν την σύγκριση των πειραματικών αποτελεσμάτων μεταξύ των ΣΠΔΕΣ και του μοντέλου νησίδων, ελέγχθηκε η απόδοση των οκτώ διαφορετικών διατάξεων *Αναπαραγωγής*, υλοποιημένες σε ανεξάρτητους ΣΓΑ, στα δύο πρώτα προβλήματα (ΑΤΝΔ και συνάρτηση Rastrigin). Και για τα δύο προβλήματα εκτελέστηκε ο αλγόριθμος για ένα σύνολο 100 πειραμάτων, για κάθε μία από τις οκτώ διατάξεις *Αναπαραγωγής*. Το μέγεθος του πληθυσμού ήταν 200 άτομα σε όλα τα πειράματα.

Πίνακας 6.1. Σύγκριση αποτελεσμάτων διαφορετικών διατάξεων *Αναπαραγωγής*

	Συνάρτηση Rastrigin		**ΑΤΝΔ**	
Αριθμός υπο-πληθυσμού	Λάθος του καλύτερου ατόμου	Μέσος όρος λάθους	Ποσοστό επιτυχίας	Μέσος όρος λάθους
1	0.994010	6.7559	52%	0.0105
2	1.990128	7.1242	41%	0.0114
3	2.984909	6.8653	40%	0.0140
4	0.994995	6.7535	43%	0.0117
5	2.984970	7.0645	20%	0.0221
6	1.989944	6.8852	46%	0.0117
7	2.984924	6.8311	35%	0.0126
8	1.990173	8.1290	50%	0.0102

Τα αποτελέσματα δείχνουν ότι η διάταξη του υπο-πληθυσμού 1, η οποία χρησιμοποιείται στο μοντέλο νησίδων, έχει καλύτερα αποτελέσματα από τις υπόλοιπες διατάξεις (Πίνακας 6.1). Έτσι η πιθανή καλύτερη απόδοση των ΣΠΔΕΣ δεν οφείλεται στην καλύτερη απόδοση κάποιου ή κάποιων από τους υπο-πληθυσμούς, αλλά στην συνεργασία διαφορετικών συμπεριφορών εξέλιξης.

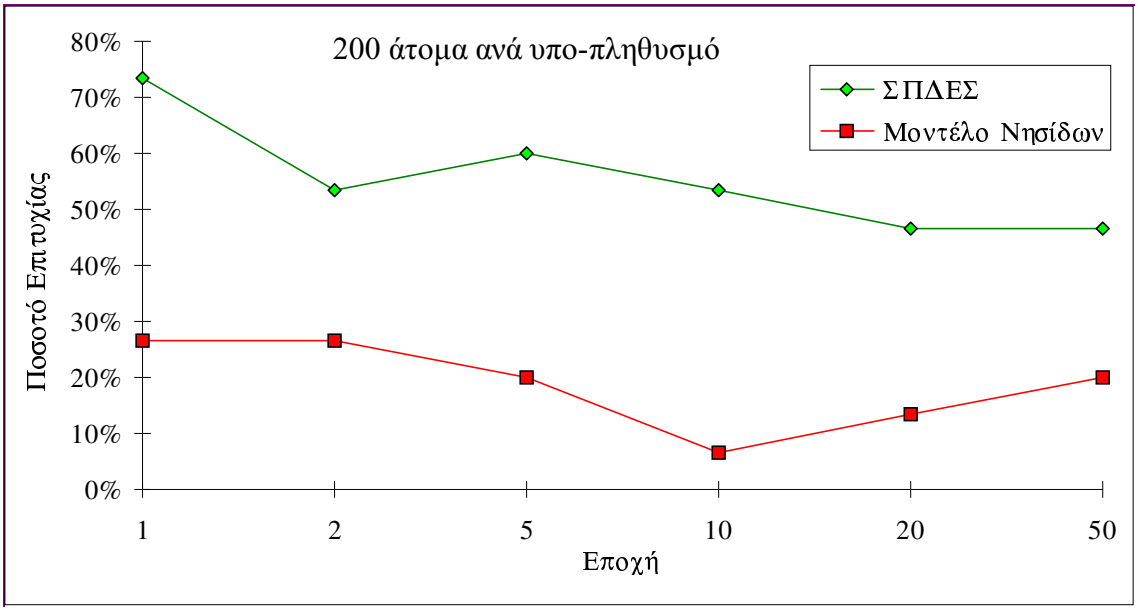
Κατά την σύγκριση των ΣΠΔΕΣ με το μοντέλο νησίδων, έγιναν δύο σειρές πειραμάτων για κάθε πρόβλημα. Η μία σειρά πειραμάτων αφορά τους ΣΠΔΕΣ και η άλλη το μοντέλο νησίδων.

6.2.1. Εκπαίδευση Τεχνητών Νευρωνικών Δικτύων

Στην περίπτωση της εκπαίδευσης του ΑΤΝΔ, κάθε σειρά αποτελείται από 90 πειράματα (συνολικά 180 πειράματα). Τα πειράματα αυτά κατανέμονται σε 6 διαφορετικές τιμές εποχής μετανάστευσης ατόμων, με 15 πειράματα για κάθε εποχή. Οι τιμές εποχής που χρησιμοποιήθηκαν είναι 1, 2, 5, 10, 20 και 50 γενιές. Το μέγεθος πληθυσμού για κάθε υπο-πληθυσμό είναι 200 άτομα.

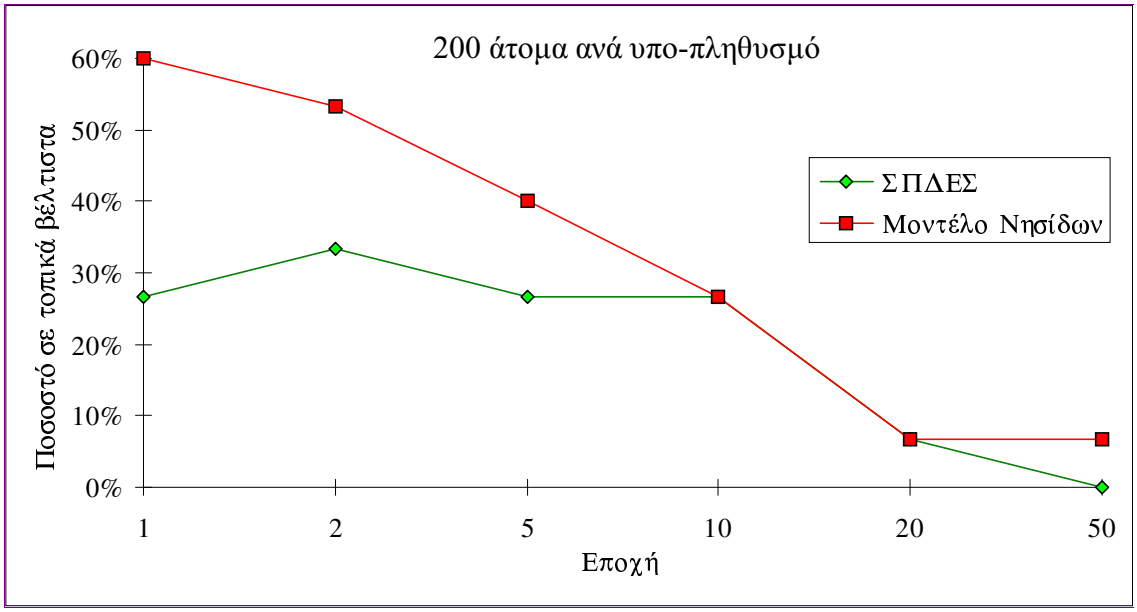
Από τα πειράματα φάνηκε η βελτιωμένη απόδοση των ΣΠΔΕΣ έναντι του μοντέλου νησίδων. (Όπως έχει ήδη αναφερθεί κάθε υπο-πληθυσμός του μοντέλου νησίδων έχει την συμπεριφορά του υπο-πληθυσμού 1 των ΣΠΔΕΣ).

Το ποσοστό επιτυχίας , στην χειρότερη περίπτωση, διπλασιάστηκε (σχ. 6.3). Δηλαδή για τιμή *εποχής* ίση με 2, το ποσοστό επιτυχίας των ΣΠΔΕΣ ήταν 53.3%, το οποίο σημαίνει ότι ο αριθμός των επιτυχών πειραμάτων ήταν 0.533x15=8, ενώ για το μοντέλο νησίδων, το ποσοστό επιτυχίας ήταν 26.7%, το οποίο σημαίνει ότι ο αριθμός των επιτυχών πειραμάτων ήταν 0.267x15=4

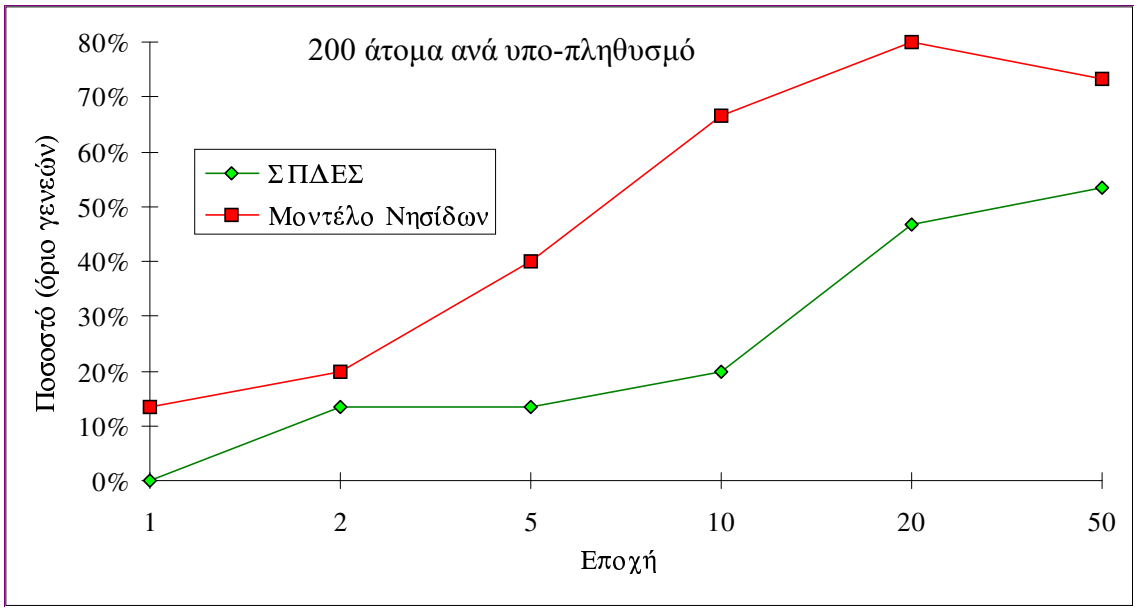


Σχήμα 6.3. Επιτυχή πειράματα για τους ΣΠΔΕΣ και το μοντέλο Νησίδων (ΑΤΝΔ)

Εκτός της βελτίωσης του ποσοστού επιτυχίας επίλυσης του προβλήματος, χρησιμοποιώντας τους ΣΠΔΕΣ παρατηρήθηκε επίσης μείωση του ποσοστού πειραμάτων που “καθηλώνονταν” σε τοπικά βέλτιστα. Το ποσοστό αυτό μειώνεται καθώς αυξάνεται η τιμή της *εποχής* (σχ. 6.4). Με την χρήση των ΣΠΔΕΣ, η αποτυχία των περισσότερων πειραμάτων οφείλεται στο όριο γενεών που χρησιμοποιήθηκε στα πειράματα. (Κάθε πείραμα τελείωνε όταν ξεπερνούσε τις 10000 γενιές ή έφτανε στο ολικό βέλτιστο ή καθηλώνονταν σε κάποιο τοπικό βέλτιστο). Στην περίπτωση του μοντέλου νησίδων, οι περισσότερες αποτυχίες οφείλονται σε αδυναμία αποφυγής των τοπικών βέλτιστων (σχ. 6.5). Εδώ γίνεται εμφανής η ικανότητα των ΣΠΔΕΣ να βελτιώνουν την διερεύνηση του χώρου λύσεων.

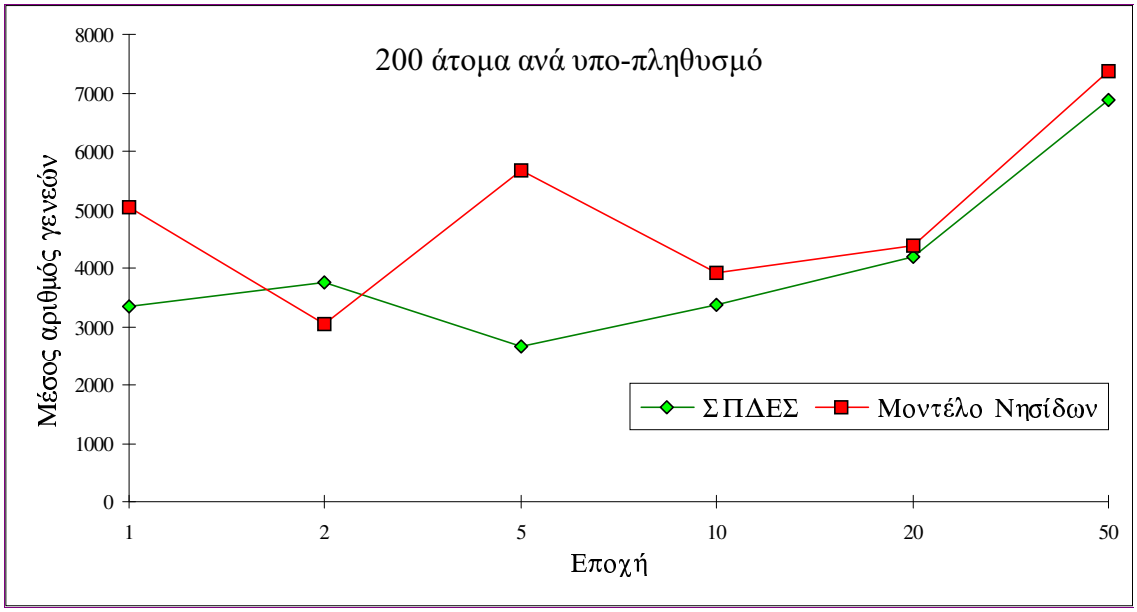


Σχήμα 6.4. Πειράματα που έμειναν σε τοπικά βέλτιστα. (ΑΤΝΔ)



Σχήμα 6.5. Πειράματα που σταμάτησαν λόγω ορίου γενεών (ΑΤΝΔ)

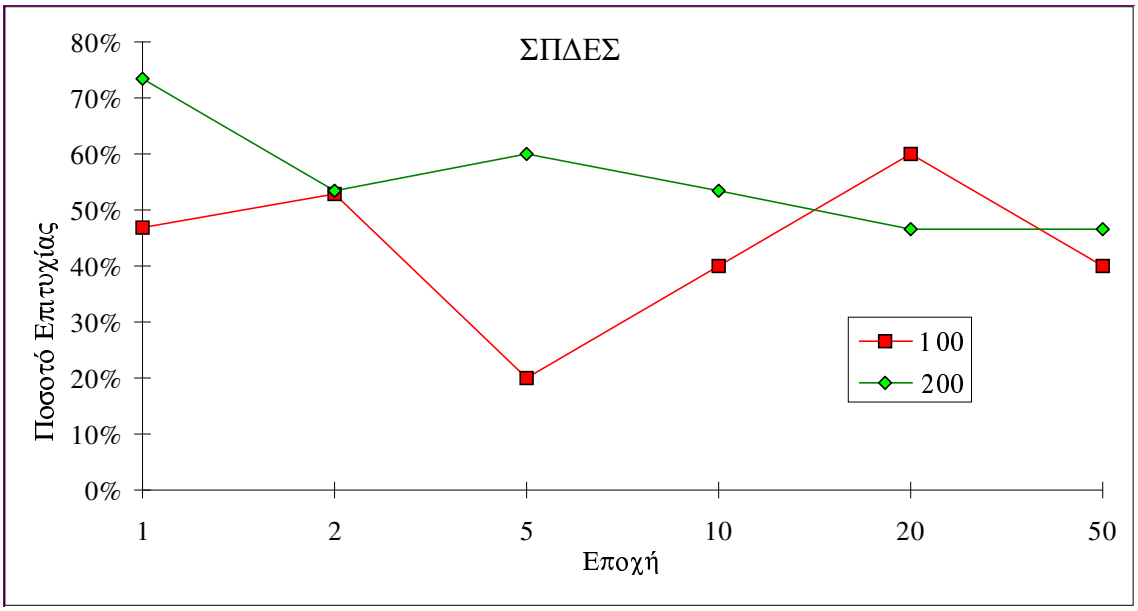
Ο αριθμός γενεών που απαιτούνται για να εκπαιδευτεί το ΑΤΝΔ βελτιώνεται και γίνεται μικρότερος με την χρήση των ΣΠΔΕΣ. Το σχήμα 6.6 δείχνει τον μέσο αριθμό γενεών για κάθε πείραμα. Παρατηρούμε ότι όσο αυξάνεται η τιμή εποχής τόσο αυξάνεται και ο αριθμός γενεών που απαιτούνται για την σύγκλιση των ΣΠΔΕΣ. Το ίδιο φαίνεται να ισχύει και για το μοντέλο νησίδων, αν και δεν είναι τόσο ξεκάθαρο όπως στους ΣΠΔΕΣ.



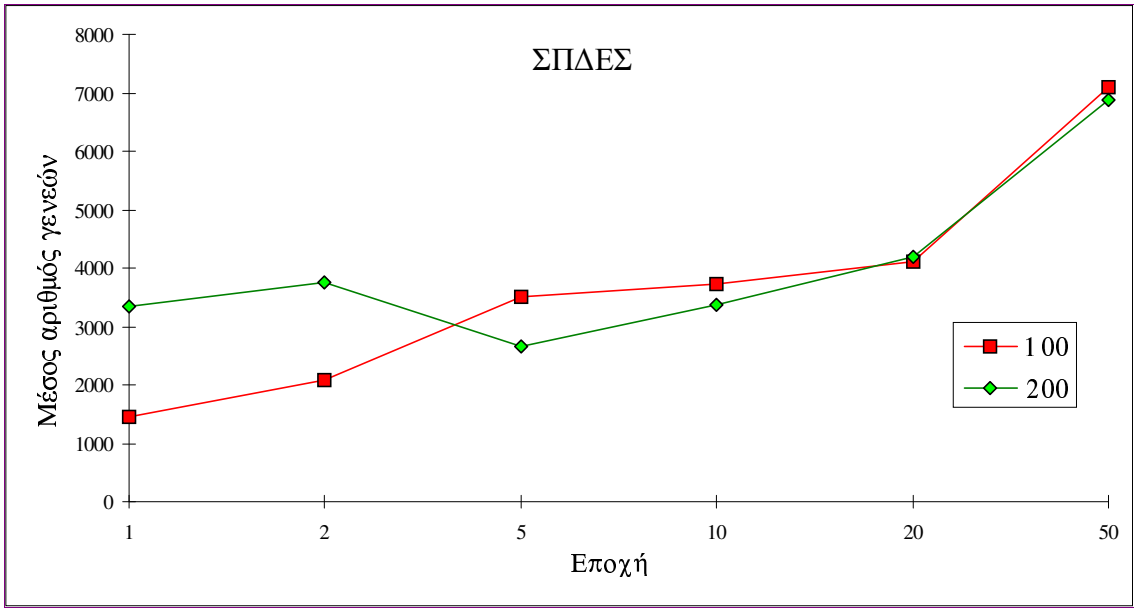
Σχήμα 6.6. Μέσος αριθμός γενεών για τα επιτυχή πειράματα (ΑΤΝΔ)

Εκτός των 180 πειραμάτων με μέγεθος υπο-πληθυσμού 200 άτομα, έγιναν ακόμη 90 πειράματα ελέγχου της απόδοσης των ΣΠΔΕΣ με μέγεθος υπο-πληθυσμού 100. Το σχήμα 6.7 δείχνει το ποσοστό επιτυχίας των ΣΠΔΕΣ για μεγέθη 100 και 200 ατόμων ανά υπο-πληθυσμό. Κάθε τιμή είναι ο μέσος όρος 15 πειραμάτων. Η απόδοση των ΣΠΔΕΣ με μέγεθος υπο-πληθυσμού 100 είναι συγκρίσιμη με το μέγεθος υπο-πληθυσμού 200.

Και εδώ ισχύει η αύξηση του αριθμού γενεών που απαιτούνται για την σύγκλιση των ΣΠΔΕΣ καθώς αυξάνεται η τιμή εποχής. Ο αριθμός γενεών που απαιτούνται για να συγκλίνουν οι ΣΠΔΕΣ με μέγεθος υπο-πληθυσμού 100 είναι συγκρίσιμος με το μέγεθος υπο-πληθυσμού 200, όπως φαίνεται στο σχήμα 6.8.



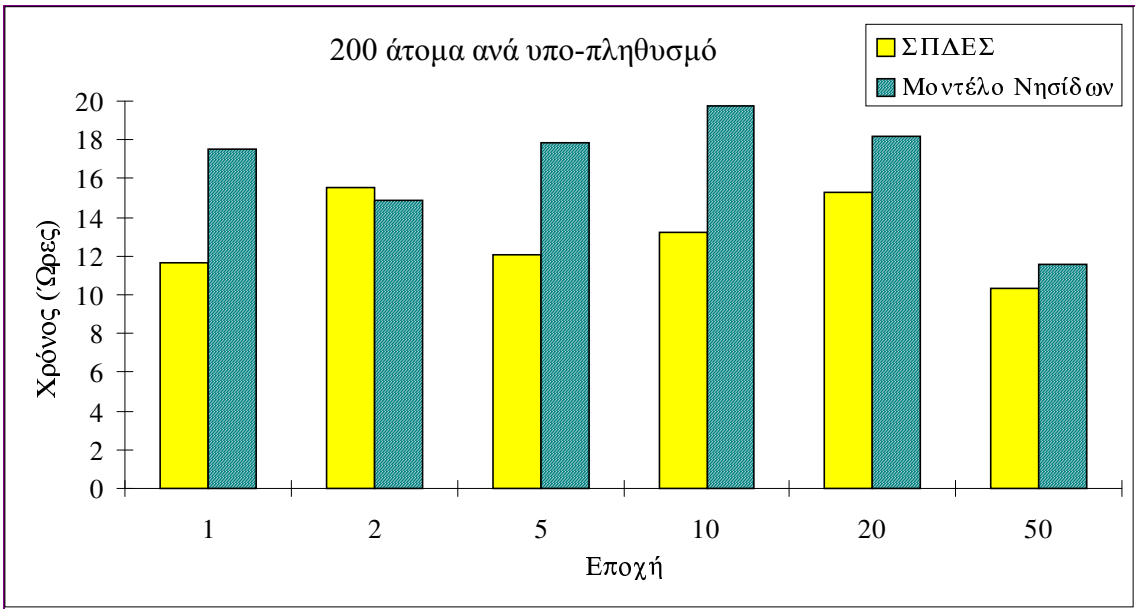
Σχήμα 6.7. Επιτυχή πειράματα των ΣΠΔΕΣ (100 και 200 άτομα ανά υπο-πληθυσμό - ΑΤΝΔ)



Σχήμα 6.8. Μέσος αριθμός γενεών των ΣΠΔΕΣ (100 και 200 άτομα ανά υπο-πληθυσμό - ΑΤΝΔ)

Περισσότερο λεπτομερής και ακριβή μελέτη της επίδρασης διαφορετικών τιμών εποχής και μεγεθών υπο-πληθυσμού, απαιτεί εκτεταμένο πειραματισμό με διαφορετικά μεγέθη υπο-πληθυσμών. Αυτό είναι πρακτικά αδύνατο να γίνει με το συγκεκριμένο πρόβλημα εκπαίδευσης των ΑΤΝΔ λόγω του μεγάλου χρόνου εκτέλεσης των πειραμάτων.

Το σχήμα 6.9 δείχνει το μέσο αριθμό ωρών εκτέλεσης ενός πειράματος. Κάθε πείραμα χρειάζεται περισσότερες από 12 ώρες. Η συνολική διάρκεια των πειραμάτων (CPU time), με πληθυσμό 200 άτομα ανά υπο-πληθυσμό, ήταν γύρω στους 4 μήνες (2 ακόμη μήνες τα πειράματα των ΣΠΔΕΣ με υπο-πληθυσμούς 100 ατόμων).



Σχήμα 6.9. Μέσος αριθμός ωρών ανά πείραμα (ΑΤΝΔ)

6.2.2. Βελτιστοποίηση Συνάρτησης Rastrigin

Η βελτιστοποίηση της συνάρτησης Rastrigin, εκτός της σύγκρισης της απόδοσης των ΣΠΔΕΣ και του μοντέλου νησίδων, χρησιμοποιήθηκε επίσης και για τον εκτενή έλεγχο της επίδρασης διαφορετικών τιμών *εποχής* και διαφορετικών μεγεθών υπο-πληθυσμού (*ΜεΠ*) στην απόδοση των δύο μοντέλων. Να επαναλάβουμε ότι κάθε υπο-πληθυσμός του μοντέλου νησίδων έχει την συμπεριφορά του υπο-πληθυσμού 1 των ΣΠΔΕΣ. Η συνάρτηση αυτή, ενώ αποτελεί καθιερωμένη συνάρτηση ελέγχου της απόδοσης (benchmark) των εξελικτικών αλγόριθμων, απαιτεί λιγότερο χρόνο βελτιστοποίησης από το προηγούμενο πρόβλημα εκπαίδευσης ενός ΑΤΝΔ.

Εδώ, κάθε σειρά πειραμάτων αποτελείται από 4500 πειράματα (σύνολο 9000 πειράματα). Οι τιμές της *εποχής* μετανάστευσης που χρησιμοποιήθηκαν κυμαίνονται μεταξύ 1 και 500 γενεών. Επίσης δοκιμάστηκαν τρεις τιμές για το μέγεθος πληθυσμού: 50, 100 και 200 άτομα ανά υπο-πληθυσμό. Τα αποτελέσματα είναι οι μέσοι όροι από 100 εκτελέσεις για κάθε πείραμα.

Τα αποτελέσματα επιβεβαιώνουν την βελτίωση της απόδοσης των ΠΓΑ με την χρήση των ΣΠΔΕΣ. Για όλες τις τιμές *εποχής* και μεγέθους πληθυσμού που χρησιμοποιήθηκαν, η απόδοση των ΣΠΔΕΣ είναι καλύτερη από την απόδοση του μοντέλου νησίδων (σχ. 6.10).

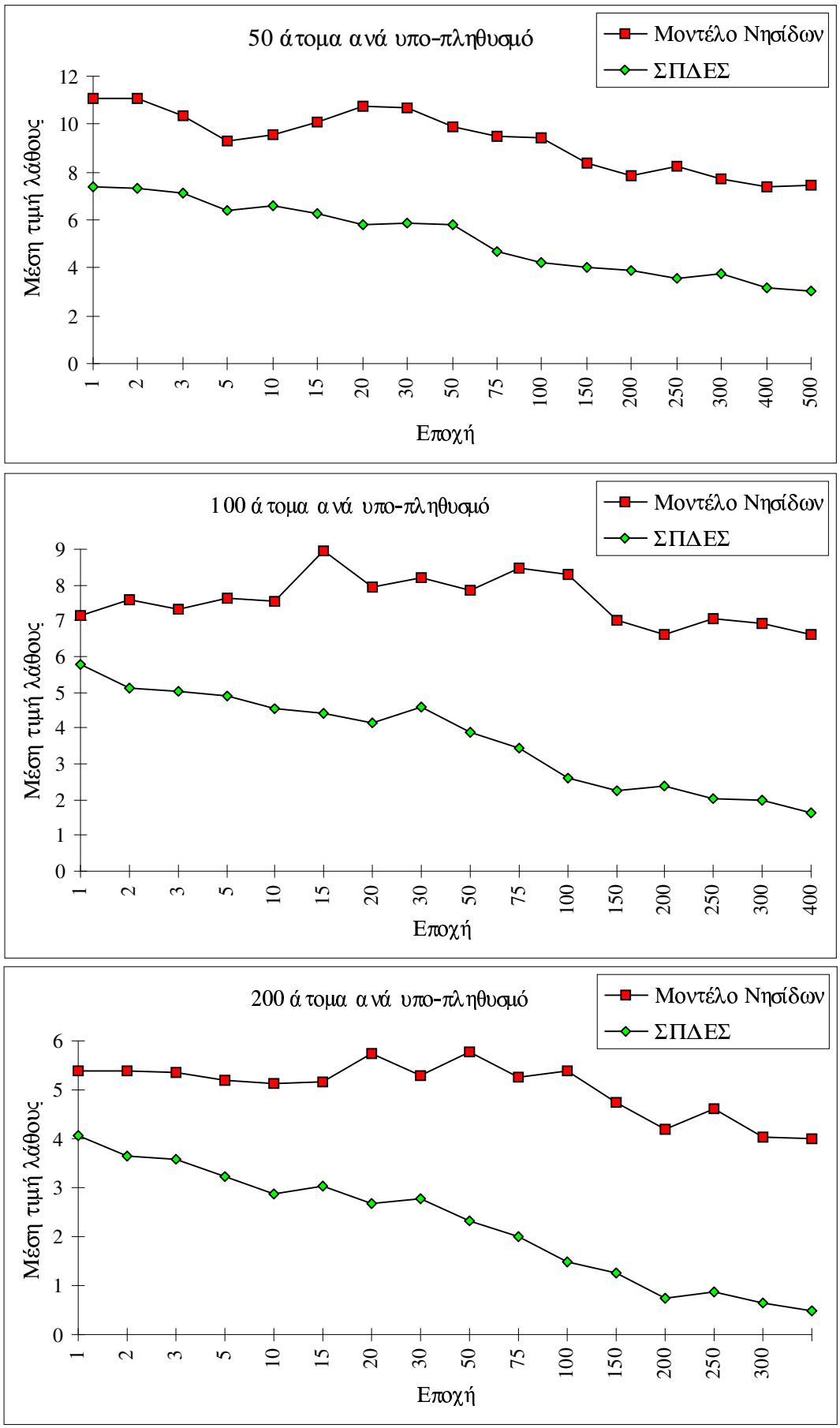
Είναι εμφανές ότι τόσο η απόδοση των ΣΠΔΕΣ όσο και του μοντέλου νησίδων, εξαρτάται από την τιμή της *εποχής* που χρησιμοποιείται. Η προσέγγιση του ελάχιστου βελτιώνεται καθώς αυξάνεται η τιμή *εποχής*, αφού η μέση τιμή του λάθους μειώνεται καθώς η τιμή *εποχής* αυξάνεται. Τα αποτελέσματα δείχνουν υψηλή γραμμική συσχέτιση μεταξύ των τιμών *εποχής* και των τιμών λάθους. Η τάση προσέγγισης της βέλτιστης τιμής καθώς αυξάνεται η τιμή *εποχής* είναι μεγαλύτερη με την χρήση των ΣΠΔΕΣ. Αυτό φαίνεται από την κλίση των ευθειών γραμμικής παλινδρόμησης (Πίνακας 6.2). Ο ρυθμός βελτίωσης των ΣΠΔΕΣ είναι μεγαλύτερος από τον ρυθμό του μοντέλου νησίδων.

Ο ρυθμός βελτίωσης της απόδοσης των ΣΠΔΕΣ καθώς αυξάνεται η τιμή *εποχής*, αυξάνεται επίσης καθώς αυξάνεται το μέγεθος πληθυσμού. Αυτό δεν συμβαίνει με το μοντέλο νησίδων.

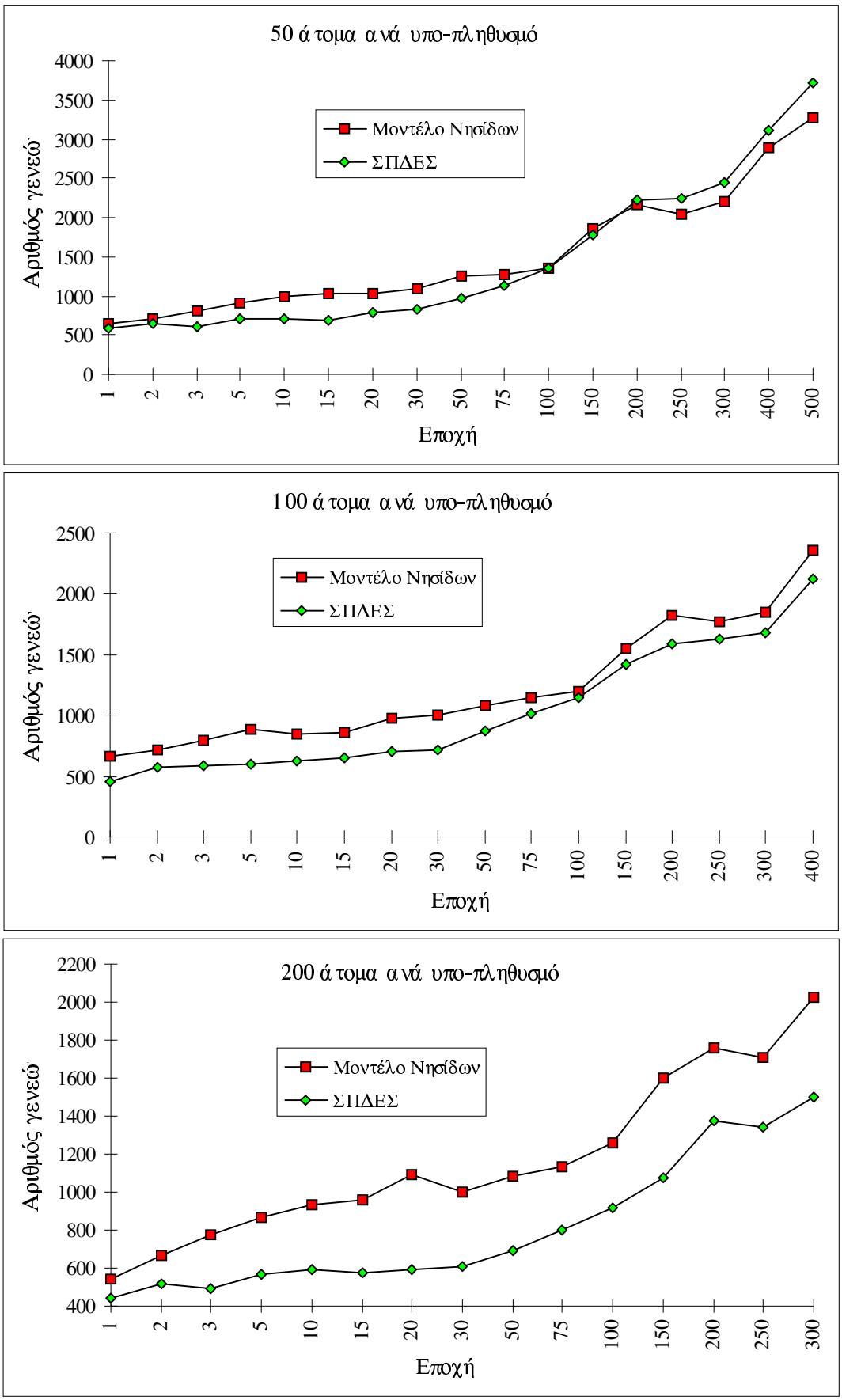
Πίνακας 6.2. Κλίση των ευθειών γραμμικής παλινδρόμησης για διάφορα μεγέθη υπο-πληθυσμού.

	50	100	200
ΣΠΔΕΣ	-0.00851	-0.00965	-0.0104
Μοντέλο νησίδων	-0.00727	-0.00336	-0.00433

Η παρατήρηση του προηγούμενου προβλήματος (εκπαίδευση ΑΤΝΔ), ότι ο αριθμός γενεών που χρειάζεται για να συγκλίνει ο αλγόριθμος αυξάνεται, καθώς η τιμή *εποχής* αυξάνεται, επιβεβαιώνεται και σε αυτό το πρόβλημα με πολύ περισσότερα πειραματικά αποτελέσματα (σχ. 6.11).

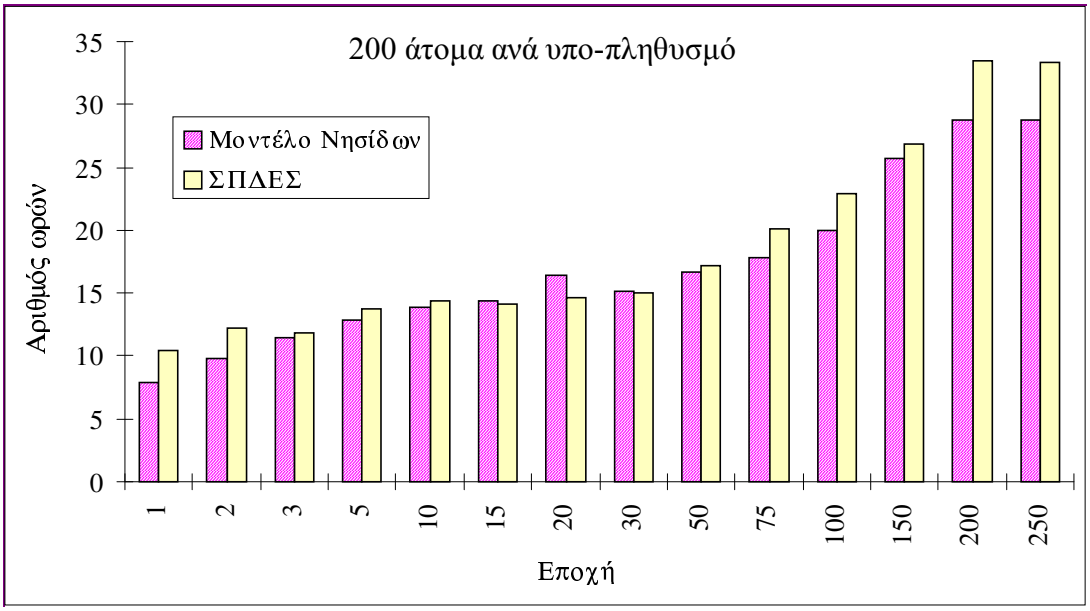


Σχήμα 6.10. Απόδοση των ΣΠΔΕΣ και του μοντέλου νησίδων για διάφορα μεγέθη υπο-πληθυσμών (Συνάρτηση Rastrigin)



Σχήμα 6.11. Μέσος αριθμός γενεών για διάφορα μεγέθη υπο-πληθυσμών (Συνάρτηση Rastrigin)

Στο σχήμα 6.12 δίνονται ενδεικτικά οι χρόνοι εκτέλεσης των πειραμάτων για μέγεθος υπο-πληθυσμού 200 ατόμων, για διάφορες τιμές εποχής.



Σχήμα 6.12. Μέσος αριθμός ωρών ανά πείραμα (Συνάρτηση Rastrigin)

Ένα άλλο θέμα έρευνας είναι η επίδραση του μεγέθους πληθυσμού στην απόδοση και των δύο μοντέλων (ΣΠΔΕΣ και μοντέλο νησίδων). Τα αποτελέσματά για το μοντέλο νησίδων συμφωνούν με την πρωτοποριακή δουλειά του Grosso [Gros85], δηλαδή. ο ρυθμός σύγκλισης είναι μεν ταχύτερος για μικρότερους πληθυσμούς, αλλά οι μικροί πληθυσμοί είναι πιο πιθανό να καθηλωθούν σε τοπικά ελάχιστα. Τα ίδια αποτελέσματα ισχύουν και για τους ΣΠΔΕΣ.

Τα σχήματα 6.13 και 6.14 δείχνουν την προσέγγιση του ελάχιστου, μέσω της βελτίωσης του μέσου λάθους (διαφορά από την ελάχιστη τιμή), για τρία διαφορετικά μεγέθη υπο-πληθυσμών, για το μοντέλο νησίδων και τους ΣΠΔΕΣ αντίστοιχα. Οι τιμές που χρησιμοποιούνται (50, 100 και 200 γενιές) επιλέχθηκαν να είναι μεγαλύτερες από την τιμή που αναφέρεται ως κρίσιμο μέγεθος πληθυσμού από τους τους Goldberg, Kargupta, Horn και Cantu-Paz[GKHC95].

Αρχικά, η απόδοση των μοντέλων με μέγεθος πληθυσμού 50 είναι καλύτερη από την απόδοση των άλλων μεγεθών πληθυσμού, αλλά συνήθως καθηλώνεται σε τοπικά ελάχιστα. Σύντομα η απόδοση με μέγεθος πληθυσμού 100 άτομα ανά υπο-πληθυσμό γίνεται καλύτερη, αλλά και αυτή ξεπερνιέται από το μέγεθος πληθυσμού 200. Αν και τα σχήματα 6.13 και 6.14 δείχνουν την απόδοση των μοντέλων μόνο για τιμές *εποχής* 1, 10 και 50, τα αποτελέσματα είναι τα ίδια και για τις υπόλοιπες τιμές *εποχής*.

Μπορούμε να συμπεράνουμε οι μικρές τιμές μεγέθους υπο-πληθυσμών είναι καλό να χρησιμοποιούνται όταν στόχος είναι η ταχεία βελτιστοποίηση και όχι τόσο η εύρεση του ολικού ελάχιστου. Αντιθέτα, η χρήση μεγάλων τιμών υπο-πληθυσμών βοηθά περισσότερο στην εύρεση του ολικού ελάχιστου, παρά στην ταχεία βελτιστοποίηση.