

**ΑΛΓΟΡΙΘΜΟΣ ΕΝΙΣΧΥΤΙΚΗΣ ΜΑΘΗΣΗΣ ΓΙΑ ΤΗ ΡΥΘΜΙΣΗ ΔΙΕΡΓΑΣΙΩΝ  
ΜΕ ΚΑΤΑΣΚΕΥΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΑΚΤΙΝΙΚΗΣ ΣΥΝΑΡΤΗΣΗΣ ΒΑΣΗΣ  
ΣΕ ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ**

**Παναγιώτης Πετσαγκουράκης<sup>1</sup>, Ευαγγελία Χονδροδήμα<sup>1,2</sup>, Αλέξανδρος  
Αλεξανδρίδης<sup>2</sup>, Χαράλαμπος Σαρίμβειος<sup>1</sup>**

<sup>1</sup>Σχολή Χημικών Μηχανικών, Εθνικό Μετσόβιο Πολυτεχνείο

<sup>2</sup>Τμήμα Ηλεκτρονικών Μηχανικών ΤΕ, Τεχνολογικό Εκπαιδευτικό Ίδρυμα Αθήνας

Η ενισχυτική μάθηση είναι μια οικογένεια τεχνικών στην επιστήμη των υπολογιστών, με κύριο χαρακτηριστικό το γεγονός ότι το σύστημα μαθαίνει από την άμεση αλληλεπίδραση με το περιβάλλον. Ειδικότερα οι μεθοδολογίες «ενέργειας-κριτικής» (actor-critic) έχουν ξεχωριστές και ανεξάρτητες δομές μνήμης για την αναπαράσταση της στρατηγικής των κινήσεων και της συνάρτησης αξίας. Η πρώτη χρησιμοποιείται για την επιλογή των ενεργειών και η δεύτερη για την αξιολόγηση των ενεργειών. Ανάμεσα στις πολλές εφαρμογές της, η ενισχυτική μάθηση έχει χρησιμοποιηθεί για το σχεδιασμό ρυθμιστών για μη γραμμικά συστήματα και διεργασίες. Στην παρούσα εργασία παρουσιάζεται πρωτότυπη μεθοδολογία ρύθμισης «ενέργειας-κριτικής», στην οποία νευρωνικά δίκτυα ακτινικής συνάρτησης βάσης (RBF neural networks) χρησιμοποιούνται για την κατασκευή και την ενημέρωση σε πραγματικό χρόνο τόσο του actor όσο και του critic μέρους του συστήματος ρύθμισης. Το πρώτο προσεγγίζει το μη γραμμικό κομμάτι του συστήματος και παράγει ένα μέρος του σήματος εισόδου, ενώ το δεύτερο προσεγγίζει τη συνάρτηση κόστους με στόχο να ανανεώσει τα βάρη του action νευρωνικού δικτύου. Η εκπαίδευση των δύο νευρωνικών δικτύων γίνεται με τη μέθοδο των ασαφών μέσων, που επιτρέπει (με βάση πληροφορίες που λαμβάνονται από το περιβάλλον) την πλήρη κατασκευή και αναπροσαρμογή των δικτύων σε πραγματικό χρόνο, συμπεριλαμβανομένης της ίδιας της δομής των δικτύων, δηλαδή του πλήθους των νευρώνων της κρυφής στοιβάδας, αλλά και των κέντρων των συναρτήσεων ενεργοποίησης. Αυτό είναι και το σημαντικό πλεονέκτημα της προτεινόμενης μεθοδολογίας σε σχέση με άλλες μεθόδους «ενέργειας-κριτικής» που υπάρχουν στη βιβλιογραφία, οι οποίες βασίζονται σε νευρωνικά δίκτυα προς τα εμπρός τροφοδότησης. Σε αυτές το πλήθος των νευρώνων και οι συνάψεις μεταξύ τους (με εξαίρεση τις συνδέσεις ανάμεσα στην κρυφή στοιβάδα και τη στοιβάδα εξόδου) επιλέγονται τυχαία, πριν ξεκινήσει η αλληλεπίδραση με το περιβάλλον, ενώ δε μεταβάλλονται κατά τη διαδικασία μάθησης. Η προτεινόμενη προσέγγιση εφαρμόστηκε σε δύο προβλήματα. Το πρώτο αποτελείται από ένα μη γραμμικό non-affine δυναμικό σύστημα, ενώ στο δεύτερο χρησιμοποιήθηκε ένα παράδειγμα μη γραμμικού αντιδραστήρα από το χώρο της χημικής μηχανικής.